

Texto para Discussão

Série Economia

TD-E 11 / 2009

**DISTRIBUIÇÃO POPULACIONAL E EVOLUÇÃO DO TAMANHO DAS
CIDADES BRASILEIRAS:**

Uma análise especial do período 1920-2000.

Prof. Dr. Alexandre Sartoris Neto



Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto
Universidade de São Paulo

Universidade de São Paulo
Faculdade de Economia, Administração e Contabilidade
de Ribeirão Preto

Reitora da Universidade de São Paulo
Suely Vilela

Diretor da FEA-RP/USP
Rudinei Toneto Junior

Chefe do Departamento de Administração
André Lucirton Costa

Chefe do Departamento de Contabilidade
Adriana Maria Procópio de Araújo

Chefe do Departamento de Economia
Walter Belluzzo Junior

CONSELHO EDITORIAL

Comissão de Pesquisa da FEA-RP/USP

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto
Avenida dos Bandeirantes, 3900
14049-900 Ribeirão Preto - SP

A série TEXTO PARA DISCUSSÃO tem como objetivo divulgar: i) resultados de trabalhos em desenvolvimento na FEA-RP/USP; ii) trabalhos de pesquisadores de outras instituições considerados de relevância dadas as linhas de pesquisa da instituição. A série foi subdividida em função das principais áreas de atuação da FEA-RP/USP: Economia, Administração e Contabilidade. Veja o site da CPq na Home Page da FEA-RP: www.fearp.usp.br. Informações: e-mail: cpq@fearp.usp.br

DISTRIBUIÇÃO POPULACIONAL E EVOLUÇÃO DO TAMANHO DAS CIDADES BRASILEIRAS: UMA ANÁLISE ESPACIAL DO PERÍODO 1920-2000

Lorena Zardo Trindade[♦]
Alexandre Sartoris^{*}

RESUMO

Este artigo objetiva um exame mais detalhado da evolução da relação entre o tamanho das cidades brasileiras e sua distribuição populacional no período de 1920 a 2000. Para isso, utilizaremos dois tipos de métodos, um estático – que envolve um indicador de convergência baseado na Lei de Zipf – e outro dinâmico, que mostra, através de uma cadeia de Markov, movimentos nas posições relativas das cidades dentro da distribuição por tamanhos. Nas duas análises, verificamos uma persistente concentração populacional em um número pequeno de áreas. Os efeitos espaciais, considerados em ambas as análises, mostraram ter importante influência nos resultados obtidos.

Palavras chave: convergência, distribuição populacional, autocorrelação espacial, SUR espacial, probabilidade de transição.

ABSTRACT

This paper aims a more detailed examination on the evolution of the relationship between Brazilian cities size and their population distribution in the period from 1920 to 2000. For this purpose, two type of methods are performed, a static one - which involves the estimation of convergence indicator given by the Zipf's Law - and a dynamic one - which shows, by means of a Markov chain, movements in cities relative positions within the distribution of their sizes. In both analyses, there is a persistent population concentration in a small number of areas, growing at a greater rate than smaller ones. Spatial effects, considered in the estimation of both analyses, is clearly an important influence in the results

Key words: convergence, population distribution, spatial autocorrelation, spatial SUR models, transition probability.

JEL: C31, J11, R12

[♦] Instituto Jones dos Santos Neves e Programa de Pós-Graduação em Economia (UNESP).

^{*} Departamento de Economia e Programa de Pós-Graduação em Economia (UNESP).

1. Introdução

As modificações do cenário econômico brasileiro no decorrer do século XX impulsionaram transformações significativas na dinâmica populacional nacional. A queda da importância relativa do setor agrícola na economia brasileira na década de 1930; a ampliação dos investimentos estatais em atividades industriais na década de 1950; o “milagre econômico” dos anos 1970; a recessão da década de 1980 e a estagnação da década de 1990; todos foram responsáveis pela dinâmica da população brasileira no decorrer do século e pela atual distribuição populacional no território nacional.

Ao longo de todo o século, os movimentos populacionais entre os estados brasileiros foram pautados pela dinâmica e pela localização das atividades industriais.

No período 1920-1950, quando a transição de um modelo econômico agro-exportador para um urbano-industrial ainda estava em curso, apenas 8% de seu total viviam nas grandes cidades¹. (Mata, 1973; Brito, 2000; Netto Jr. e Targino, 2003).

Cinquenta anos mais tarde, na década de 1970, 41% da população total do país poderia ser encontrada em regiões com mais de 500 mil habitantes. Nesse período, São Paulo e Rio de Janeiro, áreas de grande concentração populacional, consolidaram-se como os principais centros econômico-financeiros do país, gerando, respectivamente, 20,4% e 19,6% do PIB brasileiro, ou seja, quase metade da produção nacional. No que se refere à produção industrial nacional, São Paulo participava com 27,1% do valor adicionado e Rio de Janeiro com 9,5%, uma soma de 36,6% da produção industrial.

Porém, a partir dos anos 80, os obstáculos gerados pela recessão econômica, concentrada, principalmente, nas grandes metrópoles, diminuiu consideravelmente o poder de atração exercido pela indústria nos estados do Rio de Janeiro e São Paulo (Brito e Carvalho, 2006). Nesse período ocorria uma relativa desconcentração espacial da atividade industrial. São Paulo e Rio de Janeiro, que juntos geravam, em 1980, 26,4% do produto industrial brasileiro diminuiu sua participação para 16,8% em 2000. Como resultado desse processo, uma tendência de concentração populacional de ritmo mais lento tornou-se predominante, e em 2000, 56% da população encontravam nos grandes centros populacionais.

Embora este cenário forneça algumas idéias gerais sobre as mudanças na distribuição da população brasileira, certas características especiais permanecem ocultas. Por exemplo, por meio dessas estatísticas, não é possível verificar se o crescimento da população concentra-se em alguns pontos do território nacional, ou se é distribuído de maneira mais uniforme em áreas de diferentes tamanho. Neste caso, um exame mais detalhado do comportamento dos movimentos populacionais é extremamente importante para uma melhor compreensão da dinâmica populacional do país no século XX, e das mudanças ocorridas nesses movimentos durante todo este período.

Apesar da relevância do tema, existem poucos estudos que analisam como a população brasileira distribui-se entre suas cidades e como essa distribuição tem evoluído ao longo do tempo. Assim, visando preencher esta lacuna, o presente trabalho tem por objetivo realizar uma análise sobre a evolução das relações entre a distribuição populacional e o tamanho das cidades brasileiras no período de 1920 a 2000. Esse trabalho ainda pretende avançar em relação aos estudos realizados anteriormente para a distribuição do tamanho das cidades brasileiras – Oliveira (2004) e Ruiz (2005) – uma vez que pretende aplicar metodologias que permitem uma análise dinâmica de previsão da futura distribuição do tamanho das cidades, e

¹ Conforme metodologia do Instituto Brasileiro de Geografia e Estatística (IBGE), as grandes cidades são aquelas que possuem mais de 500 mil habitantes.

uma análise espacial, ao introduzir especificações e testes que consideram a localização de cada cidade na determinação dos resultados.

