

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE ADMINISTRAÇÃO, ECONOMIA E GESTÃO DE POLÍTICAS
PÚBLICAS

Matheus Devisate Borghi

Distribuição da População nos Municípios Brasileiros:
Lei de Zipf e Lei de Gibrat

BRASÍLIA – DISTRITO FEDERAL

2019

Matheus Devisate Borghi

Distribuição da População nos Municípios Brasileiros:
Lei de Zipf e Lei de Gibrat

Dissertação submetida à Faculdade de Economia,
Administração, Contabilidade e Gestão de
Políticas Públicas (FACE) da Universidade de
Brasília (UnB) como requisito parcial para a
obtenção do grau de Mestre em Economia.

Orientador: Prof. Dr. Moisés de Andrade Resende Filho

BRASÍLIA – DISTRITO FEDERAL

2019

RESUMO

Uma hierarquia urbana de um sistema de cidades é um conjunto $\{pop_1, \dots, pop_n\}$ em ordem decrescente da população (pop) de n cidades, em que o subscrito de pop é o rank r da cidade. O presente trabalho investiga, com base em um painel de 5565 dos 5570 municípios brasileiros no período 2010-2018, se a hierarquia urbana do sistema de cidades brasileiro tem distribuição de Pareto, se o expoente de Pareto é um (lei de Zipf) e se o crescimento das cidades e o seus tamanhos são independentes (lei de Gibrat). Os resultados mostram que a hierarquia urbana tem distribuição de Pareto, mas que para grande parte dos estados e Brasil como um todo, o expoente de Pareto é menor que um, indicando uma concentração na cidade de maior tamanho maior do que o previsto pela lei de Zipf. Os resultados também mostram que, na maioria dos estados e no Brasil como um todo, o processo de crescimento populacional em curso é explosivo e contribui para a concentração demográfica nas grandes cidades, o que contradiz a lei de Gibrat. Assim, como algumas cidades crescem se tornando maiores que as outras e isto tem perdurado e acentuado no tempo, parece não haver um tamanho ótimo de cidade no Brasil. Apesar disso, trabalhos futuros devem investigar a relação entre estes resultados e os baixos níveis de investimentos em transporte e tecnologias de informação no Brasil, pois estas tecnologias atenuam as forças aglomerativas.

Palavras-chave: distribuição do tamanho das cidades, lei de Zipf, lei de Gibrat, dados em painel, modelo efeitos fixos, system GMM.

ABSTRACT

An urban hierarchy of a system of cities is a set $\{pop_1, \dots, pop_n\}$ in descending order of the population (pop) of n cities, wherein the subscript of pop is the rank r of the city. In an urban hierarchy of a city system, a few cities grow larger than the others, suggesting that there is no optimal city size. This study investigates, based on a panel of 5565 out of the 5570 Brazilian municipalities in the period 2010-2018, if the urban hierarchy of the Brazilian cities system has Pareto distribution, if the Pareto exponent is one (Zipf law), and if the growth of cities and their sizes are independent (Gibrat law). The results show that the urban hierarchy has a Pareto distribution, but that for most states and Brazil as a whole, the Pareto exponent is smaller than one, indicating a concentration in the biggest city greater than the Zipf law says. The results also show that, in most states and in Brazil as a whole, the process of population growth in place is explosive and contributes to population concentration in big cities, which contradicts Gibrat's law. Thus, as some cities grow bigger than the others and this has lasted and accentuated over time, there does not seem to be an optimal city size in Brazil. In spite of this, future work should investigate the relationship between these results and the low levels of investments in transportation and information technologies in Brazil, as these technologies attenuate agglomerative forces.

Key words: size distribution of cities, Zipf law, Gibrat law, panel data, fixed effects model, GMM system.

SUMÁRIO

1. Introdução.....	7
2. Literatura.....	11
3. Métodos.....	14
3.1. Lei de Zipf	14
3.2. Lei de Gibrat.....	15
4. Dados	18
5. Resultados.....	28
5.1. Lei de Zipf	28
5.2. Lei de Gibrat.....	37
6. Conclusão.....	40

LISTA DE TABELAS

Tabela 1. Distribuição dos Municípios Brasileiros por UF's e Regiões (2018)	17
Tabela 2. Correlação com o número de municípios por estado (2018).....	19
Tabela 3. Crescimento Populacional dos Estados 2010-2018	20
Tabela 4. Crescimento Populacional da Unidades Federativas entre 2011 e 2018 ..	22
Tabela 5. Maiores Municípios de cada Região (2018)	23
Tabela 6. Modelo (5) estimado por MQO com $\alpha \equiv \alpha$ e $\gamma \equiv cs + \eta t$ (I)	26
Tabela 7. Modelo (5) estimado por MQO com $\alpha \equiv \alpha t$ e $\gamma \equiv \eta t$ (II)	27
Tabela 8. Modelo (5) estimado por MQO com $\alpha \equiv \alpha s$ e $\gamma \equiv cs$ (III).....	28
Tabela 9. Modelo (5) estimado por MQO com $\alpha \equiv \alpha s t$ e $\gamma \equiv cs + \eta t$ (IV)	29
Tabela 10. Modelo (7) estimado por system GMM com $\beta \equiv \beta$ e $\delta \equiv gs$	35
Tabela 11. Modelo (7) estimado por system GMM com $\beta \equiv \beta s$	36

1. Introdução

As aglomerações urbanas são um importante objeto de estudo na economia uma vez que constituem a unidade espacial onde a maioria das atividades econômicas ocorre. A forma como a população se distribui nos espaços geográficos não é aleatória, mas o resultado da interação de uma infinidade de incentivos e ações tomadas por milhões de indivíduos. Certamente existem motivações econômicas por trás da formação, funcionamento e desenvolvimento das cidades e investigar os aspectos espaciais da tomada de decisão significa entender porque algumas cidades são grandes e outras pequenas, quais as causas do crescimento e declínio delas, e como os governos locais afetam o processo de transformação das áreas urbanas.

O primeiro passo para responder essas questões é descrever precisamente as aglomerações e a mobilidade populacional, ou seja, identificar as relações espaciais que caracterizam a distribuição das cidades em um território e a evolução dessas relações espaciais ao longo do tempo. Duas regularidades empíricas são notáveis na literatura acerca deste tema. A primeira regularidade diz respeito à lei de Zipf: a população da n -ésima maior cidade é $1/n$ vezes a população da maior cidade. Ou seja, a segunda maior cidade de um país teria metade da população da maior cidade, a terceira maior cidade teria um terço da população da maior cidade, a quarta maior cidade, um quarto da população da maior cidade e assim por diante. Para ilustrar esse fato vale observar os dados da maior economia do mundo. Nos Estados Unidos, a maior cidade, Nova Iorque (NY), é aproximadamente duas vezes mais populosa do que a segunda maior cidade, Los Angeles (CA), aproximadamente três vezes maior do que a terceira cidade, Chicago (IL), aproximadamente quatro vezes maior do que a quarta cidade, Houston (TX) e aproximadamente cinco vezes maior do que a quinta cidade, Phoenix (AZ). A segunda regularidade empírica diz respeito à lei de Gibrat: a taxa de crescimento populacional de uma cidade é independente do tamanho da sua população. Isto é, não é possível afirmar que cidades grandes crescem mais rápido do que cidades pequenas, tampouco é possível afirmar o contrário.

As cidades emergiram em várias partes do mundo há cerca de 7 mil anos como resultado do desenvolvimento tecnológico e aumento da oferta agrícola. A divisão do trabalho em atividades especializadas também consistiu em uma mudança na estrutura social fundamental para concentração das pessoas em centros urbanos, de forma que os retornos crescentes provenientes das aglomerações populacionais

constituem uma explicação amplamente aceita acerca do processo de constituição das cidades (Krugman, 1991). A história também sugere que os detalhes da geografia que emerge – quais regiões que terminam com a população – depende sensivelmente das condições iniciais.

Em relação aos retornos crescentes, a concentração de diversas firmas em uma mesma área oferece um mercado de trabalho agrupado para os trabalhadores com habilidades industriais específicas, garantindo tanto uma menor probabilidade de desemprego para eles quanto uma menor probabilidade de escassez de mão de obra para as empresas. Outro resultado positivo é que a concentração industrial pode facilitar o fornecimento de serviços e fatores *nontradables*, como a construção civil e a geração de energia. Além disso, os transbordamentos informacionais e tecnológicos beneficiam os produtores agrupados em detrimento daqueles que se encontram isolados.

Mas somente isso não explica por que as indústrias acabam concentradas em uma ou em pouquíssimas regiões dos países com as demais regiões desempenhando o papel periférico de simples fornecedoras de produtos agrícolas para o núcleo industrial. A explicação se baseia em economias de escala mais gerais do que aquelas específicas a determinadas indústrias. A produção agrícola é caracterizada por retornos constantes de escala e pelo uso intensivo da terra. A distribuição geográfica dessa produção é, então, determinada principalmente pela oferta exógena de terras férteis. A indústria, por outro lado, é caracterizada por retornos crescentes de escala e uso modesto do fator terra. Por conta das economias de escala, a produção de cada bem manufaturado acontece em um número limitado de lugares. Tudo o mais constante, os lugares preferidos são aqueles próximos aos principais mercados consumidores a fim de minimizar os custos de transporte. Existe ainda uma circularidade nesse processo, já que é mais barato comprar os bens provenientes dos centros produtores e, portanto, é desejável viver próximo a eles. Tal circularidade, no entanto, é menos significativa se a indústria emprega apenas uma pequena parcela da população e portanto gera apenas uma pequena fração da demanda, ou devido a uma combinação de fracas economias de escalas e baixos custos de transporte. Em resumo, quando um determinado índice que leva em conta economias de escala, o percentual dos bens não agrícolas na produção total e os custos de transporte

ultrapassa um determinado nível, a população passa a se aglomerar. Uma vez iniciado, este processo se auto alimenta (Krugman, 1991).

A estabilidade e hierarquia dos sistemas urbanos tem diversas implicações econômicas: na distribuição do emprego e dos setores da economia; nos salários; no potencial de inovação das cidades como resultado do volume de pesquisa em cada localidade; na variedade de bens e serviços; na organização do mercado imobiliário e nos preços. Seguindo essa linha de raciocínio, o estudo da distribuição do tamanho das cidades é importante para a formulação de políticas públicas que promovam a otimização do processo de crescimento dos centros urbanos com a finalidade de obter uma solução ótima para a alocação dos recursos escassos que maximize o resultado positivo das economias de escala, mas que minimize os impactos negativos provenientes da concentração demográfica (poluição, congestionamentos, exclusão social, violência, problemas de saneamento e dificuldade de acesso aos serviços públicos).

O Brasil é um país altamente urbanizado: aproximadamente 85% da população reside em áreas urbanas (Censo Demográfico 2010, IBGE). Algumas peculiaridades distinguem o processo de evolução das cidades brasileiras. A evolução das cidades no Brasil é caracterizada pelo baixo grau de planejamento urbano e pela ocupação desordenada do solo nas grandes cidades, trazendo como consequência o rápido crescimento populacional das maiores regiões metropolitanas (Rolnik, 2006). Soma-se a isso a ineficiência na provisão de bens públicos e os fluxos migratórios inter-regionais (Baeninger, 2012). Há que se falar também na excessiva criação de novos municípios sem base econômica correspondente que sucedeu o processo de redemocratização e a promulgação da Constituição de 1988. A descentralização administrativa e fiscal e a flexibilização dos critérios necessários para a criação de novos municípios deram origem a um numeroso contingente de municípios totalmente dependentes de transferências constitucionais de recursos provenientes dos governos federais e estaduais (Gomes e Mac Dowell, 2000). O efeito esperado desse processo de redução de tamanho dos governos locais é o aumento das despesas per capita em decorrência da perda de economias de escala.

O objetivo deste trabalho é verificar se as regularidades empíricas mencionadas acerca da distribuição do tamanho das cidades, e que são constatadas na literatura internacional, ocorrem no Brasil mesmo diante das circunstâncias que

individualizam a distribuição da população nos municípios brasileiros. Isto é, o presente trabalho pretende investigar, para o caso das áreas urbanas brasileiras, a hipótese de independência entre o crescimento das cidades e o seu tamanho (lei de Gibrat) e também a hipótese de que o tamanho da população de cada cidade é inversamente proporcional à sua posição no ranking das maiores cidades (lei de Zipf). Nesse contexto, a análise com dados em painel pode ser útil para testar a aplicabilidade dessas leis. Dados em painel combinam diferenças entre indivíduos e diferenças do próprio indivíduo no tempo, o que torna possível aumentar o número de observações e, assim, efetuar inferências mais precisas acerca dos parâmetros do modelo devido a redução de multicolinearidade, além de permitir a construção de testes de hipóteses mais complexos, separando o efeito no tempo do efeito de seção cruzada. A aplicação de testes para dados em painel é relativamente escassa na literatura sobre esse tema, mas esse tipo de estrutura de dados possibilita a inferência simultânea para uma grande variedade de cidades mesmo para períodos mais curtos de tempo.

