

LEY DE ZIPF Y DE GIBRAT PARA COLOMBIA Y SUS REGIONES: 1835-2005*

GERSON JAVIER PÉREZ VALBUENA

Centro de Estudios Económicos Regionales-CEER, Banco de la República^a

ADOLFO MEISEL ROCA

Junta Directiva, Banco de la República

Zipf and Gibrat's Law for Colombia and her regions, 1835-2005

RESUMEN

En este documento se analiza la jerarquía urbana de Colombia y sus regiones utilizando información censal de entre 1835 y 2005. Se estudian tres aspectos: 1) la distribución del tamaño poblacional a través de las regularidades empíricas de Zipf y de Gibrat; 2) el cambio en el modelo de crecimiento poblacional, y 3) la validación empírica sobre la coincidencia de la dinámica poblacional para un país y sus regiones. Haciendo uso de la relación rango-tamaño y de técnicas no paramétricas, se encuentra coincidencia nacional y regional desde 1964 del cumplimiento de la ley de Zipf y parcialmente de la ley de Gibrat, lo cual evidencia un cambio en el modelo de crecimiento poblacional desde mediados del siglo XX en los ámbitos nacional y regional.

Palabras clave: distribución del tamaño poblacional, ley de Zipf, ley de Gibrat

* Received 17 September 2013. Accepted 24 February 2014. Se agradecen los comentarios y sugerencias de Jaime Bonet, Andrés Sánchez y Karina Acosta. De igual forma, se agradecen los comentarios y sugerencias de dos evaluadores anónimos a una versión preliminar de este documento. Simón Cháves y Sebastián Quintero también prestaron una valiosa asistencia en su elaboración.

^a Centro Calle 33 # 3-123, Plaza de Bolívar, Edificio Banco de la República. Comentarios y sugerencias a este documento son bienvenidos a los correos electrónicos gperezva@banrep.gov.co o ameisero@banrep.gov.co. El contenido de este documento es responsabilidad exclusiva del autor y no compromete al Banco de la República ni a su Junta Directiva.

ABSTRACT

By using census data from 1835 to 2005 this paper studies the urban hierarchy in Colombia and its regions. This document focuses on three issues: 1. the city size distribution by means of Zipf's law and Gibrat's law; 2. the shifts in the population growth models; 3. the empirical validation of the coincidence between national and regional population patterns. Using the adjusted rank-size relationship and non-parametric techniques, we find that city size distributions follow the Zipfian power law and that Gibrat's law holds at national level and partially at the regional level from 1964. These results are consistent with shifts in the population growth from mid 1950s at national and regional level.

Keywords: city size distribution, Zipf's law, Gibrat's law

JEL Classification: J11, R10

1. INTRODUCCIÓN

En el campo de la economía urbana, la dinámica demográfica y, en particular, la distribución del tamaño de las ciudades ha sido un tema de investigación que ha atraído mucho la atención durante las últimas décadas. La literatura se ha centrado en mostrar si, empíricamente, se cumplen las leyes de Zipf y de Gibrat. La primera establece que la distribución del tamaño de las ciudades sigue una distribución de Pareto; en la práctica, esto significa que la ciudad de mayor tamaño (Bogotá) debería ser dos veces más grande que la segunda (Medellín), tres veces más que la tercera (Cali), y así sucesivamente¹. La segunda, por su parte, especifica que la población de las ciudades crece aleatoriamente con la misma media y varianza, y que existe independencia entre el crecimiento poblacional y el tamaño de las ciudades; en otras palabras, el cumplimiento de la ley de Gibrat implica que las ciudades pequeñas crecen al mismo ritmo que las ciudades medianas o grandes.

La pregunta que surge es ¿por qué es importante, no solo en el contexto histórico sino urbano, corroborar esta regularidad empírica? La respuesta es que el estudio histórico de la dinámica poblacional permite establecer si las áreas urbanas, de una región o un país, tienen una senda de crecimiento común y si existe una única relación rango-tamaño entre ellas, lo que además permite caracterizar el desarrollo urbano y qué tanto se ve afectado por

¹ Zipf (1949) fue el primero en proponer formalmente que el tamaño poblacional se aproximaba a una distribución de Pareto. Sin embargo, en la literatura se menciona a Auerbach (1913) como uno de los primeros en notar esta regularidad empírica.

choques exógenos². Si la serie histórica del tamaño poblacional resulta ser estacionaria, es posible predecir que los choques externos son temporales y que este regresará nuevamente a su senda de largo plazo. El caso contrario ocurre si el crecimiento poblacional resulta ser aleatorio, pues los efectos ante choques exógenos serán permanentes o de largo plazo (Davis y Weinstein, 2002; Bosker *et al.*, 2007).

En términos de la caracterización del tipo de crecimiento de las ciudades, siguiendo a Schaffar y Dimou (2012), es posible distinguir dos teorías. La primera es la del crecimiento poblacional aleatorio, según la cual las ciudades crecen en forma estocástica y en el largo plazo o en el estado estacionario cumplen con las leyes de Zipf y de Gibrat. Según este tipo de crecimiento, las ciudades se caracterizan por tener mano de obra con libre movilidad y tecnologías con rendimientos constantes a escala. Los autores mencionan también que el proceso aleatorio está cercanamente ligado a choques exógenos generadores de *amenities* en forma de choques de políticas públicas o choques históricos o naturales. Duranton (2006, 2007), mediante la modelación microfundamentada de este tipo de crecimiento estocástico, mostró que el surgimiento y la desaparición de las ciudades pueden verse afectados por los choques de innovación que afrontan las firmas o las mismas ciudades.

La segunda es la teoría de crecimiento determinístico, según la cual el tamaño de las ciudades depende de las firmas y de sus decisiones de localización. De este modo, el crecimiento de las ciudades depende de aspectos como las externalidades, el capital humano y las decisiones relacionadas con la localización; en particular, según Schaffar y Dimou (2012, 709):

Las firmas se concentran geográficamente con el fin de aprovechar los efectos de la aglomeración, vinculados ya sea a la especialización [...] o a la diversificación [...]; pero sufren, por otro lado, deseconomías causadas por la congestión y los costos de desplazamiento [traducción propia]³.

Según esta teoría, las leyes de Zipf y de Gibrat no son de estricto cumplimiento⁴. El objetivo principal de este documento es caracterizar la dinámica poblacional en Colombia y en cada una de sus regiones a lo largo

² Esto, en la medida en que las regiones o los países donde se cumplen las leyes de Zipf y Gibrat están relacionadas con modelos de crecimiento urbano distintos a aquellos en los que esto no ocurre.

³ Schaffar y Dimou (2012, p. 709).

⁴ Duranton (2006, 2007) ofrece un mecanismo mediante el cual la aglomeración de firmas está relacionada con la ley de Zipf. En su argumento, el autor menciona la relación proporcional que existe entre la migración y la cantidad de bienes producidos en una ciudad o región, y la que hay entre la inversión en innovación y el número de firmas. En este escenario, las innovaciones pequeñas y discretas van a resultar, entonces, en crecimientos proporcionales en la población, lo cual genera una distribución de Pareto.

de los dos últimos siglos. En particular, se quiere explorar la hipótesis de crecimiento aleatorio en Colombia y sus regiones, y determinar si esta dinámica muestra persistencia a largo plazo. Además, se pretende establecer si, así como lo demostró teóricamente Gabaix (1999b), el cumplimiento individual de la ley de Zipf en cada región implica necesariamente observar esta regularidad empírica en el total nacional. Para el caso de Alemania, Giesen y Südekum (2011) demostraron empíricamente a nivel regional, por primera vez, lo planteado teóricamente por Gabaix (1999b). Para el caso de los Estados Unidos, a pesar de existir un amplio número de estudios, González-Val (2010) analizó recientemente por primera vez la distribución total del tamaño poblacional para todo el siglo XX. El autor encontró que para este período se cumple débilmente la ley de Gibrat y que el cumplimiento de la ley de Zipf está restringido a la parte superior de la distribución del tamaño poblacional.

Este documento hace aportes a la literatura en varias dimensiones. En primer lugar, esta es la primera vez que en Colombia se utiliza información a partir del siglo XIX en este tipo de análisis, toda vez que los documentos que se han escrito para Colombia cubren únicamente el siglo XX. La segunda dimensión es que se analiza empíricamente por primera vez para Colombia la ley de Gibrat mediante métodos no paramétricos, y se vincula directamente con los resultados de los estimadores de rango-tamaño. La tercera tiene que ver con el análisis regional de la dinámica demográfica, lo cual no se ha hecho hasta ahora para Colombia. A este respecto, se pretende establecer por primera vez para Colombia si el crecimiento poblacional sigue una dinámica según las predicciones teóricas de Gabaix (1999b), que establece que si individualmente las regiones cumplen la ley de Zipf, se cumple también para el total nacional.

La literatura se ha centrado en establecer la dinámica de la jerarquía urbana en los países desarrollados, pero poco se ha hecho en el caso de las economías emergentes. La excepción es, tal vez, el caso de la China, que ha sido objeto de varios estudios individuales (Song y Zhang, 2002; Zhou y Ma, 2003; Ye y Xie, 2012), Israel (Benguigui y Blumenfeld-Lieberthal, 2011), y parte de estudios de multi-país (Soo, 2005; Schaffar y Dimou, 2012). Otro ejemplo para países en desarrollo es el documento de Dimou y Schaffar (2009), quienes estudian la jerarquía y el crecimiento de las áreas urbanas en la península de los Balcanes⁵.

Para el caso colombiano, se destacan dos estudios que han abordado el análisis de la distribución del tamaño poblacional. Bernal y Nieto (2006), utilizando información censal de entre 1918 y 1993, encuentran que la regularidad empírica de Zipf no se cumple en ningún año, a pesar de que el

⁵ Se hace referencia también a los países ubicados en la península balcánica, en el sureste de Europa, que incluye, en parte o en su totalidad, a los siguientes países: Croacia, Eslovenia, Bosnia-Herzegovina, Rumanía, Montenegro, Serbia, Bulgaria, Albania, Macedonia, Turquía, Grecia e Italia.

coeficiente de rango-tamaño disminuye durante todo el siglo hasta ubicarse alrededor de 1 desde 1985. Por otro lado, Pérez (2006), haciendo uso de la información censal de entre 1912 y 1993, analizó la dinámica poblacional nacional y para la región Caribe. Los resultados encontrados muestran que la ley de Zipf no se cumple en la región Caribe para ninguno de los períodos analizados, y que para el total nacional existe evidencia de su cumplimiento solo a partir de 1985. Para el agregado nacional, estos resultados son consistentes con lo encontrado por Soo (2005) en su análisis multipaís, donde utiliza la información de Colombia correspondiente a 1993 y 1999 y, una vez se centra en las 16 principales aglomeraciones urbanas, encuentra un parámetro de Zipf significativamente igual a 1⁶. Los tres trabajos anteriores tienen, al menos, dos limitaciones: la primera es que no cubren en su totalidad la historia de la dinámica demográfica del país, ya que lo hacen únicamente con información del siglo XX; la segunda es que no realizan un estudio regional para todo el país, lo que limita el análisis de los resultados, en especial en un país con tan marcadas disparidades regionales.

Los resultados del presente documento muestran que, a partir de la segunda mitad del siglo XX, se presenta un cambio en la dinámica del crecimiento poblacional en Colombia. Hasta ese momento prevalecía lo que en la literatura se conoce como «modelo determinístico de crecimiento poblacional», según el cual el asentamiento de las firmas es un factor importante en las decisiones de localización de la población, al igual que la dotación y la distribución de los recursos naturales. A partir de ese momento, el país parece haber iniciado un proceso de transición a un modelo de crecimiento estocástico, como el que predicen las leyes de Zipf y de Gibrat, donde las ciudades crecen a tasas constantes, y que dicho crecimiento está ligado a choques exógenos que generan *amenities* y mejoran la calidad de vida de los habitantes⁷. Además, ejercicios regionales mostraron que estos resultados coinciden en su mayoría con lo encontrado para el total nacional, corroborando lo planteado teóricamente por Gabaix (1999b).