2. A distribuição das cidades na literatura

A literatura tem realizado debates significantes na descrição da distribuição do tamanho das cidades, pois, na busca por uma distribuição estatística que mais se assemelhasse à distribuição populacional foram encontradas duas regularidades empíricas robustas que remetem a duas abordagens distintas, a da lei de Gibrat e a da lei de Zipf.

A abordagem da lei de Gibrat, desenvolvida em 1931, se baseia, de acordo com Eeckhout (2004), na hipótese de que as cidades crescem ou se contraem por meio de um processo de crescimento proporcional, em que a taxa de crescimento populacional de uma determinada cidade não depende de seu tamanho inicial. Segundo o autor, esse processo de crescimento é válido uma vez que as taxas de crescimento entre diferentes cidades possuem a mesma variância. Contudo, mesmo que isso não ocorra, não é possível verificar algum tipo de padrão sistemático em relação ao tamanho inicial, indicando que o processo estocástico de crescimento é o mesmo para todas as cidades.

O processo de crescimento proporcional do tamanho das cidades pode ser verificado por meio da regressão de estimadores não-paramétricos², que além de indicar o formato da distribuição do tamanho das cidades – que nesse caso trata-se de uma log-normal -, estabelece uma relação entre o crescimento das cidades e seus respectivos tamanhos iniciais, permitindo testar se há uma relação de dependência entre essas duas variáveis (Rosen e Resnick, 1980; Ioannides e Overman, 2000; Eeckhout, 2004).

Essa abordagem, apesar de ter sido umas das primeiras formulações a fornecer uma descrição mais detalhada da distribuição do tamanho das cidades, começou a ser aplicada pela literatura de economia urbana dominante apenas nos últimos anos. Anteriormente, o método mais utilizado para realizar esse tipo de análise baseava-se na suposição de que a distribuição estatística mais adequada para explicar o tamanho das cidades seria a distribuição de Pareto, descrita por uma lei de potência

$$R(S) = aS^{-b}, \quad (1)$$

em que, R é a classificação das cidades por população na distribuição; S é o tamanho dessa população; e a e b são parâmetros. O expoente de Pareto - parâmetro b que assume sempre valores positivos - pode servir como um indicador de convergência da distribuição do tamanho das cidades. Assim, quando $b < 1$ tem-se uma estrutura urbana concentrada – poucas grandes cidades – e com cidades que crescem a taxas menos convergentes no que diz respeito ao número de habitantes. Mas quando $b > 1$, o inverso ocorre, tendo-se uma estrutura urbana dispersa – vários grandes centros urbanos – e com cidades que crescem a taxas mais convergentes (Rosen e Resnick., 1980).

Difundido por George Zipf, em seu livro *Human Behavior and the Principle of Least Effort* de 1949, esse método destacou-se devido à regularidade encontrada em sua aplicação. Segundo Brañas (2000) e Ruiz (2005), Zipf deparou-se com indícios de que, frequentemente, b assumia o valor 1; o que ocorria porque a partir do momento em que todas as cidades de uma determinada região possuísem o mesmo potencial de crescimento, independente de seu tamanho ou posição nos seus espaços econômicos, o produto do tamanho da população e de sua classificação na distribuição (*rank*) seria aproximadamente constante. Dessa forma, a maior

² Geralmente estimados por meio do modelo $g_i = m(S_i) + \varepsilon_i$, em que g_i é a taxa de crescimento normalizada, m é estimador não paramétrico e S_i é o logaritmo do tamanho da população de cada cidade (Eeckhout, 2004).

cidade dessa região teria uma população n vezes maior que a n -ésima, de modo que a segunda maior cidade em população teria a metade de habitantes da primeira, a terceira teria um terço, a quarta um quarto, e assim por diante. A essa regularidade foi dada o nome de lei de Zipf ou regra de ordem-tamanho.

Essas duas abordagens, a da lei de Gibrat e a da lei de Zipf, por muito tempo foram consideradas incompatíveis, contudo, em estudos recentes observou-se que, com algumas considerações, essas podem ser utilizadas como análises complementares. Dentre esses estudos cabe destacar Gabaix (1999b), Ioannides e Overman (2000) e Eeckhout (2004).

Os trabalhos de Gabaix (1999b) e Ioannides e Overman (2000), que seguem uma linha mais teórica, sugerem que a lei de Zipf seria na verdade uma consequência da lei de Gibrat. Nesse sentido, a maior contribuição de Gabaix (1999b), foi a indicação de que quando o crescimento das cidades obedece a lei de Gibrat e, concomitantemente, é submetido a choques exógenos estocásticos de mesmo desvio-padrão, a distribuição limite do seu tamanho resultará na lei de Zipf. Esse resultado ainda seria alcançado mesmo que algumas cidades possuíssem potenciais de crescimento diferentes, pois nestas circunstâncias, a partir de sugestões do próprio Gabaix (1999b) e de Ioannides e Overman (2000), basta que se considere o processo de crescimento das cidades como um movimento browniano geométrico variante no tempo para que a distribuição limite do tamanho das cidades automaticamente tenda a uma lei de potência com um expoente compatível com lei de Zipf.

O trabalho de Eeckhout (2004) chega à mesma conclusão geral de que a lei de Zipf não passa de um caso especial da lei de Gibrat, no entanto, para chegar a esse resultado o autor usa uma análise empírica da distribuição do tamanho das cidades dos Estados Unidos. Nessa análise, em que a amostra utilizada compreende a população de todas as cidades americanas no ano de 2000, verifica-se que a distribuição da população no território americano segue uma distribuição log-normal e não de Pareto. Mas, como na parte superior da amostra essas duas distribuições possuem formatos semelhantes, a lei de Zipf ainda pôde ser verificada para parte da amostra.

Os estudos que pretendem analisar a distribuição do tamanho das cidades podem, então, considerar tanto a abordagem da lei de Gibrat quanto a da lei de Zipf. Alguns autores utilizam ambas abordagens; Eeckhout (2004), Black e Henderson (2003) para os Estados Unidos e Anderson e Ge (2004) para a China. No que se refere à lei de Gibrat, com exceção de Black e Henderson (2003), todos encontram resultados que sugerem que essa se aplica ao crescimento populacional das cidades. Mas em relação à lei de Zipf, os resultados encontrados diferem entre os trabalhos, pois, ainda que certa regularidade seja verificada, observa-se uma forte sensibilidade das estimativas do parâmetro b com relação a variações na amostra e ao método de estimação empregado.

Essa sensibilidade do parâmetro b pode se vista de maneira mais clara na comparação entre as estimações feitas por Rosen e Resnick (1980), Soo (2002) e Oliveira (2004). Rosen e Resnick (1980), ao utilizar uma amostra que abrangia cidades com mais de 15 mil habitantes de diversos países, obteve um coeficiente de Pareto de 1,1341 para o Brasil, ao utilizar dados de 2000, enquanto Soo (2002), usando dados populacionais de todas as cidades do país, obteve, para o mesmo ano, 0,8690.