2. Literatura

Auerbach (1913) propôs que a distribuição do tamanho das cidades em um país pode ser aproximada por uma distribuição de Pareto. De acordo com esse autor, a distribuição teria a forma:

$$rank = A pop^{-\alpha} \quad (1)$$

ou

$$\log rank = \log A - \alpha \log pop \quad (2)$$

onde pop é a população de uma determinada cidade, $rank$ é o número de cidades com população maior ou igual a pop , e A e α são constantes positivas.

Ao longo dos anos, essa proposição foi refinada por muitos outros autores, mais notavelmente por Zipf (1949), quando o termo 'lei de Zipf' passou a ser usado com frequência para se referir à ideia de que o tamanho das cidades segue uma distribuição de Pareto. A lei de Zipf argumenta não somente que a distribuição do tamanho das cidades segue uma distribuição de Pareto, mas também estabelece que essa distribuição de Pareto tem uma forma especial com o expoente $\alpha = 1$, enquanto que A é o tamanho da população do maior município. Quando o expoente α é igual à unidade, a lei de Zipf é completamente satisfeita e implica que o tamanho da população da maior cidade é aproximadamente o dobro do tamanho da segunda maior, o triplo da terceira maior e assim por diante. Mas, além disso, os desvios em relação à lei de Zipf são considerados como evidência de distorções nos sistemas urbanos, cujas causas geralmente são atribuídas a fatores institucionais, econômicos e históricos ou estão relacionados à localização dos recursos. O valor estimado do expoente de Pareto revela a hierarquia urbana de um sistema de cidades. Quanto maior o expoente, $\alpha > 1$, mais uniformemente distribuído é o sistema urbano. Por outro lado, quanto menor o expoente, $\alpha < 1$, mais desigual é a distribuição da população entre as cidades, notando-se concentração urbana nas cidades de maior tamanho.

Do ponto de vista empírico, Rosen and Resnick (1980) realizaram uma investigação acerca do valor do expoente de Pareto para uma amostra de 44 países. A média encontrada para a amostra foi de 1,14, variando entre 0,81 para o Marrocos e 1,96 para a Austrália. O expoente excedeu a unidade para 32 dos 44 países, indicando que a população da maioria dos países é mais uniformemente distribuída do que poderia ser previsto pela lei de Zipf. Os autores encontraram ainda que o

expoente de Pareto é positivamente correlacionado com o PIB, com a população total do país e com a densidade da malha ferroviária, mas é negativamente correlacionado com a área territorial.

De acordo com Fujita, Krugman e Venable (2002) a distribuição das maiores cidades dos Estados Unidos é bem descrita por uma lei de potência, ou seja, o número de cidades com uma população maior que x é proporcional a $x^{-\alpha}$, com α próximo a 1. Plotando o logaritmo do tamanho da área metropolitana contra o logaritmo da classificação, percebe-se uma linearidade quase perfeita e a inclinação se aproxima de 45 graus. Uma análise estatística mais formal comprova a impressão visual.

Entretanto, parte da literatura é mais cética em relação à lei de Zipf. Apesar de ser muito difundida, alguns autores argumentam que a lei de Zipf não é uma boa aproximação para a distribuição do tamanho das cidades e é, muitas vezes, inadequada para descrever a classificação por ordem de tamanho. Soo (2005), por exemplo, fez uma análise empírica da distribuição do tamanho das cidades de 73 países e encontrou que a lei de Zipf é mais frequentemente rejeitada do que o esperado. Os estudos que dão suporte à lei de Zipf ainda são criticados devido ao viés de seleção, falhas metodológicas e limitação das bases de dados. Os estudos geralmente utilizam os dados das maiores cidades para examinar a questão, levando os opositores a afirmarem que a seleção de cidades da cauda superior da distribuição seria um artifício para ratificar a lei de Zipf. Os defensores da lei, por outro lado, argumentam que a sua precisão aumenta quando são empregados os métodos empíricos adequados e quando são utilizadas as definições apropriadas de cidade.

A lei de Gibrat, por sua vez, diz que a taxa de crescimento populacional de uma cidade não depende do tamanho dessa cidade, isto é, não se pode afirmar que cidades menores crescem mais rápido do que as maiores ou vice-versa. Berry e Okulicz-Kozaryn (2012) e Modica, Reggiani e Nijkamp (2017) fazem uso do seguinte modelo para explorar essa questão:

$$\log P_{i,t} = \beta_0 + \beta_1 \log P_{i,t-1} + \varepsilon_{it} \quad (3)$$

No modelo (3), os subscritos i e t representam, respectivamente, a cidade e o tempo. P representa a população da cidade e ε_{it} é o termo de erro. O coeficiente β_1 é o parâmetro de interesse e revela se a distribuição diverge ou converge para a sua média. Ou seja, o método basicamente consiste em aplicar um teste de raiz unitária sobre o processo autoregressivo primeira ordem. A lei de Gibrat é satisfeita quando

$|\beta_1| = 1$, o que configura um passeio aleatório, cujas médias e variâncias são independentes da variável explicativa. Quando $|\beta_1| > 1$, o processo é explosivo e a população diverge da média, significando que o crescimento esperado é maior nas cidades maiores. No caso oposto, quando $|\beta_1| < 1$, o processo é estacionário e a população converge para a média, significando que o crescimento esperado é maior nas cidades menores.

As duas leis são frequentemente analisadas em conjunto, pelo menos do ponto de vista teórico, devido a uma possível complementariedade entre elas. Champernowne (1953) e Simon (1955) argumentam que a regra da ordem de tamanho (lei de Zipf) surge naturalmente se a lei de Gibrat é satisfeita. Gabaix (1999), em um artigo seminal, demonstrou que, caso a média e a variância das taxas de crescimento das cidades sejam independentes do tamanho das cidades, então a lei de Gibrat valida a lei de Zipf. Enquanto que Cordoba (2003) alega que uma versão fraca da lei de Gibrat produz regras de ordem de tamanho mais gerais, onde fraca significa que o tamanho da cidade pode afetar a variância do processo de crescimento, mas não a sua média. Nesse sentido, Cordoba (2008) investigou as relações bidirecionais entre as duas leis e elucidou pela primeira vez uma relação até então desconhecida: ele mostrou que a lei de Zipf também pode, sob determinadas condições, validar na lei de Gibrat.

3. Métodos

3.1. Lei de Zipf

Considere $\{pop_1, \dots, pop_n\}$ um conjunto em ordem decrescente da população (pop) de n municípios, em que o subscrito de pop é o rank r do município no conjunto. Com base na função de distribuição acumulada (FDA), $F(pop_r)$, a probabilidade de um município de população maior que pop_r é $r/n = 1 - F(pop_r)$ (Stanley et al., 1995). No caso de $\{pop_1, \dots, pop_n\}$ ter função densidade de probabilidade de Pareto, $f(pop_r) = x_m^\alpha pop_r^{-(1+\alpha)}$, então $F(pop_r) = 1 - x_m^\alpha pop_r^{-\alpha}$ com parâmetro de escala $x_m = \min(pop)$, $pop_r \geq x_m > 0$ e parâmetro de forma ou expoente de Pareto $\alpha > 0$ (Luckstead e Devadoss, 2014).

Substituindo a FDA de Pareto em $r/n = 1 - F(pop_r)$ e manipulando, obtemos $r = nx_m^\alpha pop_r^{-\alpha}$ que, após aplicação do operador logaritmo natural em seus dois lados, resulta na equação da regra do rank tamanho ou *rank size rule*:

$$\log r = \gamma + \alpha (-\log pop_r) \quad (4)$$

em que $\gamma \equiv \log(nx_m^\alpha)$ e o parâmetro α é o expoente de Pareto, em homenagem a Vilfredo Pareto (Gabaix, 2016).

Se $\alpha=1$, a distribuição do tamanho dos municípios segue a lei de Zipf e, neste caso, o maior município é duas vezes o tamanho do segundo maior município, três vezes o tamanho do terceiro maior município e assim por diante (Luckstead e Devadoss, 2014). Se $\alpha > 1$, a distribuição do tamanho das cidades é mais igual do que prevê a lei de Zipf e, no caso extremo de α tendendo a infinito, todas as cidades teriam o mesmo tamanho. Se $\alpha < 1$, a distribuição do tamanho das cidades é mais desigual do que o previsto pela lei de Zipf e, assim, a população se concentra na cidade de maior tamanho, pois o tamanho da maior cidade é sempre maior do que r vezes o tamanho da cidade de rank r .

Com base na equação (4), o modelo econométrico base da equação da regra do posto tamanho ou *rank size rule* para dados em painel é:

$$\log r_{it} = \alpha_{st} (-\log pop_{it}) + c_s + \eta_t + u_{it} \quad (5)$$

em que para o município i no ano t , r_{it} é o posto do município ou número de municípios com população maior ou igual a pop_{it} , a população do município; α_{st} é o expoente de Pareto que pode variar por unidade da federação s e por ano. O parâmetro γ pode diferir por município e ano devido a soma do efeito fixo da unidade da federação c_s e

do efeito fixo de tempo η_t . O efeito fixo c_s captura os fatores invariantes no tempo da unidade da federação s que são correlacionados com r_{it} e pop_{it} como, por exemplo, a origem e formação histórica, relevo, localização geográfica no país, recursos naturais e clima do município. O efeito fixo de tempo η_t captura o efeito comum do ano t nos municípios brasileiros que são correlacionados com r_{it} e pop_{it} como, por exemplo, condições e choques macroeconômicos e políticas públicas comuns aos municípios brasileiros no ano t .

O modelo (5) foi estimado por Mínimos Quadrados Ordinários (MQO), com erros-padrão robustos a erros não esféricos em nível de município. Com a finalidade de encontrar a melhor especificação para os efeitos fixos de unidade federativa e de ano e também para o coeficiente de Pareto, foram levados em consideração 4 cenários distintos:

- I) o parâmetro α não varia no tempo nem entre os estados, enquanto que o parâmetro γ é potencialmente diferente ano a ano e entre as unidades federativas, isto é, $\alpha \equiv \alpha$ e $\gamma \equiv c_s + \eta_t$.
- II) os parâmetros α e γ são potencialmente diferentes ano a ano, mas não variam entre os estados, isto é, $\alpha \equiv \alpha_t$ e $\gamma \equiv \eta_t$.
- III) os parâmetros α e γ são potencialmente diferentes entre os estados, mas não variam ano a ano, isto é, $\alpha \equiv \alpha_s$ e $\gamma \equiv c_s$.
- IV) os parâmetros α e γ são potencialmente diferentes ano a ano e entre as unidades da federação, isto é, $\alpha \equiv \alpha_{st}$ e $\gamma \equiv c_s + \eta_t$.
- V) Em cada caso, foram realizados testes F de significância conjunta dos efeitos fixos de modo a determinar a melhor forma de especificá-los. Os resultados são apresentados na seção 5.1.

3.2. Lei de Gibrat

A lei de Gibrat afirma que o crescimento populacional de uma cidade não depende do tamanho de sua população (Gibrat, 1931). Em outras palavras, embora as cidades possam crescer a taxas diferentes, não existe nenhuma relação sistemática entre taxa de crescimento e número de habitantes, de modo que não se pode afirmar que cidades maiores crescem mais rápido ou vice-versa.

Analicamente, é possível escrever a seguinte expressão logarítmica (Steindl, 1968):

$$\log P(t) = \log P(0) + \varepsilon(1) + \varepsilon(2) + \dots + \varepsilon(t) \quad (6)$$

em que $P(t)$ é o tamanho da população de determinada cidade no período t , $P(0)$ é a população inicial e $\varepsilon(t)$ é um choque aleatório i.i.d. com média μ e variância σ^2 . Então, a equação (6) identifica o logaritmo da população de uma cidade como a soma da sua população inicial e das taxas passadas de crescimento.

A implicação da lei de Gibrat é que os processos de crescimento das cidades possuem uma média comum (igual a média da taxa de crescimento das cidades) e uma variância comum (Gabaix 1999), isto é, tanto a média quanto a variância são necessariamente independentes do tamanho das cidades.

