

Um estudo das metodologias e funcionalidades dos índices de segregação

Alexandre Xavier Ywata de Carvalho*

Camilo Rey Laureto**

Marina Garcia Pena***

Pedro Henrique Melo Albuquerque****

Waldery Rodrigues Junior*****

Os índices de segregação têm o intuito de avaliar a distribuição, mais especificamente, a dispersão ou a concentração de certo grupo em determinado espaço, como uma região geográfica. Enquanto uma análise mais qualitativa dos fatores definidores da mobilidade das pessoas foca em questões sociais, econômicas e culturais, os índices de segregação procuram determinar os fatores da mobilidade geográfica de forma quantitativa para, posteriormente, associar a fatores qualitativos. Este trabalho tem por objetivo tratar da metodologia dos índices de segregação, da forma como eles são construídos e a aplicação das medidas propostas. Assim, o foco é abordar os índices de segregação do tipo uniformização (*evenness*), exposição (*exposure*), concentração (*concentration*), agrupamento (*clustering*) e centralização (*centralization*) e, em seguida, construir uma aplicação empírica utilizando dados da Relação Anual de Informações Sociais (Rais). A análise consiste em avaliar a segregação espacial do rendimento anual recebido pelos trabalhadores formais nos municípios da Região Metropolitana de São Paulo (RMSP).

Palavras-chave: Segregação espacial. *Cluster*. Índices de segregação. Análise regional.

* Instituto de Pesquisa Econômica Aplicada – Ipea, Brasília-DF, Brasil (alexandre.ywata@ipea.gov.br).

** Instituto de Pesquisa Econômica Aplicada – Ipea, Brasília-DF, Brasil (camilo.laureto@ipea.gov.br).

*** Instituto de Pesquisa Econômica Aplicada – Ipea, Brasília-DF, Brasil (marina.pena@ipea.gov.br).

**** Departamento de Administração – Universidade de Brasília – UnB, Brasília-DF, Brasil (pedroa@unb.br).

***** Instituto de Pesquisa Econômica Aplicada – Ipea, Brasília-DF, Brasil (waldery@senado.gov.br).

Introdução

Segregação pode ser entendida como a distribuição irregular dos membros de um grupo num espaço territorial ou em categorias (raça, gênero, faixa salarial, etc.). Dessa forma, sua caracterização pode ser feita a partir de duas abordagens distintas: a sociológica e a geográfica. A primeira decorre da segregação entre dois grupos distintos e não correlacionados. Já no segundo caso, o foco é uma análise geográfica do território, ou seja, uma análise espacial de contiguidade ou uma medida de relação espacial entre regiões geográficas.

As medidas tradicionais não utilizam o fator geográfico, caracterizando-se como índices “aespaciais” nos quais o fator geográfico é constante no estudo. Neste artigo, serão explorados tanto os índices sem influência do fator espacial quanto aqueles que têm seus valores ponderados pelas distâncias entre os polígonos.¹

Os primeiros estudos dos índices de segregação surgiram na década de 1950, com o trabalho de Duncan e Duncan (1955a).² Segundo os autores, o problema das medidas de segregação era a sua construção a partir de noções simples e pobres, devendo ser formuladas por meio de articulações conceituais claras de seu processo. Duncan e Duncan (1955a, p. 217) caracterizavam essa simplicidade conceitual como uma “ingênua” noção de segregação:

[Segregação] é um conceito rico em subjetividade teórica e de um valor heurístico inquestionável. Claramente nós não desejamos sacrificar o capital da teorização e observação já investidos no conceito. Até o momento, isso é o que está envolvido na solução oferecida pelo operacionalismo “ingênuo”, mais ou menos uma combinação arbitrária de alguns convenientes procedimentos numéricos de um conceito verbal de segregação.

O estudo de segregação começou nos Estados Unidos no fim dos anos 1940 e início dos 1950, com vários artigos no *American Sociological Review* (BELL, 1954; COWGILL; COWGILL, 1951; JAHN, 1950, entre outros). Entretanto, os trabalhos focavam os índices de segregação baseados em dois grupos, pelo fato de o país estar em processo social e político de plena segregação racial (brancos e negros) e de gênero (homens e mulheres).

Entretanto, foi apenas na década de 1980 que os índices de dois grupos foram mais bem fundamentados, particularmente por James e Taeuber (1985) e Massey e Denton (1988). Neste segundo trabalho, os autores realizaram uma revolução metodológica na interpretação dos indicadores. A análise baseou-se em uma definição multidimensional de segregação, resultando na subdivisão das mensurações dos índices em cinco distintos conceitos das suas dimensões: uniformidade; exposição; agrupamento; centralização; e concentração.

A inclusão de elementos espaciais no cálculo dos índices mostrou-se cada vez mais necessária. Dessa forma, alguns índices espaciais de segregação foram propostos por Jakubs (1981), Morgan (1983), White (1983, 1986), Morrill (1991), Wong (1993, 1998,

¹ Ao tratar de polígonos, o referencial será regiões com limites geográficos bem definidos entre elas.

² Segundo Wong (1999), a medida de segregação de Duncan e Duncan (1955a) ainda é a melhor, pois capta efetivamente a dimensão da uniformização, a mais importante dimensão de segregação.

1999, 2003) e Reardon e O'Sullivan (2004). Entretanto, essas medidas espaciais são pouco utilizadas na literatura pelo fato de necessitarem de fatores geográficos, o que torna mais complicada a obtenção de dados. Além disso, são raros os *softwares* com a disponibilidade de índices de segregação espacial.

A literatura de índices de segregação não é completa no tema e necessita de mais estudos e aprofundamentos. Dessa forma, o presente trabalho tem por objetivo enumerar e expor as metodologias dos indicadores de segregação espacial, focando uma seleção desses índices. Também serão calculados os respectivos índices para faixas de renda em diferentes períodos nos municípios da Região Metropolitana de São Paulo (RMSP). O intuito desta análise é mostrar o comportamento da segregação das faixas de renda ao longo do tempo.

Indicadores de segregação

Os índices de segregação podem ser classificados em três grandes grupos: indicadores entre um grupo, dois grupos e de múltiplos grupos. No primeiro caso, a existência ou não de segregação é calculada em um único grupo, como, por exemplo, a ocorrência de segregação no grupo dos ricos. Já no segundo caso, este índice é obtido por meio da análise entre dois grupos distintos, sendo o Índice de Dissimilaridade de Duncan e Duncan (1955a) o mais utilizado. O último grupo é aquele em que existe abrangência maior de categorias, ou seja, calculam-se os indicadores com três ou mais grupos.

Em cada grande grupo existem ainda subdivisões dos indicadores, permitindo identificar o que, especificamente, cada indicador demonstra: uniformidade, exposição, concentração, centralização e agrupamento, para os índices de um e dois grupos; e uniformidade e exposição, para os de múltiplos grupos.

Indicadores do tipo uniformização

Os índices de segregação do tipo uniformização são utilizados quando se deseja analisar a dispersão em um grupo ou entre grupos, ou seja, a distribuição dos elementos da(s) população(ões) estudada(s) em uma área. Quanto mais desigual for esta distribuição, maior será a segregação existente. Quando a população for igualmente distribuída, os índices do tipo uniformização terão valores iguais a zero. A seguir, são descritos os índices "aespaiais" do tipo uniformização.

Índice de segregação

O índice de segregação (IS), proposto por Duncan e Duncan (1955a), é um dos utilizados para a medida de dispersão em único grupo. Ele mede a proporção da população no grupo que deveria se mudar para outras áreas para que a composição populacional fosse homogênea.

Seu cálculo é feito da seguinte forma:

$$IS = \frac{1}{2} \sum_{i=1}^n \left| \frac{x_i}{X} - \frac{t_i - x_i}{T - X} \right|$$

em que x_i é o total da população do grupo X na unidade i; t_i corresponde ao total da população na unidade i; T refere-se ao total da população; e X é o total da população no grupo X.

O IS varia entre zero e 1, sendo zero a ausência de segregação, ou seja, quando a população é distribuída homoganeamente ao longo do território. Quanto maior for seu valor, menor será a dispersão da população.

Por achar que a distância era um importante fator espacial que influía na segregação, Massey e Denton (1988) criaram uma combinação da distância entre os centroides para medir a segregação.

Outra forma adotada para incorporação do fator espacial no cálculo de segregação era por meio das matrizes de contiguidade. Tais formulações exigiam uma matriz de vizinhança formada por zeros e uns, em que zero indicava ausência de vizinhança e 1 significava que tais unidades eram vizinhas. Entretanto, na época, esta matriz era criada por meio da inspeção visual de mapas, o que foi considerado ineficiente. Dessa forma, tais índices sofreram algumas críticas. Atualmente, a abordagem espacial do índice de segregação utilizada foi proposta por Morrill (1991).

