

## DESENVOLVIMENTO ECONÔMICO E REGIONAL

### A distribuição do tamanho dos municípios do Paraná e a Lei de Zipf

Hermes Yukio Higachi\*

Thierry Molnar Prates\*\*

**RESUMO** - Este artigo estuda a evolução da distribuição do tamanho das cidades paranaenses e testa a validade da Lei de Zipf, estimando o expoente de Pareto ( $\beta$ ) por mínimos quadrados ordinários no período 1970-2000. Para verificar se a distribuição de Pareto exibe um comportamento côncavo ou convexo estima-se também uma versão não-linear. Os resultados indicam que a Lei de Zipf ou da regra da ordem do tamanho pode ser válida apenas para o ano 2000. Por sua vez, a evolução das cidades paranaenses pode ser decomposta em três períodos: no período 1970-1980, ocorre o aumento da desigualdade; no período 1980-1991 ocorre a redução da desigualdade; e no período 1991-2000, constata-se novamente o processo de concentração populacional. Outro resultado interessante é que quando são impostas restrições sobre a amostra, os coeficientes de Pareto são maiores e tendem ao declínio persistente no período 1970-2000, caracterizando um crescimento da desigualdade e da divergência ao longo do tempo e, ao mesmo tempo, indicando que a desigualdade é menor entre cidades médias e grandes do que quando se considera as cidades menores. Uma implicação é que a estrutura urbana paranaense contradiz uma das principais proposições da Nova Geografia Econômica que atribui predominância das forças centrífugas sobre as forças centrípetas, sugerindo a relevância da formulação de políticas de fortalecimento das cidades menores e médias, como a criação de infra-estrutura física, econômica, social e cultural.

Palavras-chave: Cidades do Paraná. Distribuição de Pareto. Lei de Zipf. Desigualdade e concentração populacional.

## 1 INTRODUÇÃO

A maneira pela qual a população se distribui pelos espaços territoriais não é aleatória. De fato, existe uma tendência histórica importante para a aglomeração em certos espaços, geralmente cidades, principalmente decorrente de aspectos geográficos como rios, áreas costeiras e outros incentivos comerciais e empresariais como fonte de matérias-primas e acesso a mercados.

---

\* Doutor pelo Instituto de Economia da Universidade Estadual de Campinas. É Professor Associado do Departamento de Economia da Universidade Estadual de Ponta Grossa. Endereço eletrônico: hhigachi@uepg.br.

\*\* Doutor em Economia pela Universidade Federal do Paraná. É Professor Adjunto do Departamento de Economia da Universidade Estadual de Ponta Grossa. Endereço eletrônico: thierry@uepg.br.

O interesse a respeito das aglomerações urbanas sempre permeou diversas disciplinas da economia, em especial a economia regional. Entender como algumas cidades crescem rapidamente, enquanto outras ficam estagnadas e até diminuem de tamanho, têm instigado pesquisadores desde o início do século passado até nossos dias, provocando um refinamento importante dos métodos e teorias que tratam o assunto.

Os motivos pelos quais as pessoas se aglomeram nas cidades, em um mundo de transações espaciais e custos de transporte declinantes, de acordo com Storper e Venables (2001), têm sido estudados através de três vertentes principais da economia regional: a) os efeitos de encadeamentos de insumo-produto para trás (fornecedores) e para frente (mercados); b) a aglomeração de trabalhadores, que buscam localidades onde a rotatividade é alta e a especialização é grande, e onde possam maximizar a duração de suas carreiras e salários; e c) interações localizadas, que promovem a inovação tecnológica, como transbordamentos tecnológicos, cooperação, aprendizado tecnológico decorrentes da proximidade das firmas.

A Nova Geografia Econômica (NGE) oferece explicações interessantes sobre como as pessoas tomam decisões de se fixar em certas cidades, conduzidas principalmente por retornos crescentes de escala, decorrentes de menores custos de transporte, maior qualificação de mão-de-obra e outros efeitos externos positivos. Contudo, a regra ordem do tamanho ou lei de Zipf<sup>34</sup> não é contemplada. Deficiência esta que Fujita et al (2002) admitem e tratam como um desafio para o desenvolvimento da teoria.