Com base na equação (6), o modelo econométrico base da equação da lei de Gibrat para dados em painel é:

$$\log dmp_{it} = \beta_s \log dmp_{it-1} + g_s + \varepsilon_{it} \quad (7).$$

Para o município i no ano t , $\log dmp_{it}$ corresponde ao desvio do logaritmo da cidade i em relação à média dos logaritmos das populações das cidades no ano t (Modica, Reggiani e Nijkamp, 2017). $\log dmp_{it-1}$ é o equivalente para o município i no ano anterior. β_s é o coeficiente da lei de Gibrat que pode variar por unidade da federação s . O efeito fixo da unidade da federação, g_s , captura os fatores invariantes no tempo na unidade da federação s que são correlacionados com pop_{it} e pop_{it-1} como, por exemplo, origem e formação histórica da população. ε_{it} é o erro idiossincrático. O efeito fixo de tempo foi inicialmente incluído na especificação, mas não está presente na equação (7) porque todas estimações acusaram insignificância estatística desse componente. Ele capturaria o efeito comum do ano t nos municípios brasileiros que são correlacionados com pop_{it} e pop_{it-1} como, por exemplo, choques migratórios comuns aos municípios da unidade da federação s no ano t .

A equação (7) da lei de Gibrat, de onde pretende-se testar $|\beta| = 1$, devido à presença da variável dependente defasada como variável explicativa, é estimada por Generalized Method of Moments (GMM) em suas versões difference GMM, desenvolvido em Arellano e Bond (1991), e System GMM, desenvolvido em Arellano e Bover (1995) e Blundell e Bond (1998). O estimador do GMM instrumentaliza as variáveis explicativas em diferença, que não são estritamente exógenas, com suas defasagens que cumprem os requisitos de variável instrumental: são não

correlacionadas com os termos de erro da regressão e são correlacionadas com a variável endógena. Enquanto o difference GMM trabalha apenas com defasagens em nível como instrumentos, o conjunto de instrumentos do System GMM é mais amplo, incluindo também as defasagens das diferenças. A correta estimação do coeficiente da variável dependente defasada deve ser tal que $EF < GMM < MQA$, isto é, o estimador do GMM deve ficar dentro do intervalo determinado pelas estimativas obtidas pelo Modelo de Efeitos Fixos (EF), que subestima o coeficiente, e pelo modelo de Mínimos Quadrados Ordinários Agrupado (MQA), que superestima o coeficiente. Para fins de ajustes da especificação, os estimadores foram submetidos aos testes de Hansen e de Sargan de sobreidentificação das restrições e ao teste AR de autocorrelação dos resíduos. O modelo (7) foi estimado por Generalized Method of Moments (GMM) com erros-padrão robustos a erros não esféricos em nível de município. Com a finalidade de encontrar a melhor especificação para os efeitos fixos de unidade federativa e de ano e também para o coeficiente β , foram realizados testes F de significância conjunta dos efeitos fixos de modo a determinar a melhor forma de especificá-los. Os resultados são apresentados na seção 5.2.

4. Dados

Os dados utilizados na presente análise correspondem ao número de habitantes de cada município brasileiro para o período compreendido entre os anos de 2010 e 2018. Os dados de 2010 foram extraídos do Censo Demográfico 2010 realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), em que 191 mil recenseadores visitaram 67,6 milhões de domicílios nos 5.565 municípios brasileiros. Os dados relativos aos anos subsequentes, abrangendo o período relativo aos anos de 2011 até 2018, foram extraídos da série de estimativas populacionais por municípios do IBGE. A experiência do IBGE no campo das projeções de população teve início em 1973, quando seu Centro Brasileiro de Estudos Demográficos (CBED), atual Coordenação de População e Indicadores Sociais (COPIS) da Diretoria de Pesquisas (DPE), elaborava a projeção da população do Brasil pelo método das componentes demográficas. Foi a partir de 1989, no entanto, que o IBGE consolidou seu primeiro esquema de projeções populacionais, com periodicidade anual, compreendendo os níveis Nacional, Unidades da Federação e Municípios, em cumprimento ao dispositivo constitucional, regulamentado pela Lei Complementar nº 59, de 22 de dezembro de 1988. Assim, o IBGE passou a realizar as estimativas da população residente para todos os municípios brasileiros regularmente instalados e, de acordo com o que estabelece o Artigo 102 da Lei nº 8.443, de 16 de julho de 1992, publica essas estimativas no Diário Oficial da União, até 31 de agosto de cada ano.

As projeções de população conjuntamente com os indicadores sociais, econômicos e demográficos estimados pelo IBGE, constituem um instrumento poderoso que cumpre o propósito de subsidiar o planejamento de políticas públicas que visam o atendimento das necessidades específicas de crianças, adolescentes, jovens, pessoas em idade ativa e o contingente de idosos, bem como o de fornecer parâmetros balizadores a serem considerados nos processos de monitoramento e avaliação dos diversos programas implantados na área social. Convém registrar que o total estimado de pessoas residentes em regiões do País com estruturas político-administrativa definidas é um dos parâmetros de referência para que o Tribunal de Contas da União (TCU) efetue o cálculo do Fundo de Participação dos Estados e dos Municípios (FPE e FPM), visando determinar a distribuição das respectivas quotas que compõem as transferências financeiras da União para Estados e Municípios. Tais

estimativas também são fundamentais para o cálculo de indicadores econômicos e sociodemográficos nos períodos intercensitários.

Brasil atualmente existem 5.570 municípios que estão distribuídos nas unidades federativas e nas regiões conforme a Tabela 1. O estado de Minas Gerais é o que concentra o maior número de divisões administrativas desse tipo. São 853 municípios no total. Seguido por São Paulo, com 645, e Rio Grande do Sul, com 497. São Paulo e Minas Gerais também são os estados mais populosos, embora no primeiro estado residam mais do que o dobro do número de habitantes residentes no segundo estado. Os 45,5 milhões de habitantes do estado de São Paulo correspondem a aproximadamente 22% da população brasileira.

Tabela 1. Distribuição dos Municípios Brasileiros por UF's e Regiões (2018)

	Municípios	População	População/Municípios	Área (km ²)	Área/Municípios (km ²)	Densidade demográfica (hab/km ²)
AM	62	4.080.611	65.816,31	1.559.146,876	25.147,53	2,62
PA	144	8.513.497	59.121,51	1.247.955,238	8.666,36	6,82
RO	52	1.757.589	33.799,79	237.765,293	4.572,41	7,39
AP	16	829.494	51.843,38	142.828,521	8.926,78	5,81
AC	22	869.265	39.512,05	164.123,737	7.460,17	5,30
RR	15	576.568	38.437,87	224.300,805	14.953,39	2,57
TO	139	1.555.229	11.188,70	277.720,412	1.997,99	5,60
Região Norte	450	18.182.253	40.405,01	3.853.840,882	8.564,09	4,72
BA	417	14.812.617	35.521,86	564.732,450	1.354,27	26,23
CE	184	9.075.649	49.324,18	148.887,633	809,17	60,96
PE	185	9.496.294	51.331,32	98.076,021	530,14	96,83
MA	217	7.035.055	32.419,61	331.936,949	1.529,66	21,19
AL	102	3.322.820	32.576,67	27.848,140	273,02	119,32
RN	167	3.479.010	20.832,40	52.811,107	316,23	65,88
PI	224	3.264.531	14.573,80	251.611,929	1.123,27	12,97
PB	223	3.996.496	17.921,51	56.468,435	253,22	70,77
SE	75	2.278.308	30.377,44	21.918,443	292,25	103,94
Região Nordeste	1.794	56.760.780	31.639,23	1.554.291,107	866,38	36,52
DF	1	2.974.703	2.974.703,00	5.779,997	5.780,00	514,65
GO	246	6.921.161	28.134,80	340.106,492	1.382,55	20,35
MS	79	2.748.023	34.785,10	357.145,531	4.520,83	7,69
MT	141	3.441.998	24.411,33	903.202,446	6.405,69	3,81
Região Centro-Oeste	467	16.085.885	34.445,15	1.606.234,466	3.439,47	10,01

Tabela 1. Continuação

SP	645	45.538.936	70.603,00	248.219,627	384,84	183,46
RJ	92	17.159.960	186.521,30	43.781,588	475,89	391,94
MG	853	21.040.662	24.666,66	586.520,732	687,60	35,87
ES	78	3.972.388	50.928,05	46.086,907	590,86	86,19
Região Sudeste	1.668	87.711.946	52.585,10	924.608,854	554,32	94,86
PR	399	11.348.937	28.443,45	199.307,939	499,52	56,94
RS	497	11.329.605	22.795,99	281.737,888	566,88	40,21
SC	295	7.075.494	23.984,73	95.737,954	324,54	73,90
Região Sul	1.191	29.754.036	24.982,40	576.783,781	484,29	51,59
BRASIL	5.570	208.494.900	37.431,76	8.515.759,090	1.528,86	24,48

Fonte: IBGE. Elaboração própria.

Pela Tabela 2, nota-se que não há uma correlação perfeita entre a população dos estados e o número de municípios em cada estado. De fato, o coeficiente de correlação entre essas duas variáveis é igual a 0,75 no país. Conforme elucidado na Tabela 1, o estado do Rio de Janeiro é a maior exceção a essa regra, onde a razão entre habitantes e número de municípios é muito superior à média nacional, indicando uma alta centralização administrativa por habitantes. Entre os estados com as menores quantidades de municípios, destacam-se o Acre, com 22, Amapá, com 16, e Roraima, com apenas 15. Esses três estados, nessa mesma ordem decrescente, também possuem as menores populações do país. O Tocantins registra a menor razão entre população e número de municípios, indicando uma alta descentralização administrativa por habitantes. Analisando-se a área territorial dos estados, é possível constatar que o estado com a maior extensão territorial, o Amazonas, com aproximadamente 1,5 milhão de quilômetros quadrados, é também o estado com a maior razão entre a área e o número de municípios, revelando uma altíssima centralização administrativa por área, muito superior à média nacional. O Pará, segundo maior estado em extensão territorial, com aproximadamente 1,2 milhão de quilômetros quadrados, tem uma razão entre a área e o número de municípios quase três vezes menor do que o seu vizinho Amazonas, embora a centralização por área ainda seja muito superior do que a média nacional. De fato, a correlação entre as variáveis área e número de municípios em cada estado é muito pequena no país, próxima de 0,05 (excluindo-se o Distrito Federal do cálculo), o que indica que o fator populacional está muito mais associado à lógica da distribuição dos municípios no Brasil do que o fator territorial. Roraima e Amazonas são os estados com as menores densidades demográficas, enquanto que o Distrito Federal é a unidade da federação

com o maior número de habitantes por quilometro quadrado. Rio de Janeiro e São Paulo ocupam a segunda e a terceira posições de acordo com esse critério.

Dentre as cinco regiões brasileiras definidas pelo IBGE, a região nordeste possui a maior quantidade de municípios, sendo a segunda região mais populosa do país com 56,7 milhões de pessoas habitando 1.794 municípios distribuídos em nove estados. A região mais populosa e também a mais densamente povoada, o Sudeste, com 87,7 milhões de habitantes, é a segunda região com a maior quantidade de municípios, 1.668, a grande maioria deles situados nos estados de Minas Gerais e São Paulo. A região sul é a menor em extensão territorial e também registra a menor área territorial média dos municípios, que totalizam 1.191. As regiões norte e centro-oeste, em ordem decrescente, são as maiores em área territorial, no entanto contêm poucos municípios, que são grandes em extensão territorial média e revelam alguma centralização administrativa de acordo com esse critério.

Tabela 2. Correlação com o número de municípios por estado (2018)

	População por estado	Área territorial por estado
Região Norte	0,6879	0,4493
Região Nordeste	0,7834	0,8919
Região Centro-Oeste*	0,9759	-0,1733
Região Sudeste	0,6179	0,9382
Região Sul	0,8726	0,9988
BRASIL	0,7528	0,0545

*Não leva em consideração o Distrito Federal.

Fonte: IBGE. Elaboração própria.