Este documento se divide en cuatro secciones adicionales a esta introducción. En el apartado 2 se realiza una descripción de las características teóricas y empíricas de las leyes de Zipf y de Gibrat y de la relación entre ellas. En el apartado 3 se presenta la descripción de los datos utilizados y las principales estadísticas descriptivas. En el apartado 4 se muestran y analizan

⁶ La estrategia regional en este documento se basa en la propuesta teórica de Gabaix (1999b) que, como se mencionó, fue comprobada empíricamente por primera vez para el caso de Alemania y sus regiones por Giesen y Südekum (2011). Teóricamente, el cumplimiento de la ley de Zipf en todas las regiones debería verse reflejado en el cumplimiento a nivel nacional, lo que implica, además, el cumplimiento de la ley de Gibrat.

⁷ El canal mediante el cual se relacionan el crecimiento determinístico de las ciudades y la calidad de vida proviene de la relación entre la aglomeración y el crecimiento de las ciudades. La aglomeración de firmas en ciudades o regiones específicas implica efectos positivos en el empleo e ingreso de la población.

los resultados tanto nacionales como de cada una de las regiones. La última sección, el apartado 5, corresponde a los comentarios finales.

2. CONTEXTO HISTÓRICO DE COLOMBIA

Segunda mitad del siglo XIX

Durante este periodo, la población en Colombia era eminentemente rural, dedicada a actividades agrícolas y pecuarias, tanto que, entre 1850 y 1870, no hay indicios sobre ningún proceso de urbanización en curso (Melo, 2007). Sin embargo, esto empieza a cambiar a finales del siglo XIX gracias a un mayor intercambio comercial con los países vecinos. Esto se dio principalmente en las ciudades fronterizas y costeras, así como en los principales centros de producción y en las regiones cafeteras localizadas en la parte nororiental del país. Sin embargo, estos puntos de inicio de urbanización se encontraban aislados unos de otros⁸.

En cuanto a la infraestructura vial, el país estaba conectado, fundamentalmente, por caminos de herradura y trochas. El principal medio por el cual se transportaban mercancías hacia el interior del país era el río Magdalena, el cual atraviesa el país de sur a norte hasta el mar Caribe. Como se puede suponer, el comercio afrontaba altos costos de transporte, impidiendo la especialización de las regiones. Hasta 1870 no se inicia el desarrollo ferroviario, dejando a un lado el interés por la construcción de carreteras.

Primera mitad del siglo XX

Las primeras décadas comienzan con una serie de hechos que cambian la estructura económica y social del país. El primero de ellos es una guerra civil, conocida como la «guerra de los Mil Días». Esta generó un segundo hecho: la relocalización de los cultivos de café de la parte nororiental del país a la región centro-occidental, que es donde actualmente se localiza la zona cafetera. Durante las dos primeras décadas del siglo, lo segundo significó un mayor desarrollo de la infraestructura ferroviaria, cuyo propósito fundamental fue sacar la producción cafetera a mercados externos. Además, se conectaron las zonas productoras con algunos de los principales centros de consumo nacional. Bejarano (2007) estima que, mientras que la red ferroviaria era de 600 km en 1900, 22 años más tarde era de 1.600 km, por los cuales se transportaba cerca del 90% de la producción cafetera.

El inicio de algunas actividades industriales impulsó una mayor urbanización, especialmente en Bogotá, Medellín, Cali y Barranquilla. Estos hechos,

⁸ Es importante mencionar que casi la mitad del territorio, en su parte suroriental, ha estado históricamente ocupada por selvas en las que, actualmente, solo habita cerca del 4% de la población.

junto con la inversión en infraestructura, fueron los principales determinantes de los procesos migratorios hacia las zonas urbanas durante estos años. Las siguientes dos décadas, 1930 y 1940, estuvieron en gran parte marcadas por la crisis financiera y la segunda guerra mundial. La primera, afectando a la economía exportadora de café; la segunda, con aumentos importantes de la inflación ante la intervención de los Estados Unidos en el conflicto.

Segunda mitad del siglo XX y comienzos del siglo XXI

A partir de la segunda mitad del siglo XX comienza la recuperación económica del país, que logra mantener tasas de crecimiento moderadas, pero constantes, de cerca del 5% de promedio anual. Además, se acelera el crecimiento poblacional, explicado por el desarrollo de la medicina y el mejoramiento en la calidad de vida (Ocampo *et al.*, 2007).

Durante este período, se reduce la participación de la actividad agropecuaria y se consolida la infraestructura vial y de servicios públicos, especialmente en las áreas de influencia de las zonas urbanas. Esto genera una mayor movilización poblacional de las áreas rurales a las urbanas. Además, se produce una recomposición regional en la que las ciudades industrializadas no fueron las únicas receptoras del desarrollo económico, lo cual impulsa el primer proceso de descentralización durante las décadas de 1960 y 1970.

El sector agropecuario inicia, también, un proceso de modernización, motivado en parte por la mayor demanda de insumos de la industria y por una demanda de productos del sector por parte de los centros urbanos. Las últimas décadas del siglo XX y los primeros años del siglo XXI se caracterizaron por un período de violencia e inseguridad. La coexistencia de guerrillas, paramilitares y traficantes de drogas, con mayor influencia en las áreas rurales, generaron aún mayores desplazamientos poblacionales hacia las áreas urbanas.

3. GENERALIDADES DE LAS LEYES DE ZIPF Y DE GIBRAT

La ley de Zipf establece que el tamaño poblacional de las ciudades es proporcional al inverso de la clasificación de la población cuando las ciudades se ordenan en forma descendente, según su tamaño, lo que implica que su relación se ajusta a la ley potencial con un exponente igual a -1. Eeckhout (2004) muestra que esta relación se puede ver desde otra perspectiva. Supongamos que se tiene un grupo de ciudades, cada una de tamaño s_i , ordenadas en forma descendente según su tamaño, y cada una tiene asignado un rango r_i . De este modo, se tiene que el tamaño s_i de cada ciudad será igual a $\frac{1}{r_i}$ veces el tamaño de la ciudad más grande C : $s_i = \frac{1}{r_i} C$.

Si esta relación se cumple, el gráfico que resultaría de la relación entre el tamaño poblacional y el rango, en logaritmos, sería una línea recta con pendiente negativa e igual a 1. El autor muestra también que la ley de Zipf es una representación de la distribución de Pareto. La ley potencial, o distribución de Pareto, toma la forma $r_i(s_i) = bs_i^{-a}$, la cual relaciona el tamaño poblacional s_i y el rango $r_i(\cdot)$. El cumplimiento de la ley de Zipf establece que, al estimar la versión linealizada en escala logarítmica de esta relación, se obtendrá un coeficiente para la pendiente $a = -1$:

$$\ln(r_i) = \ln(b) - a \ln(s_i) + \varepsilon_i,$$

relación que ha sido ampliamente utilizada en estudios empíricos de economía urbana⁹. Sin embargo, recientemente, Gabaix e Ibragimov (2011) demostraron que la estimación de esta relación llevaba a obtener estimadores sesgados de a en muestras pequeñas. Con el fin de solucionar este inconveniente, los autores recomiendan utilizar $r_i = \frac{1}{2}$ en vez de r_i .

$$\ln\left(r_i - \frac{1}{2}\right) = \ln(b) - a \ln(s_i) + \varepsilon_i.$$

Debido a que este cambio conlleva que los errores estándar del estimador de a no son los que usualmente se calculan por medio de mínimos cuadrados ordinarios, los autores recomiendan que sean transformados por $\left(\frac{2}{n}\right)^{\frac{1}{2}} \cdot |\hat{a}|$, donde n corresponde al número de ciudades utilizadas en la estimación.

Hasta el momento, parece que la ley de Zipf implica un análisis estático de la distribución del tamaño poblacional; sin embargo, es importante tener en cuenta que dicha ley corresponde a un equilibrio de largo plazo que implica necesariamente un análisis dinámico. Este hecho llevó a que algunos autores se centraran especialmente en el aspecto dinámico de las ciudades. En particular, Gibrat (1931) observó en un estudio de firmas que su crecimiento era independiente del tamaño y que esto implicaba que la distribución del tamaño de las firmas se aproximara a una distribución lognormal. De estos resultados surge lo que se conoce como la ley de Gibrat o ley de crecimiento proporcional, que para el caso de las ciudades establece que, independiente de su tamaño, el crecimiento es aleatorio con el mismo valor esperado y la misma varianza.

Estos hallazgos hicieron que algunos autores (Champernowne, 1953; Simon, 1955) empezaran a relacionar los resultados de la ley de Zipf con los de la ley de Gibrat. Lo que se encuentra es que existe una clara relación entre

⁹ La importancia de esta expresión es que permite estimar, de forma empírica, mediante un modelo de regresión lineal, el coeficiente a que permite probar, estadísticamente, si el tamaño poblacional sigue o no una distribución de Pareto y, por tanto, permite establecer si se cumple o no la ley de Zipf.

la tasa de crecimiento de las ciudades y la distribución de Pareto que surge en forma natural si las series de tiempo de los crecimientos poblacionales cumplen la ley de Gibrat¹⁰. Al respecto, Gabaix (1999a) argumenta que el cumplimiento de la ley de Zipf debería ser un prerrequisito para el planteamiento de un modelo de crecimiento local. El autor plantea dos caminos por los cuales se cumpliría la ley de Zipf.

El primero es que, si bien las ciudades medias y grandes son diferentes en muchos aspectos en comparación con las más pequeñas, aún existen externalidades que afectan a todas de la misma forma; es decir, que las ciudades en realidad no son tan diferentes. El segundo es que las externalidades que afectan a las ciudades son importantes en su crecimiento y, por tanto, logra diferenciarlas en su dinámica poblacional; sin embargo, la ventaja en términos de productividad de las grandes ciudades se desvanece cuando se tienen en cuenta las externalidades negativas asociadas (tráfico, polución, criminalidad), que las pone al nivel de las ciudades pequeñas.

Levy (2009) participa en la discusión mencionando que interesa analizar el comportamiento de la cola superior de la distribución del tamaño poblacional, ya que abarca una parte importante de ella y, en general, representa una alta proporción de la población total del país. Argumenta que si la cola superior de la distribución se aproxima a una distribución lognormal, esto no significa que al mismo tiempo no pueda aproximarse a una distribución de Pareto, ya que una cola superior lognormal en el límite tiende a convertirse en una distribución de Pareto¹¹.

Seguendo a Ioannides y Overman (2003) y a Eeckhout (2004), se tiene la siguiente expresión para la tasa de crecimiento de las ciudades g_i :

$$g_i = m(s_i) + \varepsilon_i,$$

donde $m(s_i) = E[g|s_i]$. Para este propósito, la tasa de crecimiento de las ciudades está normalizada, es decir, que al crecimiento de la población de cada ciudad i en un año determinado t se le resta la media y se divide por la desviación estándar del crecimiento del grupo de referencia correspondiente. El término $m(s_i)$ indica que la tasa de crecimiento de las ciudades está en función de su tamaño (en logaritmos). Con el fin de no restringir la forma funcional en m y ofrecer mayor flexibilidad, se calculan las medias y varianzas condicionales de las tasa de crecimiento de las ciudades en forma no paramétrica. Para ello se utiliza la técnica Nadaraya-Watson

¹⁰ En este caso, si el crecimiento de todas las ciudades es proporcional, tal y como predice la ley de Gibrat, la línea recta con pendiente -1 que predice la ley de Zipf debería desplazarse en el tiempo de forma paralela (Goerlich y Mas, 2010).

¹¹ Estos argumentos llevan a que el autor afirme que el poder de una prueba de hipótesis es bajo cuando intenta distinguir entre una distribución lognormal y una de Pareto cuando se analiza únicamente la parte superior de la cola de la distribución del tamaño poblacional. Esto hace parte, además, de la respuesta de Eeckhout (2009) al artículo de Levy (2009).