Este artigo tem como objetivo a verificação da Lei de Zipf para os municípios paranaenses através da estimação do expoente da distribuição de Pareto. Dividido em três seções além desta introdução, temos na segunda seção a caracterização do modelo de distribuição do tamanho das cidades, em seguida os resultados empíricos para o estado do Paraná, e por fim, as principais conclusões e implicações para políticas públicas são apresentadas na quarta seção.

## 2 MODELO DE DISTRIBUIÇÃO DO TAMANHO DAS CIDADES

Auerbach, em 1913, iniciou a trajetória dos modelos de distribuição do tamanho das cidades, estabelecendo que esta segue a distribuição de Pareto:

$$y = \alpha x^{-\beta} \quad (1)$$

---

<sup>34</sup> Cabe destacar que a distribuição de Pareto e a lei de Zipf têm sido aplicados de modo crescente para modelar fenômenos econômicos e sociais emergentes de sistemas adaptativos e complexos como é caso da distribuição do tamanho de firmas.

em que:  $x$  é população de uma determinada cidade;  $y$  é número de cidades com população maior que  $x$ ;  $\alpha$  é uma constante; e  $\beta$  é expoente de Pareto. Anos depois, Zipf aprofundou essa afirmação indicando que o expoente  $\beta$  é igual a 1 (Zipf, 1949). Essa nova premissa ficou conhecida como “Lei de Zipf”, ou regra da ordem de tamanho, que implica que o produto da população de qualquer cidade multiplicado pela sua posição na ordenação da região será igual à população da maior cidade. Isto significa dizer que a segunda maior cidade terá a metade da população da maior, a terceira terá um terço, e assim por diante.

A regra da ordem do tamanho das cidades surge da distribuição acumulada de Pareto, em que a probabilidade de encontrar uma cidade com população menor do que  $x$  é dada por:

$$\text{Prob}(X \leq x) = F(x) = 1 - \frac{\alpha}{x^\beta} \quad (2)$$

com efeito, a probabilidade de encontrar uma cidade com população maior do que  $x$  é:

$$\text{Prob}(X > x) = 1 - F(x) = \frac{\alpha}{x^\beta} \quad (3)$$

Seja  $y = 1 - F(x)$ , então, tem-se (1) e tirando os logaritmos da equação (3) chega-se a seguinte expressão:

$$\log y = \log \alpha - \beta \log x \quad (4)$$

O modelo econométrico estimado neste artigo é a equação (4) com o acréscimo do termo erro que segue uma distribuição normal com média zero e variância constante:

$$\log y_{it} = \log \alpha_{it} + \beta \log x_{it} + \varepsilon_{it} \quad (5)$$

em que  $i = 1, \dots, n$  representam os municípios,  $t = 1970, 1980, 1991$  e  $200$  representam os períodos em que a equação é estimada,  $x$  é a população de uma determinada cidade,  $y$  é número de municípios com população maior do que  $x$  representada pelo *rank* decrescente de municípios. O parâmetro  $\beta$  ou coeficiente de Pareto é sempre negativo, visto que quanto maior for a população de um município menor é a probabilidade de encontrar um município com população maior do que a sua população ou de modo equivalente menor será seu *rank* no ordenado dos municípios. Ademais, quanto maior for o valor do parâmetro  $\beta$ , menor será o grau de desigualdade da distribuição do tamanho dos municípios paranaenses, enquanto que quanto menor o valor do parâmetro  $\beta$ , maior será o grau de desigualdade e de concentração populacional. Se  $\beta = 1$ , tem-se a regra da ordem do tamanho e  $\alpha$  representa a população do maior município paranaense, no caso, Curitiba.