A Tabela 2 exhibe, para cada uma das regiões, o coeficiente de correlação entre o número de municípios por estado e a população dos estados e também o coeficiente de correlação entre o número de municípios por estado e a área territorial dos estados. Os dados do Distrito Federal não foram levados em consideração no cálculo desses indicadores, uma vez que o seu território não pode ser dividido em municípios. Vale notar que o IBGE, para fins censitários e para efeitos de contagem e estatística, não distingue Brasília do Distrito Federal, de modo que os dados relativos a Brasília correspondem ao conjunto de todas as regiões administrativas do Distrito Federal. De acordo com a Tabela 2, nas regiões Sul, Sudeste e Nordeste, a quantidade de municípios por estado está mais correlacionada com a área territorial dos estados do

que com o contingente populacional deles. Na região centro-oeste ocorre o contrário: não existe correlação com a área, mas há uma alta correlação com a população.

Haja vista que a quantidade de municípios brasileiros é razoavelmente grande, e sabendo que eles estão inseridos nos estados, e que os estados estão inseridos nas regiões, o exame dos dados demográficos desses agregados é útil para uma análise preliminar. A população brasileira apurada pelo IBGE no Censo Demográfico 2010 totalizava 190.755.799 habitantes. De acordo com os dados das estimativas populacionais, o número de habitantes no país na data de referência de 2018 era equivalente a 208.494.900 de pessoas, um crescimento acumulado de 9,30% nesse período.

A Tabela 3 apresenta as taxas de crescimento populacionais para o período de 2010 até 2018. Os estados cujas populações mais cresceram foram Roraima (27,99%), Amapá (23,89%), Acre (18,50%), Amazonas (17,12%), Distrito Federal (15,74%) e Goiás (15,28%). De fato, as regiões norte e centro-oeste exibiram as maiores taxas de crescimento. Os estados com as menores taxas de crescimento ente 2010 e 2018 foram Alagoas (6,48%), Paraíba (6,11%), Rio Grande do Sul (5,94%), Bahia (5,68%) e Piauí (4,69%). Dentre as cinco regiões, o nordeste cresceu à menor taxa (6,93%).

Tabela 3. Crescimento Populacional dos Estados 2010-2018

	População 2010	População 2018	2010-2018
AM	3.483.985	4.080.611	17,1248%
PA	7.581.051	8.513.497	12,2997%
RO	1.562.409	1.757.589	12,4922%
AP	669.526	829.494	23,8927%
AC	733.559	869.265	18,4997%
RR	450.479	576.568	27,9900%
TO	1.383.445	1.555.229	12,4171%
Região Norte	15.864.454	18.182.253	14,6100%
BA	14.016.906	14.812.617	5,6768%
CE	8.452.381	9.075.649	7,3739%
PE	8.796.448	9.496.294	7,9560%
MA	6.574.789	7.035.055	7,0005%
AL	3.120.494	3.322.820	6,4838%
RN	3.168.027	3.479.010	9,8163%
PI	3.118.360	3.264.531	4,6874%
PB	3.766.528	3.996.496	6,1056%
SE	2.068.017	2.278.308	10,1687%
Região Nordeste	53.081.950	56.760.780	6,9305%

Tabela 3 - Continuação

DF	2.570.160	2.974.703	15,7400%
GO	6.003.788	6.921.161	15,2799%
MS	2.449.024	2.748.023	12,2089%
MT	3.035.122	3.441.998	13,4056%
Região Centro-Oeste	14.058.094	16.085.885	14,4244%
SP	41.262.199	45.538.936	10,3648%
RJ	15.989.929	17.159.960	7,3173%
MG	19.597.330	21.040.662	7,3649%
ES	3.514.952	3.972.388	13,0140%
Região Sudeste	80.364.410	87.711.946	9,1428%
PR	10.444.526	11.348.937	8,6592%
RS	10.693.929	11.329.605	5,9443%
SC	6.248.436	7.075.494	13,2362%
Região Sul	27.386.891	29.754.036	8,6434%
BRASIL	190.755.799	208.494.900	9,2994%

Fonte: IBGE. Elaboração própria.

A Tabela 4, por sua vez, exhibe as taxas de crescimento populacional ano a ano das unidades federativas e das regiões brasileiras para o período de 2011 até 2018, tomando como base, para o cálculo das taxas relativas ao ano de 2011, os dados do Censo Demográfico 2010. Vale notar que, devido à crise econômica e humanitária que assola a Venezuela, apenas entre 2017 e 2018, a população de Roraima cresceu a uma taxa superior a 10%, a maior taxa anual de crescimento verificada em todo o período da amostra. Em todos os anos, as regiões norte e centro-oeste cresceram acima da média nacional, enquanto que o nordeste prevaleceu como a região de menor crescimento demográfico a partir de 2013, com alguns dos seus estados apresentando taxas negativas no último ano. Como é o caso da Bahia, cuja população diminuiu à taxa de -3,47%.

Tabela 4. Crescimento Populacional das Unidades Federativas entre 2011 e 2018

	2010-2011	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016	2016-2017	2017-2018
AM	1,56%	1,49%	6,04%	1,73%	1,67%	1,61%	1,55%	0,42%
PA	1,42%	1,74%	2,27%	1,31%	1,26%	1,20%	0,74%	1,76%
RO	0,90%	0,86%	8,69%	1,18%	1,13%	1,08%	1,04%	-2,67%
AP	2,21%	2,09%	5,21%	2,17%	2,10%	2,04%	1,97%	3,98%
AC	1,75%	1,66%	2,33%	1,76%	1,70%	1,64%	1,58%	4,78%
RR	2,15%	2,04%	3,95%	1,82%	1,76%	1,69%	1,63%	10,32%
TO	1,26%	1,21%	4,27%	1,27%	1,22%	1,17%	1,13%	0,32%
Região Norte	1,45%	1,57%	4,07%	1,46%	1,40%	1,35%	1,10%	1,37%
BA	0,57%	0,55%	6,13%	0,55%	0,51%	0,48%	0,44%	-3,47%
CE	0,92%	0,89%	2,01%	0,73%	0,70%	0,66%	0,63%	0,61%
PE	0,78%	0,75%	3,11%	0,75%	0,73%	0,70%	0,67%	0,24%
MA	1,08%	1,03%	1,19%	0,83%	0,78%	0,72%	0,66%	0,50%
AL	0,73%	0,70%	4,28%	0,63%	0,58%	0,54%	0,50%	-1,57%
RN	0,96%	0,93%	4,52%	1,02%	0,99%	0,95%	0,92%	-0,80%
PI	0,70%	0,65%	0,74%	0,33%	0,29%	0,25%	0,22%	1,41%
PB	0,66%	0,63%	2,60%	0,75%	0,72%	0,69%	0,65%	-0,72%
SE	1,05%	1,01%	4,02%	1,09%	1,05%	1,02%	0,99%	-0,43%
Região Nordeste	0,79%	0,76%	3,50%	0,70%	0,67%	0,63%	0,59%	-0,86%
DF	1,55%	1,48%	5,33%	2,24%	2,19%	2,14%	2,09%	-2,13%
GO	1,28%	1,22%	4,53%	1,39%	1,34%	1,29%	1,24%	2,10%
MS	1,16%	1,11%	3,28%	1,25%	1,21%	1,17%	1,15%	1,29%
MT	1,34%	1,28%	2,14%	1,33%	1,28%	1,23%	1,18%	2,91%
Região Centro-Oeste	1,32%	1,26%	3,95%	1,51%	1,46%	1,42%	1,37%	1,32%
SP	0,79%	0,76%	4,21%	0,85%	0,82%	0,80%	0,77%	0,98%
RJ	0,77%	0,74%	0,85%	0,56%	0,54%	0,52%	0,50%	2,64%
MG	0,67%	0,64%	3,72%	0,68%	0,65%	0,62%	0,58%	-0,37%
ES	0,91%	0,88%	7,30%	1,19%	1,15%	1,11%	1,07%	-1,09%
Região Sudeste	0,76%	0,73%	3,55%	0,77%	0,74%	0,71%	0,69%	0,88%
PR	0,65%	0,62%	3,97%	0,77%	0,73%	0,71%	0,70%	0,25%
RS	0,36%	0,35%	3,65%	0,39%	0,36%	0,34%	0,32%	0,06%
SC	1,10%	1,05%	3,93%	1,40%	1,37%	1,34%	1,31%	1,06%
Região Sul	0,64%	0,62%	3,84%	0,77%	0,74%	0,72%	0,70%	0,37%
BRASIL	0,85%	0,83%	3,65%	0,86%	0,83%	0,80%	0,75%	0,40%

Fonte: IBGE. Elaboração própria.

Tratando-se dos municípios propriamente ditos, que são o verdadeiro objeto de investigação deste trabalho, foram calculadas as taxas anuais de crescimento de todos eles. À luz da lei de Gibrat, parece não haver, num primeiro momento, qualquer relação entre o tamanho dos municípios e as suas taxas de crescimento demográfico. A Tabela 5 reproduz as taxas de crescimento dos doze maiores municípios de cada região para todo o período. À luz da lei de Zipf, por sua vez, os dados populacionais

dos municípios foram plotados contra os seus rankings em gráficos de dispersão. O Gráfico 1 ilustra a dispersão para o ano de 2018, em que a reta cujo coeficiente de inclinação é igual a unidade, em módulo, seria o resultado ideal da referida lei.

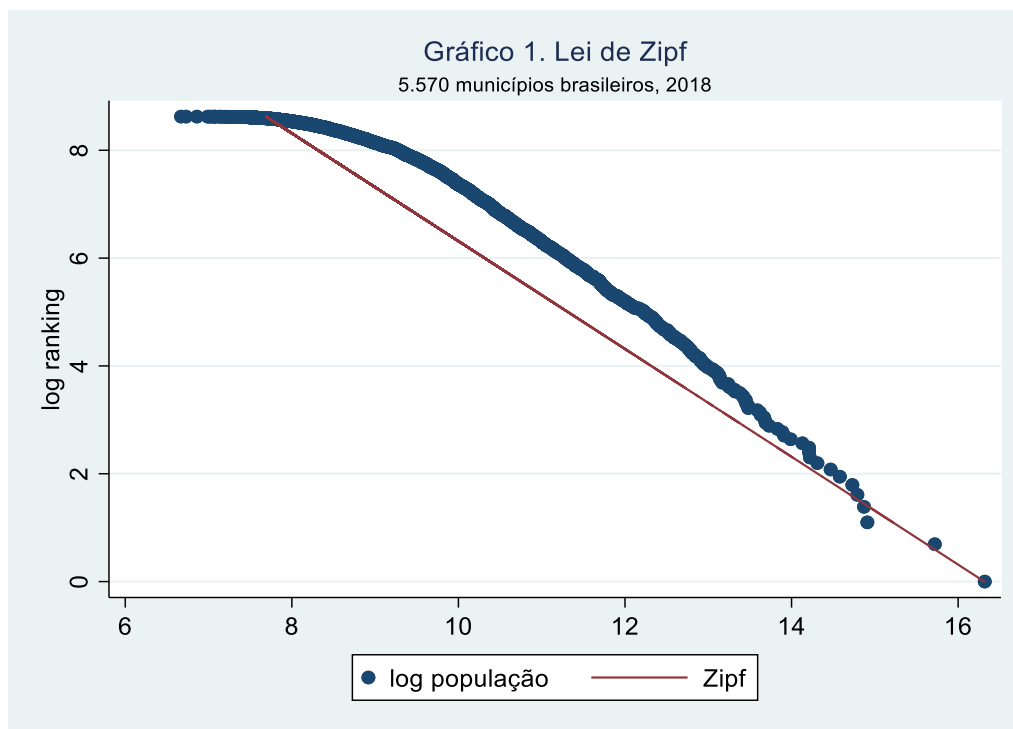
Tabela 5. Maiores Municípios de cada Região (2018)

Município	UF	População 2018	População 2010	2010-2018
Região Norte				
Manaus	AM	2.145.444	1.832.423	17,08%
Belém	PA	1.485.732	1.402.056	5,97%
Ananindeua	PA	525.566	477.999	9,95%
Porto Velho	RO	519.531	435.732	19,23%
Macapá	AP	493.634	407.023	21,28%
Rio Branco	AC	401.155	342.298	17,19%
Boa Vista	RR	375.374	290.741	29,11%
Santarém	PA	302.667	297.039	1,89%
Palmas	TO	291.855	235.315	24,03%
Marabá	PA	275.086	238.708	15,24%
Parauapebas	PA	202.882	160.228	26,62%
Castanhal	PA	198.294	176.116	12,59%
Região Nordeste				
Salvador	BA	2.857.329	2.693.605	6,08%
Fortaleza	CE	2.643.247	2.476.589	6,73%
Recife	PE	1.637.834	1.546.516	5,90%
São Luís	MA	1.094.667	1.027.429	6,54%
Maceió	AL	1.012.382	943.109	7,35%
Natal	RN	877.640	810.780	8,25%
Teresina	PI	861.442	822.363	4,75%
João Pessoa	PB	800.323	733.154	9,16%
Jaboatão dos Guararapes	PE	697.636	649.787	7,36%
Aracaju	SE	648.939	579.563	11,97%
Feira de Santana	BA	609.913	562.466	8,44%
Campina Grande	PB	407.472	387.643	5,12%
Região Centro-Oeste				

