

RETORNO À EDUCAÇÃO NA REGIÃO METROPOLITANA DE SALVADOR E EFEITOS ESPACIAIS SALARIAIS

Returns to education in the metropolitan area of Salvador and spatial effects on wages

Diana Lúcia Gonzaga da Silva

Doutoranda em Economia (IPE-USP). Economista e Mestre em Economia (UFBA)

Email: diana.gonzaga@gmail.com Gabriella Carla Leite de Vasconcelos

Resumo: O objetivo deste estudo é identificar possíveis efeitos espaciais na estimação do retorno à educação na Região Metropolitana de Salvador. A literatura em Economia do Trabalho estima o retorno à educação a partir das características observadas e não observadas dos trabalhadores. A Economia Regional e Urbana incluiu a localização como um determinante salarial, mas desconsiderou a presença de autocorrelação espacial na distribuição de salários. Interações sociais entre indivíduos pertencentes a determinados grupos (espaços) podem explicar parte dos salários. Este artigo utiliza econometria espacial para estimar uma equação salarial minceriana ampliada. O estudo utiliza uma amostra de microdados do Censo Demográfico (2010), para as áreas de ponderação da região metropolitana de Salvador. A especificação econométrica adota o *Spatial Autoregressive Model* e a estimação é realizada com o método de Mínimos Quadrados Ordinários e o *Generalized Method of Moments* em dois estágios. Os resultados mostraram um retorno à educação de 5,6% para os indivíduos da região metropolitana de Salvador. Este retorno não se alterou após a inclusão da defasagem espacial salarial. Os resultados mostraram um ganho salarial de 1,4% oriundo dos salários dos vizinhos pertencentes à mesma área de ponderação. Este resultado sugeriu que existem *Peer Effects* na determinação salarial individual.

Palavras-chave: Educação; salário; espaço.

Abstract: The aim of this study is to identify possible spatial effects in estimation of the return to education in the metropolitan area of Salvador. Literature in Labor Economics estimates the return to education from the observed and unobserved characteristics of workers. The Regional and Urban Economics included the location as a wage determinant, but disregarded the presence of spatial autocorrelation on the wage distribution. Social interactions among individuals belonging to groups (spaces) can explain a share of wages. This paper uses spatial econometrics to estimate an extended Mincerian equation. The study uses the microdata from the Census of Brazil (2010), for the minimum areas belonging the metropolitan area of Salvador. The econometric specification adopts the *Spatial Autoregressive Model* and the estimation is performed using the method of *Ordinary Least Squares* and the two stage *Generalized Method of Moments*. The results showed a return to education of 5.6% for individuals in the metropolitan region of Salvador. This return did not change after the inclusion of wage spatial lag. The results showed an individual wage gain of 1.4% coming from the wages of neighbors belonging to the same areas. This result suggested that there are Peer Effects in individual wage determination.

Keywords: Education; wage; space.

1 Introdução

A desigualdade salarial é tema recorrente em diversos estudos da Economia. Estudos mostram diferenciais salariais entre indivíduos de uma mesma localização, entre indivíduos de localizações distintas ou entre mercados de trabalhos espacialmente afastados. No entanto, poucos estudos recorrem a abordagens de econometria espacial para entender as dimensões de possíveis dependências espaciais que podem explicar os diferenciais salariais entre indivíduos.

Mincer (1974) formalizou um modelo para a determinação salarial. A equação de salários minceriana passou a ser adotada nos estudos em Economia do Trabalho, particularmente para encontrar diferenciais salariais de retorno privado à educação. Buscava-se controlar o efeito das características observáveis dos trabalhadores sobre os salários e, em seguida, das características não observáveis. Entretanto, permanecia um diferencial salarial favorável aos trabalhadores dos grandes centros urbanos. Isso motivou o surgimento de estudos que integram Economia Urbana e Economia do Trabalho para explicar o diferencial de salários no espaço. Neste cenário, surgem os estudos envolvendo economias de aglomeração urbana e capital humano, a partir da inclusão da localização urbana como determinante salarial. No entanto, embora a localização fosse incluída como elemento explicativo para a determinação salarial, esses estudos ainda utilizavam métodos econométricos tradicionais, sem recorrer às abordagens específicas de econometria espacial.

A presença de externalidades locais específicas pode garantir retornos diferenciados para cada ano adicional de escolaridade. No entanto, a importância do espaço como elemento explicativo para diversas relações econômicas pode decorrer do fato de que, não somente a localização do indivíduo pode afetar suas escolhas e seus resultados econômicos, como também a de seus vizinhos mais próximos. Interações sociais de indivíduos pertencentes a determinados grupos (espaços) podem explicar comportamentos individuais e resultados decorrentes destes. Neste sentido, os efeitos de interações sociais no espaço podem ser relevantes para a determinação dos salários e dos ganhos oriundos da educação.

Portanto, supondo a existência de dependência espacial na equação minceriana clássica, o objetivo deste estudo é identificar possíveis efeitos espaciais existentes na estimação econométrica da equação salarial e do retorno à educação dos indivíduos residentes na Região Metropolitana (RM) de Salvador. Será utilizada a amostra de microdados do Censo Populacional do Instituto Brasileiro de Geografia e Estatística (IBGE, 2010) para as áreas de ponderação da Região Metropolitana de Salvador. Uma equação salarial minceriana será estimada ao utilizar a modelagem de econometria espacial. Além das variáveis explicativas comuns na literatura que utiliza a equação salarial, serão incluídas defasagens espaciais da variável dependente para corrigir a possível autocorrelação espacial existente dentro da área de ponderação em que o indivíduo reside. A especificação econométrica será baseada no *Spatial Autoregressive Model - SAR*.

Além desta introdução, o artigo é composto de mais quatro seções. A seção dois trata do referencial teórico adotado nesta pesquisa, baseado na Economia Espacial, Economia do Trabalho e Interações Sociais (*Peer Effects*). A seção três descreve a metodologia do estudo. A seção quatro apresenta a Análise Exploratória de Dados Espaciais (AEDE) e os respectivos resultados das estimações econométricas. Finalmente, na seção cinco serão apresentadas as considerações finais deste artigo.

2 Literatura

Esta seção faz uma revisão da literatura de determinação salarial, economias de aglomeração (prêmio salarial urbano), capital humano e *Peer Effects*, integrando a Economia Espacial, a Economia do Trabalho e as abordagens de Interações Sociais no Espaço.

2.1 Determinação salarial e retorno à educação no mercado de trabalho

Na literatura do capital humano, um trabalho seminal e pioneiro na medição dos ganhos individuais no mercado de trabalho foi o de Mincer (1974). O autor apresentou uma equação de determinação salarial que ficou conhecida como equação minceriana. A equação minceriana define a escolaridade e a experiência como os componentes do investimento individual em capital humano que explicam os salários. Estudos em Economia do Trabalho encontra-

ram significativos retornos à educação no mercado de trabalho e diferenciais entre os países, com base na equação minceriana. Como resultado, a educação tem ocupado posição de destaque nos estudos sobre a determinação salarial.