Para verificar a possibilidade de uma relação côncava ou convexa entre a ordem e o tamanho dos municípios, acrescentam-se termos não-lineares na forma de potências de  $\log x$ , do tipo:

$$\log y_{it} = \log \alpha_{it} + \beta \log x_{it} + \delta (\log x)^2 + \varepsilon_{it} \quad (6)$$

Se a distribuição de Pareto exibe um comportamento côncavo ( $\delta < 0$ ), indica que as cidades médias possuem níveis populacionais superiores ao ideal previsto pela Lei de Zipf, ao passo que se exibir um comportamento convexo ( $\delta > 0$ ), indica que as cidades médias possuem níveis populacionais inferiores ao ideal para a Lei de Zipf. Outra interpretação é que se exibir um comportamento côncavo o crescimento é convergente e as menores cidades crescem mais do que as maiores cidades e vice-versa.

### 3 RESULTADOS EMPÍRICOS PARA O PARANÁ

As equações (5) e (6) são estimadas por mínimos quadrados ordinários. Como o teste de White com termos cruzados rejeita a hipótese de homocedasticidade para todas as regressões, foram usados os erros-padrões robustos de White. O modelo da distribuição de Pareto apresenta um grau de ajuste aos dados relativamente elevado, conforme revela o coeficiente de determinação próximo de 90% (Tabela 1). Os resultados da estimação da distribuição de Pareto para o caso das cidades paranaenses mostra que apenas no ano 2000 existe a possibilidade de ocorrer a Lei de Zipf. A hipótese<sup>35</sup> de  $\alpha = \text{Ln}(\text{População de Curitiba})$  e  $\beta = 1$  é aceita com 95% de confiança.

No início da década de 70, a distribuição do tamanho das cidades paranaenses começa mais igualitária do que prediz a Lei de Zipf. No entanto, ao longo das décadas o coeficiente de Pareto tende a apresentar um resultado inferior a um ( $\beta < 1$ ), caracterizando uma estrutura urbana com cidades bastante assimétricas e polarizadas. Verifica-se também que há três períodos na evolução da distribuição do tamanho das cidades. No período de 1970-1980, ocorre um aumento na desigualdade na distribuição do tamanho das cidades, enquanto no período 1980-1991, ocorre uma desconcentração ou dispersão populacional, provavelmente devido ao aumento do número de municípios. Por último, no período 1991-2000, verifica-se novamente um processo de concentração populacional, embora tenha havido ampliação no número de municípios. A elevação da concentração e da desigualdade é resultante da ação de forças centrípetas (custos de transporte, mercado de trabalho e eficiência

---

<sup>35</sup> No teste empírico da Lei de Zipf foi imposta a restrição linear na matriz de parâmetros ( $R\beta = r$ ) de que o parâmetro  $\alpha = \text{Ln}(\text{População da maior cidade do Paraná})$  e o parâmetro  $\beta = -1$ . A estatística de teste segue a distribuição F.

na transferência de conhecimento) e do processo de urbanização do Paraná ao longo das décadas.

TABELA 1 – RESULTADO DA ESTIMAÇÃO DE EXPOENTES DE PARETO- 1970-2000

Anos	$\alpha$	$\beta$	R <sup>2</sup>	N
1970	15,665 (35,958)*	1,1335 (-25,208)	0,8877	288
1980	13,726 (38,516)	-0,94371 (-25,329)	0,8926	290
1991	13,9216 (54,108)	-0,9635 (-35,885)	0,9367	323
2000	13,7475 (90,33)	-0,9444 (-57,782)	0,9655	399

FONTE: Elaboração própria a partir de dados do IPEA.

NOTA: (\*) Números entre parêntese são as estatísticas t-Student computado pelo erro-padrão robusto de White. Os regressores são significativos ao nível de 1%.

Os resultados da estimação mostram também que quando são impostas restrições sobre a amostra, o coeficiente de Pareto é maior: a desigualdade é menor entre cidades médias e grandes do que quando se acrescenta as cidades menores. Outro fato interessante é que a desigualdade ou concentração populacional em cidades maiores aumentou em todos os níveis, no período 1970-2000. Esses resultados revelam que as cidades paranaenses não seguem as contribuições da nova geografia econômica que postula o predomínio das forças centrífugas (níveis indesejados de congestionamento, grandes distâncias e crime) sobre as forças centrípetas (Tabela 2).