Índice de dissimilaridade

O índice de dissimilaridade (ID) proposto por Duncan e Duncan (1955a) é o mais citado na literatura sobre segregação. Ele fornece informação similar à proposta do IS, mas, nesse caso, para dois grupos. Seu valor é dado por:

$$ID = \frac{1}{2} \sum_{i=1}^n \left| \frac{x_i}{X} - \frac{y_i}{Y} \right|$$

em que y_i é o total da população do grupo Y na unidade i; e Y corresponde ao total da população do grupo Y.

O ID varia entre zero e 1. Ele pode, também, ser derivado da curva de Lorenz, na qual a proporção acumulada do grupo X é esboçada contra a proporção acumulada da população por meio das unidades, as quais são ordenadas em ordem crescente de proporção da população dos grupos.

O ID é usado como indicador da relação existente entre a população das diversas áreas ou a oportunidade da interação intrazonal entre grupos.

Índice de Gini (G)

Outro indicador do tipo uniformização é o índice de Gini, que representa a igualdade de distribuição de um grupo ou característica, sendo utilizado no ramo da economia para o cálculo de igualdade de renda. Ele pode ser obtido da seguinte forma:

$$G = \sum_{i=1}^n \sum_{j=1}^n \left[t_i t_j |p_i - p_j| / 2T^2 P(1-P) \right]$$

em que p_i é a proporção do grupo na unidade i , ou seja, x_i/t_i ; p_j corresponde à proporção do grupo na unidade j , ou seja, x_j/t_j ; e P refere-se à proporção do grupo no total, ou seja, X/T .

Outra forma de se calcular o índice de Gini é por meio da curva de Lorenz, o que faz com que ele seja intimamente relacionado ao ID (já que este também pode ser obtido da mesma forma).

O índice de Gini assume valores entre zero e 1, sendo que zero representa uma distribuição uniforme dos membros do grupo ou característica estudada e o 1 equivale a uma concentração total destes.

Índice de entropia

O índice de entropia, também conhecido como índice de informação ou índice de Shannon, é outra medida de dispersão, sendo comumente utilizado no ramo da comunicação como uma medida de incerteza. Essa funcionalidade é proposta por Shannon (1948).

A entropia é dada por:

$$E = (P) \ln \left[\frac{1}{P} \right] + (1 - P) \ln \left[\frac{1}{1 - P} \right]$$

O valor máximo atingido pelo índice é 1 e o mínimo é zero, que ocorre quando a distribuição é igualitária, ou seja, uma divisão entre 50 e 50 no caso de dois grupos.

Indicadores do tipo exposição

Os índices de segregação classificados como exposição são empregados para medir o grau de possível contato ou interação entre os membros de um mesmo grupo (indicadores de um grupo) ou entre membros de outros grupos (indicadores de dois grupos e de múltiplos grupos) nas unidades. Pode-se dizer que eles medem a probabilidade de um indivíduo encontrar membros do seu próprio grupo ou de grupos distintos.

Índices de isolamento e de interação

O índice de isolamento (${}_x P_x$) mede o grau de isolamento de um grupo em relação aos outros grupos; similarmente o índice de interação (${}_x P_y$) mede a interação de membros do grupo X com membros do grupo Y e o quanto um está exposto ao outro. Eles podem ser interpretados como a probabilidade de um membro do grupo X dividir espaço com outro membro do mesmo grupo (isolamento) ou a probabilidade de membros do grupo X dividir espaço com membros do grupo Y (interação).

Seus cálculos são dados por:

$${}_x P_x = \sum_{i=1}^n [x_i / X] \cdot [x_i / t_i]$$

$${}_x P_y = \sum_{i=1}^n [x_i / X] \cdot [y_i / t_i]$$

Ambos os índices variam entre zero e 1, já que representam uma probabilidade. A composição do índice de interação mostra que este não é simétrico, ou seja, ${}_xP_y$ não é igual a ${}_yP_x$. A única exceção que faz ambos os índices serem iguais é quando os dois grupos têm a mesma proporção da população.

Razão de correlação

O indicador de razão de correlação, chamado também de *ETA quadrado* (ETA^2), é uma medida entre a dispersão em cada categoria e a dispersão por todo o conjunto da população ou amostra. Ele é obtido por meio da seguinte fórmula:

$$ETA^2 = \left[\left(\sum_x P_x - P \right) / \left(1 - P \right) \right]$$

O índice de correlação varia entre zero e 1. Tal medida pode ser considerada um ajustamento do índice de interação.

Indicadores do tipo agrupamento

Os índices de agrupamento indicam o quão semelhante é certa área. As técnicas de agrupamento exploram semelhanças entre padrões e constroem grupos que tenham indivíduos com características semelhantes dentro dele, mas que sejam diferentes entre eles. Assim, quanto mais semelhante é a distribuição da população em um grupo em determinada área, mais ela estará conglomerada e, portanto, maior será a segregação ali presente.

Índice de agrupamento absoluto (ACL)

Preocupados com o problema de contiguidade, Massey e Denton (1988) elaboraram uma série de indicadores que pudessem incorporar tal fator ao seu cálculo. Assim, os autores desenvolveram um índice adaptado de Dacey (1968) e Geary (1954) que media a aglomeração absoluta em um espaço, o ACL:

$$ACL = \frac{\left[\sum_{i=1}^n \left(\frac{x_i}{X} \right) \sum_{j=1}^n (c_{ij} x_j) \right] - \left[\frac{X}{n^2} \sum_{i=1}^n \sum_{j=1}^n c_{ij} \right]}{\left[\sum_{i=1}^n \left(\frac{x_i}{X} \right) \sum_{j=1}^n (c_{ij} t_j) \right] - \left[\frac{X}{n^2} \sum_{i=1}^n \sum_{j=1}^n c_{ij} \right]}$$

em que c_{ij} representa os valores da matriz de contiguidade (matriz de vizinhança) e n é o total de unidades de área.

O agrupamento absoluto tem uma variação entre zero, que representa uma baixa segregação, e 1, que equivale à máxima contiguidade do grupo ao longo da unidade. O índice também expressa o número médio de membros de um grupo nas áreas próximas como sendo a proporção entre esses membros de grupos em áreas próximas em relação à população total nessas áreas.

Proximidade média entre membros do grupo X e proximidade média entre membros do grupo X e Y

Quando se deseja medir a segregação em municípios ou países, por exemplo, por meio de índices que levam à contiguidade como fator relevante, a obtenção de uma matriz de vizinhança é relativamente simples quando se dispõe de uma malha digital desta área e um *software* computacional que o faça. Entretanto, muitas vezes o foco da análise de segregação são residências, o que faz com que a criação da matriz de contiguidade seja mais complicada. Dessa forma, uma saída foi não mais utilizar as matrizes de vizinhança como fator espacial e sim a distância entre os centroides das unidades de interesse. Portanto, White (1986) propôs os índices Proximidade Média entre Membros do Grupo X, P_{xx} , (um grupo) e Proximidade Média entre Membros do Grupo X e Y, P_{xy} , (dois grupos):

$$P_{xx} = \frac{1}{X^2} \sum_{i=1}^n \sum_{j=1}^n x_i x_j d_{ij}$$

$$P_{xy} = \frac{1}{XY} \sum_{i=1}^n \sum_{j=1}^n x_i y_j d_{ij}$$

em que d_{ij} é a distância entre os centroides da unidade i e da j .

Esses índices medem simplesmente a proximidade média intragrupos (P_{xx}) ou a média das proximidades entre grupos (P_{xy}). Seu menor valor é zero, porém ele não possui valor máximo.

Existem derivações desses índices também bastante usuais, em que, em vez da utilização comum de d_{ij} , usa-se o exponencial negativo da distância, ou seja, $d^*_{ij} = \exp(-d_{ij})$.

Índice de agrupamento relativo (RCL)

O RCL é uma derivação dos índices de média aproximada e utilizado para o caso em que existem dois grupos, podendo ser calculado da seguinte forma:

$$RCL = \frac{P_{xx}}{P_{yy}} - 1$$

sendo P_{xx} e P_{yy} os índices citados anteriormente. Ele mede a aglomeração relativa dos grupos, já que compara as distâncias médias entre os indivíduos de um grupo com outro.

O índice de agrupamento relativo não possui limites, variando de $-\infty$ a $+\infty$. Quando o valor obtido é igual a zero ($P_{xx} / P_{yy} = 1$), as aglomerações dos dois grupos são iguais. Quando o valor é positivo ($P_{xx} / P_{yy} > 1$), o grupo X possui aglomeração maior do que o grupo Y. No caso de um valor negativo ($P_{xx} / P_{yy} < 1$), o grupo Y apresenta maior aglomeração quando comparado ao grupo X.

Indicadores do tipo concentração

Os indicadores de concentração referem-se ao espaço físico que certo grupo ocupa. Quanto maior a área de ocupação desse grupo, menos concentrado ele será e, portanto,

menos segregado. Por outro lado, se o grupo está em um pequeno espaço geográfico, considera-se que existe uma grande segregação.

Índice delta (DEL)

Na literatura foram criados relativamente poucos índices para medir a concentração em uma região. Um deles foi o DEL, que é uma aplicação específica do ID. Seu valor é dado por:

$$DEL = \frac{1}{2} \sum_{i=1}^n \left| \frac{x_i}{X} - \frac{a_i}{A} \right|$$

em que A é a área total e a_i é a área da unidade espacial i.