O retorno à educação é o ganho salarial obtido pelo indivíduo a cada ano adicional de escolaridade alcançado, ou seja, é a contribuição da educação para o salário individual. A equação minceriana clássica, disseminada nos diversos estudos em Economia do Trabalho, identifica os retornos à educação definidos por:

$$\ln E_t = \ln E_s + \beta_1 t + \beta_2 t^2. \quad (1)$$

Na equação (1), E_t são os ganhos no mercado de trabalho, ou salários, E_s são os ganhos após completar a escolaridade e t mede os anos de experiência de trabalho. No modelo, $\ln E_s = \ln E_0 + rs$, sendo E_0 a capacidade original de ganhos e s os anos de escolaridade obtidos pelo indivíduo. A experiência de trabalho (t) inicia após o período de aquisição da escolaridade (s), onde $t = A - s - b$, em que A é a idade atual do trabalhador e b é a idade em que começou a adquirir escolaridade. A equação salarial de Mincer é, então, definida como:

$$\ln E_t = \ln E_0 + rs + \beta_1 (A - s - b) + \beta_2 (A - s - b)^2. \quad (2)$$

A equação de Mincer (2) geralmente é apresentada como: $\ln w_t = \beta_0 + rs + \beta_1 x + \beta_2 x^2$. Nessa equação, w_t é o salário individual, s os anos de escolaridade e x é a experiência de trabalho.

A literatura empírica em Economia do Trabalho identifica o capital humano individual a partir de dois componentes mensuráveis: a escolaridade e a experiência. Os desenvolvimentos na Teoria do Capital Humano permitiram disseminar evidências da influência da educação (ou habilidade observada) na determinação dos salários. Estimativas do retorno privado a anos adicionais de educação variam entre 8 e 12%. Estudos recentes têm considerado as externalidades de capital humano. Uma questão ainda ambígua é a identificação de medidas de capital humano para a estimação das externalidades (HALFDANARSON, HEUERMAN; SÜDEKUM et al., 2008). Estudos mais recentes em Economia do Trabalho e Espacial construíram versões ampliadas da equação minceriana, considerando a presença de características não obser-

váveis dos trabalhadores, que poderiam explicar a diferenciação salarial. No entanto, após o controle, permanecia um diferencial de retornos entre trabalhadores em diferentes localizações.

Variações espaciais nos custos de vida, na produtividade, nas dotações de amenidades locais e nas estruturas produtivas são alguns dos fatores que podem explicar a manutenção dos diferenciais salariais no espaço. Os diferenciais de produtividade, no entanto, podem estar associados às características específicas locais. Nesse contexto, surgem as análises em Economia Espacial que relacionam os ganhos de produtividade e salários às economias de aglomeração urbana, com baseadas nos fatores locais específicos. Abordagens com baseadas nas integrações entre a Economia Urbana e a Economia do Trabalho relacionam o prêmio salarial urbano e as externalidades de capital humano. Essas abordagens empregam versões ampliadas de equações mincerianas, com a inclusão de variáveis de localização.

2.2 Economias de aglomeração (Prêmio Salarial Urbano) e educação

A partir do desenvolvimento teórico sobre as economias de aglomeração surgiu uma literatura que integra elementos da Economia Urbana à Economia do Trabalho. Essa abordagem avalia os efeitos das economias de aglomeração sobre os diferenciais de salários. Os trabalhos pioneiros de Glaeser e Maré (1994; 2001) encontraram uma relação positiva entre os salários e o tamanho das cidades definida como prêmio salarial urbano. O prêmio salarial está associado, dentre outros fatores, ao capital humano acumulado pelos trabalhadores nas cidades densas. A partir desses trabalhos seminais, Yankow (2006), Glaeser e Resseger (2010) e outros pesquisadores desenvolveram estudos para avaliar os efeitos da interação entre as habilidades, a educação e a aglomeração sobre a produtividade e os salários.

A literatura sobre as economias de aglomeração prevê maiores salários e produtividade em áreas urbanas densas. Os ganhos de aglomeração no mercado de trabalho podem emergir das externalidades positivas dessas áreas. São exemplos dessas externalidades, o acesso a maior diversidade de ocupações e mercados de trabalho amplos, a concentração e o acúmulo de capital humano, os

spillovers de conhecimento e outros. O benefício da proximidade é um suposto subjacente ao prêmio salarial urbano. Os maiores salários recebidos pelos trabalhadores nas cidades representam um prêmio salarial urbano, cuja magnitude varia em torno de 5 a 10% na literatura internacional. Quando possível, é descontado desse prêmio o viés de seleção associado à heterogeneidade individual ou espacial não observada (GLAESER; MARÉ, 2001; HALFDANARSON; HEUERMANN; SÜDEKUM, 2008; ABEL; DEY; GABE, 2011).

Os trabalhos de Glaeser e Maré (1994; 2001) encontraram evidências de salários 32-33% maiores para trabalhadores em grandes cidades de áreas metropolitanas dos Estados Unidos (EUA). O prêmio salarial urbano foi maior para os trabalhadores mais experientes e mais educados. Yankow (2006) identificou um prêmio salarial urbano de 19% para os trabalhadores das grandes áreas urbanas dos EUA.

As abordagens recentes da Economia Urbana sugerem a existência de externalidades de capital humano como uma explicação fundamental para a maior produtividade e os maiores salários nas grandes cidades. O capital humano acumulado nas cidades amplia a produtividade média dos trabalhadores, através dos *spillovers* de conhecimento e das experiências através das interações em ambientes densos. Os salários devem ser mais altos nas cidades densas com maiores níveis de educação. Os efeitos positivos da densidade sobre a produtividade devem ser ampliados pelo estoque de capital humano das áreas metropolitanas. A interação da densidade com a habilidade foi definida por Abel, Dey e Gabe (2011) como a densidade de capital humano.

As abordagens das externalidades localizadas de capital humano e do prêmio salarial urbano mostraram, dentre outros aspectos, que: (a) existe um prêmio salarial urbano que difere de acordo com o trabalhador; (b) os trabalhadores ganham salários mais altos em áreas intensivas em capital humano; e (c) a externalidade do capital humano é uma das causas da existência do prêmio salarial urbano, mas não a única. Portanto, o prêmio salarial urbano tende a ser maior para os trabalhadores mais qualificados (HALFDANARSON; HEUERMANN; SÜDEKUM, 2008).