Oliveira (2004, p. 7), ao verificar variações do coeficiente de Zipf para diferentes tamanhos de amostras compostas por cidades brasileiras ressalta que a sensibilidade do parâmetro b sugere que “quando se estabelecem restrições sobre a amostra, o coeficiente de Pareto é maior, o que indica que a desigualdade é menor entre cidades médias e grandes do que quando se acrescenta as cidades menores”.

Até então, os métodos utilizados para examinar a relação entre distribuição populacional e o tamanho das cidades têm revelado características essenciais para um melhor entendimento da dinâmica populacional ao longo do tempo e do espaço. Contudo, como a

verificação da lei de Zipf se trata de uma análise estática, não é possível identificar como são as mudanças nas posições relativas das cidades na distribuição de seus tamanhos. Lanaspa et al. (2003) afirma que essa é uma deficiência que deve ser corrigida, dada a dificuldade de se aceitar que a estrutura urbana não passou por mudanças profundas ao longo de um século. Em outras palavras, Lanaspa et al. (2003) explica que mesmo que a distribuição tenha permanecido essencialmente a mesma por um longo período de tempo, mudanças significantes podem ter ocorrido na posição relativa das cidades na distribuição.

Para suprir essa lacuna presente nos estudos referentes à relação existente entre distribuição populacional e o tamanho das cidades, Eaton e Eckstein (1997), Black e Henderson (2003), Lanaspa et al. (2003), Le Gallo e Chasco (2008) aplicaram, - respectivamente, às cidades da França e do Japão, dos Estados Unidos e da Espanha - uma metodologia que permite a descrição do comportamento das cidades na distribuição populacional ao longo do tempo.

Essa metodologia, até então ignorada pelos estudos aplicados ao Brasil, se baseia na suposição de que a dinâmica interna da distribuição do tamanho das cidades entre os diferentes períodos pode ser modelada por uma cadeia de Markov que produz uma matriz de probabilidade de transição. Nessa matriz, cada elemento (i, j) indica a probabilidade de uma cidade que estava na classe i no período t , passar para a classe j no período seguinte, o que permite uma previsão da futura distribuição do tamanho das cidades e a descrição do comportamento da amostra inteira, ao invés de apenas um elemento médio, como no caso da análise estática (Lanaspa et al., 2003).

A possibilidade da distribuição do tamanho das cidades sofrer efeitos espaciais é outra característica ignorada pelos estudos aplicados ao caso brasileiro. Devido à natureza geográfica dos dados utilizados nesse tipo de análise, o uso de uma metodologia econométrica que considere essas características foi essencial para que uma inferência estatística adequada pudesse ser realizada. Le Gallo e Chasco (2008) introduziram em sua análise, testes e especificações de dependência espacial, tanto na parte estática quanto na dinâmica, para captar a influência do espaço na convergência e nas probabilidades de transição, realizando dessa forma, um exame mais completo da velocidade do movimento e da forma de convergência na distribuição do tamanho das cidades. Na análise da convergência, os resultados encontrados pelo método dos Mínimos Quadrados Ordinários (MQO) indicaram a existência de resíduos não-normais, que por sua vez, apresentaram heteroscedasticidade e autocorrelação espacial. E na análise da cadeia de Markov foram encontradas evidências de que a probabilidade de um movimento dentro da distribuição ocorrer difere em diferentes áreas urbanas. Ambos os resultados confirmam a suposição de que há influência espacial na distribuição do tamanho das cidades.

Diante dessas evidências, o presente estudo trata a distribuição do tamanho das cidades brasileiras por meio de abordagem estáticas, dinâmica e espacial

3. Metodologia e resultados

3.1 Descrição dos dados

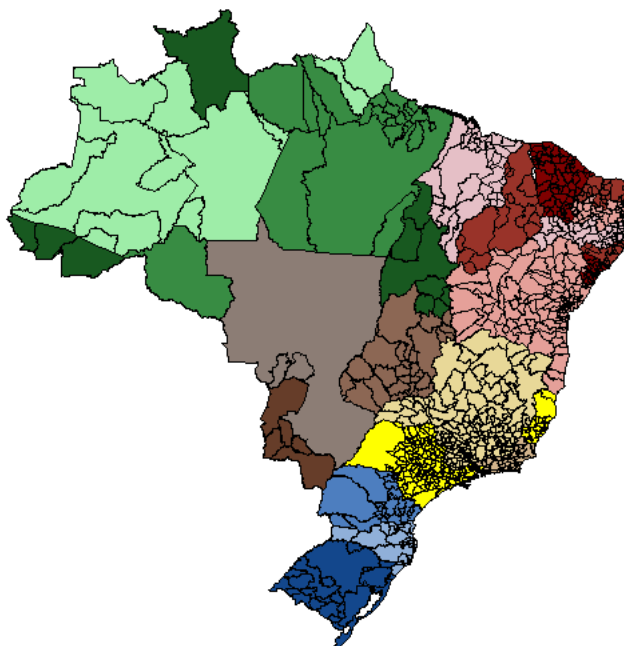
As alterações ocorridas no número, área e fronteira das cidades brasileiras ao longo do século XX tornam inconsistentes as comparações intertemporais em nível geográfico estritamente municipal. Portanto, para possibilitar comparações consistentes no tempo, o Instituto de Pesquisa Econômica Aplicada (IPEA) agregou os dados municipais em áreas geográficas mais abrangentes denominadas de áreas mínimas comparáveis (AMC).

De acordo com Reis *et al.* (2008), as AMCs são áreas geográficas que compatibilizam as divisões político-administrativas apresentadas nos vários censos para possibilitar a construção de painéis de dados econômicos estaduais e municipais no período que se estende de 1872 a 2000³.

A Figura 1 é um mapa das AMCs no período 1920-2000, distinguindo-se por meio de cores as Unidades da Federação atuais.

Figura 1 Brasil: Áreas Mínimas Comparáveis, 1920-2000.

BRASIL: Areas Mínimas Comparáveis 1920-2000

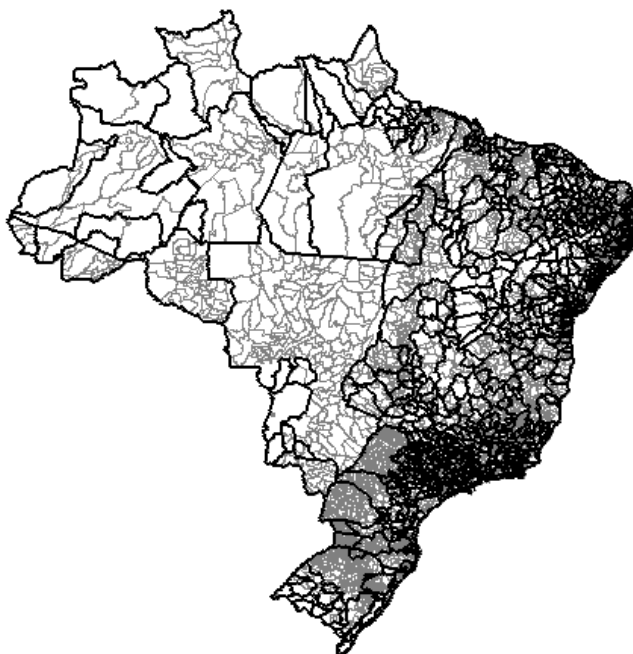


Elaboração: IPEA/DIMAC utilizando a Malha Municipal Digital do Brasil 1997 (IBGE/DGC/DECAR)

A Figura 2 apresenta em linhas fortes o contorno das AMCs para o período 1920-2000, comparando-as com a malha municipal utilizada no Censo de 2000 - em linhas esmaecidas.