Tabela 5. Continuação

Região Centro-Oeste				
Brasília	DF	2.974.703	2.609.997	13,97%
Goiânia	GO	1.495.705	1.318.148	13,47%
Campo Grande	MS	885.711	796.252	11,24%
Cuiabá	MT	607.153	556.298	9,14%
Aparecida de Goiânia	GO	565.957	465.092	21,69%
Anápolis	GO	381.970	338.544	12,83%
Várzea Grande	MT	282.009	255.448	10,40%
Rio Verde	GO	229.651	181.020	26,86%
Rondonópolis	MT	228.857	198.949	15,03%
Dourados	MS	220.965	198.421	11,36%
Águas Lindas de Goiás	GO	207.070	163.495	26,65%
Luziânia	GO	205.023	177.098	15,77%
Região Sudeste				
São Paulo	SP	12.176.866	11.316.149	7,61%
Rio de Janeiro	RJ	6.688.927	6.355.949	5,24%
Belo Horizonte	MG	2.501.576	2.385.639	4,86%
Guarulhos	SP	1.365.899	1.233.436	10,74%
Campinas	SP	1.194.094	1.088.611	9,69%
São Gonçalo	RJ	1.077.687	1.008.064	6,91%
Duque de Caxias	RJ	914.383	861.157	6,18%
São Bernardo do Campo	SP	833.240	770.253	8,18%
Nova Iguaçu	RJ	818.875	799.047	2,48%
Santo André	SP	716.109	678.485	5,55%
São José dos Campos	SP	713.943	636.876	12,10%
Osasco	SP	696.850	667.826	4,35%
Região Sul				
Curitiba	PR	1.917.185	1.764.540	8,65%
Porto Alegre	RS	1.479.101	1.413.094	4,67%
Joinville	SC	583.144	520.905	11,95%
Londrina	PR	563.943	511.278	10,30%
Caxias do Sul	RS	504.069	441.332	14,22%
Florianópolis	SC	492.977	427.298	15,37%
Maringá	PR	417.010	362.329	15,09%
Blumenau	SC	352.460	312.634	12,74%
Ponta Grossa	PR	348.043	314.527	10,66%
Canoas	RS	344.957	325.188	6,08%
Pelotas	RS	341.648	328.864	3,89%
Cascavel	PR	324.476	289.339	12,14%

Fonte: IBGE. Elaboração própria.



A análise econométrica com os dados em painel é detalhada no capítulo 4. Do total de 5.570 municípios existentes no Brasil atualmente, foram utilizados, no painel, os dados de 5.565 municípios, compreendendo o período de 2010 a 2018. Os municípios de Mojuí dos Campos (PA), Balneário Rincão (SC), Pescaria Brava (SC), Paraíso das Águas (MS) e Pinto Bandeira (RS) foram retirados da amostra para fins de balanceamento do painel, uma vez que os processos de emancipação desses cinco municípios só foram concluídos após 1º de julho de 2012, resultando na ausência de estimativas para os anos de 2011 e 2012. A soma da população desses cinco municípios, na data de referência de 2018, foi estimada em apenas 46.997 habitantes. O painel é, portanto, balanceado e constituído por 5.565 observações analisadas durante 9 anos.

5. Resultados

5.1. Lei de Zipf

A estimação da equação (5) utilizando os dados do painel nacional com 50.085 observações em 9 períodos, referentes à população anual dos 5.565 municípios brasileiros entre 2010 e 2018, revelou os seguintes resultados acerca da aderência da lei de Zipf os dados da distribuição da população brasileira nas cidades.

No cenário (I) em que os efeitos fixos de unidade federativa e de ano são explicitamente discriminados em γ e o parâmetro α é mantido constante no tempo e entre os estados da federação, o expoente de Pareto é estimado em 0,8424 com intervalo de 95% de confiança variando de 0,8295 a 0,8552. Rejeita-se, então, a hipótese de que o coeficiente é igual a unidade e de que a distribuição da população pode ser descrita, no nível nacional, pela lei de Zipf. O resultado é reportado na tabela 5.1. A distribuição do tamanho das cidades é mais desigual, com as grandes cidades maiores do que o previsto pela lei de Zipf. Há concentração urbana nas cidades de maior tamanho, com o tamanho da maior cidade mais do que r vezes o tamanho da cidade de posto r . A hipótese de insignificância conjunta dos efeitos fixos de ano, ou seja, de que o parâmetro γ não varia no tempo, foi rejeitada no teste F com estatística $F(7, 5564) = 2007,32$ e p-valor igual a zero, da mesma forma que a hipótese nula de insignificância conjunta dos efeitos fixos de UF também foi descartada com $F(25, 5564) = 230,80$ e p-valor igual a zero. A regressão obteve $R^2 = 0,9425$.

Tabela 6. Modelo (5) estimado por MQO com $\alpha \equiv \alpha$ e $\gamma \equiv c_s + \eta_t$ (I)

	α	σ	IC(α) 95%
Brasil 2010 - 2018	0,8424	(0,0066)***	[0,8295 ; 0,8552]

* p<0.05; ** p<0.01; *** p<0.001

No cenário (II) em que apenas os efeitos fixos de ano são discriminados em γ e o parâmetro α é estimado ano a ano, os resultados para todos os anos são bastante próximos daquele verificado para o período 2010-2018. O resultado é reportado na tabela 5.2. A lei de Zipf é sempre descartada rejeitando-se a hipótese de que o coeficiente é igual a unidade. Vale notar que existe uma tendência decrescente no parâmetro estimado, possivelmente evidenciando que a distribuição da população

urbana brasileira tem se tornado mais concentrada nas grandes cidades. A hipótese de insignificância conjunta dos efeitos fixos de ano é rejeitada com $F(7, 5564) = 144,74$ e p-valor zero, validando a necessidade de inclusão desse componente na especificação do modelo. A hipótese nula de que o coeficiente α é constante no tempo, isto é, $\alpha_{2010} = \alpha_{2011} = \dots = \alpha_{2018}$, é igualmente rejeitada com estatística $F(8, 5564) = 160,27$ e p-valor nulo, validando também a especificação ano a ano do parâmetro α . A regressão apresentou $R^2 = 0,9366$.

Tabela 7. Modelo (5) estimado por MQO com $\alpha \equiv \alpha_t$ e $\gamma \equiv \eta_t$ (II)

	α	σ	IC(α) 95%
Brasil 2010	0,8382	(0,0065)***	[0,8255 ; 0,8509]
Brasil 2011	0,8359	(0,0065)***	[0,8232 ; 0,8486]
Brasil 2012	0,8338	(0,0065)***	[0,8211 ; 0,8465]
Brasil 2013	0,8313	(0,0065)***	[0,8186 ; 0,8440]
Brasil 2014	0,8290	(0,0065)***	[0,8163 ; 0,8417]
Brasil 2015	0,8268	(0,0065)***	[0,8141 ; 0,8395]
Brasil 2016	0,8246	(0,0065)***	[0,8119 ; 0,8373]
Brasil 2017	0,8225	(0,0065)***	[0,8098 ; 0,8352]
Brasil 2018	0,8171	(0,0065)***	[0,8043 ; 0,8298]

* p<0.05; ** p<0.01; *** p<0.001

No cenário (III) os parâmetros α e γ são potencialmente diferentes entre os estados, mas não variam ano a ano. Como pode ser observado na tabela 5.3, o resultado é diverso: treze unidades da federação parecem confirmar a lei de Zipf. Para duas unidades da federação, o coeficiente estimado é maior do que a unidade, indicando uma distribuição mais igual do tamanho das cidades. A primeira delas é o distrito federal, cuja contagem populacional e divulgação dos dados pelo IBGE considera todo o seu território como um único município devido à peculiaridade de sua divisão administrativa, lembrando o caso extremo em que o coeficiente tende ao infinito quando todas as cidades têm o mesmo tamanho. A segunda é o Rio de Janeiro, que se destaca por ser um estado com uma quantidade relativamente pequena de municípios face ao grande número de habitantes do estado, o que resulta em uma distribuição pouco desigual. Nos demais estados, doze no total, a rejeição da lei de Zipf é caracterizada pelo coeficiente menor do que unidade. A regressão obteve $R^2 = 0,953$. No entanto, esses resultados devem ser analisados com ressalva, uma vez que os efeitos fixos de ano não foram considerados na estimação. Conforme demonstrado

no cenário (II), os efeitos fixos de ano são altamente significantes e devem ser discriminados no modelo, bem como o coeficiente α deve variar de ano pra ano. A utilidade de estimação do modelo (5), de acordo com o presente cenário (III), é justamente testar a validade dos efeitos fixos de unidade federativa (c_s) e também a hipótese de que o expoente de Pareto varia entre os estados. O primeiro teste rejeita a hipótese nula de insignificância conjunta dos efeitos fixos de UF com estatística $F(25, 5564) = 1720,55$ e p-valor zero, resultado que valida a inclusão do componente c_s no modelo. O segundo teste, cuja hipótese nula é $\alpha_{AC} = \alpha_{AM} = \dots = \alpha_{TO}$, valida a especificação do coeficiente na forma α_s , com $F(26, 5564) = 884,53$ e p-valor zero.

Tabela 8. Modelo (5) estimado por MQO com $\alpha \equiv \alpha_s$ e $\gamma \equiv c_s$ (III)

	α	σ	IC(α) 95%
AC	0,9936	(0,0542)***	[0,8874 ; 1,0999]
AL	0,9176	(0,0520)***	[0,8156 ; 1,0196]
AM	1,0780	(0,0547)***	[0,9707 ; 1,1853]
AP	0,9537	(0,0591)***	[0,8378 ; 1,0696]
BA	0,9841	(0,0268)***	[0,9316 ; 1,0367]
CE	1,0168	(0,0319)***	[0,9543 ; 1,0792]
DF	1,7318	(0,0000)***	[1,7318 ; 1,7318]
ES	1,0346	(0,0215)***	[0,9925 ; 1,0767]
GO	0,7775	(0,0304)***	[0,7180 ; 0,8370]
MA	0,9564	(0,0276)***	[0,9022 ; 1,0106]
MG	0,7959	(0,0172)***	[0,7622 ; 0,8296]
MS	0,9207	(0,0539)***	[0,8151 ; 1,0263]
MT	0,7645	(0,0422)***	[0,6818 ; 0,8472]
PA	1,0210	(0,0278)***	[0,9666 ; 1,0754]
PB	0,7398	(0,0448)***	[0,6521 ; 0,8276]
PE	1,0270	(0,0263)***	[0,9754 ; 1,0785]
PI	0,7435	(0,0531)***	[0,6395 ; 0,8475]
PR	0,8222	(0,0240)***	[0,7752 ; 0,8692]
RJ	1,1014	(0,0168)***	[1,0683 ; 1,1344]
RN	0,7810	(0,0493)***	[0,6843 ; 0,8778]
RO	0,8999	(0,0555)***	[0,7910 ; 1,0088]
RR	1,0138	(0,0390)***	[0,9374 ; 1,0903]
RS	0,7398	(0,0195)***	[0,7016 ; 0,7779]
SC	0,7538	(0,0235)***	[0,7076 ; 0,7999]
SE	0,8562	(0,0515)***	[0,7552 ; 0,9572]
SP	0,8729	(0,0146)***	[0,8443 ; 0,9016]
TO	0,6399	(0,0557)***	[0,5308 ; 0,7491]

* p<0.05; ** p<0.01; *** p<0.001

À luz dos resultados anteriores, a melhor especificação do modelo (5) é, portanto, aquela correspondente ao cenário (IV) em que o expoente de Pareto varia no tempo e também entre as unidades da federação, isto é, $\alpha \equiv \alpha_{st}$, e os efeitos fixos, tanto de ano quanto de UF, são discriminados no modelo, isto é, $\gamma \equiv c_s + \eta_t$. O resultado é reportado na tabela 5.4. A lei de Zipf é confirmada novamente para os mesmos treze estados brasileiros do cenário (III): Acre, Alagoas, Amazonas, Amapá, Bahia, Ceará, Espírito Santo, Maranhão, Mato Grosso do Sul, Pará, Pernambuco, Rondônia e Roraima. A lei de Zipf é refutada, com expoente de Pareto menor que um, para os estados de Goiás, Minas Gerais, Mato Grosso, Paraíba, Piauí, Paraná, Rio Grande do Norte, Rio Grande do Sul, Santa Catarina, Sergipe, São Paulo e Tocantins e, com expoente maior que um, para Rio de Janeiro e Distrito Federal. O maior coeficiente foi encontrado no Rio de Janeiro, enquanto que o menor, no Tocantins. A regressão apresentou $R^2 = 0,9542$. As hipóteses de insignificância dos efeitos fixos de ano e de UF foram rejeitadas, respectivamente, com $F(6, 5564) = 124,11$ e $F(24, 5564) = 23,92$, ambas com p-valor zero. A hipótese de insignificância estatística conjunta do coeficiente α para todos os estados ao longo de todos os anos foi igualmente rejeitada com $F(242, 5564) = 2,3e + 05$ e p-valor zero.