Glaeser e Resseger (2010) incluíram a proporção de trabalhadores com ensino superior e identificaram uma queda de 25% no prêmio urbano,

mostrando que grande parte do prêmio seria um retorno à escolaridade elevada. Rauch (1993) estimou as externalidades de capital humano sobre os salários e a produtividade. Os resultados mostraram que cada ano adicional na escolaridade média das áreas metropolitanas dos EUA levou a um aumento de 2,8% na produtividade dos fatores. Abel, Dey e Gabe (2011) identificaram um efeito positivo da interação entre a densidade e o capital humano em áreas metropolitanas dos EUA. A duplicação da densidade gerou um aumento de 9,7% na produtividade. Essa relação variou de 13,4%, em as áreas de capital humano acima da média, para 4,8%, em áreas de baixo capital humano.

A revisão dos trabalhos teóricos e empíricos da literatura que integra Economia Espacial (prêmio salarial urbano) e Economia do Trabalho (externalidades do capital humano) sugere a existência de ganhos de aglomeração sobre os salários em áreas urbanas densas. Esses ganhos são mais evidentes em áreas com maior proporção de trabalhadores de alta escolaridade. Portanto, existem evidências teóricas de que a educação atua ampliando os ganhos salariais associados à localização. Neste sentido, as interações sociais entre os indivíduos no espaço podem gerar retornos salariais diferenciados espacialmente para cada ano de educação.

2.3 Evidências empíricas na literatura nacional: retorno à educação e Prêmio Salarial Urbano

Os estudos recentes no Brasil reforçam as evidências sobre o papel da educação na determinação salarial, nos diferenciais no mercado de trabalho e na redução das desigualdades de renda. Diversos estudos já estimaram o retorno à educação para o Brasil, a partir de uma equação salarial minceriana. No período mais recente, uma nova literatura sobre as economias de aglomeração urbana no Brasil identificou um prêmio salarial em áreas densas do país.

Campos e Silveira Neto (2009) estimaram um retorno à educação de 11,8% sobre os salários nos municípios brasileiros acima de cem mil habitantes, em 2000. O impacto de um ano adicional de educação sobre os salários chega a 16% no Brasil, de acordo com o estudo de Barros et al. (2000). Falcão e Silveira Neto (2007) revelaram um ganho de 1,7% sobre os salários, associado a um aumento na proporção de trabalhadores com ensino supe-

rior nos municípios com população acima de cem mil habitantes. O estudo sugeriu a existência de externalidades positivas da concentração espacial de capital humano. A educação apresentou a maior contribuição marginal para as desigualdades salariais entre as maiores áreas metropolitanas do Brasil, entre 17,6% e 31%, de 1992 a 1997 (SERVO; AZZONI, 2002).

A persistência dos diferenciais salariais, após considerar o custo de vida, as habilidades observadas e não observadas e as características ocupacionais, sugere a existência de um efeito de localização sobre os salários no Brasil. Entre os estudos existentes na literatura nacional, que tratam desse efeito estão: Servo e Azzoni (2002); Campos e Silveira Neto (2009) e Rocha, Silveira Neto e Gomes (2011). Servo e Azzoni (2002) revelaram uma contribuição marginal da residência em áreas metropolitanas, para o diferencial salarial no Brasil, entre 3% e 7,4%. Campos e Silveira Neto (2009) mostraram que, os indivíduos ocupados das áreas metropolitanas, em 2000, apresentaram ganhos de 16% sobre os salários. Rocha, Silveira Neto e Gomes (2011) mostraram um diferencial salarial de 9,4% para os trabalhadores das áreas metropolitanas do Brasil, entre 2000 e 2008. Os resultados sugerem a existência de ganhos de aglomeração urbana (prêmio urbano) no Brasil.

Por sua vez, alguns estudos estimaram o retorno à educação na Bahia. Lacerda (2008) encontrou um retorno de 9% para cada ano adicional de educação na Bahia, em 2005. Suliano e Siqueira (2010) estimaram um retorno à educação entre 12% e 19% para os trabalhadores do sexo masculino na Bahia, entre 2001 e 2006. Dantas e Figueiredo (2010) encontrou um retorno à educação entre 12% e 20% para os trabalhadores da Bahia do sexo masculino, entre 1986, 1996 e 2005. Para a região metropolitana de Salvador, Freitas e Freitas (2008) encontraram estimativas de retorno à educação variando entre 12,7% e 16%, para o ano de 2005. Por fim, Silva (2012) estimou o retorno à educação dos trabalhadores do mercado de trabalho formal das regiões metropolitanas do Brasil, entre 1995-2008, encontrando um ganho salarial de 3,2% por ciclo de escolaridade concluído na região metropolitana de Salvador, após controlar pelas habilidades não observadas desses trabalhadores.

Neste sentido, embora alguns estudos recentes na literatura empírica nacional tenham considerado

a importância da localização para a determinação salarial e dos retornos à educação, não foram encontrados estudos que utilizassem uma abordagem de econometria espacial para considerar possíveis dependências espaciais associadas às interações. Portanto, este estudo busca identificar possíveis interações espaciais existentes na determinação do retorno à educação, na região metropolitana de Salvador, utilizando uma abordagem de econometria espacial, com base na equação salarial minceriana.

2.4 Interações sociais na Economia (Peer Effects)

Uma literatura relativamente recente, na teoria econômica, trata das interações sociais influenciando os comportamentos e as escolhas individuais. A questão central nessa literatura é inferir se o comportamento médio de um grupo influencia o comportamento individual de seus membros. A nova abordagem trouxe para a teoria econômica elementos e conceitos dos cientistas sociais e da psicologia social. Essa literatura tem como objetivo principal entender se a interação pode explicar porque os agentes pertencentes ao mesmo grupo se comportam de modo similar. Além disso, busca resolver os problemas de identificação que surgem na inferência dos efeitos sociais que explicam os resultados econômicos.

As perguntas fundamentais que norteiam os estudos de interações sociais são: quais são as unidades que interagem entre si e como elas interagem? (MANSKI, 2000). O trabalho seminal de Manski (1993) definiu como “efeitos sociais endógenos” os canais através dos quais a sociedade pode afetar o comportamento dos indivíduos. De acordo com o autor, o “problema de reflexão” existente é entender se a sociedade (grupo) afeta o comportamento individual ou se o comportamento do grupo é uma agregação de comportamentos individuais. A propensão de um indivíduo a se comportar de determinada maneira varia de acordo com a intensidade de tal comportamento no seu grupo de referência. Diversos termos são usados para denotar os efeitos sociais endógenos na literatura: *social norms*, *peer influences*, *neighbourhood effects*, *imitation*, *contagion*, *epidemics*, *social interactions*, *conformity* ou *interdependent preferences*.

Na perspectiva da teoria de equilíbrio geral, interações que não ocorrem através do mercado, assumidas como problemas de mercado incompleto ou

externalidades, não seriam de interesse econômico. Economistas ainda não têm clareza sobre quais interações sociais constituem um domínio próprio da teoria econômica. Em termos econômicos, os agentes são as unidades que interagem entre si através de ações escolhidas. Uma ação escolhida por um agente pode afetar as ações de outros agentes através de três canais de interações: restrições, expectativas e preferências (MANSKI, 2000).