³ Entre 1940 e 1975, os estados do Amazonas e Mato Grosso tiveram alguns de seus municípios desmembrados, originando, respectivamente, os estados de Roraima e Mato Grosso do Sul. O mesmo ocorreu com Goiás, que deu origem ao estado de Tocantins (Reis *et al.*, 2008).

Figura 2 Brasil: Áreas mínimas comparáveis, 1920-2000 e área dos municípios em 2000.



Elaboração: IPEA/DIMAC utilizando a Malha Municipal Digital do Brasil 1997 (IBGE/DGC/DECAR)

Portanto, em face das mudanças nas divisões municipais e estaduais, a evolução da distribuição populacional será analisada por meio de dados censitários das AMCs do período 1920-2000, em que oito décadas são analisadas; 1920, 1940, 1950, 1960, 1970, 1980, 1991 e 2000.

3.2 Aspectos metodológicos

Os trabalhos aplicados na área de ciências regionais geralmente baseiam-se em amostras de dados recolhidos com referência a sua localização no espaço. Na década de 1970, frente à necessidade de integrar esses aspectos de localização da amostra de dados multiregionais com estimações e testes de modelos econométricos, Jean Paelinck introduziu uma nova forma analítica denominada de Econometria Espacial, que associada a instrumentais da econometria tradicional, contemplava a importância da questão espacial nos estudos regionais (Anselin, 1988).

A preocupação fundamental da Econometria espacial originou na hipótese de dados distribuídos no espaço apresentaram (i) dependência espacial entre as observações; ou (ii) heterogeneidade espacial nas relações estimadas pelo modelo (Anselin, 1988; Le Sage, 1998).

Estas duas questões, ignoradas pela econometria tradicional, violam as propriedades básicas do teorema de Gauss-Markov, essenciais para modelos clássicos de regressão. No que diz respeito à violação pela existência de dependência espacial, as propriedades do teorema de Gauss-Markov exigem que as variáveis explicativas sejam fixas em amostragem repetidas. Do mesmo modo, a heterogeneidade espacial viola a propriedade referente à existência de uma única relação linear em toda a amostra de dados observados.

Portanto, a existência de dependência e heterogeneidade espaciais em um conjunto de dados amostrais modelados exige a abordagem mais adequada da Econometria Espacial.

Existem duas formas de autocorrelação espacial; defasagem espacial (*spatial lag*) e o erro espacial (*spatial error*). O modelo que incorpora a defasagem espacial (SAR, *spatial autoregressive model*) é apresentado por

$$y_{it} = \rho_{\tau}W + X\beta_{\tau} + \varepsilon_{it}, \quad (2)$$

onde y é o vetor de observações da variável endógena, W é a matriz das vizinhanças, X é a matriz das observações das variáveis exógenas, β é o vetor dos coeficientes, ρ é o coeficiente espacial autoregressivo e ε é o vetor dos erros. O coeficiente ρ é uma medida que explica como as observações vizinhas afetam a variável dependente. O modelo de erro espacial (SEM, *spatial error model*) é expresso por

$$y_{it} = x_{it}\beta_{\tau} + \varepsilon_{it}, \quad (3)$$

onde a dependência espacial está considerada no termo de erro

$$\varepsilon_{it} = \rho_{\tau}W\varepsilon_{it} + \mu_{it}, \quad (4)$$

A existência, ou não, destes efeitos pode ser determinada através de um conjunto de técnicas desenvolvidas pela econometria espacial. Para evitar este tipo de erros de especificação os resultados são geralmente complementados com testes de especificação; Jarque-Bera, que testa a estabilidade dos parâmetros; Koenker-Basset, que verifica a presença de heterocedasticidade; e Kelejian-Robinson, o mais apropriado para verificar a existência de heterocedasticidade nos parâmetros quando a normalidade é rejeitada pelo teste Jarque-Bera.

3.3 Testes para dependência espacial dos erros

O teste Kelejian-Robinson (K-R) é o teste robusto para identificar dependência espacial dos erros utilizando amostras grandes que não apresentam normalidade e nem linearidade dos resíduos. Essa estatística é derivada de uma regressão auxiliar que utiliza produtos cruzados de observações que são potencialmente espacialmente correlacionadas e produtos cruzados das variáveis explicativas correspondentes. Especificamente, a variável dependente na regressão auxiliar é

$$C_h = e_i e_j, \quad (5)$$

Em que h é um índice para cada produto cruzado, e é o resíduo e i e j são observações contínuas. As variáveis explicatórias na regressão auxiliar, Z_h , são formadas como produtos de X_i e X_j . Sendo γ o vetor de coeficientes obtido na regressão de C em Z por MQO, e α o vetor de resíduos correspondente. Os resultados de K-R são

$$K - R = \frac{Y'Z'ZY}{\alpha'\alpha/h_R}, \quad (6)$$

em que h_R é o número de observações do vetor auxiliar (5). A estatística do teste segue uma distribuição χ^2 , com K^4 graus de liberdade, sob a hipótese nula de não existência de dependência espacial.

3.4 Modelos espaciais no tempo e no espaço

Os modelos compostos por observações dispostas em duas dimensões, geralmente tempo e espaço, tem se tornado cada vez mais relevante nos trabalhos empíricos das ciências regionais. Quando efeitos espaciais podem ser verificados em modelos com dados de tempo e espaço, a econometria espacial constitui a perspectiva mais adequada para inferência estatísticas.

⁴ K é o número de variáveis explicativas em Z .

Os modelos do tipo Regressões Aparentemente Não Relacionadas (SUR, *Seemingly Unrelated Regression*) é um daqueles que podem contemplar a dependência espacial. Originalmente sugerido por Arnold Zellner em 1962, foi designado para situações empíricas onde existe um limitado grau de simultaneidade na forma de dependência entre os erros de diferentes equações. Se as equações pertencem a uma série temporal de diferentes regiões a dependência resultante pode ser considerada como uma forma de autocorrelação espacial (Anselin, 1988).