(Nadaraya, 1964; Watson, 1964), cuyos estimadores para la media y varianza condicionales están dados por¹²:

$$\widehat{m}_h(s) = \frac{n^{-1} \sum_{i=1}^n K_h(s-s_i)g_i}{n^{-1} \sum_{i=1}^n K_h(s-s_i)},$$

$$\widehat{\sigma}_h^2(s) = \frac{n^{-1} \sum_{i=1}^n K_h(s-s_i)(g_i - \widehat{m}(s))^2}{n^{-1} \sum_{i=1}^n K_h(s-s_i)};$$

donde K_h se refiere al kernel (Epanechnikov), cuyo ancho de banda h es igual a $0,5^{13}$. En términos del análisis regional de las regularidades empíricas descritas, Gabaix (1999b) demostró teóricamente que si en un país compuesto por regiones heterogéneas cada una de ellas sigue individualmente lo planteado por la ley de Zipf, entonces esta se cumplirá también para el total nacional.

4. DATOS Y ESTADÍSTICAS DESCRIPTIVAS

La fuente principal de los datos son los censos poblaciones correspondientes a los años 1835, 1843, 1851, 1870, 1905, 1912, 1918, 1938, 1951, 1964, 1973, 1985, 1993 y 2005¹⁴. El censo de 1928 no se incluye en el análisis, ya que la información obtenida no contó con la aprobación oficial de las autoridades correspondientes¹⁵. Es importante observar que la periodicidad

¹² Härdle (1990) presenta una descripción detallada del cálculo de los estimadores no paramétricos de la media y la varianza.

¹³ Con el fin de corroborar la robustez de los resultados ante diferentes anchos de banda, se realizaron cálculos adicionales utilizando el criterio de ancho de banda óptimo de Silverman (1986) y se obtuvieron resultados similares. Giesen y Südekum (2011), en un ejercicio para Alemania, encontraron también resultados similares al utilizar estos dos criterios de ancho de banda.

¹⁴ En el censo de 1905, la población de los departamentos de la región Caribe, exceptuando Atlántico, La Guajira y Magdalena, y algunos otros municipios individuales (Tamalameque, Pueblo Nuevo y Valledupar), se encontraba agregada por provincias. Para obtener una aproximación de la población de los municipios faltantes, se calculó la población tomando como referencia la población de los municipios correspondientes del censo de 1912 y proyectándola hacia atrás por medio de la tasa de crecimiento intercensal. Teóricamente, puede existir la posibilidad de contabilizar dos veces la población de algunos municipios que pudieron crearse entre 1905 y 1912. Sin embargo, consideramos que esto no conlleva un problema mayor, por dos razones. La primera es que la dinámica de creación de municipios no se da de forma masiva. En segundo lugar, esto se hizo solo para un número reducido de municipios de una región en particular, pero no para todo el país.

¹⁵ Los datos se obtuvieron, en todos los casos, de información censal oficial tomada del Archivo General de la Nación y del Departamento Nacional de Estadística (DANE).

de los censos es heterogénea, especialmente durante los períodos 1870-1905 y 1918-1938.

En el caso de las unidades geográficas, se consideran las definidas administrativamente en la Constitución como municipio o entidad fundamental de la división político-administrativa del Estado¹⁶. Es importante aclarar que en Colombia, especialmente en el caso de las grandes ciudades, las ampliaciones de frontera de estos municipios se han dado por segregación o fusión con otros centros poblacionales. Los efectos potenciales de esta situación pueden ser tasas de crecimiento diferenciales (muy altas o muy bajas) durante el período en el que ocurrió la segregación. Sin embargo, los efectos son menores y no afectan a las estimaciones para los casos en los que el número de segregaciones ocurridas sea pequeño.

A efectos del presente documento, las regiones se han definido de la siguiente forma: Caribe (La Guajira, Magdalena, Atlántico, Bolívar, Cesar, Córdoba y Sucre); Central (Antioquia, Caldas, Caquetá, Huila, Quindío, Risaralda y Tolima); Oriental (Boyacá, Cundinamarca, Meta, Norte de Santander, Santander); Pacífico (Cauca, Chocó, Nariño y Valle); Nuevos Departamentos (Amazonas, Arauca, Casanare, Guainía, Guaviare, Putumayo, Vaupés y Vichada).

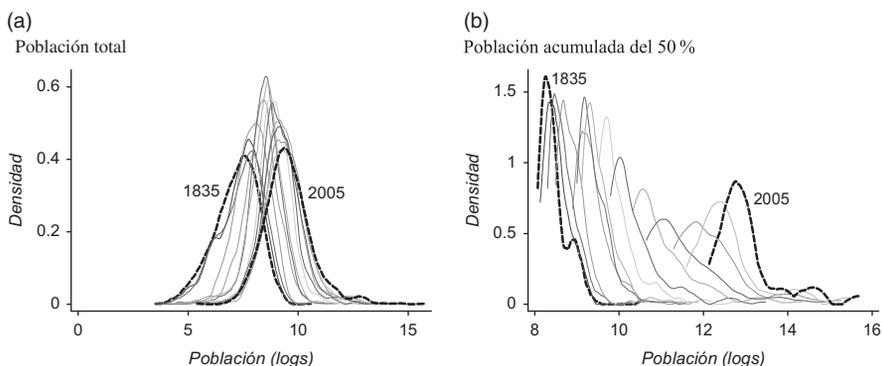
Desde el inicio del estudio de la dinámica poblacional mediante la distribución de su tamaño, ha existido controversia sobre la definición de ciudad o sobre qué nivel de agregación o qué unidad geográfica se debería considerar. Dentro de las posibilidades está considerar aquellas unidades administrativas con más de 100.000 habitantes (Rosen y Resnick, 1980; Chesire, 1999; Soo, 2005; Giesen y Südekum, 2011) o las definidas como áreas metropolitanas como tal (Dobkins y Ioannides, 2001; Ioannides y Overman, 2003), o tomar aquellas unidades administrativas en la parte superior de la distribución del tamaño. Algunos autores han tomado la distribución completa del tamaño poblacional (Anderson y Ge, 2005; Nota y Song, 2006; González-Val, 2010). Al respecto se presentaron algunas críticas tales como que así se desvanece la distinción entre lo urbano y lo rural. Además, esto hace aún más compleja la aproximación al término «ciudad» en estudios cuyo principal objetivo es analizar la dinámica poblacional urbana¹⁷. Estos argumentos estarían mostrando que, si se tiene en cuenta el proceso de urbanización en el mundo moderno, sería razonable establecer cierto umbral del tamaño poblacional que describiría mejor la jerarquía urbana (Schaffar y Dimou, 2012)¹⁸. En términos

¹⁶ Otras unidades, frecuentemente utilizadas en este tipo de análisis, son las áreas metropolitanas; sin embargo, para el caso colombiano, no se consideran debido a que actualmente solo son diez. Además, en varios de estos casos, en la práctica no funcionan administrativa o territorialmente como tales.

¹⁷ Sin embargo, estudios recientes han argumentado que trabajar con una distribución poblacional truncada podría generar sesgos en la estimación de los coeficientes de Pareto (Giesen *et al.*, 2010; Ioannides y Skouras, 2013; González-Val *et al.*, 2013).

¹⁸ En Colombia, en el año 2012, el 76% de la población vivía en las áreas urbanas y se espera que en 2050 este porcentaje haya aumentado hasta cerca del 85%.

GRÁFICO 1
DENSIDADES KERNEL PARA EL TAMAÑO DE LAS CIUDADES COLOMBIANAS,
1835-2005



Fuentes: Calculado a partir de información del Archivo General de la Nación y del DANE.

de los resultados que se obtendrían al analizar una distribución con truncamiento izquierdo, Giesen y Südekum (2011) mencionan:

Es bien sabido que el truncamiento izquierdo no cambia las propiedades de la ley potencial de esta distribución; por ejemplo, si las cien ciudades más grandes [...] siguen la ley de Zipf de la potencia, también deberían hacerlo las cincuenta más grandes [traducción propia]¹⁹.

Con el fin de aproximar el concepto de «área urbana», en este documento se consideran como punto de comparación para cada uno de los años censales los municipios de mayor tamaño que acumulen el 50 % de la población. Esta definición se utiliza de igual forma para seleccionar los municipios en cada una de las regiones definidas anteriormente²⁰. Para dar una primera idea de la dinámica de este grupo poblacional, el Gráfico 1 presenta la evolución de la distribución del tamaño de los municipios. En el panel *a* se puede observar la dinámica para la población total, es decir, sin restringirla a ningún criterio de selección. Como se mencionó, esto implica que en cada uno de los años censales se considera al total de la población (urbana y rural). Por otro lado, el panel *b* muestra el cambio en el tiempo de las

¹⁹ Giesen y Südekum (2011, p. 682).

²⁰ Con el fin de establecer la robustez de los resultados, se utilizaron dos criterios alternativos. Aquellos que incluyeran los municipios dentro de los percentiles 90 y 95 de la distribución poblacional. Los resultados encontrados fueron similares en cuanto a la magnitud y significancia de los parámetros.

distribuciones del tamaño poblacional nacional cuando se aplica el criterio de selección del 50% de población acumulada.

En dicho gráfico se pueden observar algunas características interesantes. Si se consideraran únicamente el año inicial (1835) y el final (2005), podría afirmarse que la población total del país evolucionó de modo que hubo solo un desplazamiento de la distribución, producto del crecimiento natural de la población, donde la única diferencia evidente, a parte de la media y de un aumento pequeño en la concentración alrededor de esta, es la existencia de algunas ciudades de gran tamaño (Bogotá, Medellín, Cali, Barranquilla y Cartagena), como lo muestra la cola derecha de la distribución en el panel *a*. Sin embargo, es mucho más interesante la evolución en el tiempo, donde la distribución se desplaza a la derecha mientras que aumenta la concentración hasta alcanzar su máximo en 1912. En los años posteriores, la distribución continúa desplazándose a la derecha, a la vez que la concentración disminuye paulatinamente, hasta que en 2005 alcanza niveles similares a los de 1835. La razón de este comportamiento parece estar fundamentada en la entrada de nuevos municipios a la distribución poblacional, lo que parece corroborarse cuando se observan los municipios de la cola superior de la distribución (panel *b*). En este caso, se espera que sea menor la entrada de ciudades de mayor tamaño, en comparación con las ciudades del resto de la distribución y, por esta razón, la concentración se reduce durante todo el período.

Para el grupo de ciudades de mayor tamaño (panel *b*), es posible observar una mayor dinámica relativa a lo largo de los casi doscientos años. Al comparar los dos años extremos, se puede observar que en 1835 la población no solamente era de menor tamaño, sino que estaba más concentrada en ciudades de tamaño medio (con una media de 5.186 habitantes y una desviación de 3.321). A lo largo del tiempo, la distribución de la población urbana cambió su dinámica a una donde se destaca la preponderancia de los grandes centros urbanos. El Cuadro 1 y el Cuadro 2 resumen la información de la población seleccionada con el criterio que se va a utilizar en este documento: para el total nacional y para cada una de las regiones²¹.

Tanto en el nivel nacional como en el regional es posible observar la dinámica de primacía de las grandes ciudades, que coincide con los resultados descritos y que es evidente con el decreciente número de ciudades que acumulan el 50% de la población, con excepción de la región Central. Este patrón de comportamiento urbano no puede explicarse por un factor en particular, sino por un conjunto de ellos, tales como los históricos, geográficos, institucionales y políticos, entre otros. De todos ellos depende la dinámica de urbanización, que puede ser del tipo de las megaciudades,

²¹ En el apéndice, los cuadros A-1 and A-2 presentan las estadísticas descriptivas para el caso del total de municipios en el total nacional y para cada una de las regiones, respectivamente.