Tabela 9. Modelo (5) estimado por MQO com $\alpha \equiv \alpha_{st}$ e $\gamma \equiv c_s + \eta_t$ (IV)

	α	σ	IC(α) 95%
AC 2010	0,9996	(0,0540)***	[0,8938 ; 1,1054]
AC 2011	1,0046	(0,0537)***	[0,8993 ; 1,1099]
AC 2012	1,0020	(0,0537)***	[0,8967 ; 1,1074]
AC 2013	0,9987	(0,0536)***	[0,8937 ; 1,1038]
AC 2014	0,9960	(0,0536)***	[0,8909 ; 1,1010]
AC 2015	0,9933	(0,0536)***	[0,8883 ; 1,0984]
AC 2016	0,9905	(0,0535)***	[0,8856 ; 1,0955]
AC 2017	0,9879	(0,0535)***	[0,8831 ; 1,0927]
AC 2018	0,9899	(0,0533)***	[0,8854 ; 1,0944]
AL 2010	0,9199	(0,0523)***	[0,8174 ; 1,0225]
AL 2011	0,9256	(0,0522)***	[0,8232 ; 1,0280]
AL 2012	0,9235	(0,0522)***	[0,8212 ; 1,0258]
AL 2013	0,9210	(0,0520)***	[0,8191 ; 1,0229]
AL 2014	0,9187	(0,0520)***	[0,8168 ; 1,0206]
AL 2015	0,9165	(0,0520)***	[0,8146 ; 1,0184]
AL 2016	0,9140	(0,0520)***	[0,8121 ; 1,0159]
AL 2017	0,9117	(0,0520)***	[0,8098 ; 1,0135]
AL 2018	0,9139	(0,0520)***	[0,8119 ; 1,0159]

Tabela 9 - continuação

AM 2010	1,0834	(0,0550)***	[0,9756 ; 1,1911]
AM 2011	1,0885	(0,0548)***	[0,9811 ; 1,1959]
AM 2012	1,0861	(0,0547)***	[0,9789 ; 1,1934]
AM 2013	1,0830	(0,0544)***	[0,9764 ; 1,1895]
AM 2014	1,0805	(0,0543)***	[0,9741 ; 1,1868]
AM 2015	1,0781	(0,0542)***	[0,9719 ; 1,1843]
AM 2016	1,0755	(0,0541)***	[0,9695 ; 1,1815]
AM 2017	1,0730	(0,0540)***	[0,9672 ; 1,1789]
AM 2018	1,0753	(0,0539)***	[0,9696 ; 1,1809]
AP 2010	0,9589	(0,0593)***	[0,8427 ; 1,0752]
AP 2011	0,9641	(0,0589)***	[0,8486 ; 1,0795]
AP 2012	0,9614	(0,0588)***	[0,8462 ; 1,0767]
AP 2013	0,9584	(0,0588)***	[0,8432 ; 1,0737]
AP 2014	0,9556	(0,0587)***	[0,8406 ; 1,0706]
AP 2015	0,9530	(0,0587)***	[0,8379 ; 1,0681]
AP 2016	0,9503	(0,0587)***	[0,8351 ; 1,0654]
AP 2017	0,9476	(0,0586)***	[0,8326 ; 1,0625]
AP 2018	0,9492	(0,0585)***	[0,8346 ; 1,0638]
BA 2010	0,9872	(0,0269)***	[0,9344 ; 1,0400]
BA 2011	0,9928	(0,0270)***	[0,9399 ; 1,0457]
BA 2012	0,9907	(0,0270)***	[0,9377 ; 1,0436]
BA 2013	0,9877	(0,0268)***	[0,9351 ; 1,0403]
BA 2014	0,9853	(0,0268)***	[0,9327 ; 1,0379]
BA 2015	0,9832	(0,0268)***	[0,9306 ; 1,0358]
BA 2016	0,9806	(0,0266)***	[0,9286 ; 1,0327]
BA 2017	0,9784	(0,0266)***	[0,9263 ; 1,0304]
BA 2018	0,9810	(0,0268)***	[0,9286 ; 1,0335]
CE 2010	1,0200	(0,0320)***	[0,9573 ; 1,0826]
CE 2011	1,0252	(0,0319)***	[0,9627 ; 1,0877]
CE 2012	1,0231	(0,0319)***	[0,9606 ; 1,0855]
CE 2013	1,0202	(0,0318)***	[0,9578 ; 1,0827]
CE 2014	1,0178	(0,0318)***	[0,9554 ; 1,0802]
CE 2015	1,0156	(0,0318)***	[0,9532 ; 1,0780]
CE 2016	1,0130	(0,0318)***	[0,9506 ; 1,0754]
CE 2017	1,0106	(0,0318)***	[0,9482 ; 1,0730]
CE 2018	1,0125	(0,0318)***	[0,9502 ; 1,0748]
DF 2010	1,0706	(0,0356)***	[1,0009 ; 1,1404]
DF 2011	1,0736	(0,0356)***	[1,0039 ; 1,1434]
DF 2012	1,0716	(0,0355)***	[1,0019 ; 1,1412]
DF 2013	1,0682	(0,0354)***	[0,9988 ; 1,1376]
DF 2014	1,0655	(0,0354)***	[0,9962 ; 1,1348]
DF 2015	1,0628	(0,0353)***	[0,9936 ; 1,1321]
DF 2016	1,0794	(0,0353)***	[1,0102 ; 1,1485]
DF 2017	1,0766	(0,0352)***	[1,0076 ; 1,1456]
DF 2018	1,0794	(0,0353)***	[1,0103 ; 1,1486]

Tabela 9 - continuação

ES 2010	1,0384	(0,0216)***	[0,9962 ; 1,0807]
ES 2011	1,0438	(0,0215)***	[1,0017 ; 1,0859]
ES 2012	1,0417	(0,0215)***	[0,9995 ; 1,0838]
ES 2013	1,0384	(0,0214)***	[0,9964 ; 1,0804]
ES 2014	1,0360	(0,0214)***	[0,9940 ; 1,0780]
ES 2015	1,0337	(0,0214)***	[0,9917 ; 1,0758]
ES 2016	1,0312	(0,0214)***	[0,9892 ; 1,0732]
ES 2017	1,0288	(0,0214)***	[0,9867 ; 1,0708]
ES 2018	1,0309	(0,0213)***	[0,9892 ; 1,0726]
GO 2010	0,7783	(0,0306)***	[0,7184 ; 0,8383]
GO 2011	0,7846	(0,0305)***	[0,7248 ; 0,8443]
GO 2012	0,7826	(0,0305)***	[0,7229 ; 0,8423]
GO 2013	0,7806	(0,0304)***	[0,7210 ; 0,8402]
GO 2014	0,7784	(0,0304)***	[0,7189 ; 0,8379]
GO 2015	0,7763	(0,0303)***	[0,7168 ; 0,8357]
GO 2016	0,7739	(0,0304)***	[0,7143 ; 0,8335]
GO 2017	0,7717	(0,0304)***	[0,7121 ; 0,8312]
GO 2018	0,7747	(0,0304)***	[0,7152 ; 0,8343]
MA 2010	0,9594	(0,0278)***	[0,9049 ; 1,0138]
MA 2011	0,9649	(0,0277)***	[0,9106 ; 1,0191]
MA 2012	0,9627	(0,0277)***	[0,9085 ; 1,0169]
MA 2013	0,9600	(0,0276)***	[0,9058 ; 1,0142]
MA 2014	0,9576	(0,0276)***	[0,9034 ; 1,0118]
MA 2015	0,9552	(0,0276)***	[0,9011 ; 1,0094]
MA 2016	0,9527	(0,0276)***	[0,8985 ; 1,0068]
MA 2017	0,9504	(0,0276)***	[0,8963 ; 1,0045]
MA 2018	0,9523	(0,0276)***	[0,8983 ; 1,0064]
MG 2010	0,7971	(0,0173)***	[0,7631 ; 0,8310]
MG 2011	0,8032	(0,0173)***	[0,7693 ; 0,8370]
MG 2012	0,8012	(0,0173)***	[0,7673 ; 0,8350]
MG 2013	0,7990	(0,0172)***	[0,7653 ; 0,8328]
MG 2014	0,7968	(0,0172)***	[0,7630 ; 0,8305]
MG 2015	0,7946	(0,0172)***	[0,7609 ; 0,8284]
MG 2016	0,7922	(0,0172)***	[0,7585 ; 0,8260]
MG 2017	0,7899	(0,0172)***	[0,7561 ; 0,8236]
MG 2018	0,7925	(0,0172)***	[0,7588 ; 0,8263]
MS 2010	0,9228	(0,0541)***	[0,8168 ; 1,0289]
MS 2011	0,9286	(0,0539)***	[0,8229 ; 1,0343]
MS 2012	0,9266	(0,0540)***	[0,8207 ; 1,0324]
MS 2013	0,9241	(0,0539)***	[0,8186 ; 1,0297]
MS 2014	0,9219	(0,0539)***	[0,8162 ; 1,0276]
MS 2015	0,9197	(0,0539)***	[0,8141 ; 1,0253]
MS 2016	0,9173	(0,0538)***	[0,8118 ; 1,0228]
MS 2017	0,9149	(0,0538)***	[0,8095 ; 1,0203]
MS 2018	0,9171	(0,0537)***	[0,8118 ; 1,0224]

Tabela 9 - continuação

MT 2010	0,7655	(0,0425)***	[0,6822 ; 0,8489]
MT 2011	0,7716	(0,0424)***	[0,6885 ; 0,8548]
MT 2012	0,7697	(0,0424)***	[0,6865 ; 0,8528]
MT 2013	0,7677	(0,0423)***	[0,6848 ; 0,8507]
MT 2014	0,7656	(0,0423)***	[0,6827 ; 0,8484]
MT 2015	0,7635	(0,0422)***	[0,6807 ; 0,8463]
MT 2016	0,7611	(0,0422)***	[0,6784 ; 0,8438]
MT 2017	0,7589	(0,0422)***	[0,6762 ; 0,8416]
MT 2018	0,7610	(0,0420)***	[0,6786 ; 0,8434]
PA 2010	1,0249	(0,0280)***	[0,9701 ; 1,0797]
PA 2011	1,0300	(0,0279)***	[0,9753 ; 1,0846]
PA 2012	1,0278	(0,0278)***	[0,9732 ; 1,0824]
PA 2013	1,0250	(0,0278)***	[0,9705 ; 1,0795]
PA 2014	1,0226	(0,0278)***	[0,9682 ; 1,0771]
PA 2015	1,0203	(0,0278)***	[0,9659 ; 1,0747]
PA 2016	1,0177	(0,0277)***	[0,9634 ; 1,0719]
PA 2017	1,0154	(0,0277)***	[0,9612 ; 1,0696]
PA 2018	1,0170	(0,0277)***	[0,9628 ; 1,0713]
PB 2010	0,7409	(0,0450)***	[0,6526 ; 0,8291]
PB 2011	0,7471	(0,0449)***	[0,6591 ; 0,8352]
PB 2012	0,7450	(0,0449)***	[0,6569 ; 0,8330]
PB 2013	0,7431	(0,0448)***	[0,6553 ; 0,8310]
PB 2014	0,7408	(0,0448)***	[0,6530 ; 0,8287]
PB 2015	0,7386	(0,0448)***	[0,6508 ; 0,8264]
PB 2016	0,7362	(0,0448)***	[0,6483 ; 0,8241]
PB 2017	0,7338	(0,0448)***	[0,6460 ; 0,8217]
PB 2018	0,7363	(0,0448)***	[0,6485 ; 0,8240]
PE 2010	1,0303	(0,0264)***	[0,9786 ; 1,0821]
PE 2011	1,0356	(0,0263)***	[0,9839 ; 1,0872]
PE 2012	1,0334	(0,0263)***	[0,9818 ; 1,0850]
PE 2013	1,0305	(0,0263)***	[0,9790 ; 1,0820]
PE 2014	1,0281	(0,0263)***	[0,9766 ; 1,0796]
PE 2015	1,0258	(0,0263)***	[0,9744 ; 1,0773]
PE 2016	1,0233	(0,0263)***	[0,9718 ; 1,0747]
PE 2017	1,0209	(0,0263)***	[0,9694 ; 1,0724]
PE 2018	1,0228	(0,0262)***	[0,9713 ; 1,0742]
PI 2010	0,7439	(0,0533)***	[0,6394 ; 0,8485]
PI 2011	0,7504	(0,0533)***	[0,6459 ; 0,8549]
PI 2012	0,7483	(0,0533)***	[0,6439 ; 0,8528]
PI 2013	0,7468	(0,0532)***	[0,6425 ; 0,8510]
PI 2014	0,7445	(0,0531)***	[0,6404 ; 0,8486]
PI 2015	0,7423	(0,0531)***	[0,6382 ; 0,8464]
PI 2016	0,7400	(0,0531)***	[0,6358 ; 0,8441]
PI 2017	0,7376	(0,0531)***	[0,6335 ; 0,8417]
PI 2018	0,7398	(0,0530)***	[0,6358 ; 0,8438]