Em seu formato tradicional, o SUR apresenta coeficientes (β_i) que variam de acordo com a unidade espacial, mas são constantes ao longo do tempo. Os erros, espacialmente correlacionados, exibem covariância constante entre si para diferentes unidades espaciais, medidas no mesmo ponto do tempo. De maneira formal, o modelo é expresso pela equação

$$y_{it} = x_{it}\beta_i + \varepsilon_{it}, \quad (7)$$

com

$$E[\varepsilon_{it} \cdot \varepsilon_{jt}] = \sigma_{ij}. \quad (8)$$

Em seu formato espacial, o SUR exhibe coeficientes (β_i) constantes no espaço, mas variáveis ao longo do tempo. Os erros (ε) são temporalmente correlacionados, i.e., existe uma covariância constante entre resíduos de diferentes períodos para as mesmas unidades espaciais. De maneira formal, o modelo é expresso pela equação

$$y_{it} = x_{it}\beta_t + \varepsilon_{it}, \quad (9)$$

com

$$E[\varepsilon_{it} \cdot \varepsilon_{it'}] = \sigma_{it'}. \quad (10)$$

Como o SUR-espacial consiste em equações para cada período de tempo, estimadas para dados de cortes transversais de unidades espaciais, a autocorrelação espacial dos erros, inerentes às equações, são um problema em potencial. Portanto, de maneira análoga à autocorrelação serial dos erros tem-se que os erros seguem um processo autoregressivo em cada equação, com parâmetros diferentes para cada (t), é correlacionado entre si.

$$y_{it} = x_{it}\beta_t + \varepsilon_{it}, \quad (11)$$

com

$$\varepsilon_t = \rho_t W \varepsilon_t + \mu_t, \quad (12)$$

e

$$E[\mu_t \mu_t'] = \sigma_{t\varepsilon} \cdot I. \quad (13)$$

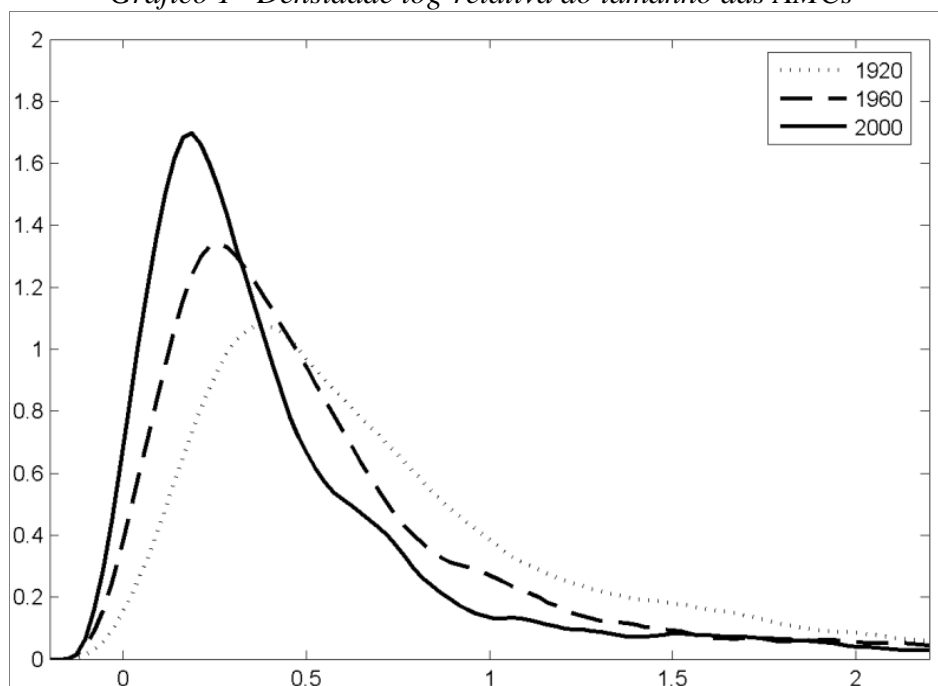
A forma de estimação mais apropriada dos coeficientes do SUR-espacial é por meio do Método dos Mínimos Quadrados Generalizados (MQG) ou por Máxima Verossimilhança (MV).

3.5 Evolução do formato da distribuição populacional

A partir do conjunto de dados descrito acima, analisou-se a evolução do formato da distribuição populacional nas AMCs brasileiras por meio de estimadores não-paramétricos de densidade de kernel, estimados para cada década com o objetivo de analisar seus aspectos monomodais ou multimodais.

O gráfico 1 apresenta o gráfico do logaritmo da distribuição relativa da população dos municípios ao longo de 1920, 1960 e 2000. Tem-se por definição que o 1 (um) do eixo horizontal indica o tamanho médio das AMCs brasileiras, e 2 (dois) indica o valor da média dobrado.

Gráfico 1 Densidade log-relativa do tamanho das AMCs



As distribuições nas três décadas ilustradas são muito semelhantes uma vez que todas apresentam formato unimodal. A principal diferença parece ser a perda de densidade de 1920 em relação a 1960 e 2000, o que sugere a existência de uma mudança da massa de densidade para a esquerda em direção ao valor 0,25, alcançado no ano de 2000. Este movimento pode ter sido causado por um grupo considerável de AMCs que inicialmente apresentavam maior importância relativa, mas ao longo das décadas perderam grande número de residentes.

Portanto, a distribuição da população brasileira parece assumir uma tendência crescente da quantidade de municípios que apresentam uma quantidade de população abaixo da média e que convergem para um nível populacional mais baixo do que o resto das AMCs.

3.6 Resultados da estimação

Empiricamente, a relação entre a distribuição da população e a dimensão das 952 AMCs (i) em cada década (t) é dada pela equação (1) linearizada

$$\ln R_{it} = b_t \ln S_{it} + \varepsilon_{it}, \quad (14)$$

estimada para os dados de cada uma das 8 décadas considerado pelo estudo.

Os parâmetros de cada modelo são estimados pelo método dos Mínimos Quadrados Ordinários (MQO).

Tabela 1 Coeficiente de Pareto por MQO

Year	α^*	β^*	JB	KB	KR
1920	17,101	-1,132	620,573	13,456	9,535
1940	16,635	-1,064	658,713	0,347	2,864
1950	16,005	-0,988	2021,666	3,040	6,729
1960	15,579	-0,930	983,936	4,579	5,105
1970	15,061	-0,867	1285,525	3,672	8,283
1980	14,670	-0,817	2075,140	1,931	13,982
1991	14,535	-0,792	2836,998	0,257	12,908
2000	14,502	-0,781	3330,278	0,045	16,080

Notes: $\alpha = \ln a$. * Hipótese nula rejeitada a 1%. A coluna *JB* refere-se aos resultados da aplicação do teste Jarque-Bera de não normalidade dos resíduos. A *KB* indica os resultados do teste Koenker-Basset para heteroscedasticidade. A *KR* apresenta os resultados do teste Kelejian-Robinson para autocorrelação espacial dos erros. Para *KR* e *KB*, o valor crítico de significância a 5% é de 3,8415.

De acordo com a Tabela 1, o expoente Pareto por MQO assume valores decrescentes durante o período 1920-2000, sugerindo que, em 1920 a população das AMCs brasileiras crescia de maneira mais uniforme do que em 2000. Portanto, estes resultados reafirmam o cenário descrito anteriormente, onde a população brasileira tende a se concentrar em um pequeno número de grandes AMCs e com crescimento superior àquelas de menor tamanho populacional.