Restrições de mercados e de congestão no consumo de um bem ou atividade são exemplos de interações negativas por restrições. Agentes formando expectativas podem aprender ao observar as escolhas e resultados experimentados por outros, o que seria o *learning* observacional gerando interações por expectativas. Por sua vez, as interações por preferências ocorrem quando as preferências dos agentes sobre escolhas alternativas dependem das ações escolhidas por outros agentes. Na interação por preferência um agente pode imitar o outro simplesmente porque prefere agir igual, enquanto na interação por expectativa a imitação decorre da crença de que o outro tem informação superior (MANSKI, 2000).

Os economistas têm se preocupado com a identificação dos efeitos endógenos canalizados através dos mercados. Porém, a identificação de outros efeitos endógenos permaneceu relativamente pouco examinada e compreendida. Um problema associado à inferência dos efeitos sociais endógenos é que esta só será possível se o pesquisador possuir informação prévia especificando a composição dos grupos de referência. Se estas informações estão disponíveis, a inferência dependerá da relação entre as variáveis que definem os grupos de referência e aquelas que afetam diretamente os resultados de interesse. A inferência será limitada se estas variáveis forem funcionalmente dependentes ou estatisticamente independentes. Se estas variáveis são moderadamente relacionadas, as perspectivas de inferência são ampliadas (MANSKI, 1993).

Três hipóteses salientadas na literatura empírica tentam explicar a influência das interações sociais sobre o comportamento dos indivíduos pertencentes ao mesmo grupo. A primeira, denominada de Efeito Endógeno afirma que a propensão de um indivíduo a se comportar de certa maneira varia com o comportamento do grupo. A hipótese do Efeito Exógeno (contextual) diz que a propensão de um indivíduo a se comportar de algum modo

varia com as características exógenas do grupo. Por fim, a hipótese dos Efeitos Correlacionados assume que os indivíduos do mesmo grupo tendem a se comportar da mesma forma porque eles apresentam características ou ambientes institucionais semelhantes (MANSKI, 1993; MANSKI, 2000).

Geralmente, os Efeitos Exógenos e os Efeitos Correlacionados não geram um efeito multiplicador social, ou seja, se uma política for realizada com o foco sobre um indivíduo (ou alguns) pertencente a um grupo, seus resultados serão restritos a este indivíduo. Por sua vez, aconteceria o contrário se o Efeito Endógeno dominasse, cujo grupo também seria afetado pela política. O conceito de Efeitos Endógenos engloba os três canais de interação econômica, através dos quais o comportamento do grupo pode afetar o comportamento individual: as preferências, as expectativas e as restrições de interação (exemplo: bens de uso comum).

Manski (1993) definiu o “problema de reflexão” como uma consequência da presença de uma variável média do grupo, $E(y | x)$, como um regressor para a variável de resultado individual (y). Se a variável média impacta o resultado individual, existem Efeitos Sociais Endógenos sobre o resultado de interesse, ou seja, o indivíduo sofre influência do comportamento do grupo de referência. O modelo apresentado por Manski (1993) para a inferência dos Efeitos Sociais Endógenos é detalhado no Apêndice A deste artigo. Os estudos empíricos geralmente assumem que os Efeitos Exógenos e Correlacionados não estão presentes. Esses estudos utilizam métodos de estimação em dois estágios para estimar o Efeito Endógeno, a partir de um modelo de correlação espacial. Deste modo, as estimativas definidas na literatura de correlação espacial podem ser interpretadas como estimativas dos Efeitos Sociais Endógenos.

A literatura recente sobre interações sociais tem mostrado que o comportamento do indivíduo pode ser influenciado pelo grupo de referência no qual está inserido. Os desenvolvimentos teóricos relativamente recentes, em alguns ramos da teoria econômica, tais como *Labor Economics*, Teoria dos Jogos Dinâmica e Teoria do Crescimento Endógeno, já reconhecem aspectos das interações sociais. No entanto, a literatura empírica nessa área ainda é bastante limitada e apresenta pouco avanço (MANSKI, 2000). Portanto, ainda existe bastante espaço na teoria econômica para as pesquisas

que, reconhecendo a existência de interações sociais, possam explicar como esses efeitos afetam os resultados econômicos, bem como mensurar tais efeitos.

3 Estratégia Empírica

Assumindo que os salários dos indivíduos não são espacialmente distribuídos de forma homogênea, ou seja, de maneira independente no espaço, é preciso considerar a possível presença de autocorrelação espacial na estimação dos retornos à educação. Para tanto, este artigo vai recorrer aos métodos específicos de econometria espacial para estimar um *Spatial Autoregressive Model – SAR*, a partir da equação salarial minceriana. Neste sentido, esta seção fará uma exposição da estratégia empírica adotada neste estudo.

3.1 Modelo de Regressão Espacial

A existência de interações sociais entre indivíduos próximos estabelece a necessidade de considerar o processo de *feedback*, em que os indivíduos interagem e influenciam uns aos outros e as interações são mais fortes se o contato é frequente. Wrigley et al. (1996) mostrou que os processos de *feedback* podem determinar a autocorrelação espacial. A proximidade geralmente aumenta a frequência das interações, de modo que essas podem ser intensificadas entre indivíduos nas suas vizinhanças. A presença dessas interações mostra a necessidade de especificar um modelo que controle tais efeitos.

Este trabalho assumirá um modelo de *lag* (defasagem) espacial para tratar os efeitos de interações es-

paciais na equação minceriana. O modelo *SAR* assume que a natureza da correlação espacial se encontra na variável dependente. O modelo inclui a variável dependente defasada espacialmente como variável explicativa na regressão. Será assumida a hipótese de existência de interação entre vizinhos, de modo que os valores da variável dependente de um indivíduo serão diretamente dependentes, através de λW , dos valores da variável dependente dos indivíduos vizinhos. A matriz W ($n \times n$) de pesos espaciais define a estrutura de vizinhança espacial e λ é o parâmetro de *lag* espacial. Neste sentido, o modelo econométrico espacial é definido por:

$$\begin{aligned} y &= \lambda Wy + X\beta + \varepsilon \\ y &= (I_n - \lambda W)^{-1} (X\beta + \varepsilon) \end{aligned} \quad (3)$$

Em que onde Wy é um vetor ($n \times 1$) de *lags* espaciais da variável dependente y , o qual representa a média ponderada da variável dependente para os indivíduos vizinhos e X representa as variáveis explicativas que afetam y diretamente. A segunda equação representa a forma reduzida do modelo SAR.