Esse índice possui valores entre zero e 1. Sua metodologia mostra a proporção de membros de um grupo residente em uma área que estão acima da densidade média deste grupo. Ele pode representar a proporção de membros deste grupo que deveriam se mudar para outra área para que a densidade fosse uniforme ao longo do território.

Índice de concentração absoluta (ACO)

Outra medida utilizada para calcular a concentração em um espaço é o ACO, que compara o máximo e o mínimo de áreas que poderiam ser habitadas por um grupo com a área total ocupada por este. Sendo assim:

$$ACO = 1 - \frac{\left(\sum_{i=1}^n \left(\frac{x_i A_i}{X} \right) - \sum_{i=1}^{n_1} \left(\frac{t_i A_i}{T_1} \right) \right)}{\left(\sum_{i=n_2}^n \left(\frac{t_i A_i}{T_2} \right) - \sum_{i=1}^{n_1} \left(\frac{t_i A_i}{T_1} \right) \right)}$$

em que T_1 é a soma de todos os t_i da área 1 à área n_1 e T_2 corresponde à soma de todos os t_i da área n_2 à área n .

Para a utilização dessa formulação, as áreas devem estar em ordem crescente de tamanho (territorial).

O primeiro termo do numerador indica a área média que está ocupada pelos membros do grupo em questão, enquanto o segundo corresponde à área média de habitação quando a concentração é máxima. Já no denominador, o primeiro termo indica a área média que seria ocupada sobre condições de concentração mínima.

O índice atinge valores que variam entre zero, concentração mínima do grupo que está sendo analisado, e 1, concentração máxima.

Índice de concentração relativa (RCO)

Como visto anteriormente, o ACO mede a concentração absoluta em uma área geográfica. Para uma análise da concentração relativa de certo grupo, foi criado o RCO,³ que mede a proporção de espaço ocupado por um grupo quando comparado com outro.

³ Para mais detalhes, ver Massey e Denton (1988).

$$RCO = \frac{\left(\frac{\sum_{i=1}^n \left(\frac{x_i A_i}{X} \right) - 1}{\sum_{i=1}^n \left(\frac{y_i A_i}{Y} \right) - 1} \right)}{\left(\frac{\sum_{i=1}^n \left(\frac{t_i A_i}{T_1} \right) - 1}{\sum_{i=n_2}^n \left(\frac{t_i A_i}{T_2} \right) - 1} \right)}$$

O RCO é uma razão que compara dois grupos com o caso de uma máxima concentração do primeiro contra o de uma mínima concentração do segundo. Dessa forma, se existem dois grupos, X e Y, primeiro é feita uma razão entre a concentração de X e a de Y. Depois essa razão é comparada com aquela entre os mesmos grupos, em que X seria maximamente concentrado e Y minimamente concentrado.

O coeficiente é padronizado para que varie entre -1 e 1, sendo que zero indica que os dois grupos estão igualmente concentrados. O valor -1 mostra que o segundo grupo excede ao máximo possível a concentração do primeiro, enquanto o 1 representa que o primeiro grupo é o que excede.

Indicadores do tipo centralização

Os indicadores do tipo centralização medem o grau de proximidade de certo grupo com os centros das unidades de área, por exemplo, o centro da cidade. Grupos que se instalam perto do centro da cidade tendem a ser concentrados. A importância deste estudo pode ser entendida na sua aplicação na área econômica, mais especificamente estudos sobre as economias do centro e da periferia de uma área.

Proporção na área central

Assim como no caso dos indicadores de concentração, também existem poucos indicadores de segregação para a centralização. Um deles, bastante usado, é a proporção na área central, que pode ser calculada por:

$$PCC = \frac{X_{cc}}{X}$$

em que X_{cc} é a população total residente no centro.

Esse índice é uma proporção de habitantes de uma área que vivem na zona central sobre a população total da área. Por ser uma proporção, seu limite está entre zero e 1. Um dos problemas desse índice é a delimitação da abrangência espacial da região, já que esta é, muitas vezes, uma definição política ou econômica, e não geográfica.

Índice de centralização absoluta

Outra forma de se computar a centralização em uma região é por meio do índice de centralização absoluta (ACE), que indica a proporção de moradores que deveriam se mudar para que a distribuição da população de certo grupo fosse homogênea.

$$ACE = \left(\sum_{i=1}^n X_{i-1} S_i \right) - \left(\sum_{i=1}^n X_i S_{i-1} \right)$$

em que S_i é a proporção acumulada da área na unidade i .

Esse índice varia entre 1 e -1, sendo que valores positivos indicam uma tendência de o grupo X residir perto do centro da cidade e valores negativos representam uma tendência a morar em outras áreas. Quando a medida for zero, o grupo tem uma distribuição uniforme por toda a cidade.

Índice de centralização relativa (RCE)

Por fim, existe um índice que mede a centralização relativa em uma área, o RCE. Essa medida é utilizada para o caso de dois grupos, da seguinte forma:

$$RCE = \left(\sum_{i=1}^n X_{i-1} Y_i \right) - \left(\sum_{i=1}^n X_i Y_{i-1} \right)$$

em que X_i e Y_i são as proporções acumuladas dos grupos X e Y na unidade i .

O RCE pode ser interpretado como a cota de membros do grupo X que deveriam mudar de local de residência para atingir o grau de centralidade do grupo Y. Este índice varia entre -1 e 1, sendo que os valores positivos indicam que os membros do grupo X estão localizados mais próximos do centro da cidade do que os do grupo Y, enquanto os negativos mostram que os membros do grupo X estão distribuídos mais distantes do centro. Quando o índice é zero, os dois grupos têm a mesma distribuição espacial acerca do centro da cidade.

Na Figura 1, observam-se as diferenças entre os diversos tipos de indicadores e como cada localidade se comporta com um alto e um baixo nível de segregação.

Quanto mais escuro o polígono, maior a quantidade de pessoas nesta área. Na baixa segregação do tipo uniformidade, tem-se uma população igualmente distribuída ao longo do território. Na alta segregação, a população está distribuída de forma desigual, aglomerando-se em alguns polígonos específicos. Para os índices do tipo exposição, a baixa segregação mostra que o grupo X não está isolado, ou seja, ele interage com outros grupos. Na alta segregação, o grupo X está isolado nos polígonos em que se encontra. No caso dos índices de concentração, na baixa segregação, a população está espalhada ao longo do território, enquanto a alta segregação apresenta concentração da população/grupo em certa área geográfica. Os índices de agrupamento que indicam baixa segregação mostram que não existe um *cluster* no território. Já uma alta segregação neste grupo representa conglomerado no território. Por fim, os índices de centralização determinam a existência de concentração no centro da cidade. A baixa segregação indica ausência de centralização e a alta mostra concentração no centro da região.

FIGURA 1
Representação da alta e baixa segregação para os diferentes tipos de índice
 Baixa segregação Alta segregação



Fonte: Adaptado de Aparicio (2000).

Análise empírica dos indicadores de segregação

De modo a elucidar com mais clareza as funcionalidades dos índices de segregação enumerados anteriormente, esta seção desenvolve uma aplicação empírica. O objetivo é gerar

os índices de segregação para faixas de renda nos municípios da RMSP. O período escolhido foi 1994, 1999, 2004 e 2009. Essa definição de intervalos de cinco anos na elaboração dos resultados teve como intuito a expansão da série de tempo em período de 15 anos. Com isso, quanto maior o intervalo da série, maior será as variações na renda ao longo do tempo nas regiões e, portanto, melhores serão as conclusões sobre o comportamento da segregação da renda.

O cálculo da renda das pessoas foi baseado na Relação Anual de Informações Sociais (Rais),⁴ disponibilizada pelo Ministério do Trabalho e Emprego (MTE).

A renda é calculada pela variável remuneração de dezembro em salários mínimos recebidos pelos trabalhadores no ano de referência. Entretanto, para um estudo da renda ao longo do tempo, é necessária a transformação dos rendimentos em valores reais. O valor-base foi definido para dezembro de 2009, ou seja, o deflator apresenta valor igual a 1 nessa data. O deflator utilizado foi o Índice Nacional de Preços ao Consumidor (INPC)⁵ (Tabela 1). Além disso, para transformar os valores para a moeda corrente brasileira (real), o número de salários mínimos recebidos pelo trabalhador foi multiplicado pelo valor do salário mínimo da correspondente data. Os salários mínimos estão de acordo com os valores estabelecidos pelo MTE⁶ (Tabela 1).

TABELA 1
Deflator baseado no INPC para adequação das rendas (1) e valor dos salários mínimos (2)
Brasil – 1994-2009

Anos	Valor do deflator (3)	Valor do salário mínimo (R\$)
1994	0,3323	70,00
1999	0,5128	136,00
2004	0,7942	260,00
2009	1,0000	465,00

Fonte: Tabelas históricas do INPC fornecidas pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

(1) Em valores de dezembro de 2009.