A Tabela 1 mostra também as estatísticas dos testes Jarque-Bera, Koenker-Basset e Kelejian-Robinson para cada modelo. Os resultados verificados sugerem, respectivamente, a existência de não-normalidade dos erros; heterocedasticidade - mas apenas para os modelos de 1920 e de 1960 - e autocorrelação espacial - exceto para a distribuição de 1940. Consequentemente, os parâmetros obtidos por MQO podem ser considerados tendenciosos.

A exemplo de Le Gallo e Chasco (2008), uma especificação econométrica mais adequada foi utilizada com o intuito de obter estimativas mais robustas. Assim, uma vez que os modelos são compostos por dados de cortes transversais de diferentes pontos no tempo, o SUR – erro espacial (SSUERR) foi utilizado.

Tabela 2 Coeficiente de Pareto por SUR - erro espacial

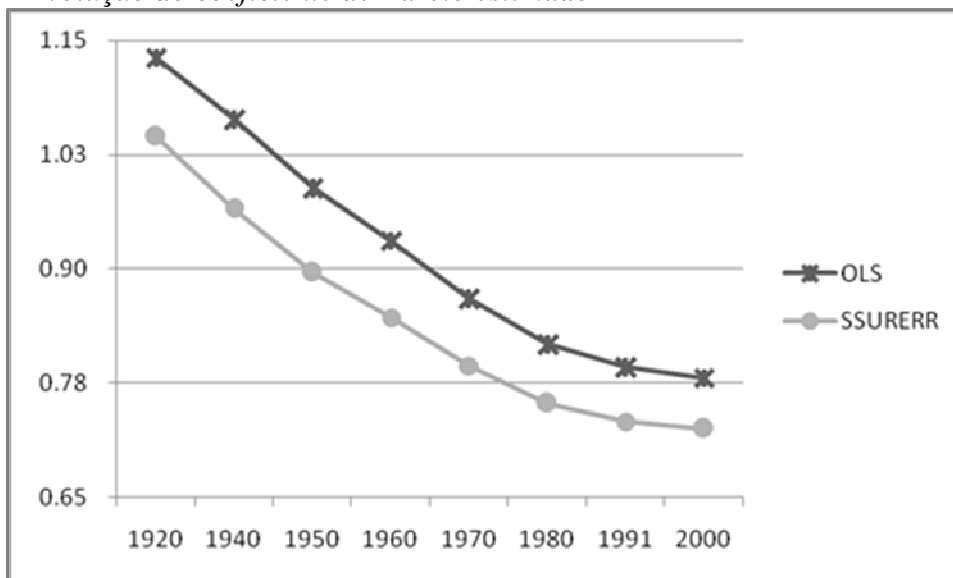
Year	α^*	β^*	λ^*
1920	16,230	-1,045	-0,014
1940	15,639	-0,966	0,001
1950	15,064	-0,896	0,034
1960	14,722	-0,847	0,126
1970	14,291	-0,794	0,153
1980	13,980	-0,753	0,182
1991	13,890	-0,733	0,179
2000	13,883	-0,725	0,178

Nota: * Hipótese nula rejeitada a 5%.

Os resultados estatisticamente significantes das estimativas por SUR – erro espacial, apresentados na tabela 2, confirmam a influência espacial na distribuição da população.

Para comparar a magnitude da diferença entre os expoentes de Pareto estimados por MQO e por SSUER, a evolução de seus valores ao longo do tempo é apresentada pelo gráfico 2.

Gráfico 2 *Evolução do coeficiente de Pareto estimado*



Nota: N=951.

Os estimadores encontrados por MQO são sempre superiores aos do SSURERR, mas ambos seguem mesma evolução decrescente, permitindo mais uma vez verificar que a distribuição da população brasileira em suas 952 AMCs tornou-se cada vez mais divergente nos últimos 80 anos.

Comparando essas estimativas com os resultados obtidos por Oliveira (2004), tendências similares foram encontradas; 0,0234 foi o coeficiente estimado para uma amostra composta por todos os municípios do Brasil; 0,005 para uma amostra composta pelos municípios com população superior a 50 mil habitantes; e (-) 0,0312 para uma amostra dos 100 maiores municípios. No entanto, Soo (2002) observou um comportamento diverso dado por um coeficiente de Pareto 1,134, que implica um cenário caracterizado por municípios de tamanho populacional convergente, em vez de divergente.

3.7 Análise dinâmica espacial da distribuição do tamanho das AMCs

As estruturas locais de divisão territorial dos países são frequentemente expostas a mudanças ao longo de um século. No Brasil, em um período de 80 anos, essas mudanças resultaram em redistribuição da população entre as AMCs.

Tabela 3 *Mudanças na distribuição do tamanho dos municípios brasileiros por classe de tamanho relativo*

Estado (m) (%)	1920	2000
$\bar{P}_t < 25$	124	373
$25 < \bar{P}_t < 40$	172	190
$40 < \bar{P}_t < 55$	132	94
$55 < \bar{P}_t < 70$	116	72
$70 < \bar{P}_t < 100$	152	67
$100 < \bar{P}_t < 150$	113	46
$\bar{P}_t > 150$	142	109

Nota: \bar{P}_t representa a média populacional brasileira para cada década (t) analisada.

No ano de 1920, 124 municípios tinham tamanho populacional correspondente a um quarto (25%) da média brasileira, enquanto 142 apresentaram uma população 1,5 vezes maior que a média. Em 2000, os dados referentes a estas proporções modificara-se, respectivamente, para 373 e 109, confirmando a existência de uma tendência à concentração populacional em áreas já concentradas visto que as AMCs menores perderam população para as AMCs maiores, sem alterações significantes no número absoluto das primeiras.

Para um exame mais detalhado da dinâmica dos dados, aplicou-se mesma metodologia utilizada por Eaton e Eckstein (1997), Black e Henderson (2003), Lanaspa et al. (2003) e Le Gallo e Chasco (2008). Inicialmente elaborada por Qua (1993), essa metodologia baseia-se no pressuposto de que a distribuição interna da dinâmica populacional de determinada área entre dois pontos diferentes no tempo pode ser modelada por um processo estocástico com espaço paramétrica de tempo discreto e um espaço finito de estados. Este processo, também conhecido como cadeia de Markov, deve contemplar as seguintes propriedades

$$P\left(M_t = \frac{j}{M_{t-1}} = i, M_{t-2} = c, \dots, X_1 = a\right) = P\left(M_t = \frac{j}{M_{t-2}} = i\right) = p_{ij}(t), \quad (15)$$

onde M é a denotação da cadeia de Markov; j , i , c e a se referem ao estado espacial; e p_{ij} é a probabilidade de transição estacionária, que especifica as chances do estado i no período t a passar para o estado j em $t + 1$.

Quando todas as probabilidades de transição são agrupados em uma matriz estocástica, temos

$$P = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{pmatrix} \quad (16)$$

com $p_{ij} \geq 0$ e $\sum p_{ij} = 1$.

Seguindo a sugestão de Qua (1993), o estado de transição necessário para a formulação de processo de Markov é obtido por meio da divisão do tamanho relativo da distribuição das AMCs referente a um ponto do tempo em intervalos discretos com pontos de corte definidos por N_{it}/\bar{N}_t onde $N_t = \sum_i^{m_t} N_{it}/m_t$.