A literatura de *Peer Effects* assume a existência de efeitos dos grupos em que os indivíduos estão inseridos sobre os resultados desses indivíduos. Este artigo assume a existência de efeitos de vizinhança dentro da mesma área de ponderação em que o indivíduo de referência está inserido. Para determinar tais efeitos, foi construída uma matriz de pesos para a determinação dos *lags* espaciais. A matriz de vizinhança W_i identifica os indivíduos que compartilham a mesma área de ponderação do indivíduo de referência, captando os efeitos de vizinhança imediatos. Um exemplo de parte da matriz W_i é apresentado abaixo.

Matriz W_i : Indivíduos na mesma área de ponderação					
(I)	(II)	(III)	(IV)	(V)	(VI)
		Indivíduo 1	Indivíduo 2	Indivíduo 3	Indivíduo 4
Indivíduo 1	Área de ponderação 1	0	1	0	0
Indivíduo 2	Área de ponderação 1	1	0	0	0
Indivíduo 3	Área de ponderação 2	0	0	0	1
Indivíduo 4	Área de ponderação 2	0	0	1	0

A primeira coluna mostra o identificador do indivíduo e a segunda, o identificador da área de ponderação em que o indivíduo reside. Neste exemplo, os dois primeiros indivíduos residem na área de ponderação 1, enquanto os dois indivíduos seguintes residem na área de ponderação 2. Nas colunas três, quatro, cinco e seis são definidas as relações de vizinhança. Se o indivíduo na linha es-

pecificada é vizinho dos indivíduos em cada coluna, ou seja, reside na mesma área de ponderação do indivíduo de referência, ele receberá o valor um, e caso contrário, receberá o valor zero. Portanto, considerando a matriz de pesos W_i , o modelo (3) será reescrito como:

$$y = (I - \lambda_1 W_i)^{-1} (X\beta + \varepsilon) \quad (3.1)$$

A especificação (3.1) controla as possíveis interações entre os indivíduos que compartilham a mesma área de ponderação. Como em (3), a equação (3.1) representa a forma reduzida do modelo, em que $(I - \lambda_j W_j)^{-1}$ é a Inversa de Leontief, a qual relaciona a variável y com todas as variáveis explicativas, através da matriz de pesos espaciais.

3.2 Base de dados e especificação econométrica

O banco de dados deste estudo é construído a partir dos microdados amostrais do Censo Demográfico do IBGE (2010). A amostra de microdados do Censo Demográfico constitui uma fração da população total investigada (e domicílios), a qual contém informações socioeconômicas e demográficas dos indivíduos e domicílios. Com base nesses microdados é possível agregar geograficamente as informações dos indivíduos e domicílios nas áreas de ponderação¹ e municípios de interesse.

Neste estudo, serão utilizadas as informações referentes aos indivíduos residentes em algum dos 13 municípios da região metropolitana da Salvador². A RM de Salvador é constituída por 170 áreas de ponderação no Censo de 2010, as quais serão utilizadas para a construção de uma matriz de vizinhança do modelo a ser estimado. Para a estimação dos retornos à educação para os indivíduos da RM de Salvador, o banco de dados de interesse foi composto das seguintes variáveis: 1) Identificação: identificador do indivíduo; 2) Geográficas: localização do indivíduo por município e área de ponderação da RM de Salvador; 3) Características do indivíduo: idade, sexo e escolaridade; e 4) Características do emprego: setor de atividade econômica e ocupação. Além dessas variáveis, também foi extraída a remuneração do trabalhador. A remuneração utilizada será o rendimento em todos os trabalhos³.

1 A área de ponderação é uma unidade geográfica formada por um agrupamento de setores censitários, os quais definem a unidade territorial mínima de coleta das informações do Censo Demográfico do IBGE.

2 De acordo com o Censo Demográfico do IBGE (2010), a região metropolitana de Salvador é composta pelos seguintes municípios e respectivas áreas de ponderação: Camaçari (12), Candeias (4), Dias d'Ávila (3), Itaparica (1), Lauro de Freitas (8), Madre de Deus (1), Mata de São João (2), Pojuca (2), Salvador (62), São Francisco do Conde (2), São Sebastião do Passé (2), Simões Filho (6) e Vera Cruz (2).

3 O mês de referência para os rendimentos no Censo Demográfico é julho de 2010 (IBGE, 2010).

O Índice Nacional de Preços ao Consumidor Amplo (IPCA) do IBGE foi utilizado para deflacionar os salários. Foi utilizado o IPCA calculado pelo IBGE para a região metropolitana de Salvador (ano base 2012). Foram retirados da amostra os trabalhadores com idade menor do que 18 anos e maior do que 65 anos⁴. O banco de dados final utilizado é composto por 153.020 indivíduos. Esses trabalhadores estão distribuídos espacialmente em todos os municípios e áreas de ponderação da região metropolitana de Salvador.

No Censo Demográfico (IBGE, 2010), a variável educação é revelada pelo grau de instrução que o indivíduo frequenta ou frequentou. Para a inclusão da educação no modelo foi definida uma variável discreta assumindo valores de 0 a 15 (anos de educação), a partir da concatenação das informações educacionais dos microdados da amostra⁵. A categoria "15" da variável educação representará os indivíduos que possuíam 15 anos ou mais de educação. Foi definida uma variável *dummy* para cada setor de atividade⁶ em que os indivíduos trabalhavam, de acordo com Classificação Nacional de Atividades Econômicas Domiciliar (CNAE-Domiciliar 2.0), que é uma adaptação da Classificação de Nacional de Atividades Econômicas (CNAE 2.0), do Censo do IBGE (2010). O setor assumido como categoria base na análise é o de comércio, reparação de veículos automotores e motocicletas. Para a variável ocupação, foi definida uma classificação por grau de exigência de qualificação, a partir da COD⁷, gerando as seguintes *dummies*: (ai) Categoria 0: ocupações mal definidas; (bii) Categoria 1: trabalhos manuais; (ciii) Categoria 2: trabalhos técnicos; e (div) Categoria 3: trabalhos especializados. A categoria base de análise para a ocupação é a categoria 1, de trabalhos manuais.

A estratégia econométrica aplicada é tem baseada na modelagem de dados espaciais. Para a estimação do modelo será adotado o *Generalized*

4 Optou-se por trabalhar apenas com os indivíduos na faixa etária considerada produtiva, entre 18 e 65 anos, seguindo alguns estudos da literatura, tais como Glaeser e Maré (2001), Servo e Azzoni (2002).

5 Para a definição dos anos de escolaridade dos trabalhadores foram utilizadas as informações contidas nas variáveis V0627, V0628, V0629, V0630, V0631, V0632, V0633, V0634, V0635 e V6400, do Censo Demográfico (IBGE, 2010).

6 Ver Tabela 1 no Apêndice B.

7 Classificação de Ocupações para Pesquisas Domiciliares – COD, desenvolvida pelo IBGE para as pesquisas domiciliares (IBGE, 2010).