(2) Com base no Ministério do Trabalho e Emprego – MTE.

(3) O deflator foi apresentado na tabela com quatro casas decimais por motivos ilustrativos. No cálculo da renda foi utilizado o deflator com quinze casas decimais.

As faixas de renda foram definidas de acordo com os percentis da variável em questão. Os valores dos cortes são apresentados na Tabela 2 e as faixas das rendas foram estabelecidas da seguinte forma: pessoas com renda menor ou igual ao valor de corte do percentil de 20% (grupo 1); pessoas com renda maior que o valor de corte do percentil de 20% e menor ou igual ao valor de corte do percentil de 50% (grupo 2); pessoas com renda maior que o valor

⁴ A Rais foi instituída pelo Decreto n. 76.900, de 2 de dezembro de 1975. Em 14 de dezembro de 2000 foi publicada a Portaria n. 945. A Rais é um registro administrativo, de periodicidade anual, criada com a finalidade de suprir as necessidades de controle, de estatísticas e de informações às entidades governamentais da área social.

⁵ Tabelas e metodologia disponíveis em: <http://www.ibge.gov.br/home/estatistica/indicadores/precos/inpc_ipca/defaultinpc.shtml>.

⁶ Para mais esclarecimentos: <<http://www.mte.gov.br/rais/default.asp>>.

de corte do percentil de 50% e menor ou igual ao valor de corte do percentil de 80% (grupo 3); e pessoas com renda maior que o valor de corte do percentil de 80% (grupo 4).

TABELA 2
Valores dos percentis (1) (de rendimento em reais) em relação à remuneração dos trabalhadores formais Brasil – 1994-2009

Anos	Percentil 20%	Percentil 50%	Percentil 80%
1994	528	998	2.333
1999	678	1.193	2.622
2004	635	1.034	2.235
2009	702	1.111	2.483

Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994 a 2009.

(1) Os cortes foram apresentados na tabela sem casas decimais por motivos ilustrativos. No cálculo da renda, foram utilizadas nos cortes sete casas decimais.

Os indicadores de segregação foram calculados para a RMSF pelo fato de a região ser composta pelos mesmos municípios nos quatro períodos estudados. Isto evita a distorção dos índices ao se introduzir, na análise, a variação da organização dos municípios pertencentes à região metropolitana em cada período.

O nível geográfico atingido foi o municipal, baseado na estrutura de municípios estabelecida pelo IBGE. Essa estrutura geográfica é importante para o cálculo dos índices que englobam fatores espaciais, comentados anteriormente.

Para o cálculo dos índices foi utilizado o *software* gratuito IpeaGEO, desenvolvido por pesquisadores do Ipea, que permite o cálculo de vários índices de segregação tanto para a categoria de um grupo quanto para as categorias de dois e de múltiplos grupos. Daqueles apresentados na metodologia, são calculados pelo programa os índices de: segregação; segregação ajustada; entropia; Gini; isolamento; razão de correlação; agrupamento absoluto; proximidade média entre membros de um grupo (tanto o índice quanto o exponencial dele); dissimilaridade; dissimilaridade ajustada; interação; proximidade média entre membros do grupo X; proximidade média entre membros dos grupos X e Y; e agrupamento relativo. Entretanto, o programa não calcula índices do tipo centralização e concentração.

Por fim, para uma análise mais completa da segregação da renda, foram elaborados mapas do índice de concentração, ou seja, a razão entre a população pertencente a certa faixa de renda (percentis 20%, 50% e 80%) e a população total para cada município da referida região. Os valores de corte do método quebras naturais (*natural breaks*, de Jenks)⁷ foram fixados para os cortes de 1994. Isso possibilita a visualização no mapa do comportamento do índice ao longo do período. Os mapas estão no Anexo.

⁷ Para mais informações sobre o método, consultar o *site* do *software* ArcGis: <<http://webhelp.esri.com/arcgisSDEsktop/9.3>>.

Resultados

A seguir, são apresentados os índices de segregação para três categorias: um grupo, dois grupos e múltiplos grupos. Todos os resultados foram gerados pelo *software* IpeaGEO e os mapas pelo *software* ArcGis.

Resultados para os índices de um grupo

Primeiramente, o estudo da renda na RMSP baseia-se nos indicadores de um grupo para todos os grupos criados. Cada índice é calculado para um grupo individualmente.

Na Tabela 3, nota-se que há uma tendência de diminuição da segregação da classe mais pobre da população. Todos os índices referentes ao grupo 1 apresentam menores resultados em 2009 quando comparados aos valores obtidos pelo grupo 1 em períodos anteriores. Nas outras classes também é possível perceber que a segregação também seguiu uma tendência decrescente ao longo dos anos, porém, de forma mais branda do que no primeiro grupo.

TABELA 3
Indicadores de um grupo para a distribuição de rendimento dos empregados formais
Região Metropolitana de São Paulo – 1994-2009

Grupos	Índice de segregação	Índice de segregação ajustado pela contiguidade	Indicador de entropia	Coefficiente de Gini
1994				
Grupo 1	0,068	-0,057	0,017	0,119
Grupo 2	0,042	-0,043	0,005	0,061
Grupo 3	0,028	-0,028	0,002	0,042
Grupo 4	0,074	-0,001	0,016	0,111
1999				
Grupo 1	0,104	0,008	0,017	0,121
Grupo 2	0,039	-0,019	0,003	0,052
Grupo 3	0,051	-0,002	0,004	0,058
Grupo 4	0,077	0,009	0,013	0,107
2004				
Grupo 1	0,110	0,016	0,015	0,131
Grupo 2	0,038	-0,013	0,003	0,052
Grupo 3	0,051	0,004	0,003	0,059
Grupo 4	0,082	0,020	0,012	0,109
2009				
Grupo 1	0,066	-0,026	0,008	0,084
Grupo 2	0,038	-0,014	0,002	0,051
Grupo 3	0,025	-0,026	0,002	0,041
Grupo 4	0,083	0,022	0,009	0,100

(continua)

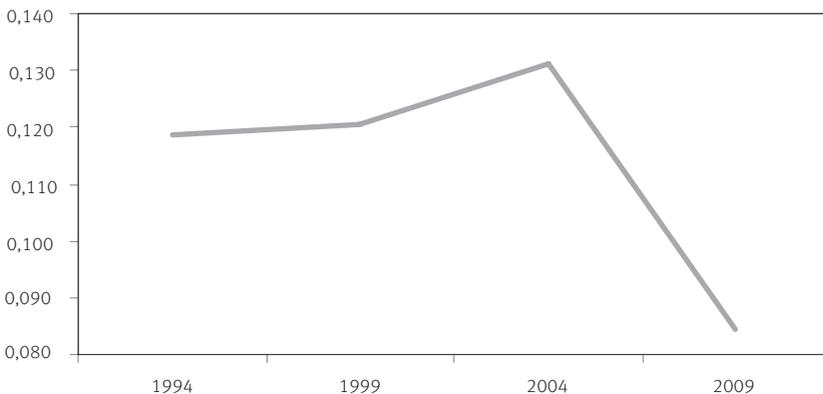
(continuação)

Variáveis	Índice de isolamento	Razão de correlação	Índice de agrupamento absoluto	Proximidade média entre membros do grupo X	Proximidade média entre membros de um grupo (exp)
1994					
Grupo 1	0,213	0,019	0,135	13,654	0,001
Grupo 2	0,305	0,006	0,204	11,857	0,002
Grupo 3	0,304	0,003	0,216	11,459	0,003
Grupo 4	0,213	0,017	0,155	11,504	0,002
1999					
Grupo 1	0,217	0,020	0,156	17,840	0,001
Grupo 2	0,302	0,003	0,226	14,712	0,002
Grupo 3	0,303	0,005	0,208	12,201	0,002
Grupo 4	0,209	0,011	0,137	11,134	0,002
2004					
Grupo 1	0,215	0,017	0,155	18,517	0,001
Grupo 2	0,302	0,003	0,228	16,069	0,002
Grupo 3	0,302	0,004	0,216	13,771	0,002
Grupo 4	0,209	0,010	0,140	12,395	0,002
2009					
Grupo 1	0,208	0,009	0,147	16,697	0,001
Grupo 2	0,300	0,003	0,225	15,697	0,002
Grupo 3	0,303	0,002	0,227	14,585	0,002
Grupo 4	0,207	0,008	0,135	12,005	0,002

Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994 a 2009.

Analisando o índice de Gini no primeiro grupo, nota-se que, após uma sequência crescente ao longo dos três primeiros períodos, em 2009 este indicador teve uma queda considerável em relação aos anos anteriores, principalmente na primeira classe (passando de 0,13 para 0,08). Como a diferença entre o índice do grupo 1 e os dos outros grupos ficou menor, é possível concluir que há maior igualdade na distribuição de renda na RMSP.

GRÁFICO 1
Índice de Gini para o grupo 1 com relação à distribuição de rendimento dos empregados formais
Região Metropolitana de São Paulo – 1994-2009



Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994 a 2009.