Nas tabelas 3 e 4 verificam-se todos os 7 intervalos definidos pelos pontos de corte; (1) AMCs cuja população está abaixo dos 25 por cento da média; (2) entre 25 e 40 por cento da média; (3) entre 40 e 55 por cento da média; (4) entre 55 e 70 por cento da média; (5) entre 70 por cento da média e da média; (6) entre a média e 1,5 vezes a média; e (7) mais de 1,5 vezes a média.

Tabela 4 Matriz de probabilidade de transição, 1920-2000

	1 < 25%	2 25-40%	3 40-55%	4 55-70%	5 70-100%	6 100-150%	7 > 150%
1	0.766	0.097	0.040	0.032	0.016	0	0.048
2	0.674	0.186	0.052	0.017	0.023	0.012	0.035
3	0.500	0.303	0.053	0.061	0.045	0.023	0.015
4	0.388	0.250	0.207	0.052	0.034	0.052	0.017
5	0.217	0.283	0.191	0.099	0.112	0.033	0.066
6	0.142	0.221	0.097	0.150	0.133	0.097	0.159
7	0.014	0.063	0.063	0.134	0.134	0.134	0.458

A Tabela 4 contém a matriz de probabilidade de transição de primeira ordem para o período 1920-2000 em que cada elemento (i, j) indica a probabilidade de uma AMC se deslocar do estado i no momento t para o estado j no período seguinte $(t + 1)$. Algumas conclusões interessantes foram encontradas a partir do conteúdo desta matriz;

- a probabilidade de ficar no estado inicial, dado pelos elementos da coluna diagonal, é muito mais elevada nos dois estados extremos (76,6% das AMCs com população inferior a 0,25 da média em 1920 permaneceu neste intervalo até o final do período, e 45,8 % das AMCs com população 1,5 superior à média em 1920 manteve esse tamanho);
- a probabilidade de passagem para intervalos inferiores excede a probabilidade de transição para intervalos superiores, exceto no intervalo 6;
- na maioria dos intervalos, as chances de descer mais de um intervalo são maiores do que as chances de descer apenas um intervalo; a probabilidade de subir mais de um intervalo é inferior à probabilidade de subir apenas um intervalo;
- nos intervalos medianos de 2 a 6, as chances de passar para o estado 1 são mais elevadas (de 14,2% a 50%) do que as chances de mover-se ao estado 7 (de 4,8% a 6,6%);
- a probabilidade de subir um intervalo tendo como estado inicial 1 (9,7%) é superior aos demais estados 2 (4%), 3 (1,7%), 4 (4,5%) e 5 (5,2%);
- a possibilidade de crescimento paralelo ou uniforme entre as AMCs não pode ser aceita pois vários elementos da diagonal da matriz não assumem valor 1, o que indica que a distribuição do tamanho da população brasileira sofreu mudanças estruturais durante o período 1920-2000.

A fim de captar a influência do espaço nos resultados da cadeia de Markov, a probabilidade de movimentos na distribuição populacional das AMCs brasileiras foi verificada considerando a distribuição média da população de AMCs vizinhas, como em Le Gallo e Chasco (2008). A correlação envolvendo a direção do movimento das AMCs na distribuição populacional e a média populacional de seus vizinhos é medida na tabela 5.

Tabela 5 Probabilidade de transição com efeitos espaciais

Defasagem espacial	Movimento		
	<i>Sobe</i>	<i>Permanece</i>	<i>Desce</i>
<i>População menor</i>	0,003	0,046	0,042
<i>População maior</i>	0,120	0,199	0,590

Pelos elementos de probabilidade de movimento (sobe, permanece e desce) relatados na tabela, é possível notar que a probabilidade de subir, permanecer ou descer difere de acordo com o contexto local da área. Por exemplo, a probabilidade de uma AMC subir na hierarquia de tamanho é 0,03% quando contém, em média, população menor, e 12% quando contém, em média, população maior que a de seus vizinhos. No caminho oposto, a probabilidade de uma AMC se descolar para uma posição inferior na hierarquia é de 4,2% quando seus vizinhos apresentam, em média, população menor, e 59% quando apresentam em média população maior. Portanto, a influência do espaço na probabilidade de transição das AMCs entre os intervalos da distribuição populacional parece mais relevante para movimentos ascendentes.

4 Considerações finais

Ao analisar a evolução dos expoentes de Pareto no período de 1920 a 2000 verificou-se que, embora haja uma ligeira mudança de declividade nas últimas duas décadas, a tendência geral, durante todo o período é de divergência, ou seja, há uma contínua tendência à concentração populacional.

Este resultado também é enfatizado pelo fato de que, AMCs menores tem maiores probabilidades de perda de população relativa, enquanto AMCs com o dobro ou mais da média populacional nacional tendem a atrair relativamente mais população.

Considerando efeitos espaciais, se uma AMC está cercada por vizinhos com número populacional superior, maior é a probabilidade de perder população.

A mudança de declividade, porém, sugere que, talvez, o século 21, será caracterizado por uma nova fase de convergência, ou seja, dispersão, com pessoas que se deslocam de AMCs maiores para AMCs menores. Existe uma controvérsia aqui visto que as dimensões continentais fazem do Brasil um país com diferenças significativas em termos de concentração populacional em seu território.

Bibliografia

1. Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.
2. Anselin, L. and Florax, J. G. M. (1995). "Directions in Spatial Econometrics: Introduction". In: Anselin, L. and Florax, R. J. G. M. (eds) *New Direction in Spatial Econometrics*, Springer, New York.
3. _____ (1995). "Small Sample of Tests for Spatial Dependence in Regression Models: Some Further Results". In: Anselin, L. and Florax, R. J. G. M. (eds) *New Direction in Spatial Econometrics*, Springer, New York.
4. Anderson G; Ge Y. (2005). "The size distribution of Chinese cities". *Regional Science and Urban Economics*, v. 35, n. 6, p. 756-776.
5. Brito, F. (2000). *Brasil, Final de século: a transição para um novo padrão migratório?* Anais da ABEP, Caxambu, Minas Gerais.
6. Brito, F.; Garcia, R.; Souza, R. (2004). *As tendências recentes das migrações Interestaduais e o padrão migratório*. Trabalho apresentado no XIV Encontro Nacional de Estudos Populacionais, ABEP, Caxambu, Minas Gerais, setembro.
7. Brito, F.; Carvalho, J. (2006). "As migrações internas no Brasil: as novidades sugeridas pelos censos demográficos de 1991 e 2000 e pelas PNADs recentes". *Parcerias estratégicas*, n. 22, junho.
8. Black, D.; Henderson, V. (1999a) "A Theory of Urban Growth" *Journal of Political Economy*, v. 107 n. 2, p. 252-284.
9. Black, D.; Henderson, V. (1999b) "Spatial Evolution of Population and Industry in the United States". *The American Economic Review*, v. 89 n. 2, p. 321-327.
10. _____ (2003). "Urban Evolution in the USA", *Journal of Economic Geography*, v. 3, p. 343-372.
11. Brañas, P.; Alcalá, F. (2000). "Entropía, aglomeración urbana y la ley del '1': evidencia para las regiones españolas". *III Encuentro de Economía Aplicada*, Valença.
12. _____ (s. d.). "Aglomeración urbana española (1900-2000): estimaciones "rank-size" vs. tests LM". *Working paper 0101/Nº 19*, Departamento de Economía Aplicada, Universidad de Jaén.
13. Camarano, A. A.; Beltrão, K. I. (2000). "Distribuição espacial da população brasileira: mudanças na segunda metade deste século". *Texto para discussão nº 766*. Rio de Janeiro: IPEA.
14. Carvalho, A. Y. et al. (2007). *Ensaio sobre economia regional*. Instituto de Pesquisa Econômica Aplicada, Brasília.
15. Cordoba J. C. (2004b). *On the Distribution of City Sizes*. Economics Department, Rice University.
16. Dobkins, L. H. and Ioannides, Y. M. (2000). "Dynamic evolution of the size distribution of U.S. cities". In: J. M. Huriot and J. F. Thisse (ed.) *Economics of Cities*, p. 217-260. New York: Cambridge University Press.
17. Eaton, J. and Eckstein, Z. (1997). "Cities and growth: theory and evidence from France and Japan". *Regional Science and Urban Economics*, v. 27, p. 443-474.