CUADRO 1
ESTADÍSTICAS DESCRIPTIVAS DE LAS CIUDADES EN COLOMBIA, 1835-2005
(MUNICIPIOS QUE ACUMULAN EL 50% DE LA POBLACIÓN)

AÑO	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	149	5.186	3.321	4.376	3.199	39.442	0,213
1843	161	5.486	3.201	4.782	3.402	40.086	0,196
1851	166	6.309	2.772	5.446	4.034	29.649	0,189
1870	175	7.625	3.701	6.823	4.878	40.883	0,185
1905	163	13.094	9.298	10.858	7.304	100.000	0,254
1912	180	13.800	10.700	10.935	8.110	121.257	0,248
1918	178	15.932	13.050	12.757	9.032	143.994	0,262
1938	161	26.969	33.202	19.110	13.722	355.502	0,350
1951	124	46.487	77.596	26.225	17.975	715.250	0,460
1964	82	106.461	214.350	45.060	30.637	1.697.311	0,560
1973	60	190.391	405.843	75.219	42.131	2.861.913	0,606
1985	42	357.279	687.004	167.832	79.893	4.236.490	0,581
1993	33	500.291	881.159	248.525	109.115	4.945.448	0,554
2005	28	762.797	1.292.902	380.061	187.249	6.840.116	0,535

Fuentes: Calculado a partir de información del Archivo General de la Nación y del DANE.

donde un número reducido de centros urbanos son los que acumulan el mayor porcentaje de la población; por ejemplo, Tokio representa cerca del 30% de la población de su país, Seúl, por su parte, cerca del 50%, y para el caso latinoamericano está Buenos Aires, con cerca del 30% de la población argentina. En Colombia se observa un proceso tendente al desarrollo y crecimiento de las grandes urbes, con Bogotá representando poco menos del 20% de la población.

El efecto de las grandes ciudades en la distribución poblacional puede verse también cuando se observa una media poblacional que es creciente, pero cada vez mucho mayor que la mediana. Al calcular el indicador de primacía nacional, se observa un aumento, aunque no muy considerable, al pasar de 1,22 en 1835 a 1,25 en 2005²². Sin embargo, cuando este indicador se calcula para cada región, se observa que la primacía de las grandes

²² Medida útil para determinar el grado de urbanización de las principales ciudades, en el cual se relaciona el tamaño poblacional de la ciudad más habitada con las menos pobladas,

CUADRO 2
ESTADÍSTICAS DESCRIPTIVAS DE LAS REGIONES DE COLOMBIA, 1835-2005
(MUNICIPIOS QUE ACUMULAN EL 50% DE LA POBLACIÓN)

REGIÓN CARIBE							
AÑO CENSAL	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	58	2.711	1.879	2.132	1.240	11.929	0,312
1843	61	2.713	1.701	1.996	1.234	10.145	0,299
1851	61	2.905	1.725	2.300	1.278	9.896	0,294
1870	52	4.357	2.140	3.416	2.383	11.595	0,245
1905	54	6.823	6.062	4.916	3.138	40.115	0,342
1912	49	10.507	7.938	7.798	5.311	48.907	0,313
1918	45	12.976	10.922	10.001	5.797	64.543	0,328
1938	40	24.989	26.051	15.631	10.111	152.348	0,396
1951	39	34.567	46.277	19.064	12.713	283.238	0,437
1964	40	56.854	82.268	31.685	20.834	498.301	0,464
1973	47	68.796	109.939	34.125	24.086	703.488	0,495
1985	43	101.955	158.681	50.298	32.142	927.233	0,512
1993	40	119.867	181.150	53.968	36.474	993.759	0,526
2005	40	158.421	228.043	70.611	42.542	1.146.498	0,533

CUADRO 2 (Cont.)

REGIÓN ORIENTAL							
AÑO CENSAL	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	124	4.590	3.493	3.894	2.571	39.442	0,251
1843	128	4.883	3.552	4.279	2.794	40.086	0,234
1851	131	5.703	3.044	5.022	3.172	29.649	0,224
1870	140	6.553	3.750	5.842	3.709	40.883	0,223
1905	141	8.524	8.391	6.639	4.840	100.000	0,266
1912	146	9.057	9.847	7.435	5.011	121.257	0,260
1918	144	10.320	11.833	8.368	5.672	143.994	0,268
1938	136	14.835	30.202	10.516	6.896	355.502	0,360
1951	119	21.030	65.479	11.958	8.026	715.250	0,477
1964	85	45.215	183.941	16.467	11.677	1.697.311	0,639
1973	60	85.360	368.247	22.516	15.436	2.861.913	0,734
1985	37	186.449	689.286	41.464	23.874	4.236.490	0,760
1993	20	382.818	1.080.676	89.766	45.696	4.945.448	0,732
2005	14	715.672	1.770.276	172.068	107.417	6.840.116	0,714

CUADRO 2 (Cont.)

REGIÓN CENTRAL							
AÑO CENSAL	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	51	3.994	1.888	3.390	2.261	10.280	0,237
1843	59	4.241	1.775	3.494	2.509	9.118	0,217
1851	64	4.869	2.156	4.106	2.738	13.755	0,223
1870	67	6.557	3.485	5.538	3.872	29.765	0,211
1905	76	12.148	6.983	10.224	6.312	53.936	0,252
1912	81	13.063	8.420	10.586	6.834	71.004	0,257
1918	86	14.262	9.439	11.586	7.122	79.146	0,269
1938	88	21.809	19.633	16.193	11.603	168.266	0,298
1951	82	31.285	42.221	20.516	14.429	363.865	0,375
1964	78	47.446	90.933	25.695	17.831	772.887	0,476
1973	79	56.543	134.003	27.940	19.042	1.163.868	0,526
1985	74	76.487	177.522	32.298	24.264	1.480.382	0,556
1993	65	93.312	209.432	38.660	27.038	1.630.009	0,581
2005	53	143.159	310.588	53.507	2.214.494	2.214.494	0,597

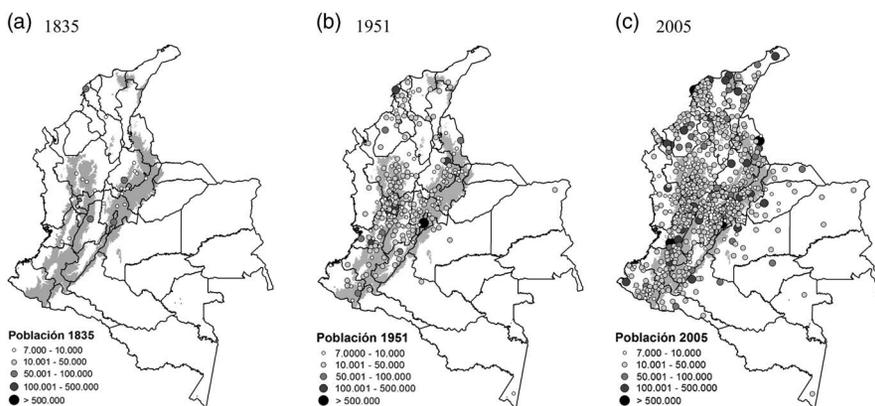
CUADRO 2 (Cont.)

REGIÓN PACÍFICA							
AÑO CENSAL	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	41	3.445	1.443	3.027	1.987	8.173	0,212
1843	43	4.204	1.790	3.709	2.392	10.376	0,209
1851	45	4.664	2.034	4.034	2.754	11.848	0,218
1870	54	5.151	2.499	4.324	2.543	12.743	0,248
1905	43	11.234	6.825	9.500	5.695	30.835	0,291
1912	53	10.301	5.276	8.918	5.659	27.760	0,248
1918	55	11.637	7.112	9.109	6.114	45.525	0,276
1938	56	19.202	14.226	14.084	10.294	101.883	0,291
1951	48	32.204	40.897	19.093	14.844	284.186	0,400
1964	44	51.271	94.717	25.500	17.836	637.929	0,499
1973	39	75.877	155.686	35.401	22.007	991.549	0,541
1985	30	122.421	254.493	48.511	30.803	1.429.026	0,583
1993	32	129.778	288.482	50.903	28.978	1.666.468	0,603
2005	30	172.622	379.397	63.131	34.710	2.119.843	0,620

Fuentes: Calculado a partir de información del Archivo General de la Nación y del DANE.

Nota: las regiones se definieron de la siguiente forma: Caribe (La Guajira, Magdalena, Atlántico, Bolívar, Cesar, Córdoba y Sucre); Central (Antioquia, Caldas, Caquetá, Huila, Quindío, Risaralda y Tolima); Oriental (Boyacá, Cundinamarca, Meta, Norte de Santander, Santander); Pacífico (Cauca, Chocó, Nariño y Valle); Nuevos Departamentos (Amazonas, Arauca, Casanare, Guainía, Guaviare, Putumayo, Vaupés y Vichada).

FIGURA 2
DISTRIBUCIÓN POBLACIONAL EN LOS MUNICIPIOS DE COLOMBIA



Fuentes: Calculado a partir de información del Archivo General de la Nación y del DANE.

ciudades ha ido aumentando de forma considerable. La única excepción es la región Caribe, donde el indicador se mantuvo relativamente estable, al pasar de 0,59 a 0,65 entre 1835 y 2005. Las demás regiones se han caracterizado, cada vez más, por seguir una redistribución poblacional concentradora relativa en favor de las principales ciudades, si se comparan los años 1835 y 2005, respectivamente: región Central (0,42 a 1,68), región Oriental (1,37 a 4,54) y región Pacífica (0,43 a 2,13). Estos resultados se corroboran al calcular el índice de Gini (últimas columnas en los cuadros 1 y 2)²³. Es posible observar que, tanto para el agregado nacional como para cada una de las regiones, la desigualdad ha venido aumentando; en otras palabras: los municipios de mayor tamaño se han ido alejando cada vez de los de menor tamaño en términos de su población.

El

FIGURA 2 muestra los mapas con la distribución espacial de los municipios de más de 7.000 habitantes en Colombia entre los siglos XIX y XX²⁴.

(Fnote continued)

generalmente las tres siguientes. Se calcula como el cociente entre la población de mayor tamaño y la sumatoria de las tres ciudades siguientes.

²³ El índice de Gini es uno de los más utilizados para medir la desigualdad, generalmente para el caso del ingreso. Este índice toma valores entre 0 y 1, donde la desigualdad aumenta a medida que aumenta el indicador.

²⁴ Se toma un tamaño poblacional arbitrario con el fin de facilitar la visualización del cambio en el número y tamaño de las ciudades en el tiempo.

Los resultados dejan ver al menos tres características interesantes. Las dos más evidentes tienen que ver con el incremento en el tiempo del número de ciudades y el correspondiente aumento de tamaño. Como anota González-Val (2010), estas son las características recurrentes de los países jóvenes, incluso en el caso de los Estados Unidos, en comparación con los países europeos, los cuales parecen haber alcanzado desde hace varias décadas su senda de largo plazo, no solo en cuanto al número de ciudades sino en lo referente a su tamaño poblacional.

La tercera característica que destaca en el desarrollo demográfico colombiano es la ubicación espacial de los principales centros urbanos. Se puede observar que los factores geográficos han sido determinantes para conformar los asentamientos poblacionales, en particular a lo largo de la cordillera de los Andes, que atraviesa toda Suramérica y se divide en tres cadenas montañosas en el territorio colombiano. Históricamente, los principales centros urbanos han estado concentrados alrededor de las tres cordilleras y cerca de los puertos marítimos, especialmente los del mar Caribe. Estos hallazgos son consistentes con lo mencionado: las características geográficas son importantes a la hora de explicar no solo la distribución de los centros urbanos sino su evolución en el tiempo.