A razão de correlação foi outro índice que diminuiu para todas as classes ao longo do período analisado. Esta diferença é mais nítida para o grupo 4, já que o índice caiu de 0,017 para 0,008 entre 1994 e 2009, ou seja, seu valor reduziu-se quase pela metade.

Os índices que mostraram maior existência de segregação para todas as classes foram os de isolamento e de agrupamento absoluto. Para o primeiro, a probabilidade de membros dos grupos 1 e 4 (valores próximos a 0,20) interagirem com membros do seu próprio grupo é menor do que nos grupos 2 e 3 (aproximadamente 0,30). O mesmo acontece no segundo índice, no qual os grupos 1 e 4 possuem valores de aproximadamente 0,15 e os grupos 2 e 3, de 0,22.

Em 1999 e 2004, o índice de segregação apresentou valores quase duas vezes maiores para os grupos 1 e 3. Isso indica que, nesses anos, houve uma distribuição menos homogênea da população dessas classes no território. Por outro lado, os grupos 2 e 4 se mantiveram quase constantes, com uma variação mínima ao longo dos anos (os índices passaram de 0,042 para 0,038 e de 0,074 para 0,083, respectivamente). Pelos resultados obtidos para os índices de um grupo, pode-se concluir que, em todos os grupos analisados, houve diminuição da segregação na RMSP, em 2009.

Resultados para os índices de dois grupos

O propósito da análise de grupos é verificar se os resultados se comportam de forma semelhante entre dois grupos específicos quando comparados com os resultados de um grupo com todo o resto. Ao se considerarem os índices de dois grupos é possível saber entre quais grupos existe tal variação. Esta análise é pertinente para saber se o percentil de renda mais baixa está se aproximando do percentil de renda mais alta. Com isso, pode-se supor uma diminuição da segregação entre grupos diferentes. Da mesma forma, os resultados da abordagem de múltiplos grupos serão comparados com aqueles das categorias um grupo e dois grupos.

Na categoria de dois grupos, o primeiro índice a ser explorado é o de dissimilaridade. De acordo com essa medida, de 1999 para 2009, a segregação entre os grupos 1 e 4 diminuiu (passando de 0,137 para 0,103), ou seja, a menor segregação apresentada pelos índices de um grupo decorre do fato de os grupos 1 e 4 serem menos segregados no período. Entretanto, o índice em 1994 é muito semelhante ao de 2009 para todos os pares de grupos. Isso mostra que o cenário da segregação na RMSP era semelhante nesses dois períodos. Portanto, houve alguma mudança que fez com que, a partir de 1999, os índices aumentassem, voltando a diminuir posteriormente em 2009. Os grupos 2 e 3 também apresentaram redução na segregação no período anteriormente citado (de 0,055 para 0,028), o que corrobora os resultados dos índices de um grupo. Tais fatos podem ser observados no Gráfico 2.

TABELA 4
Indicadores de dois grupos para a distribuição de rendimento dos trabalhadores formais
Região Metropolitana de São Paulo – 1994-2009

Classe 1	Classe 2	Índice de dissimilaridade entre XY	Índice de dissimilaridade ajustado para contiguidade entre XY	Índice de interação
1994				
Grupo 1	Grupo 2	0,062	-0,006	0,299
Grupo 1	Grupo 3	0,070	0,006	0,297
Grupo 1	Grupo 4	0,106	8,863	0,192
Grupo 2	Grupo 3	0,024	3,501	0,303
Grupo 2	Grupo 4	0,082	0,161	0,196
Grupo 3	Grupo 4	0,064	0,132	0,201
1999				
Grupo 1	Grupo 2	0,073	0,054	0,301
Grupo 1	Grupo 3	0,116	0,141	0,292
Grupo 1	Grupo 4	0,137	23,428	0,190
Grupo 2	Grupo 3	0,055	24,558	0,299
Grupo 2	Grupo 4	0,086	0,201	0,197
Grupo 3	Grupo 4	0,038	0,061	0,203
2004				
Grupo 1	Grupo 2	0,082	0,073	0,301
Grupo 1	Grupo 3	0,121	0,153	0,292
Grupo 1	Grupo 4	0,143	57,690	0,193
Grupo 2	Grupo 3	0,049	19,130	0,299
Grupo 2	Grupo 4	0,090	0,221	0,198
Grupo 3	Grupo 4	0,047	0,096	0,202
2009				
Grupo 1	Grupo 2	0,054	0,018	0,298
Grupo 1	Grupo 3	0,069	0,044	0,298
Grupo 1	Grupo 4	0,103	61,940	0,196
Grupo 2	Grupo 3	0,028	2,473	0,302
Grupo 2	Grupo 4	0,092	0,226	0,197
Grupo 3	Grupo 4	0,068	0,150	0,200

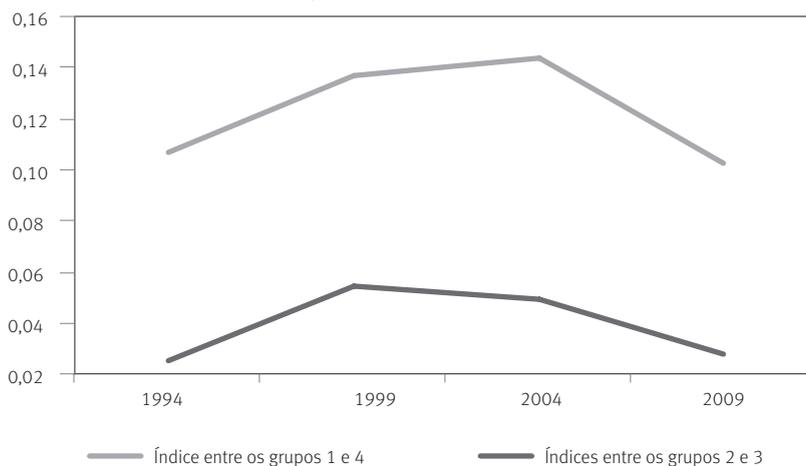
(continua)

(continuação)

Classe 1	Classe 2	Proximidade média entre os membros dos grupos X e Y	Proximidade média entre os membros de um grupo X e grupo Y	Índice de agrupamento relativo
1994				
Grupo 1	Grupo 2	12,932	0,553	0,240
Grupo 1	Grupo 3	12,735	0,550	0,239
Grupo 1	Grupo 4	12,825	0,526	0,251
Grupo 2	Grupo 3	11,743	0,565	0,250
Grupo 2	Grupo 4	11,814	0,541	0,241
Grupo 3	Grupo 4	11,574	0,538	0,240
1999				
Grupo 1	Grupo 2	16,532	0,448	0,241
Grupo 1	Grupo 3	15,372	0,477	0,242
Grupo 1	Grupo 4	14,907	0,480	0,252
Grupo 2	Grupo 3	13,667	0,515	0,250
Grupo 2	Grupo 4	13,162	0,518	0,241
Grupo 3	Grupo 4	11,810	0,553	0,240
2004				
Grupo 1	Grupo 2	17,538	0,415	0,241
Grupo 1	Grupo 3	16,461	0,440	0,242
Grupo 1	Grupo 4	15,844	0,449	0,252
Grupo 2	Grupo 3	15,099	0,477	0,250
Grupo 2	Grupo 4	14,456	0,487	0,241
Grupo 3	Grupo 4	13,206	0,517	0,240
2009				
Grupo 1	Grupo 2	16,381	0,449	0,240
Grupo 1	Grupo 3	15,822	0,461	0,240
Grupo 1	Grupo 4	14,594	0,491	0,251
Grupo 2	Grupo 3	15,293	0,474	0,250
Grupo 2	Grupo 4	14,059	0,504	0,241
Grupo 3	Grupo 4	13,439	0,517	0,240

Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994 a 2009.

GRÁFICO 2
Índices de dissimilaridade de dois grupos para a distribuição do rendimento dos empregados formais
Região Metropolitana de São Paulo – 1994-2009



Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994 a 2009.

No caso do índice de interação, pelo fato de ser uma medida mais estável ao longo do tempo, é mais pertinente focar a interpretação nas diferenças entre os grupos, ou seja, tratar dos grupos em um mesmo ano em vez de comparar os grupos entre os anos. De acordo com o índice, os membros do grupo 4 apresentam a menor probabilidade de interação entre eles quando comparados aos membros dos outros grupos. Dessa forma, enquanto a probabilidade dos grupos 1, 2 e 3 de interagirem entre si possui valor aproximado de 0,3, a do quarto grupo diminui para 0,2. Esta tendência permanece constante em todos os anos estudados.

Analizando todos os índices da categoria dois grupos, nota-se que há uma diminuição da segregação da renda na RMS, sendo mais importante quando considerada a classe mais pobre, ou seja, pessoas do grupo 1. Existindo uma segregação baixa ou decrescente deste grupo em relação aos outros, é possível dizer que este grupo está sendo menos excluído e que a região tem uma tendência de ser cada vez mais igualitária.

Para concluir a análise dos índices de segregação, é preciso verificar os resultados dos índices de múltiplos grupos. Esse estudo será desenvolvido no próximo tópico com o índice de dissimilaridade e o de Gini.