Method of Moments (GMM) em dois estágios. A estratégia de identificação deste estudo consiste na adoção de modelagem de dados espaciais para controlar a heterogeneidade espacial presente na equação salarial dos trabalhadores. O objetivo será captar o retorno à educação para os trabalhadores da região metropolitana de Salvador, controlando as possíveis dependências espaciais existentes na equação salarial. Estatísticas espaciais, como o I de Moran, foram utilizadas para testar a existência de autocorrelação espacial.

O modelo adotado para as estimações econométricas segue a equação salarial de Mincer (1974), apresentada na segunda seção. Essa equação foi ampliada com variáveis explicativas adicionais. Portanto, o modelo econométrico adotado para as estimações é dado por:

$$\ln(w_i) = \lambda W \ln(w_j) + \text{educ}_i \beta_1 + X \beta_2 + \varepsilon_i \quad (4)$$

Na equação (4), $\ln(w_i)$ é o logaritmo natural do salário real do trabalhador i e $W \ln(w_j)$ é o logaritmo natural do salário real dos trabalhadores na vizinhança de i (mesma área de ponderação), onde W é a matriz de peso espacial, ou de vizinhança, adotada. O vetor de controles X inclui as seguintes características do indivíduo:

- i. exp_i - experiência de trabalho, calculada a partir da definição de Mincer (1974), em que $\text{exp}_i = \text{age}_i - \text{educ}_i - 6$, assumindo que o indivíduo inicia a escolarização aos 6 anos;
- ii. exp_i^2 - experiência ao quadrado;
- iii. dsexo - *dummy* de gênero (=1 se mulher);
- iv. dsetor_i - vetor de dummies para o setor de atividade;
- v. docup_i - vetor de dummies para as categorias de ocupação.

A variável explicativa de interesse é educ_i , a escolaridade individual, a qual identificará o retorno à educação dos indivíduos na região metropolitana de Salvador. As demais variáveis servirão de controle para a equação salarial. O termo de erro do modelo é definido por ε_i . Por se tratar de uma amostra complexa, além dessas variáveis, todas as estatísticas básicas e estimações serão geradas

com o peso amostral definido no Censo Demográfico do IBGE (2010). O peso amostral define o número de observações na população que cada indivíduo representa. Além do retorno à educação, a equação (4) permitirá a obtenção do efeito de interações espaciais sobre os salários dos indivíduos.

O estimador de Mínimos Quadrados Ordinários (MQO) para os modelos do tipo *SAR* é enviesado e geralmente inconsistente, pois a variável dependente defasada é correlacionada com o termo de erro (KELEJIAN; PRUCHA, 1998). A presença da variável dependente defasada, $\lambda W y$, gera endogeneidade no estimador de MQO decorrente da simultaneidade. A estimação da equação (4) adotará o método dos momentos generalizados (GMM) em dois estágios. Na estimação por GMM as defasagens espaciais (WX e W^2X) podem ser usadas como instrumentos para a variável dependente defasada espacialmente, na tentativa de corrigir a endogeneidade presente.

O primeiro estágio do modelo estima o *lag* espacial da variável dependente ($W \ln(w_j)$), representando os salários dos vizinhos, o qual será utilizado como instrumento no segundo estágio. A estimação no primeiro estágio utilizará *lags* espaciais das variáveis explicativas ($W \text{educ}$, $W^2 \text{educ}$, WX e $W^2 X$) para estimar o *lag* espacial de $\ln(w^j)$. No segundo estágio o valor predito $W \ln(w_j)$ será incluído na equação salarial (4) para identificar o parâmetro de interesse (β_1), o retorno à educação. O *lag* espacial da variável dependente $\ln(w_j)$ será definido a partir dos indivíduos pertencentes à mesma área de ponderação do indivíduo de referência. As equações do primeiro e segundo estágio da estimação econométrica deste estudo serão:

1º estágio

$$W_1 \ln(w_j) = \rho_{11} W_1 \text{educ}_i + \rho_{12} W_1^2 \text{educ}_i + \rho_{13} W_1 X + \rho_{14} W_1^2 X + \rho_{15} \text{educ}_i + \rho_{16} X + e_1 \quad (5)$$

2º estágio

$$\ln(w_i) = \lambda_1 W_1 \ln(w_j) + \text{educ}_i \beta_1 + X \beta_2 + \varepsilon_i \quad (6)$$

A equação (5) contempla as defasagens espaciais para a mesma área de ponderação do indivíduo i e será estimada por Mínimos Quadrados em Dois Estágios (*Two Stage Least Squares-2SLS*). O valor estimado será inserido como variável explicativa na equação (6). A equação (6) será, então, estimada de

maneira robusta pelo Método dos Momentos Generalizados (GMM), no segundo estágio.

4 Resultados

Esta seção apresenta as estatísticas básicas da amostra, uma breve análise exploratória dos dados espaciais e os principais resultados encontrados com as estimações.

4.1 Análise Exploratória de Dados Espaciais

O banco de dados total é composto por 153.020 indivíduos pertencentes à região metropolitana de Salvador. A Tabela 1 apresenta a distribuição dos indivíduos na amostra por município da RM de Salvador. O município de Salvador, capital do estado da Bahia, possui 62,39% dos indivíduos na amostra deste estudo.

Tabela 1 – Observações por municípios: Amostra Censo 2010

Município	Frequência	Participação (%)
Camaçari	14.687	9,60
Candeias	5.438	3,55
Dias d'Ávila	4.152	2,71
Itaparica	1.269	0,83
Lauro de Freitas	10.662	6,97
Madre de Deus	2.233	1,46
Mata de São João	2.521	1,65
Pojuca	2.264	1,48
São Francisco do Conde	2.065	1,35
São Sebastião do Passe	2.686	1,76
Salvador	95.468	62,39
Simões Filho	7.325	4,79
Vera Cruz	2.250	1,47
Total	153.020	100,00

Fonte: Elaborado pela autora com base nos dados do Censo do IBGE (2010).

A Tabela 2, a seguir, apresenta as estatísticas descritivas básicas das variáveis na amostra. Os dados da amostra para a RM de Salvador mostram que 53% dos trabalhadores eram do sexo feminino, com idade média de 37 anos, escolaridade média de 8,5 anos e salário real médio (em logaritmo) de R\$ 4,28, a preços de 2012.