Para entender mejor la distribución poblacional en Colombia y su evolución, es importante mencionar su posible relación con la distribución de los recursos naturales y, en general, con las características geográficas. Colombia tiene acceso al mar Caribe en su parte noroccidental y al océano Pacífico en su parte occidental. Además, parte de la riqueza hídrica del país está representada por los ríos Magdalena y Cauca, que lo atraviesan de sur a norte entre los valles de las cordilleras. Estos cuerpos de agua determinaron la evolución de la distribución poblacional, no solo por ofrecer recursos hídricos para las actividades agropecuarias, sino porque durante varias décadas fueron el principal medio de transporte de carga desde el interior del país hasta los puertos exportadores en el mar Caribe. Durante muchos años, estos factores resultaron atractivos para las empresas, las cuales se localizaron cerca de estas fuentes hídricas porque facilitaban la movilización de sus productos hacia los mercados externos. Así es como las características geográficas y la dotación de recursos naturales de Colombia han resultado determinantes en la distribución del tamaño de sus ciudades.

5. RESULTADOS

Desde el punto de vista metodológico, en este documento se analizan las características de la relación rango-tamaño nacional y de cada una de las regiones, y se establece si estas cumplen con lo planteado por la ley de Zipf. Posteriormente, se analiza en detalle la dinámica de la distribución poblacional por decenios, con el fin de caracterizar el crecimiento urbano

CUADRO 3
COEFICIENTES DE PARETO PARA LA DISTRIBUCIÓN POBLACIONAL: TOTAL NACIONAL

AÑO CENSAL	NÚMERO DE CIUDADES	EXPONENTE DE PARETO	DESVIACIÓN ESTÁNDAR	ESTADÍSTICO $t_{H_0: \beta = 1}$
1835	149	2,767	0,321	5,511 (**)
1843	161	2,946	0,328	5,926 (**)
1851	166	3,074	0,337	6,147 (**)
1870	175	3,104	0,332	6,340 (**)
1905	163	2,397	0,266	5,262 (**)
1912	180	2,479	0,261	5,659 (**)
1918	178	2,363	0,250	5,441 (**)
1938	161	1,878	0,209	4,194 (**)
1951	124	1,488	0,189	2,583 (**)
1964	82	1,217	0,190	1,140
1973	60	1,092	0,199	0,463
1985	42	1,104	0,241	0,433
1993	33	1,133	0,279	0,476
2005	28	1,168	0,312	0,539

Nota: La hipótesis nula se rechaza al 5% (*) y al 1% (**)

Fuente: Calculado a partir de la información censal de los años correspondientes.

(ley de Gibrat) y la dinámica de crecimiento de las ciudades en términos agregados para Colombia y en cada una de sus regiones.

Total nacional

El Cuadro 3 muestra los resultados de los coeficientes de la relación rango-tamaño (o Pareto) para los censos poblacionales correspondientes al período 1835-2005²⁵. En todos los casos las correcciones de sesgo por muestra pequeña sobre el modelo MCO, sugeridos por Gabaix e Ibragimov (2011), fueron implementados en el coeficiente y en los errores estándar.

Los resultados muestran evidencias del cumplimiento de la ley de Zipf únicamente desde la segunda mitad del siglo XX; en concreto, desde 1964.

²⁵ El GRÁFICO A-1 muestra además los gráficos de Zipf de 1835 y de 2005, en los que se relacionan el logaritmo del rango y el logaritmo del tamaño poblacional.

Aunque el coeficiente ha disminuido históricamente, a partir de dicho año parece haber un patrón de estabilización alrededor del valor óptimo. Es posible observar que para los casos en los que no se rechaza estadísticamente la hipótesis nula (exponente de Pareto = 1), los coeficientes presentan desviaciones del valor unitario. Giesen y Südekum (2011) mencionan, al respecto, que existe esta posibilidad sin que signifique el no cumplimiento de la ley de Zipf, y hacen la salvedad de que el valor unitario óptimo se obtiene solo en el límite. En general, uno de los factores que influyen más en las desviaciones del coeficiente de Pareto es el tamaño de la muestra. Además, González-Val (2010) menciona que algunos estudios empíricos han demostrado la alta sensibilidad de los resultados a dos factores: la unidad geográfica que se considere y el tamaño de la muestra. Eeckhout (2004) resume esta caracterización en un solo aspecto: el punto de truncamiento de la distribución del tamaño poblacional.

Es importante mencionar que, debido a la relación entre el cumplimiento de las leyes de Zipf y de Gibrat, una aproximación a 1 del coeficiente de la relación rango-tamaño estaría indicando también el cumplimiento de la ley de Gibrat, caso en el cual las ciudades, sin importar su tamaño, crecerían a la misma tasa²⁶.

Una mirada al patrón de comportamiento temporal del coeficiente y al número de municipios incluidos parecería mostrar una relación de causalidad. Es decir, parece que el coeficiente se va reduciendo en el tiempo debido a la reducción en el número de municipios incluidos. Sin embargo, ejercicios adicionales, utilizando los criterios de selección alternativos ya mencionados, mostraron coeficientes similares que se reducen en el tiempo, a la vez que el número de municipios fue aumentando. Esto ofrece evidencias sobre la robustez de los resultados y la tranquilidad sobre la escogencia del criterio de selección de los municipios.

Si consideramos las teorías del crecimiento urbano, el país como un todo estaría atravesando por un período de transición de un modelo de crecimiento determinístico a uno estocástico. En otras palabras, el crecimiento de las ciudades ya no estaría determinado fundamentalmente por las decisiones de localización de las firmas y de las externalidades sobre el capital humano. El nuevo modelo de crecimiento urbano se estaría caracterizando por una mayor movilidad de la mano de obra y por los *amenities* que se ofrecen en los centros urbanos, generados muchos de ellos por la implementación de políticas públicas (Duranton, 2006, 2007).

Un segundo concepto que vale la pena explorar es el de la ley de Gibrat, la cual está ligada a la ley de Zipf. Bajo ciertas condiciones estadísticas, el cumplimiento de una de ellas significa el cumplimiento de la otra. Según la ley de Gibrat, el crecimiento de las ciudades es aleatorio con el mismo valor

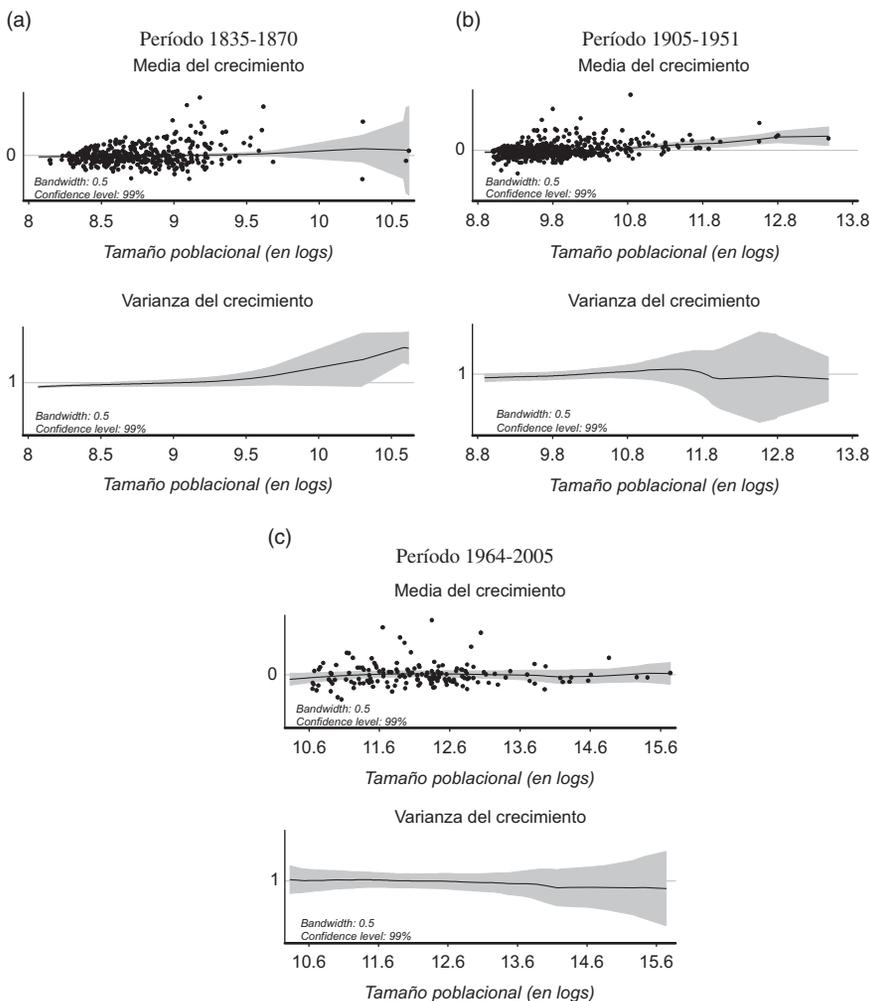
²⁶ Esta relación se mantiene siempre que se cumplan las condiciones estadísticas que pueden consultarse en Gabaix (1999b) y Skouras (2010).

esperado y varianza, independientemente de su tamaño. La verificación de su existencia sería, entonces, una prueba de hipótesis del crecimiento aleatorio de las ciudades. Si se establece que el crecimiento poblacional de un grupo de ciudades obedece a esta regla, significa que los efectos de escala no desempeñan un papel fundamental en la senda de crecimiento de las ciudades (Schaffar y Dimou, 2012). Giesen y Südekum (2011) también ofrecen algunos argumentos a favor de la prueba de Gibrat y en contraposición de la ley de Zipf. El primero es que, al realizarse esta prueba con métodos no paramétricos, no se condiciona a una forma funcional específica y, por tanto, se ofrece una mayor flexibilidad en la prueba de hipótesis. La segunda bondad tiene que ver con que el resultado de la prueba no se basa en un único estadístico, sino que es posible observar los resultados a lo largo de la distribución del tamaño de las ciudades.

Con el fin de realizar la verificación empírica, no solo de la ley de Gibrat sino de su relación con la ley de Zipf, el Gráfico 3 presenta los resultados de los estimadores no paramétricos para la media y la varianza del crecimiento poblacional en Colombia a lo largo de los dos últimos siglos. Con el fin de facilitar la lectura de los resultados, el período de análisis (1835-2005) se dividió en tres subperíodos: el primero, que corresponde a buena parte del siglo XIX (1835-1870); el segundo, para la primera mitad del siglo XX (1905-1951), y el tercero para el resto del siglo XX y hasta el año 2005 (1964-2005). Para la lectura de los resultados, conviene recordar que el cumplimiento de la ley de Gibrat implica que la población de las ciudades debe crecer con la misma media y varianza, y que dicho comportamiento debe ser independiente de su tamaño. Debido a que las series del tamaño poblacional fueron estandarizadas previamente (media 0 y varianza 1), se espera, entonces, que cualquier desviación de la media de 0 y de la varianza de 1, implique el no cumplimiento de la ley de Gibrat.

Los resultados son interesantes en varios aspectos. El primero es que las poblaciones urbanas de las primeras décadas de la República del siglo XIX (panel *a*) estaban caracterizadas por un tipo de crecimiento determinístico, con unas pocas ciudades de gran tamaño relativo que estaban creciendo a ritmos diferentes de las demás, lo que contraviene uno de los principios de Gibrat. El segundo es que el ritmo de crecimiento poblacional heredado del siglo XIX parece estar aún presente en la primera década del siglo XX. El panel *b* muestra un patrón similar al del siglo XIX, donde claramente el crecimiento poblacional medio y su varianza dependen del tamaño poblacional; sin embargo, la varianza parece mantenerse a lo largo de casi toda la distribución del tamaño de las ciudades dentro de los intervalos de confianza, mostrando cierto patrón de estabilización. El tercer factor es que la historia cambia a partir la segunda mitad del siglo XX y los primeros años del siglo XXI. Durante este período, es posible observar un paso hacia un crecimiento poblacional como el que predice la ley de Gibrat. Tanto para el caso de la media como para la varianza, no es posible rechazar la hipótesis de crecimiento urbano

GRÁFICO 3
ESTIMADORES NO PARAMÉTRICOS NADARAYA-WATSON PARA LA MEDIA Y
VARIANZA DEL CRECIMIENTO POBLACIONAL: TOTAL NACIONAL



Fuentes: Calculado a partir de la información censal de los años correspondientes.

independiente del tamaño de las ciudades, resultados que son consistentes con lo encontrado mediante los coeficientes de rango-tamaño o de Zipf.