Resultados para os índices de múltiplos grupos

No presente estudo, o foco da categoria de múltiplos grupos correspondeu aos índices de dissimilaridade e de Gini. Essa categoria gera um valor para cada índice com o propósito de mostrar a segregação no âmbito do grupo como um todo, ou seja, os índices mostram a segregação com apenas um grupo que engloba todas as classes em apenas uma. Dessa forma, os índices serão gerados para o grupo de municípios da Região Metropolitana de São Paulo.

Com relação ao ID, seu valor foi de 0,0505 em 1994 e 0,0500 em 2009. Desse modo, constata-se que ele se comportou da mesma forma que as categorias de dois grupos e de um grupo, em que, para algumas classes específicas, alguns índices se comportaram de maneira estacionária. Esses indicadores foram: o índice de segregação (de um grupo) nas classes 1 e 3; o índice de Gini (de um grupo) na classe 3; e a razão de correlação (de um grupo) na classe 3, entre outros. De 1994 a 1999, o índice aumentou em 1%, mantendo-se estável em 2004, com uma posterior queda de 1% em 2009.

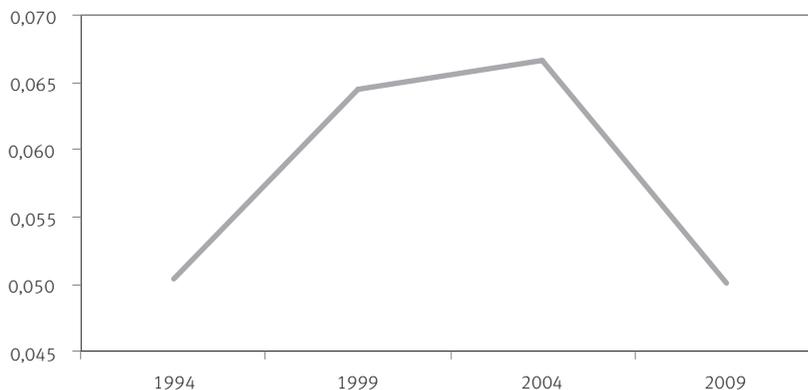
TABELA 5
Indicadores de múltiplos grupos para a distribuição do rendimento dos trabalhadores formais
Região Metropolitana de São Paulo – 1994-2009

Anos	Índice de dissimilaridade	Índice de Gini
Todos os grupos		
1994	0,0505	0,079
1999	0,0646	0,081
2004	0,0666	0,083
2009	0,0500	0,066

Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994 a 2009.

O ID mostra a proporção de pessoas que deveriam se mudar para a classe se tornar mais igualitária. Dessa forma, em 1999, aproximadamente 6,46% da população da RMSP deveria se mudar para que nessa área a distribuição da renda fosse igualitária. Para 2009, esta proporção diminuiu para 5%, valor encontrado também em 1994. O Gráfico 3 ilustra a evolução do índice ao longo dos anos.

GRÁFICO 3
Índices de dissimilaridade de múltiplos grupos para a distribuição do rendimento dos trabalhadores formais
Região Metropolitana de São Paulo – 1994-2009



Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994 a 2009.

O índice de Gini apresentou trajetória de queda entre 1994 e 2009, passando de 0,079 para 0,066. Os valores aumentaram de 1994 a 1999, com estabilização em 2004 e posterior queda em 2009. A variação de 2004 até 2009 foi semelhante à verificada entre 1994 e 1999, com valor próximo a 2%.

Essa diminuição do índice mostra que a segregação reduziu-se de forma generalizada em todos os municípios estudados. Assim, a distribuição da renda, segundo os percentis, tornou-se mais igualitária entre 1994 e 2009.

Considerações finais

A segregação espacial pode ser definida como uma distribuição desigual em um espaço territorial, ou de categorias, dos membros de determinado grupo. Os índices de segregação são classificados em três grandes grupos: de um grupo, de dois grupos e de múltiplos grupos.

Em cada classificação existem subdivisões dos indicadores para dividi-los segundo sua funcionalidade. Para as categorias de um grupo e de dois grupos, os indicadores são subdivididos em uniformidade, exposição, concentração, centralização e agrupamento. Já a categoria múltiplos grupos pode ser subdividida em uniformidade e exposição.

Este artigo teve por objetivo classificar e discorrer sobre as metodologias dos indicadores de segregação espacial, bem como calcular os índices por faixas de renda em quatro períodos distintos. A área escolhida foi a RMSP, pelo fato de a organização geográfica dos municípios

desta região se manter constante ao longo do tempo. O período da amostra compreende de 1994 a 2009, com intervalos de cinco anos, ou seja, 1994, 1999, 2004 e 2009. O intuito desta análise foi mostrar o comportamento da segregação das faixas de renda ao longo do tempo.

A amostra de dados utilizou como variável de interesse o rendimento do trabalho formal para os municípios da RMSP. Os dados foram divididos em percentis de 20%, 50% e 80%. O rendimento dos trabalhadores formais foi obtido a partir da Rais.

Afim de empiricamente explorar de forma mais eficaz a distribuição da renda entre esses municípios, foram calculados, além dos índices de segregação, os de concentração, que posteriormente foram colocados em mapas para visualização geográfica do comportamento da concentração da renda ao longo dos períodos.

Os principais resultados da categoria de um grupo mostram que, para o índice de segregação – o mais confiável segundo a literatura –, os valores para os grupos 1 e 3 permaneceram praticamente constantes entre 1994 e 2009. Entretanto, o índice de Gini para o grupo 1 cresceu entre 1994 e 2004, com posterior queda em 2009.

Na categoria de dois grupos fica mais evidente que a segregação entre o grupo 1 e todos os outros grupos está diminuindo segundo o índice de dissimilaridade. Isso mostra uma aproximação entre os grupos de renda mais baixa com relação aos de renda mais elevada. Além disso, ocorreu também uma redução de segregação entre as classes 2 e 3, ou seja, uma aproximação de renda entre estes dois grupos. Já em relação aos períodos de 1999 e 2004, os índices apresentaram valores praticamente estacionários.

Por fim, na categoria de múltiplos grupos, o ID apresentou valor constante entre 1994 e 2009. Entretanto, o índice de Gini teve uma queda de aproximadamente 1%. Isso mostra um decréscimo das desigualdades de renda entre os trabalhadores formais.

É pertinente ressaltar que os índices das categorias de um grupo, de dois grupos e de múltiplos grupos apresentaram valores com tendências semelhantes. Isso significa que os índices são complementares e o estudo de uma categoria não substitui a análise das outras. Uma lacuna no artigo que deve ser preenchida refere-se à estimação da variância para determinar intervalos de confiança para os indicadores de segregação expostos, com o intuito de analisar a robustez dos indicadores utilizados.

Referências

- APPARICIO, P. Les indices de ségrégation résidentielle: un outil intégré dans un système d'information géographique, Cybergeo. **European Journal of Geography**, n. 134, Jun. 2000.
- APPARICIO, P.; LELOUP, X.; RIVET, P. La diversité montréalaise à l'épreuve de la ségrégation: pluralisme et insertion résidentielle des immigrants. Revue de l'intégration et de la migration internationale. **Journal of International Migration and Integration**, v. 8, p. 63-87, 2007.
- APPARICIO, P.; SÉGUIN, A. M. La division de l'espace résidentiel montréalais en fonction de la langue maternelle: apport des indices de ségrégation résidentielle. **Canadian Journal of Urban research/Revue Canadienne de Recherché Urbaine**, v. 11, p. 265-297, 2002.
- ATKINSON, A. B. On the measurement of inequality. **Journal of Economic Theory**, v. 43, p. 865-880, 1970.

BELL, W. A probability model for the measurement of ecological segregation. **Social Forces**, n. 32, p. 357-364, 1954.

CARLSON, S. M. Trends in race/sex occupational inequality: conceptual and measurement issues. **Social Problems**, n. 39, p. 269-290, 1992.

COWGILL, D.; COWGILL, M. S. An index of segregation based on block statistics. **American Sociological Review**, n. 16, p. 825-831, 1951.

DACEY, M. F. A review on measures of contiguity for two and K-color maps. In: BERRY, B. J. L.; MARBLE, D. F. M. (Eds.). **Spatial analysis: a reader in statistical geography**. Englewood Cliff: Prentice-Hall, 1968, p. 479-495.

DUNCAN, O. D.; CUZZORT, R. P.; DUNCAN, B. **Statistical Geography: problems in analyzing areal data**. Illinois: The free press of Glencoe, 1961.

DUNCAN, O. D.; DUNCAN, B. A methodological analysis of segregation indexes. **American Sociological Review**, n. 20, p. 210-217, 1955a.

_____. Residential distribution and occupational stratification. **American Journal of Sociology**, n. 60, p. 493-503, 1955b.

GEARY, R. C. The contiguity ratio and statistical mapping. **Incorporated Statistician**, n. 5, p. 115-141, 1954.