18. Eeckhout, J. (2004). "Gibrat's Law for (All) Cities". *The American Economic Review*, v. 94, n. 5, December, p. 1429-1451.
19. Fan, C. C.; Casetti, E. (1994). "The spatial and temporal dynamics of US regional income inequality, 1950-1989". *The Annals of Regional Science*, n. 28, p. 177-196.
20. Gabaix, X. (1999a). "Zipf's Law and the Growth of Cities". *The American Economic Review*, v. 89, n. 2, May, p. 129-132.
21. _____ (1999b), "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, CXIV, August, p.739 – 767.
22. Ioannides, Y.M.; Gabaix, X. (2003). "The Evolution of City Size Distributions". *Handbook of Urban and Regional Economics, Volume IV: Cities and Geography*.
23. Glaeser, E. L. *et al.* (1992) "Growth in Cities". *The Journal of Political Economy*, v. 100, n. 6, p. 1126-52.
24. Glaeser, E. L.; Scheinkman, J. A.; Shleifer, A. (1995). "Economic growth in a cross-section of cities". *Journal of Monetary Economics*, v. 36, p. 117-143.
25. Guérin-Pace, F. (1995). "Rank-size distribution and the process of urban growth". *Urban Studies*, v. 32, n. 3, p. 551-562.
26. Haessel, W. (1978). "Macroeconomic Policy, Investment, and Urban Unemployment in Less Developed Countries". *American Journal of Agricultural Economics*, v. 60, n. 1, Feb., p. 29-36.
27. Henderson, J. V. (1974). "The Sizes and Types of Cities". *The American Economic Review*, v. 64, n. 4, September, p. 640-656.
28. Henderson, J. V. (1988). *Urban Development. Theory, Fact, and Illusion*. Oxford: Oxford University Press.
29. Henderson, J. V. (2003). "The urbanization process and economic growth: the so-what question". *Journal of Economic Growth*, v. 8, p. 47-71.
30. Henderson, J. V. (2004). *Urbanization and Growth*. Brown University.
31. Ioannides, Y.; Overman, H. (2003). "Zipf's Law for Cities: an Empirical Examination", *Regional Science and Urban Economics*, v. 33, p. 127-137.
32. Lanaspá L, Pueyo F, Sanz F. (2003). "The evolution of Spanish urban structure during the Twentieth Century". *Urban Studies* v. 40, n. 3, p. 567–580.
33. Lanaspá L; Perdiguero A.M.; Sanz F. (2004). *La distribución del tamaño de las ciudades en España, 1900-1999*. Mimeo, Departamento de Análisis Económico, Universidad de Zaragoza.
34. Le Gallo, J.; Chasco, C. (2008). "Spatial analysis of urban growth in Spain (1900 – 2001)". *Empirical Economics*, v. 34, n. 1, p. 59 – 80.
35. Le Sage, J. (1998) *Spatial Econometrics* (circulated for review). Department of Economics, University of Toledo. December.
36. Mata, M.; Carvalho, E.W.; Castro e Silva, M. T. (1973). *Migrações internas no Brasil: aspectos econômicos e demográficos*. IPEA, coleção Relatórios de Pesquisa, n. 19, Rio. Janeiro.
37. Miranda, R. A. de; Badia, B. D. (s. d.). *A evolução da distribuição do tamanho das MUNICÍPIOS de Minas Gerais: 1920 – 2000*. [artigo científico]. Disponível em: http://www.cedeplar.ufmg.br/seminarios/seminario_diamantina/2006/D06A075.pdf. Acesso em: 5 maio 2008.
38. Netto Junior, J. L. da S.; Targino, I. (2003). *Migrações e Diferenciais de Renda Estaduais: uma análise por dados em painel no período de 1950 - 2000*. Trabalho apresentado no Encontro Transdisciplinar Espaço e População, Campinas.
39. Oliveira, C.A e Guimarães Neto, L. (1997). "Emprego Organizado e Regiões nos Anos 90: Quem Perdeu Mais?" *Estudos Econômicos*, v. 27, n. especial, p. 37-64, São Paulo.

40. Oliveira, C. A. de (2004). *A Evolução da Distribuição do Tamanho das Cidades Brasileiras: 1936-2000*. [artigo científico]. Disponível em: <http://www.bnb.gov.br/content/aplicacao/ETENE/Anais/docs/2004-aevolucao.pdf>. Acesso em: 20 fev. 2008.
41. Quah, D., 1993. “Empirical cross-section dynamics in economic growth”. *European Economic Review* v. 37, p. 426–434.
42. Reis, E.; Pimentel, M.; Alvarenga, A.I. (2008). *Áreas mínimas comparáveis para os períodos intercensitários de 1872 a 2000*. Mimeo, IPEA/DIMAC, Rio de Janeiro.
43. Resende, M. (2004). “Gibrat's Law and the Growth of Cities in Brazil: A Panel Data Investigation”. *Urban Studies*, v. 41, n. 8, p. 1537-1549.
44. Resende, M. (2005). “O crescimento econômico dos municípios mineiros: externalidades importam?”. *Cadernos BDMG*, n. 11, p. 5-33.
45. Rosen, K. T.; Resnick, M. (1980), “The Size Distribution of Cities: An Examination of the Pareto Law and Primacy”, *Journal of Urban Economics*, v. 8, p. 165-186.
46. Ruiz, R. M. (2005). “Estruturas Urbanas Comparadas: Estados Unidos e Brasil”. *Pesquisa e Planejamento Econômico*, julho, p. 715-737.
47. Soo, K. (2002). “Zipf's Law for Cities: A Cross Country Investigation”, *Working Paper*, *Centre for Economic Performance*, London School of Economics.