Una posible explicación del cambio en la dinámica poblacional a partir de la segunda mitad del siglo XX es el desarrollo de la infraestructura del

transporte en el país²⁷. Precisamente, entre 1920 y 1960, Colombia se enfrenta a una de las mayores transformaciones en su infraestructura vial hasta ese momento (Bonet y Meisel, 1999). Durante este período, el país logró por primera vez interconectar las principales regiones mediante una red ferroviaria y carretera con grandes efectos económicos y demográficos. Algunas de las teorías del crecimiento económico mencionan que uno de los efectos esperados ante una mayor integración regional es el aumento en las disparidades económicas y la consolidación de un modelo de organización territorial centro-periferia. En este esquema existe una clara diferenciación en términos económicos, sociales o demográficos, entre una ciudad o centro urbano predominante (centro) y las demás que de alguna manera dependen de ella (periferia). En la siguiente sección se corroboran estos hechos con el análisis de la dinámica poblacional en las regiones colombianas.

Nivel regional

Esta subsección persigue dos objetivos principales. El primero es caracterizar la dinámica poblacional de las regiones colombianas desde la conformación de la República, observando si los patrones de crecimiento se rigen por las regularidades empíricas que predicen las leyes de Zipf y de Gibrat. El segundo objetivo está relacionado con la predicción teórica desarrollada por Gabaix (1999b), donde si las regiones cumplen la ley de Zipf individualmente, también lo hará el país como un todo. Como se mencionó, se definieron cinco regiones: Caribe, Central, Oriental, Pacífica y Nuevos Departamentos. Sin embargo, y debido al limitado número de observaciones para el caso de los Nuevos Departamentos, esta región no se incluye en el análisis, ya que no fue posible realizar algunas estimaciones. El Cuadro 4 presenta los resultados de las estimaciones de rango-tamaño para cada una de las regiones²⁸.

Los resultados permiten observar varias características. La primera y más evidente es que todas las regiones siguen un patrón similar al nacional en el sentido de que, durante todo el siglo XIX y la primera década del siglo XX, la dinámica del crecimiento poblacional parece haber sido diferente a la que ha prevalecido desde la segunda mitad del siglo XX y principios del siglo XXI. En particular, las regiones a su propio ritmo iniciaron también una etapa de cambio en su dinámica de crecimiento poblacional, de un modelo determinístico condicionado por la localización geográfica de los sectores productivos a otro donde las políticas públicas, a partir del mejoramiento del

²⁷ Estos hallazgos son coincidentes a partir de los resultados de las pruebas de Zipf y de Gibrat.

²⁸ En el GRÁFICO A- se presenta, gráficamente y para cada región, una comparación entre 1835 y 2005 de la relación entre el logaritmo del rango y el logaritmo del tamaño poblacional.

CUADRO 4
COEFICIENTES DE PARETO PARA LA DISTRIBUCIÓN POBLACIONAL: REGIONES DE COLOMBIA

(a) REGIÓN CARIBE				
AÑO CENSAL	NÚMERO DE CIUDADES	EXPONENTE DE PARETO	DESVIACIÓN ESTÁNDAR	ESTADÍSTICO $t_{Ho}: \beta = 1$
1835	58	1,852	0,344	2,476 (*)
1843	61	1,901	0,344	2,616 (*)
1851	61	1,897	0,343	2,611 (*)
1870	52	2,287	0,449	2,869 (**)
1905	54	1,780	0,342	2,276 (*)
1912	49	1,889	0,382	2,329 (*)
1918	45	1,816	0,383	2,132 (*)
1938	40	1,546	0,346	1,579
1951	39	1,459	0,330	1,389
1964	40	1,391	0,311	1,258
1973	47	1,322	0,273	1,181
1985	43	1,247	0,269	0,918
1993	40	1,183	0,264	0,691
2005	40	1,139	0,255	0,545

CUADRO 4 (Cont.)

(b) REGIÓN ORIENTAL				
AÑO CENSAL	NÚMERO DE CIUDADES	EXPONENTE DE PARETO	DESVIACIÓN ESTÁNDAR	ESTADÍSTICO t Ho: $\beta = 1$
1835	124	2,390	0,304	4,579 (**)
1843	128	2,526	0,316	4,832 (**)
1851	131	2,605	0,322	4,986 (**)
1870	140	2,586	0,309	5,131 (**)
1905	141	2,323	0,277	4,781 (**)
1912	145	2,383	0,280	4,942 (**)
1918	144	2,341	0,276	4,860 (**)
1938	136	1,954	0,237	4,026 (**)
1951	119	1,634	0,212	2,991 (**)
1964	85	1,269	0,195	1,382
1973	60	1,065	0,194	0,333
1985	37	0,925	0,215	-0,348
1993	20	0,852	0,270	-0,547
2005	14	0,806	0,305	-0,638

CUADRO 4 (Cont.)

(c) REGIÓN CENTRAL				
AÑO CENSAL	NÚMERO DE CIUDADES	EXPONENTE DE PARETO	DESVIACIÓN ESTÁNDAR	ESTADÍSTICO $t_{Ho}: \beta = 1$
1835	51	2,343	0,464	2,894 (**)
1843	59	2,536	0,467	3,289 (**)
1851	64	2,521	0,446	3,412 (**)
1870	67	2,676	0,462	3,625 (**)
1905	76	2,287	0,371	3,469 (**)
1912	81	2,291	0,360	3,586 (**)
1918	86	2,186	0,333	3,557 (**)
1938	88	2,092	0,315	3,462 (**)
1951	82	1,754	0,274	2,752 (**)
1964	88	1,456	0,233	1,955
1973	79	1,350	0,215	1,628
1985	74	1,251	0,206	1,220
1993	65	1,159	0,203	0,783
2005	53	1,089	0,211	0,419

CUADRO 4 (Cont.)

(d) REGIÓN PACÍFICA				
AÑO CENSAL	NÚMERO DE CIUDADES	ExPONENTE DE PARETO	DESVIACIÓN ESTÁNDAR	ESTADÍSTICO $t_{Ho: \beta = 1}$
1835	41	2,574	0,569	2,768 (**)
1843	43	2,666	0,575	2,897 (**)
1851	45	2,563	0,540	2,892 (**)
1870	54	2,215	0,426	2,850 (**)
1905	43	1,926	0,415	2,229 (*)
1912	53	2,286	0,444	2,896 (**)
1918	55	2,116	0,403	2,765 (**)
1938	56	2,052	0,388	2,712 (**)
1951	48	1,578	0,322	1,794
1964	44	1,333	0,284	1,172
1973	39	1,237	0,280	0,846
1985	30	1,106	0,286	0,371
1993	32	1,077	0,269	0,286
2005	30	1,010	0,261	0,039

Nota: La hipótesis nula se rechaza al 5% (*) y al 1% (**)

Fuente: Calculado a partir de la información censal de los años correspondientes.

entorno y mayor oferta de *amenities* y calidad de vida, parecen ser determinantes en el crecimiento poblacional²⁹.

La segunda característica tiene que ver con el valor de los estimadores. A pesar de que estadísticamente no puede rechazarse la hipótesis nula ($\beta = 1$), es posible observar que en algunos casos los coeficientes se alejan sustancialmente de 1. Como se vio, esto es posible sin que se afecte el cumplimiento de la regularidad empírica de Zipf, y dentro de los factores que pueden causar estas desviaciones se han mencionado el tamaño de la muestra, la unidad geográfica que se considere y el punto de truncamiento de la distribución. Otra de las razones que se menciona en la literatura es el surgimiento de nuevos centros urbanos, siempre y cuando la tasa de creación de nuevas ciudades no sea mayor que la tasa de crecimiento urbano de las ya existentes (Gabaix, 1999b). Esta posibilidad coincide con el hecho de que Colombia es un país en desarrollo, donde la dinámica poblacional está aún en transición, con la aparición de nuevos centros urbanos y con un crecimiento importante en el tamaño de los ya existentes. Esta situación contrasta con la de varios países europeos, donde el número y tamaño de los centros urbanos parece haber alcanzado el nivel de largo plazo desde hace varias décadas (González-Val, 2010).

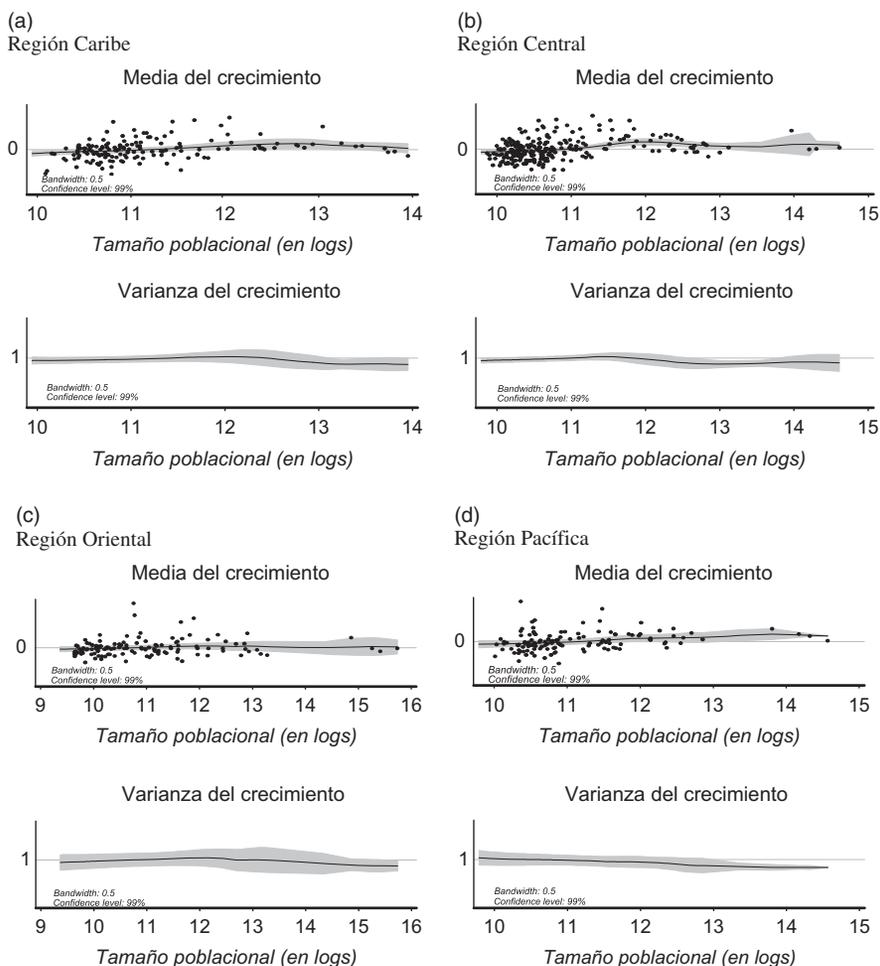
Una tercera característica, que debe ser analizada junto con los resultados del total nacional, es que parece haber indicios acerca del cumplimiento empírico para Colombia de lo planteado teóricamente por Gabaix (1999b). Los resultados muestran una coincidencia en el tiempo del cumplimiento de la ley de Zipf en cada una de las regiones, individualmente y también en el total nacional. La primera evidencia empírica acerca de esta relación está documentada en Giesen y Südekum (2011) para el caso de Alemania, donde los autores muestran que existe coincidencia en el cumplimiento de la ley de Zipf, en lo regional y en lo nacional.