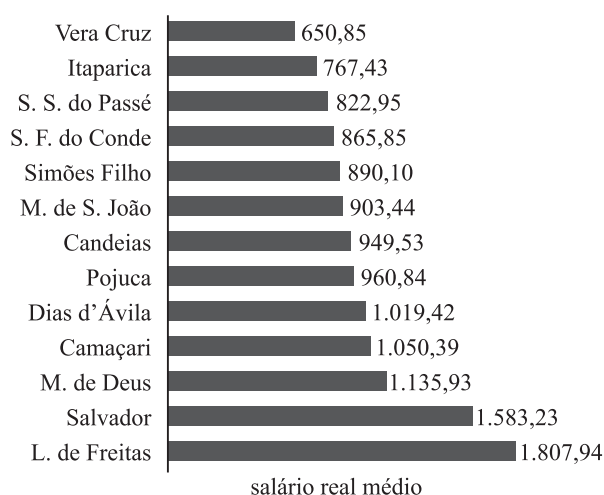
Tabela 2 – Estatísticas descritivas de variáveis da amostra

Variável	Obs.	Média	Desvio	Mín.	Máx.
Log. natural salário real	153.020	4,28	3,35	0	12,7
Educação	153.020	8,55	4,38	0	15
Dummy de sexo	153.020	0,53	0,50	0	1
Experiência	153.020	22,28	13,82	0	59
Setor	153.020	7,56	7,48	0	22
Idade	153.020	36,83	12,44	18	65

Fonte: Elaborado pela autora com base nos dados do Censo do IBGE (2010).

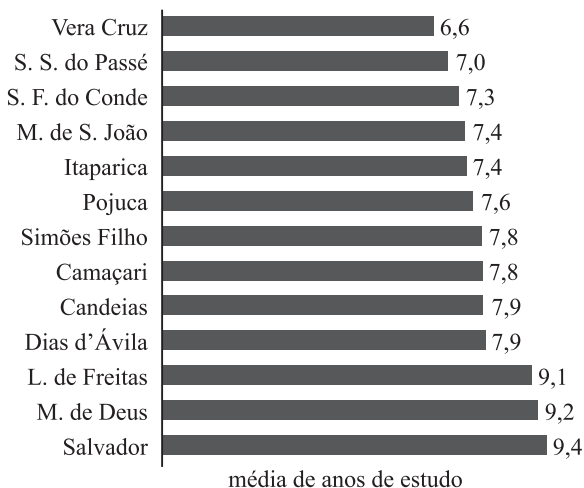
O Gráfico 1 apresenta o salário real médio, a preços de 2012, por município da região metropolitana de Salvador. O maior salário real médio esteve no município de Lauro de Freitas, seguido por Salvador. Esse dado pode refletir o comportamento dos indivíduos que, em decorrência das desamenidades urbanas e do custo habitacional da capital, preferem morar fora de Salvador, mesmo trabalhando nesse município. O menor salário real médio esteve nos municípios (ou ilhas) de Vera Cruz e Itaparica. O Gráfico 2 apresenta a escolaridade média por município da RM de Salvador. No município de Salvador estão concentrados os indivíduos com maior escolaridade média, 9,4 anos de estudo, enquanto no município de Vera Cruz estão concentrados os indivíduos com menor escolaridade média, 6,6 anos de estudo.

Gráfico 1 – Salário real médio por município da RM de Salvador, em 2010 (2012=100)



Fonte: Elaborado pela autora com base nos dados do Censo do IBGE (2010).

Gráfico 2 – Escolaridade média por município da RM de Salvador

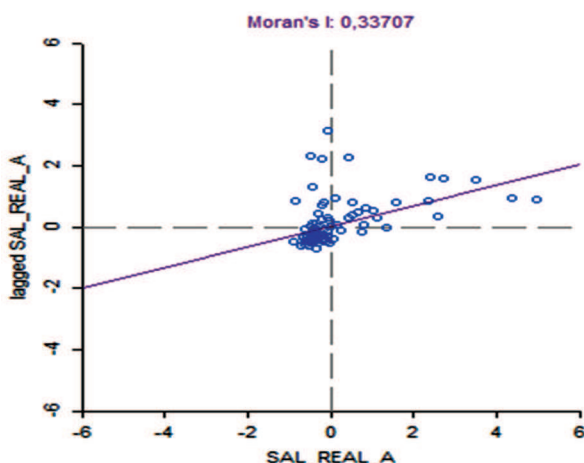


Fonte: Elaborado pela autora com base nos dados do Censo do IBGE (2010).

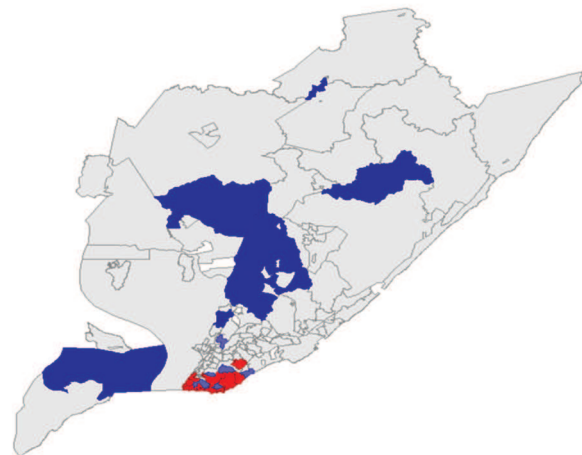
A Figura 1 mostra o I de Moran e o Local *Indicator of Spatial Association* (LISA) para os salários reais médios das áreas de ponderação da região metropolitana de Salvador. O valor do índice de Moran sugere a possível existência de autocorrelação espacial positiva na distribuição dos salários da região metropolitana de Salvador. Por sua vez, o indicador LISA mostra os clusters espaciais a partir da distribuição de salários reais da RM de Salvador. Como pode ser observado, há uma forte presença de clusters de baixos salários na RM de Salvador, representados na cor azul.

Figura 1 – I de Moran e *Local Indicator of Spatial Association* (LISA) dos salários reais médios das áreas de ponderação da RM de Salvador: 2010 (R\$ de 2012)

I de Moran



LISA¹



Fonte: Elaborado pela autora com base nos dados do Censo do IBGE (2010) e do software GEODA.

Nota: 1- A cor azul representa os clusters low-low (9), a cor vermelha representa os clusters high-high (8) e a cor lilás representa os clusters low-high (4).

4.2 Retorno à educação na Região Metropolitana de Salvador

A Tabela 3 apresenta os principais resultados encontrados com as estimações econométricas da equação (4), por MQO, e das equações (5) e (6) por GMM em dois estágios. É possível notar que a maioria das estimativas foi estatisticamente significativa ao nível de 1%, na estimação por MQO. No geral, os resultados apresentaram os sinais esperados de acordo com a literatura apresentada neste trabalho. O gênero e a experiência dos trabalhadores da região metropolitana de Salvador foram significantes. A experiência apresentou efeitos positivos e não lineares sobre os salários reais, sugerindo a tendência dos salários crescerem a taxas decrescentes, com a experiência, o que é compatível com as evidências da literatura.