GRANNIS, R. Discussion: segregation indices and their functional inputs. In: STOLZENBERG, R. M. (Ed.). **Sociological methodology**. Boston, MA: Blackwell Publishing, 2002, p. 69-84. v. 32.

HOOVER, E. M. Interstate redistribution of population, 1850-1940. **Journal of Economic History**, n. 1, p. 199-205, 1941.

JAHN, J. A. The measurement of ecological segregation: derivation of an index based on the criterion of reproducibility. **American Sociological Review**, n. 15, p. 100-104, 1950.

JAKUBS, J. F. A distance-based segregation index. **Journal of Socio-Economic Planning Sciences**, n. 15, p. 129-141, 1981.

JAMES, D. R.; TAEUBER, K. E. Measures of segregation. **Sociological Methodology**, n. 14, p. 1-32, 1985.

LIBERSON, S. Measuring population diversity. **American Sociological Review**, n. 34, p. 850-862, 1969.

MASSEY, D. S.; DENTON, N. A. The dimensions of residential segregation. **Social Forces**, n. 67, p. 281-315, 1988.

MORGAN, B. S. The segregation of socioeconomic groups in urban areas: a comparative analysis. **Urban Studies**, n. 12, p. 47-60, 1975.

_____. An alternate approach to the development of the distance-based measure of racial segregation. **American Journal of Sociology**, n. 88, p. 1237-1249, 1983.

MORRILL, R. On the measure of geographic segregation. **Geography Research Forum**, n. 11, p. 25-36, 1991.

REARDON, S. F. Measures of racial diversity and segregation in multigroup and hierarchical structured populations. In: ANNUAL MEETING OF THE EASTERN SOCIOLOGICAL SOCIETY, Philadelphia, PA, 1998.

REARDON, S. F.; FIREBAUGH, G. Measures of multigroup segregation. **Sociological Methodology**, n. 32, p. 33-67, 2002.

REARDON, S. F.; O'SULLIVAN, D. Measures of spatial segregation. **Sociological Methodology**, n. 34, p. 121-162, 2004.

SAKODA, J. N. A generalized Index of dissimilarity. **Demography**, n. 18, p. 245-250, 1981.

SHANNON, C. E. A Mathematical theory of communication. **Bell System Technical Journal**, n. 27, p. 379-422, 623-656, 1948.

THEIL, H.; FINEZZA, A. J. A note on the measurement of racial integration of schools by means of informational concepts. **Journal of Mathematical Sociology**, n. 1, p. 187-194, 1971.

WHITE, M. J. The measurement of spatial segregation. **American Journal of Sociology**, n. 88, p. 1008-1018, 1983.

_____. Segregation and diversity measures in population distribution. **Population Index**, n. 52, p. 198-221, 1986.

WONG, D. S. Spatial indices of segregation. **Urban Studies**, n. 30, p. 559-572, 1993.

_____. Measuring multiethnic spatial segregation. **Urban Geography**, n. 19, p. 77-87, 1998.

_____. Geostatistics as measures of spatial segregation. **Urban Geography**, n. 20, p. 635-647, 1999.

_____. Spatial measures of segregation and GIS. **Urban Geography**, n. 23, p.85-92, 2002.

_____. Implementing measures of spatial segregation in GIS. **Computers, Environment and Urban Systems**, n. 27, p.53-70, 2003.

WONG, D. W. S. Enhancing segregation studies using GIS. **Computers, Environment, and Urban Systems**, n. 20, p. 99-109, 1996.

Autores

Alexandre Xavier Ywata de Carvalho é bacharel em Engenharia Mecânica-Aeronáutica pelo Instituto Tecnológico de Aeronáutica, mestre em Estatística pela Universidade de Brasília e doutor em Estatística pela Northwestern University. Técnico de pesquisa e planejamento do Instituto de Pesquisa Econômica Aplicada – Ipea.

Camilo Rey Laureto é bacharel em Ciências Econômicas pela Universidade de Brasília e pesquisador visitante do Instituto de Pesquisa Econômica Aplicada – Ipea.

Marina Garcia Pena é bacharel em Estatística pela Universidade de Brasília e bolsista de pesquisa do Instituto de Pesquisa Econômica Aplicada – Ipea.

Pedro Henrique Melo Albuquerque é bacharel em Estatística pela Universidade de Brasília, mestre em Estatística pela Universidade Federal de Minas Gerais e doutor em Administração pela Universidade de Brasília. Professor do Departamento de Administração da Universidade de Brasília e pesquisador visitante do Instituto de Pesquisa Econômica Aplicada – Ipea.

Waldery Rodrigues Junior é bacharel em Engenharia pelo Instituto Tecnológico de Aeronáutica, mestre em Economia pela University Of Michigan e pela Universidade de Brasília e doutor em Economia pela Universidade de Brasília. Técnico de pesquisa e planejamento do Instituto de Pesquisa Econômica Aplicada – Ipea.

Abstract

A study of methods and functionality of segregation indicators

Segregation indicators measure the evenness of distribution, that is, the contrasting dispersion or concentration of a certain group in a given space, such as a geographical area. Whereas the qualitative analysis of people's mobility focuses on social, economic and cultural factors, segregation indicators aim at determining geographic mobility by quantitative factors than can later connect to qualitative factors. This article deals with segregation indicators and describes how they are derived and how each one functions. The following types of indicators are described: evenness, exposure, concentration, clustering and centralization. The empirical analysis is based on data from the Annual (Brazilian) Report on Social Information (RAIS) and the analysis is based on the spatial segregation of annual formal labor income in the São Paulo Metropolitan Region.

Keywords: Spatial Segregation. Cluster. Segregation. Index. Regional Analysis.

Resumen

Un estudio de las metodologías y funcionalidades de los índices de segregación

Los índices de segregación tienen el propósito de evaluar la distribución o, más específicamente, la dispersión o concentración de un grupo dado en un determinado espacio, como una región geográfica. Mientras un análisis más cualitativo de los factores definidores de la movilidad de las personas aborda cuestiones sociales, económicas y culturales, los índices de segregación intentan determinar los factores de la movilidad geográfica de forma cuantitativa para, posteriormente, asociarlos a factores cualitativos. Este trabajo tiene como objetivo tratar de la metodología de los índices de segregación, de la forma como son construidos, y de la aplicación de las medidas propuestas. De este modo, el enfoque es abordar los índices de segregación del tipo uniformización (evenness), exposición (exposure), concentración (concentration), agrupamiento (clustering) y centralización (centralization), para luego construir una aplicación empírica utilizando datos de la Relación Anual de Informaciones Sociales (Rais). El análisis consiste en evaluar la segregación espacial del ingreso anual recibido por los trabajadores formales en los municipios de la Región Metropolitana de São Paulo (RMSP).

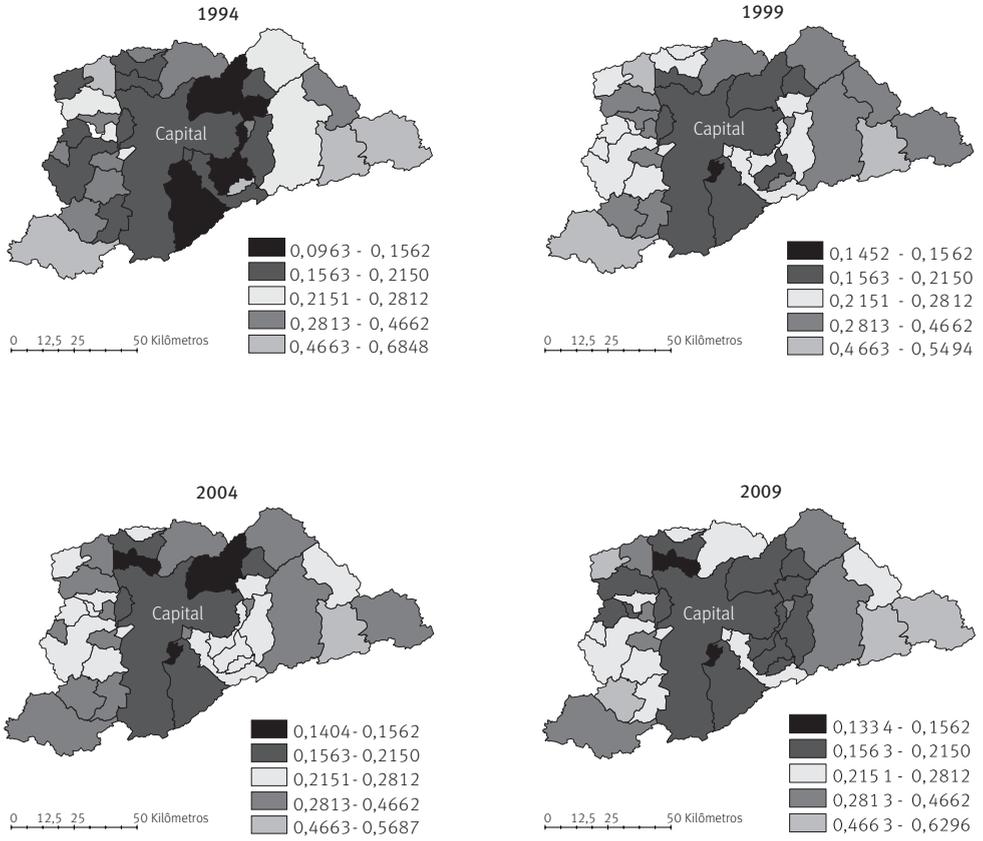
Palabras clave: Segregación espacial. Cluster. Índices de segregación. Análisis regional.