TABELA 2 - EXPOENTES DE PARETO COM RESTRIÇÕES NA AMOSTRA - 1970-2000

Anos	1970	1980	1991	2000
≥50.000 hab.	-1,321	-1,195	-1,152	-1,12
50 maiores	-1,586	-1,361	-1,219	-1,089
≥20.000 hab.	-1,701	-1,44	-1,204	-1,094

FONTE: Elaboração própria a partir de dados do IPEA.

Por sua vez, a estimação da lei de Zipf na forma não-linear mostra que no ano de 1970, a distribuição de Pareto exibe um comportamento côncavo, indicando que as cidades médias possuem níveis populacionais superiores ao ideal para a Lei de Zipf e que o crescimento foi convergente.

No entanto, a evolução do tamanho da distribuição do tamanho das cidades paranaenses no período 1970-2000, em direção ao aumento da desigualdade e da concentração populacional, alterou esse padrão não-linear entre a ordem e o tamanho de cidades. No ano 2000, a distribuição de Pareto exibe um comportamento convexo, indicando que as cidades médias estão subdimensionadas (apesar do crescimento das cidades médias

paranaenses, essas ainda possuem populações inferiores ao ideal para a Lei de Zipf), e que o crescimento foi divergente em que as cidades maiores crescem mais do que as cidades menores.

#### 4 CONCLUSÕES

Os resultados indicam que a Lei de Zipf ou da regra da ordem do tamanho pode ser válida apenas para o ano 2000. Por sua vez, a evolução da distribuição do tamanho das cidades paranaenses pode ser decomposta em três fases: no período 1970-1980 o grau de desigualdade aumentou, ao passo que no período 1980-1991 o grau de desigualdade diminuiu e, no período 1991-2000, o processo de concentração da população voltou a elevar-se, ou seja, as cidades menores cresceram menos do que as cidades maiores.

Outro resultado interessante é que quando são impostas restrições sobre a amostra o coeficiente de Pareto é maior e tende a declinar no período compreendido entre 1970 a 2000: a desigualdade é menor entre cidades médias e grandes do que quando se acrescenta as cidades menores e, ao mesmo tempo, a desigualdade ou concentração populacional em cidades maiores aumentou em todos os níveis

Por seu turno, a estimação de uma versão não-linear da relação entre a ordem e tamanho de cidades paranaenses também revela que no período 1970-2000, o comportamento convexo acabou predominando, indicando que apesar do crescimento das cidades médias, essas ainda possuem populações inferiores ao ideal para a Lei de Zipf. Além disso, revela que as cidades maiores acabaram crescendo mais do que as cidades menores.

A dinâmica da estrutura urbana paranaense ao não seguir uma das principais proposições da nova geografia econômica que atribui predominância das forças centrífugas sobre as forças centrípetas, sugere a necessidade da formulação de políticas de fortalecimento das redes de cidades menores e médias, com a criação de infra-estrutura física, econômica, social e cultural.

#### REFERÊNCIAS

- FUJITA, M.; KRUGMAN, P.; VENABLES, A. J. **Economia Espacial**: urbanização, prosperidade econômica e desenvolvimento humano no mundo. São Paulo: Futura, 2002.
- GABAIX, X. Zipf's Law for cities: an explanation. **The Quarterly Journal of Economics**, Cambridge, Massachusetts, v. 114, n. 2, p. 739-767, August 1999.
- GIBRAT, R. **Les Inegalit'es Economiques**. Paris: Librairie du Recueil Sirey, 1931.

---

RESENDE, M. Lei de Gibrat na indústria brasileira: evidência empírica. **Economia**, Brasília, v. 5, n. 2, p. 221-268, jul./dez. 2004.

SIMON, H. A. On a class of skew distribution functions. **Biometrika**, Londres, v. 42, n. 3/4, p. 425-440, Dec. 1955.

STORPER, M.; VENABLES, J. O burburinho: a força econômica da cidade. In: DINIZ, C. C.; LEMOS, M. B. (Orgs.) **Economia e território**. Belo Horizonte: UFMG, 2005.

ZIPF, G. K. **Human behavior and the principle of least effort**. Cambridge: Addison-Wesley Press, 1949.