Tabela 9 - continuação

PR 2010	0,8230	(0,0241)***	[0,7758 ; 0,8703]
PR 2011	0,8292	(0,0241)***	[0,7820 ; 0,8764]
PR 2012	0,8273	(0,0241)***	[0,7801 ; 0,8745]
PR 2013	0,8251	(0,0240)***	[0,7780 ; 0,8721]
PR 2014	0,8229	(0,0240)***	[0,7759 ; 0,8700]
PR 2015	0,8208	(0,0240)***	[0,7738 ; 0,8679]
PR 2016	0,8185	(0,0240)***	[0,7715 ; 0,8655]
PR 2017	0,8162	(0,0240)***	[0,7692 ; 0,8632]
PR 2018	0,8192	(0,0240)***	[0,7722 ; 0,8662]
RJ 2010	1,1048	(0,0169)***	[1,0718 ; 1,1379]
RJ 2011	1,1095	(0,0168)***	[1,0766 ; 1,1424]
RJ 2012	1,1073	(0,0168)***	[1,0744 ; 1,1402]
RJ 2013	1,1045	(0,0168)***	[1,0715 ; 1,1375]
RJ 2014	1,1021	(0,0169)***	[1,0690 ; 1,1351]
RJ 2015	1,0998	(0,0169)***	[1,0667 ; 1,1329]
RJ 2016	1,0973	(0,0169)***	[1,0641 ; 1,1304]
RJ 2017	1,0948	(0,0169)***	[1,0617 ; 1,1280]
RJ 2018	1,0960	(0,0168)***	[1,0631 ; 1,1289]
RN 2010	0,7822	(0,0496)***	[0,6849 ; 0,8795]
RN 2011	0,7884	(0,0495)***	[0,6913 ; 0,8856]
RN 2012	0,7864	(0,0495)***	[0,6894 ; 0,8835]
RN 2013	0,7843	(0,0494)***	[0,6875 ; 0,8812]
RN 2014	0,7821	(0,0494)***	[0,6853 ; 0,8789]
RN 2015	0,7800	(0,0494)***	[0,6832 ; 0,8767]
RN 2016	0,7775	(0,0493)***	[0,6808 ; 0,8742]
RN 2017	0,7752	(0,0493)***	[0,6785 ; 0,8719]
RN 2018	0,7778	(0,0493)***	[0,6812 ; 0,8744]
RO 2010	0,9020	(0,0560)***	[0,7922 ; 1,0118]
RO 2011	0,9077	(0,0559)***	[0,7982 ; 1,0172]
RO 2012	0,9058	(0,0558)***	[0,7964 ; 1,0152]
RO 2013	0,9032	(0,0556)***	[0,7943 ; 1,0121]
RO 2014	0,9010	(0,0555)***	[0,7921 ; 1,0098]
RO 2015	0,8989	(0,0555)***	[0,7900 ; 1,0077]
RO 2016	0,8966	(0,0555)***	[0,7877 ; 1,0055]
RO 2017	0,8943	(0,0555)***	[0,7855 ; 1,0031]
RO 2018	0,8977	(0,0556)***	[0,7888 ; 1,0066]
RR 2010	1,0197	(0,0377)***	[0,9459 ; 1,0936]
RR 2011	1,0250	(0,0374)***	[0,9516 ; 1,0983]
RR 2012	1,0223	(0,0374)***	[0,9490 ; 1,0957]
RR 2013	1,0194	(0,0375)***	[0,9459 ; 1,0929]
RR 2014	1,0165	(0,0374)***	[0,9431 ; 1,0899]
RR 2015	1,0139	(0,0374)***	[0,9406 ; 1,0871]
RR 2016	1,0112	(0,0376)***	[0,9375 ; 1,0849]
RR 2017	1,0084	(0,0375)***	[0,9348 ; 1,0820]
RR 2018	1,0100	(0,0378)***	[0,9360 ; 1,0841]

Tabela 9 - continuação

RS 2010	0,7402	(0,0196)***	[0,7018 ; 0,7785]
RS 2011	0,7466	(0,0196)***	[0,7082 ; 0,7850]
RS 2012	0,7447	(0,0196)***	[0,7064 ; 0,7831]
RS 2013	0,7427	(0,0195)***	[0,7044 ; 0,7809]
RS 2014	0,7405	(0,0195)***	[0,7022 ; 0,7787]
RS 2015	0,7384	(0,0195)***	[0,7001 ; 0,7766]
RS 2016	0,7360	(0,0195)***	[0,6978 ; 0,7742]
RS 2017	0,7337	(0,0195)***	[0,6955 ; 0,7719]
RS 2018	0,7369	(0,0194)***	[0,6989 ; 0,7750]
SC 2010	0,7539	(0,0237)***	[0,7074 ; 0,8005]
SC 2011	0,7603	(0,0237)***	[0,7138 ; 0,8067]
SC 2012	0,7584	(0,0237)***	[0,7120 ; 0,8048]
SC 2013	0,7567	(0,0236)***	[0,7105 ; 0,8029]
SC 2014	0,7546	(0,0235)***	[0,7084 ; 0,8008]
SC 2015	0,7527	(0,0235)***	[0,7066 ; 0,7988]
SC 2016	0,7505	(0,0235)***	[0,7044 ; 0,7966]
SC 2017	0,7483	(0,0235)***	[0,7023 ; 0,7944]
SC 2018	0,7509	(0,0235)***	[0,7049 ; 0,7969]
SE 2010	0,8584	(0,0519)***	[0,7566 ; 0,9601]
SE 2011	0,8641	(0,0518)***	[0,7626 ; 0,9657]
SE 2012	0,8621	(0,0518)***	[0,7606 ; 0,9635]
SE 2013	0,8597	(0,0516)***	[0,7586 ; 0,9608]
SE 2014	0,8573	(0,0515)***	[0,7563 ; 0,9584]
SE 2015	0,8552	(0,0515)***	[0,7542 ; 0,9561]
SE 2016	0,8527	(0,0515)***	[0,7518 ; 0,9536]
SE 2017	0,8503	(0,0514)***	[0,7494 ; 0,9511]
SE 2018	0,8525	(0,0514)***	[0,7517 ; 0,9533]
SP 2010	0,8752	(0,0147)***	[0,8463 ; 0,9041]
SP 2011	0,8808	(0,0147)***	[0,8520 ; 0,9095]
SP 2012	0,8786	(0,0147)***	[0,8499 ; 0,9074]
SP 2013	0,8761	(0,0146)***	[0,8474 ; 0,9048]
SP 2014	0,8737	(0,0146)***	[0,8450 ; 0,9024]
SP 2015	0,8715	(0,0146)***	[0,8428 ; 0,9002]
SP 2016	0,8690	(0,0147)***	[0,8403 ; 0,8978]
SP 2017	0,8667	(0,0147)***	[0,8379 ; 0,8954]
SP 2018	0,8686	(0,0146)***	[0,8400 ; 0,8972]
TO 2010	0,6407	(0,0560)***	[0,5309 ; 0,7506]
TO 2011	0,6473	(0,0559)***	[0,5378 ; 0,7569]
TO 2012	0,6453	(0,0558)***	[0,5358 ; 0,7548]
TO 2013	0,6436	(0,0557)***	[0,5344 ; 0,7528]
TO 2014	0,6413	(0,0557)***	[0,5320 ; 0,7505]
TO 2015	0,6390	(0,0557)***	[0,5298 ; 0,7483]
TO 2016	0,6365	(0,0557)***	[0,5273 ; 0,7458]
TO 2017	0,6341	(0,0557)***	[0,5250 ; 0,7432]
TO 2018	0,6367	(0,0556)***	[0,5277 ; 0,7457]

* p<0.05; ** p<0.01; *** p<0.001

O modelo (5) também foi estimado enquanto modelo Efeitos Aleatórios (EA) com erros-padrão robustos para erros não esféricos em nível de município, que é um procedimento MQO factível, no qual se admite que c_s e η_t são não correlacionados com pop_{it} . A contraposição das estimações caso a caso, no entanto, não demonstrou diferenças significativas nos resultados.

5.2. Lei de Gibrat

A estimação da equação (7) utilizando os dados do painel nacional com 50.085 observações em 9 períodos, referentes à população anual dos 5.565 municípios brasileiros entre 2010 e 2018 revelou os seguintes resultados acerca da aderência da lei de Gibrat aos dados da distribuição da população brasileira nas cidades.

Na primeira estimação, o parâmetro β foi mantido constante entre os estados da federação a fim de se obter um resultado nacional. Nesse caso, β é estimado em 1,0261 pelo *system GMM* com intervalo de 95% de confiança variando de 1,0237 a 1,02849. Rejeita-se, então, a esse nível de significância, a hipótese de que o coeficiente é igual a unidade e de que o crescimento populacional dos municípios brasileiros independe do tamanho da sua população. Com $|\beta| > 1$, o processo é explosivo e a população diverge da média, significando que o crescimento esperado é maior nas cidades maiores. O teste AR(1) de Arellano-Bond rejeita, com p-valor de 0,038, a hipótese nula de ausência de correlação serial de primeira ordem da primeira diferença dos resíduos, conforme esperado por construção. O teste AR(2), por sua vez, com p-valor de 0,551, não rejeita a hipótese nula de ausência de correlação serial de segunda ordem dos resíduos. O método utilizou 62 variáveis instrumentais, cuja validade não foi reprovada nos testes de Sargan e de Hansen de sobreidentificação das restrições. A hipótese nula de insignificância conjunta dos efeitos fixos das unidades federativas foi rejeitada no teste F com p-valor zero.

Tabela 10. Modelo (7) estimado por *system GMM* com $\beta \equiv \beta$ e $\delta \equiv g_s$

	β	σ	IC(α) 95%
Brasil	1,026127	(0,001208)***	[1,0237 ; 1,0285]

* p<0.05; ** p<0.01; *** p<0.001

A estimação com o parâmetro β potencialmente diferente entre os estados é apresentada na Tabela 5.6. Em relação à análise do modelo *system GMM*, o teste AR(1) de Arellano-Bond novamente rejeita, com p-valor de 0,038, a hipótese nula de ausência de correlação serial de primeira ordem da primeira diferença dos resíduos, enquanto que o teste AR(2), por sua vez, com p-valor de 0,551, não rejeita a hipótese nula de ausência de correlação serial de segunda ordem dos resíduos. O método utilizou 907 variáveis instrumentais, cuja validade não foi reprovada nos testes de Sargan e de Hansen de sobreidentificação das restrições. A hipótese nula de insignificância conjunta dos efeitos fixos das unidades federativas foi rejeitada no teste F com p-valor zero.