Recebido para publicação em 08/05/2012

Aceito para publicação em 21/06/2012

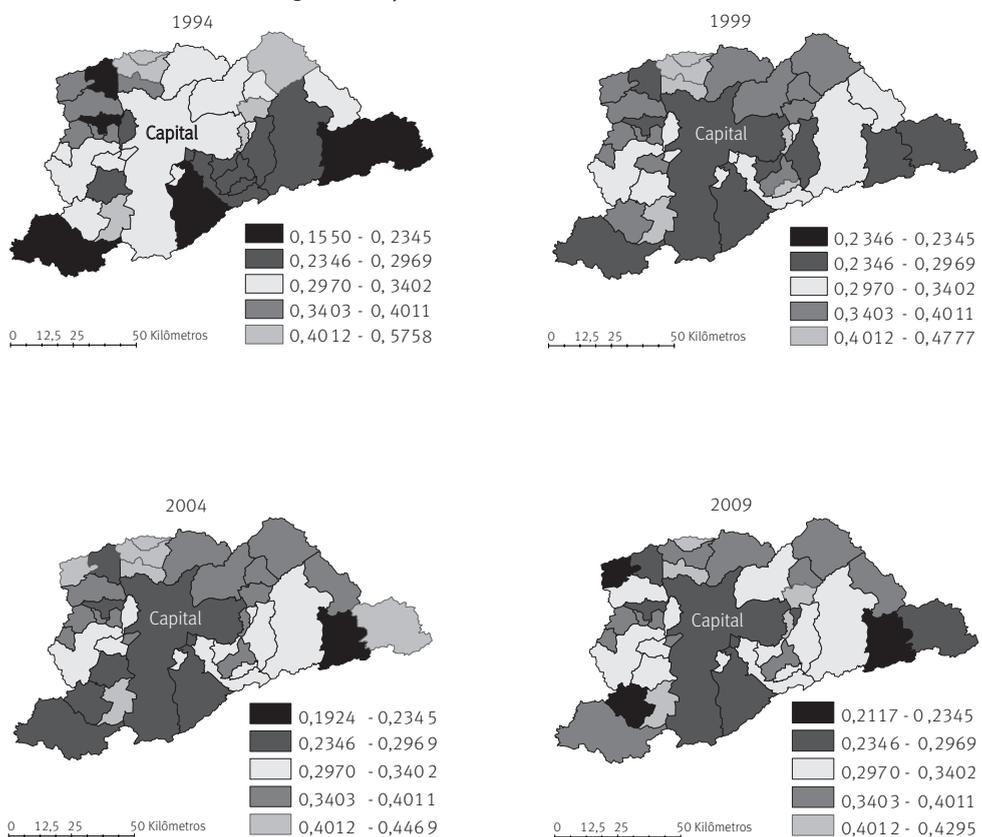
Anexo

MAPA 1
Proporção de empregados com renda menor ou igual ao percentil de 20%
Região Metropolitana de São Paulo – 1994-2009



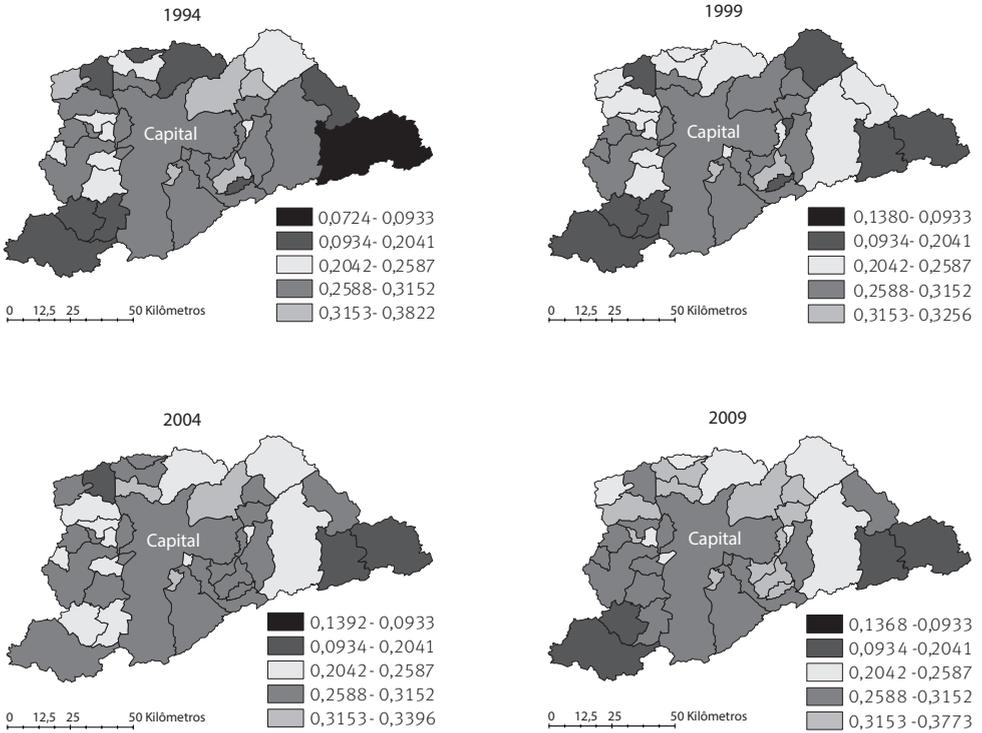
Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994, 1999, 2004, 2009.

MAPA 2
Proporção de empregados com renda maior ou igual ao percentil de 20% e menor que o percentil de 50%
Região Metropolitana de São Paulo – 1994-2009



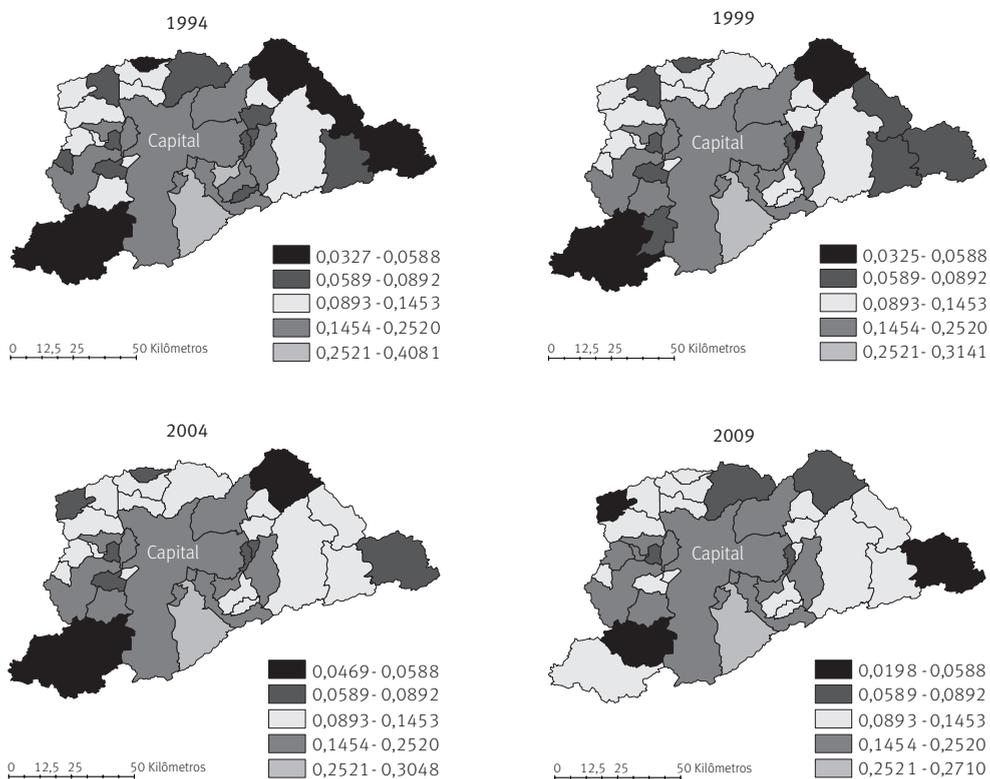
Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994, 1999, 2004, 2009.

MAPA 3
Proporção de empregados com renda maior ou igual ao percentil de 50% e menor que o percentil de 80%
Região Metropolitana de São Paulo – 1994-2009



Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994, 1999, 2004, 2009.

MAPA 4
Proporção de empregados com renda maior ou igual ao percentil de 80%
Região Metropolitana de São Paulo – 1994-2009



Fonte: MTE. Relação Anual de Informações Sociais – Rais 1994, 1999, 2004, 2009.