El siguiente ejercicio consistió en calcular los estimadores no paramétricos de la media y la varianza del crecimiento poblacional para cada una de las cuatro regiones definidas. Los resultados se pueden observar en el Gráfico 4 y muestran cómo solo en dos de los cuatro casos (la región Caribe y la Oriental) existe evidencia para afirmar que las áreas urbanas dentro de cada una de las regiones crecieron con una media y una varianza constante entre 1964 y 2005³⁰. Para las regiones Central y Pacífica, al menos con un

²⁹ Es importante observar que, por ejemplo, en la región Caribe, el cumplimiento de la ley de Zipf se inicia antes que en las demás regiones. Aunque este documento no pretende responder a esta pregunta en particular, una posible explicación es que, históricamente, fue una de las primeras regiones del país, debido a su importancia como puerto y punto clave en el comercio internacional.

³⁰ Se presentan los resultados para este período en particular, por dos razones: la primera es que, en todas las regiones y para todos los períodos, se rechaza la hipótesis de media y varianza constantes, así como la de independencia de escala del crecimiento poblacional. La segunda es que se quisieron contrastar los resultados de las pruebas de Gibrat con las de Zipf solamente para el período en el que esta última se cumplía.

GRÁFICO 4
ESTIMADORES NO PRAMÉTRICOS NADARAYA-WATSON PARA LA MEDIA Y
VARIANZA DEL CRECIMIENTO POBLACIONAL (1964-2005): REGIONES



Fuente: Calculado a partir de la información censal de los años correspondientes.

nivel de confianza del 99% y para el período completo (1964-2005), existen evidencias de crecimiento demográfico diferencial y dependiente del tamaño en algunos segmentos de la distribución. Es posible que estos resultados puedan estar siendo afectados por las razones mencionadas acerca de las desviaciones de los parámetros de rango-tamaño.

Una posible explicación a este comportamiento diferencial de las regiones la ofrecen Bonet y Meisel (1999). Los autores encuentran evidencias de que a partir de 1960 las regiones se ven enfrentadas a lo que en la literatura se conoce como disparidades económicas regionales. Esto significa que las diferencias entre las distintas regiones del país, en términos del producto, no se redujeron sino que aumentaron o se mantuvieron constantes entre 1960 y 1995. Dentro de las posibles explicaciones que ofrecen los autores están: 1) el impacto de las políticas de industrialización mediante el modelo de sustitución de importaciones; 2) el fortalecimiento de las ciudades predominantes, especialmente Bogotá, en términos demográficos y económicos, y 3) el debilitamiento de la importancia relativa de los departamentos de la región Caribe. Estos factores contribuyeron, entonces, a favorecer a algunas regiones más que a otras, y el consecuente crecimiento económico y demográfico diferencial de algunas ciudades con respecto a las demás.

CONCLUSIONES

Los resultados encontrados en este documento muestran que el crecimiento demográfico de las ciudades cambió a partir de la segunda mitad del siglo XX. Desde el inicio de la República, a comienzos del siglo XIX y durante la primera mitad del siglo XX, el crecimiento urbano estaba caracterizado por patrones de dependencia del tamaño poblacional; es decir, las ciudades pequeñas crecían a ritmos diferentes de las medianas y las grandes, en contra de lo que predecían las leyes de Zipf y de Gibrat. En la teoría, este tipo de comportamiento lo describe el crecimiento poblacional determinístico, donde la dinámica demográfica está influenciada, en especial, por la localización de los sectores productivos. A partir de la segunda mitad del siglo XX, el modelo de crecimiento poblacional colombiano inició su transición hacia uno que coincide con las predicciones de las leyes de Zipf y de Gibrat, según las cuales las ciudades crecen al mismo ritmo sin importar su tamaño. Este tipo de modelo de crecimiento poblacional se conoce como crecimiento estocástico, y se caracteriza por verse afectado por choques exógenos, especialmente políticas públicas que mejoran la calidad de vida mediante lo que se conoce en la literatura como *amenities*.

El segundo resultado es que existe una coincidencia, en términos de la ley de Zipf, entre la dinámica nacional y la de cada una de las regiones estudiadas. En este documento se demostró que las distribuciones del tamaño de las ciudades exhiben una clara relación lineal de rango-tamaño en cada una de las regiones, individualmente y en el total nacional, tal como lo planteó teóricamente Gabaix (1999b). Sin embargo, cuando se comparan los resultados de la prueba de Gibrat, la tendencia nacional coincide solo en dos de las cuatro regiones consideradas, lo que puede explicarse por la consolidación de la infraestructura de transporte de la primera mitad del siglo XX.

Los principales mensajes derivados, a partir de los resultados encontrados, son que, en contra de lo que ocurría antes de mediados del siglo XX, el Estado, representado por los gobiernos locales de los centros urbanos, tiene una mayor incidencia en la dinámica urbana. La razón es que Colombia y sus regiones, tal como predice el modelo de crecimiento urbano estocástico, depende mucho más de las políticas públicas, entendidas como choques exógenos que pueden cambiar el ritmo de crecimiento de las ciudades. Es importante tener en cuenta que así como los choques positivos pueden traer grandes beneficios, lo contrario puede ocurrir a través de políticas públicas inadecuadas. Es aquí donde las autoridades locales, mediante políticas públicas, pueden afectar la senda futura del tamaño de las ciudades, según el nuevo modelo de crecimiento urbano. Si los gobiernos locales no adoptan ninguna medida tendente a atraer una mayor población, empleando políticas públicas que hagan a sus ciudades más atractivas, los resultados predicen que reducirán su tamaño relativo hasta formar parte del grupo de ciudades pequeñas. Todo dependerá, por supuesto, del modelo de ciudad que el gobierno local y sus habitantes establezcan como el más adecuado en cada caso en particular.

BIBLIOGRAFÍA

- ANDERSON, G. y GE, Y. (2005): «The Size Distribution of Chinese Cities». *Regional Science and Urban Economics* 35 (6), pp. 756-776.
- AUERBACH, F. (1913): «Das Gesetz der Bevölkerungskonzentration». *Petermanns Geographische Mitteilungen* 59, pp. 74-76.
- BEJARANO, J. (2007): «El despegue cafetero (1900-1928)», en J. A. Ocampo (comp.), *Historia económica de Colombia*. Bogotá: Planeta/Fedesarrollo, pp. 195-232.
- BENIGUI, L. y BLUMENFELD-LIEBERTHAL, E. (2011): «The End of a Paradigm: Is Zipf's Law Universal?». *Journal of Geographical Systems* 13 (1), pp. 87-100.
- BERNAL, G. y NIETO, C. (2006): «Evolución del coeficiente de Zipf para Colombia en el siglo XX». *Documentos de economía* 5. Bogotá: Universidad javeriana/Departamento de Economía.
- BONET, J. y MEISEL, A. (1999): «La convergencia regional en Colombia: una visión de largo plazo, 1926-1995». *Documentos de trabajo sobre economía regional* 8. Cartagena: Banco de la República.
- BOSKER, E. M.; BRAKMAN, S.; GARRETSEN, H. y SCHRAMM, M. (2007): «Looking for multiple equilibria when geography matters: German city growth and the WWII shock». *Journal of Urban Economics* 61, pp. 152-169.
- CHAMPERNOWNE, D. G. (1953): «A Model of Income Distribution». *Economic Journal* 63 (250), pp. 318-351.
- CHESIRE, P. (1999): «Trends in Sizes and Structures of Urban Areas», en P. C. Chesire y E. S. Mills (eds.), *Handbook of Regional and Urban Economics*, vol. 3. Amsterdam: North-Holland.

- DAVIS, D. R. y WEINSTEIN, D. E. (2002): «Bones, Bombs, and Break Points: The Geography of Economic Activity». *American Economic Review* 92 (5), pp. 1.269-1.289.
- DIMOU, M. y SCHAFFAR, A. (2009): «Urban Hierarchies and City-growth in the Balkans». *Urban Studies* 46 (13), pp. 2.891-2.906.
- DOBKINS, L. y IOANNIDES, Y. (2001): «Spatial Interactions among US Cities: 1900-1990». *Regional Science and Urban Economics* 31 (6), pp. 701-731.
- DURANTON, G. (2006): «Some Foundations for Zipf's Law: Product Proliferation and Local Spillovers». *Regional Science and Urban Economics* 36 (4), pp. 542-563.
- DURANTON, G. (2007): «Urban Evolutions: the Fast, the Slow, and the Still». *American Economic Review* 97 (1), pp. 197-221.
- ECKHOUT, J. (2004): «Gibrat's Law for (All) Cities». *American Economic Review* 94 (5), pp. 1.429-1.451.
- ECKHOUT, J. (2009): «Gibrat's Law for (All) Cities: Reply». *American Economic Review* 99 (4), pp. 1.676-1.683.
- GABAIX, X. (1999a): «Zipf's Law and the Growth of Cities». *American Economic Review* 89 (2), pp. 129-132.
- GABAIX, X. (1999b): «Zipf's Law for Cities: An Explanation». *Quarterly Journal of Economics* 114 (3), pp. 739-767.
- GABAIX, X. y IBRAGIMOV, R. (2011): «Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents». *Journal of Business Economics and Statistics* 29 (1), pp. 24-39.
- GIBRAT, R. (1931): *Les inégalités économiques*. París: Recueil Sirey.
- GIESEN, K. y SÜDEKUM, J. (2011): «Zipf's Law for Cities in the Regions and the Country». *Journal of Economic Geography* 11 (4), pp. 667-686.
- GIESEN, K.; ZIMMERMANN, A. y SUEDEKUM, J. (2010): «The size distribution across all cities - double Pareto lognormal strikes!». *Journal of Urban Economics* 68, pp. 129-137.
- GOERLICH, F. y MAS, M. (2010): «La distribución empírica del tamaño de las ciudades en España, 1900-2001. ¿Quién verifica la ley de Zipf?». *Revista de Economía Aplicada XVIII* (54), pp. 133-159.
- GONZÁLEZ-VAL, R. (2010): «The evolution of US city size distribution from a long-term perspective (1900-2000)». *Journal of Regional Science* 50 (5), pp. 952-972.
- GONZÁLEZ-VAL, R.; RAMOS, A.; SANZ-GARCÍA, F. y VERA-CABELLO, M. (2013). «City Distribution FOR All Cities: Which One is the Best?». *Papers in Regional Science*, forthcoming.
- HÄRDLE, W. (1990): *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.
- IOANNIDES, Y. y OVERMAN, H. (2003): «Zipf's Law for Cities: An Empirical Examination». *Regional Science and Urban Economics* 33 (2), pp. 127-137.
- IOANNIDES, Y. y SKOURAS, S. (2013): «US City Size Distribution: Robustly Pareto, but only in the Tail». *Journal of Urban Economics* 73, pp. 18-29.

- LEVY, M. (2009): «Gibrat's Law for (All) Cities: Comment». *American Economic Review* 99 (4), pp. 1.672-1.675.
- MELO, J. (2007): «Las vicisitudes del modelo liberal (1850-1899)», en J. A. Ocampo (comp.), *Historia económica de Colombia*. Bogotá: Planeta/Fedesarrollo, pp. 135-194.
- NADARAYA, E. (1964): «On Estimating Regression». *Theory of Probability and its Applications* 9, pp. 141-142.
- NOTA, S. y, SONG, F. (2006). «Further Analysis of the Zipf's Law: Does the Rank-Size Rule Really Exist». *UNR Joint Economics Working Paper Series* 07-2006. Reno: University of Nevada Press.
- O CAMPO, J.; BERNAL, J.; AVELLA, M. y ERRÁZURIZ, M. (2007): «La industrialización y el intervencionismo estatal (1945-1980)», en J. A. Ocampo (comp.), *Historia económica de Colombia*. Bogotá: Planeta/Fedesarrollo, pp. 271-340.
- PÉREZ, G. J. (2006): «Población y ley de Zipf en Colombia y la Costa Caribe, 1912-1993». *Documentos de trabajo sobre economía regional* 71. Cartagena de Indias: Banco de la República/CEER.
- ROSEN, K. y RESNICK, M. (1980): «The size distribution of cities: an examination of the Pareto law and primacy». *Journal of Urban Economics* 8 (2), pp. 165-186.
- SCHAFFAR, A. y DIMOU, M. (2012): «Rank-size City Dynamics in China and India, 1981-2004». *Regional Studies* 46 (6), pp. 707-721.
- SILVERMAN, B. (1986): *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.
- SIMON, H. (1955): «On a Class of Skew Distribution Functions». *Biometrika* 42 (3-4), pp. 425-440.
- SKOURAS, S. (2010). «Explaining Zipf's law for US Cities». *SSRN Working Paper*.
- SONG, S. y ZHANG, K. (2002): «Urbanization and city-size distribution in China». *Urban Studies* 39 (12), pp. 2.317-2.327.
- SOO, K. (2005): «Zipf's Law for Cities: A Cross-Country Investigation». *Regional Science and Urban Economics* 35 (3), pp. 239-263.
- WATSON, G. (1964): «Smooth Regression Analysis». *Sankhya: The Indian Journal of Statistics Series A* 26, pp. 359-376.
- YE, X. y XIE, Y. (2012): «Re-examination of Zipf's Law and Urban Dynamic in China: A Regional Approach». *Annals of Regional Science* 49 (1), pp. 135-156.
- ZHOU, Y. y MA, L. (2003): «China's urbanization levels: reconstructing a baseline from the Fifth Population Census». *China Quarterly* 173, pp. 176-196.
- ZIPF, G. (1949): *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.