Tabela 11. Modelo (7) estimado por *system GMM* com $\beta \equiv \beta_s$

	β	σ	IC(α) 95%
AC	1,0309	(0,003718)***	[1,0236 ; 1,0382]
AL	1,0082	(0,0022)***	[1,0040 ; 1,0126]
AM	1,0827	(0,00936)***	[1,0643 ; 1,1011]
AP	0,9969	(0,00334)***	[0,9903 ; 1,0035]
BA	1,0300	(0,00429)***	[1,0216 ; 1,0385]
CE	1,0140	(0,00266)***	[1,0095 ; 1,0201]
DF	0,8720	(5,04e-)***	[0,8721 ; 0,8721]
ES	1,0183	(0,00385)***	[1,0107 ; 1,0259]
GO	1,0279	(0,00386)***	[1,0204 ; 1,0355]
MA	1,0013	(0,02046)***	[0,9612 ; 1,0415]
MG	1,0197	(0,00164)***	[1,0165 ; 1,0230]
MS	1,0239	(0,00445)***	[1,0152 ; 1,0327]
MT	1,0210	(0,00406)***	[1,0131 ; 1,0290]
PA	1,0212	(0,00669)***	[1,0081 ; 1,0343]
PB	1,0137	(0,00769)***	[0,9986 ; 1,0288]
PE	1,0251	(0,00646)***	[1,0124 ; 1,0378]
PI	1,0007	(0,0015)***	[0,9978 ; 1,0037]
PR	1,0216	(0,0023)***	[1,0171 ; 1,0262]
RJ	1,0209	(0,01123)***	[0,9989 ; 1,0430]
RN	1,0227	(0,00571)***	[1,0116 ; 1,0340]
RO	1,0374	(0,00881)***	[1,0202 ; 1,0548]
RR	0,9489	(0,06759)***	[0,8164 ; 1,0815]
RS	1,0228	(0,0019)***	[1,0189 ; 1,0267]
SC	1,0186	(0,00145)***	[1,0158 ; 1,0215]
SE	1,0164	(0,00581)***	[1,0050 ; 1,0278]
SP	1,0146	(0,00154)***	[1,0116 ; 1,0176]
TO	1,0019	(0,00174)***	[0,9986 ; 1,0054]

* p<0.05; ** p<0.01; *** p<0.001

Nos estados do Amapá, Maranhão, Paraíba, Piauí, Rio de Janeiro, Roraima e Tocantins, a lei de Gibrat é aceita e o crescimento populacional dos municípios desses estados é independente do tamanho da população. Os demais estados, entretanto, compoem o grupo majoritário com dezenove unidades da federação em que o coeficiente estimado é maior do que a unidade, significando que o processo é explosivo e a população diverge da média, de modo que o crescimento esperado é maior nas cidades maiores. Nesse grupo, destacam-se o Amazonas, com o maior coeficiente entre todas as unidades federativas, e também Rondônia, com o segundo maior resultado. Apenas no Distrito Federal o coeficiente é menor do que a unidade, o que significaria que o processo é estacionário e a população converge para a média, de modo que o crescimento esperado seria maior nas cidades menores. Contudo, esse resultado não tem significado econômico porque a população do Distrito Federal é contabilizada como um único município na amostra (Brasília).

6. Conclusão

O estudo da distribuição da população de um país ao longo do seu território e das suas cidades é o primeiro passo para compreender a hierarquia dos sistemas urbanos e as suas implicações econômicas - distribuição do emprego e dos setores da economia por região; salários; potencial de inovação das cidades como resultado do volume de pesquisa em cada localidade; variedade de bens e serviços; organização e preços do mercado imobiliário. Somente a partir desse primeiro passo, é possível conceber políticas públicas que promovam a otimização do processo de crescimento dos centros urbanos com a finalidade de obter uma solução ótima para a alocação dos recursos escassos que maximize o resultado positivo das economias de escala, mas que minimize os impactos negativos provenientes da concentração demográfica (poluição, congestionamentos, exclusão social, violência, problemas de saneamento e dificuldade de acesso aos serviços públicos).

Partindo desse princípio, o objetivo deste trabalho foi verificar a aplicabilidade, para o caso dos municípios brasileiros, de duas regularidades empíricas fundamentais da economia urbana: as leis de Zipf e de Gibrat. Para esse propósito, foram utilizados dados em painel referentes ao total de habitantes dos 5.565 municípios existentes no país em 2010 numa frequência anual até o ano de 2018.

A lei de Zipf, que enuncia que o tamanho da população das cidades está relacionada com a sua posição no ranking das maiores cidades através de uma distribuição de Pareto, foi estimada por Mínimos Quadrados Ordinários (MQO), levando em consideração os efeitos fixos de tempo e de unidade federativa e erros-padrão robustos a erros não esféricos em nível de município. A finalidade disto foi obter uma estimativa para o expoente α de Pareto que pudesse indicar não apenas se a lei de Zipf se aplica ($\alpha = 1$), mas também se a distribuição da população é mais igual ($\alpha > 1$), com as pessoas pouco concentradas nas maiores cidades, ou mais desigual ($\alpha < 1$), com grande concentração de habitantes nas maiores cidades.

Para averiguar a lei de Gibrat, por sua vez, que enuncia que o crescimento populacional de uma cidade não depende do tamanho de sua população, um modelo autoregressivo dinâmico foi estimado pelo *Generalized Method of Moments (GMM)* levando em consideração os efeitos fixos de unidade federativa e erros-padrão robustos a erros não esféricos em nível de município com a finalidade de obter uma estimativa do parâmetro β que pudesse indicar não apenas se a lei de Gibrat se aplica

($|\beta_1| = 1$), mas também se o processo de crescimento das cidades é explosivo ($|\beta_1| > 1$), com maior crescimento nas maiores cidades, ou estacionário, ($|\beta_1| < 1$), com maior crescimento nas menores cidades.

Em relação aos resultados, a lei de Zipf foi confirmada para treze das vinte e sete unidades da federação. Para as quatorze restantes, o resultado apontou distribuição desigual e concentração urbana nas grandes cidades, com exceção de Rio de Janeiro, onde a distribuição é mais igual do que o esperado com base na lei de Zipf. O resultado nacional também acusou forte concentração demográfica nas maiores cidades. A lei de Gibrat foi confirmada para apenas sete unidades da federação. Na grande maioria dos estados, o processo de crescimento populacional em curso é explosivo e contribui para a concentração demográfica, resultado similar ao encontrado quando considerados em conjunto os dados de todo o território nacional.

Intuitivamente, é possível conjecturar como os resultados da lei de Gibrat impactam o coeficiente α da lei de Zipf. Quando o processo de crescimento das cidades é explosivo ($|\beta_1| > 1$), a população diverge da média e as cidades grandes crescem mais rápido. Isso favorece a concentração urbana e torna a distribuição da população mais desigual, fazendo com que o coeficiente α se torne menor. No caso oposto, quando o processo de crescimento das cidades é estacionário ($|\beta_1| < 1$), a população converge para a média e as cidades pequenas crescem mais rápido. Isso favorece a desconcentração urbana e torna a distribuição da população mais igual, fazendo com que o coeficiente α se torne maior. De fato, nota-se que, nos dezenove estados nos quais o processo de crescimento populacional é explosivo, o coeficiente α apresentou tendência de queda no período considerado.

Em linhas gerais, pode-se dizer que os resultados deste trabalho descrevem o Brasil como um país em que a população, em boa parte dos estados, está distribuída de maneira mais desigual ou mais concentrada do que a prevista pela regularidade empírica constatada internacionalmente na forma da lei de Zipf. Ainda, este trabalho indica que o processo de crescimento demográfico em curso nas cidades, como constatado na grande maioria dos estados, reforça essa característica na medida em que ocorre no sentido de aumentar a concentração de pessoas nos maiores centros urbanos.

Como se sabe, as explicações mais amplamente aceitas na literatura para justificar a concentração populacional dizem respeito às economias de escala decorrentes das aglomerações urbanas: transbordamentos tecnológicos, minimização dos custos de transporte, maior oferta de fatores de produção *nontradables* e redução da probabilidade de escassez de mão de obra para as empresas e de desemprego para os trabalhadores. Mais recentemente, no entanto, novos elementos, como o progresso das tecnologias de informação e de transportes de alta velocidade, têm influenciado a lógica da distribuição populacional no sentido de atenuar essas forças aglomerativas na medida em que as distâncias geográficas perdem importância. Como o Brasil investe pouco em infraestrutura para a disseminação do uso de tecnologias de informação e de meios de transportes modernos, os resultados obtidos no presente estudo podem estar associados a isto, o que deve ser investigado em trabalhos futuros. Nesse sentido, políticas de investimento em infraestrutura poderiam servir para diminuir desigualdades regionais e espalhar o desenvolvimento no território nacional promovendo a racionalização do crescimento econômico com melhor aproveitamento das economias de escala e das vantagens comparativas de cada localidade. Ao mesmo tempo, políticas públicas dessa natureza contribuiriam sobremaneira para frear o inchaço das grandes cidades e para atenuar o impacto das externalidades negativas provenientes da concentração demográfica exagerada no bem-estar social (poluição, congestionamentos, exclusão social, violência, problemas de saneamento e dificuldade de acesso aos serviços públicos).

Referências

- ARELLANO, M., BOND, S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58: 277–297, 1991.
- ARELLANO, M., BOVER, O. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* 68: 29–51, 1995.
- AUERBACH, F. Das Gesetz der Bevolkerungs Konzentration. *Petermanns Geogr Mitt*, 1913.
- BAENINGER, R. Rotatividade Migratória: um novo olhar para as migrações internas no Brasil. *Rev. Inter. Mob. Hum.*, Brasília, Ano XX, Nº 39, p. 77-100, jul./dez. 2012.
- BERRY, B.J., OKULICZ-KOZARYN, B.J., The city size distribution debate: Resolution for US urban regions and megalopolitan areas, *Cities* 29 (2012) S17–S23.
- BLUNDELL, R., BOND, S. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87: 115–143, 1998.
- CHAMPERNOWNE, D.G. A model of income distribution. *Econ J* 63(250):318–351, 1953.
- CORDOBA, J. On the distribution of city sizes. *Urban/Regional*. 0302002, EconWPA, 2003.
- CORDOBA, J. A generalized Gibrat's law. *Int Econ Rev* 49(4):1463–1468, 2008.
- FUJITA, M., KRUGMAN, P., VENABLES, A.J. *The Spatial Economy*. MIT Press, 2002.
- GABAIX, X. Zipf's law for cities: an explanation. *Q J Econ* 114(3):739–767, 1999.
- GABAIX, X. Power Laws in Economics: An Introduction. *Journal of Economic Perspectives*, v. 30, n. 1, p. 185-206, 2016.
- GOMES, G.M., DOWELL, M., CRISTINA, M. Descentralização política, federalismo fiscal e criação de municípios: o que é mau para o econômico nem sempre é bom para o social. *Texto para Discussão Nº 706*, IPEA, 2000.
- GIBRAT, R. *Les Inégalités Economiques*. Libraire du Recueil Siray, Paris, 1931.
- KRUGMAN, P. Increasing Returns and Economic Geography. *The Journal of Political Economy*, v. 99, Issue 3, 1991.
- LUCKSTEAD, J., DEVADOSS, S.. Do the world's largest cities follow Zipf's and Gibrat's laws? *Economics Letters*, v. 125, n. 2, p. 182–186, 2014.
- MODICA, M., REGGIANI, A., NIJKAMP, P. Methodological advances in Gibrat's and Zipf's laws: A comparative empirical study on the evolution of urban systems, in: H.

- SHIBUSAWA, K., SAKURAI, T., MIZUNOYA, S. UCHIDA (Eds.), Socioeconomic Environmental Policies and Evaluations in Regional Science: Essays in Honor of Yoshiro Higano, Springer, Singapore, pp. 37–59, 2017.
- ROLNIK, R., A Construção de uma política fundiária e de planejamento urbano para o país: avanços e desafios. Políticas sociais – acompanhamento e análise. IPEA, 2006.
- ROSEN KT, RESNICK M. The size distribution of cities: an examination of the Pareto law and primacy. *J Urban Econ* 8(2):165–186, 1980.
- SIMON, H. On a class of skew distribution functions. *Biometrika* 42(3/4):425–440, 1955.
- SOO, K.T. Zipf's law for cities: a cross-country investigation. *Reg Sci Urban Econ* 35(3): 239–263, 2005.
- STANLEY, M. H. R. et al. Zipf plots and the size distribution of firms. *Economics Letters*. v. 49, n. 4, p. 453-457, 1995.
- STEINDL, J. Size distributions in economics, in *International Encyclopedia of the Social Sciences*, vol 14. Macmillan and the Free Press, New York, 1968.