APÉNDICE

CUADRO A-1
ESTADÍSTICAS DESCRIPTIVAS DE LAS CIUDADES EN COLOMBIA, 1835-2005
(MUNICIPIOS QUE ACUMULAN EL 100% DE LA POBLACIÓN)

AÑO	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	749	2.070	2.270	1.512	34	39.442	0,468
1843	751	2.360	2.355	1.846	63	40.086	0,445
1851	802	2.617	2.461	2.010	37	29.649	0,459
1870	739	3.624	3.045	2.845	58	40.883	0,392
1905	762	5.616	6.004	4.175	97	100.000	0,423
1912	767	6.486	6.759	5.054	33	121.257	0,387
1918	806	7.056	7.955	5.408	160	143.994	0,411
1938	809	10.745	17.079	7.309	399	355.502	0,442
1951	827	13.944	33.148	8.330	347	715.250	0,508
1964	879	19.875	71.058	10.093	294	1.697.311	0,581
1973	1020	22.403	106.622	10.118	85	2.861.913	0,635
1985	1024	29.352	153.862	12.000	797	4.236.490	0,669
1993	1061	31.206	175.391	11.343	78	4.945.448	0,701
2005	1113	38.524	233.820	12.626	225	6.840.116	0,726

Fuentes: Calculado a partir de información del Archivo General de la Nación y del DANE.

CUADRO A-2
ESTADÍSTICAS DESCRIPTIVAS DE LAS REGIONES DE COLOMBIA, 1835-2005
(MUNICIPIOS QUE ACUMULAN EL 100% DE LA POBLACIÓN)

(a) REGIÓN CARIBE							
AÑO	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	182	1.236	1.480	765	74	11.929	0,512
1843	187	1.269	1.416	761	70	10.145	0,497
1851	189	1.344	1.475	791	93	9.896	0,498
1870	126	2.595	2.051	1.950	390	11.595	0,387
1905	145	3.641	4.466	2.506	97	40.115	0,462
1912	118	6.277	6.270	4.473	184	48.907	0,403
1918	119	7.018	8.202	4.472	726	64.543	0,448

CUADRO A-2 (Cont.)

1938	113	12.707	17.967	8.177	1.412	152.348	0,481
1951	117	16.541	29.526	10.663	1.970	283.238	0,512
1964	128	25.490	50.441	14.614	2.318	498.301	0,535
1973	161	28.693	64.513	15.206	2.624	703.488	0,562
1985	161	38.933	89.988	18.962	3.676	927.233	0,580
1993	162	42.404	99.832	19.631	3.840	993.759	0,595
2005	194	46.861	117.517	19.065	2.721	1.146.498	0,628
(b) REGIÓN CENTRAL							
AÑO	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	127	2.300	1.895	1.911	79	10.280	0,418
1843	135	2.661	1.887	2.276	66	9.118	0,370
1851	156	2.876	2.229	2.344	198	13.755	0,397
1870	159	3.949	3.258	3.336	281	29.765	0,388
1905	176	7.516	6.191	5.566	464	53.936	0,375
1912	184	8.257	7.109	6.396	547	71.004	0,368
1918	203	8.665	7.896	6.454	276	79.146	0,390
1938	211	13.004	14.809	9.755	474	168.266	0,408
1951	217	16.970	28.290	11.615	1.255	363.865	0,461
1964	236	22.435	55.056	13.303	2.297	772.887	0,526
1973	270	23.692	75.289	12.897	918	1.163.868	0,566
1985	273	29.656	96.443	14.922	2.005	1.480.382	0,593
1993	273	31.748	107.418	14.858	2.329	1.630.009	0,623
2005	278	39.136	143.948	16.079	2.690	2.214.494	0,663
(c) RREGIÓN ORIENTAL							
AÑO	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	315	2.590	2.813	2.006	71	39.442	0,425
1843	306	2.923	2.891	2.454	63	40.086	0,399
1851	321	3.338	2.847	2.766	51	29.649	0,407
1870	314	4.174	3.358	3.446	58	40.883	0,362
1905	322	5.350	6.268	4.280	262	100.000	0,379
1912	323	5.870	7.261	4.732	567	121.257	0,359
1918	324	6.555	8.618	5.163	524	143.994	0,373
1938	334	8.656	19.937	6.111	681	355.502	0,428

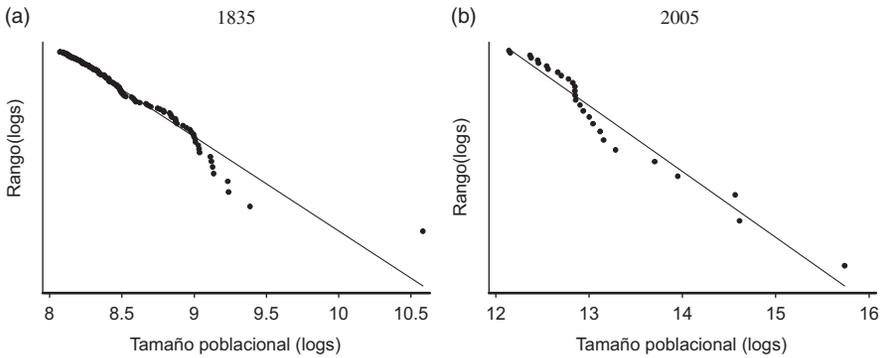
CUADRO A-2 (Cont.)

1951	336	10.666	39.638	6.379	785	715.250	0,500
1964	345	15.937	92.455	7.396	965	1.697.311	0,615
1973	383	19.140	147.547	7.029	833	2.861.913	0,691
1985	384	25.672	217.821	7.767	797	4.236.490	0,743
1993	391	28.064	252.483	6.982	270	4.945.448	0,786
2005	396	36.264	346.825	7.503	885	6.840.116	0,822
(d) REGIÓN PACÍFICA							
AÑO	NÚMERO DE CIUDADES	MEDIA	DESVIACIÓN ESTÁNDAR	MEDIANA	TAMAÑO MÍNIMO	TAMAÑO MÁXIMO	ÍNDICE GINI
1835	98	2.067	1.546	1.675	156	8.173	0,389
1843	94	2.782	1.832	2.320	179	10.376	0,335
1851	103	2.948	2.080	2.377	199	11.848	0,359
1870	130	3.059	2.423	2.176	480	12.743	0,391
1905	107	6.481	5.888	4.656	325	30.835	0,414
1912	118	6.669	4.893	5.272	938	27.760	0,351
1918	127	7.247	6.128	5.634	261	45.525	0,379
1938	131	11.804	11.393	9.209	1.109	101.883	0,391
1951	136	16.388	26.951	10.184	954	284.186	0,485
1964	141	22.892	55.991	13.374	2.149	637.929	0,544
1973	153	27.733	82.912	13.040	1.938	991.549	0,593
1985	153	34.534	119.513	14.119	3.661	1.429.026	0,643
1993	163	36.422	134.530	15.329	2.063	1.666.468	0,648
2005	177	41.955	165.143	15.696	3.481	2.119.843	0,672

Fuentes: Calculado a partir de información del Archivo General de la Nación y del DANE.

Nota: las regiones se definieron de la siguiente forma: Caribe (La Guajira, Magdalena, Atlántico, Bolívar, Cesar, Córdoba y Sucre); Central (Antioquia, Caldas, Caquetá, Huila, Quindío, Risaralda y Tolima); Oriental (Boyacá, Cundinamarca, Meta, Norte de Santander, Santander); Pacífico (Cauca, Chocó, Nariño y Valle); Nuevos Departamentos (Amazonas, Arauca, Casanare, Guainía, Guaviare, Putumayo, Vaupés y Vichada).

GRÁFICO A-1
GRÁFICOS DE RANGO-TAMAÑO (ZIPF): NIVEL NACIONAL



Fuente: Calculado a partir de la información censal de los años correspondientes.

GRÁFICO A-2
GRÁFICOS DE RANGO-TAMAÑO (ZIPF): REGIONES DE COLOMBIA

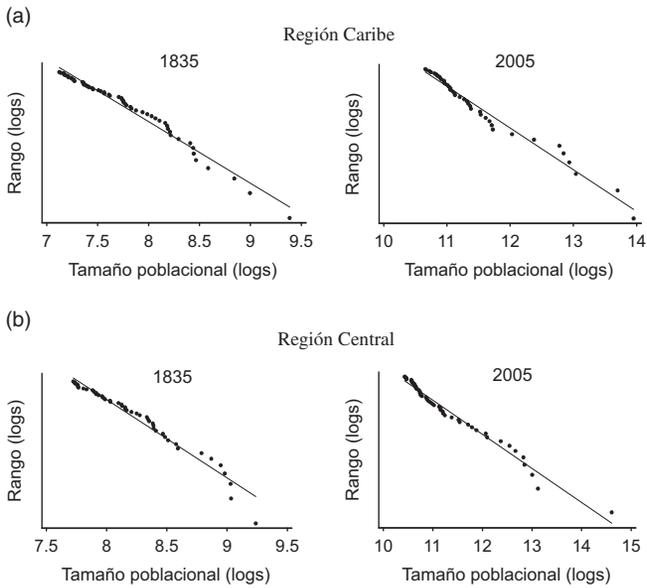
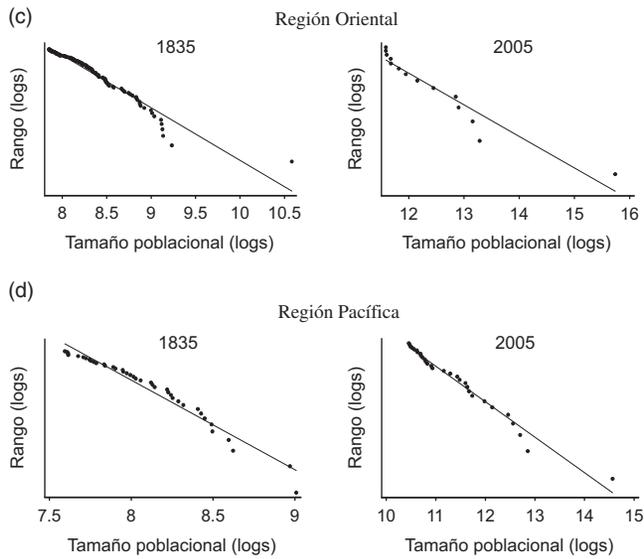


GRÁFICO A-2 (Continued)



Fuente: Calculado a partir de la información censal de los años correspondientes.
 Nota: las regiones se definieron de la siguiente forma: Oriental (Boyacá, Cundinamarca, Meta, Norte de Santander, Santander); Pacífica (Cauca, Chocó, Nariño y Valle).