As mulheres receberam salário real 71,5% abaixo dos salários reais dos homens no mercado de trabalho da região metropolitana de Salvador. Para as características do emprego, setor de atividade e ocupação, os resultados também foram estatisticamente significantes. O trabalho em uma ocupação técnica ofereceu salários 201,3% mais altos em relação às ocupações manuais, enquanto o trabalho especializado ofereceu salários reais 297,2% mais elevados, em relação à mesma categoria.

Por outro lado, tomando o setor de comércio, reparação de veículos automotores e motocicletas como referência, todos os 21 setores apresentaram efeitos positivos sobre os salários reais. O setor com maior efeito positivo sobre os salários reais, em relação ao setor base, foi o de serviços domésticos, com ganhos de 461,8%. O setor com menor ganho salarial (171,5%), em relação ao setor de referência, foi o da agricultura, pecuária, produção florestal, pesca e aquicultura.

A educação apresentou estimativa significativa e efeito positivo sobre os salários reais. O retorno à educação na região metropolitana de Salvador foi de 5,6%, sugerindo o ganho médio que um trabalhador deve obter para cada ano adicional de escolaridade, nessa localidade. A magnitude da estimativa foi compatível com os estudos existentes na literatura nacional para essa região metropolitana. Por fim, resta saber se essa estimativa se altera com a possível existência de autocorrelação espacial na distribuição salarial da região metropolitana de Salvador, a qual tornaria o parâmetro estimado enviesado. Essa análise será abordada com a inclusão de defasagens espaciais da variável dependente como um controle na regressão, a qual foi estimada através do método de GMM em dois estágios.

Tabela 3 – Resultados das estimações (Variável dependente: ln salário real)

Variáveis explicativas	MQO	GMM-2SLS
<i>W_1 ln(salário real)</i>	-	0,0124**
	-	(0,0045)
<i>educação</i>	0,0557***	0,0557***
	(0,0014)	(0,0016)
<i>dummy de sexo(=1 se mulher)</i>	-0,715***	-0,715***
	(0,0108)	(0,0129)
<i>experiência</i>	0,0680***	0,0679***
	(0,0013)	(0,0014)
<i>experiência²</i>	-0,00124***	-0,00124***
	(0,0000)	(0,00003)
<i>dummy setor22</i>	3,002***	3,004***
	(0,0260)	(0,0280)
<i>dummy setor21</i>	3,163*	3,081***
	(1,612)	(0,0431)
<i>dummy setor20</i>	4,618***	4,620***
	(0,0217)	(0,0171)
<i>dummy setor19</i>	4,138***	4,140***
	(0,0339)	(0,0307)

Variáveis explicativas	MQO	GMM-2SLS
<i>dummy setor18</i>	2,899***	2,900***
	(0,0555)	(0,0486)
<i>dummy setor17</i>	3,586***	3,588***
	(0,0294)	(0,0259)
<i>dummy setor16</i>	2,990***	2,990***
	(0,0292)	(0,0273)
<i>dummy setor15</i>	3,567***	3,569***
	(0,0287)	(0,0247)
<i>dummy setor14</i>	3,346***	3,346***
	(0,0297)	(0,0250)
<i>dummy setor13</i>	3,240***	3,243***
	(0,0371)	(0,0288)
<i>dummy setor12</i>	3,644***	3,644***
	(0,0639)	(0,0506)
<i>dummy setor11</i>	3,441***	3,445***
	(0,0526)	(0,0352)
<i>dummy setor10</i>	3,087***	3,087***
	(0,0514)	(0,0331)
<i>dummy setor9</i>	3,871***	3,875***
	(0,0286)	(0,0247)
<i>dummy setor8</i>	3,156***	3,157***
	(0,0303)	(0,0249)
<i>dummy setor6</i>	4,135***	4,136***
	(0,0219)	(0,0190)
<i>dummy setor5</i>	3,868***	3,870***
	(0,0653)	(0,0433)
<i>dummy setor4</i>	3,451***	3,461***
	(0,119)	(0,0630)
<i>dummy setor3</i>	3,560***	3,561***
	(0,0245)	(0,0226)
<i>dummy setor2</i>	3,711***	3,718***
	(0,0764)	(0,0459)
<i>dummy setor1</i>	1,715***	1,715***
	(0,0525)	(0,0714)
<i>dummy ocupação4</i>	2,972***	2,973***
	(0,0213)	(0,0247)
<i>dummy ocupação3</i>	2,013***	2,012***
	(0,0162)	(0,0189)
<i>dummy ocupação1</i>	2,135***	2,137***
	(0,0240)	(0,0299)
<i>Constante</i>	0,757***	0,708***
	(0,0214)	(0,0294)
<i>Número de observações</i>	153.020	153.020
<i>R²</i>	0,680	0,680
<i>R² Ajustado</i>	0,680	0,680
<i>Prob.> F</i>	0,000	-
<i>Prob.> chi2</i>	-	0,000

Fonte: Elaborado pela autora com base nos dados do Censo do IBGE (2010).

Notas: Erro padrão entre parênteses; * p<0,05; ** p<0,01; *** p<0,001.

Os resultados estimados por GMM, com o controle de possível autocorrelação espacial na distribuição de salários dos indivíduos pertencentes à mesma área de ponderação (W1), apresentaram alterações marginais, sem grandes variações em relação às estimativas de MQO. Em geral, as estimativas foram significantes a 1% e apresentaram sinais similares aos encontrados na regressão de MQO. As características individuais e do trabalho apresentaram o mesmo comportamento em relação à regressão de MQO. O retorno à educação se manteve no mesmo patamar da estimação por MQO, mesmo após o controle para a dependência espacial existente na distribuição de salários.

O efeito de interação social entre os indivíduos no espaço, conhecido na literatura como Peer Effect, pode ser representado pela defasagem espacial dos salários para os indivíduos vizinhos pertencentes à mesma área de ponderação. Os resultados apresentados mostraram que a distribuição de salários dos vizinhos importa para a determinação do salário de um indivíduo pertencente a este grupo, ou espaço. A estimativa do parâmetro de defasagem espacial salarial foi positiva e estatisticamente significativa. Os resultados encontrados sugerem que os indivíduos residentes na região metropolitana de Salvador compartilham um ganho salarial de 1,24% em decorrência dos salários de seus vizinhos, pertencentes à mesma área de ponderação. Esse efeito sugere que a determinação salarial pode depender dos salários que os indivíduos espacialmente próximos recebem, ou seja, o salário de um indivíduo pode ser uma função não somente de suas características e das características de seu trabalho, mas, também, da distribuição de salários de seus vizinhos. Esses efeitos podem evidenciar um comportamento compatível com a realidade, uma vez que os indivíduos tendem a se localizar espacialmente próximos aos seus semelhantes (pares) em termos de características socioeconômicas e isso pode reforçar os padrões de desigualdade existentes.

Portanto, os resultados encontrados sugerem a existência de efeitos espaciais na equação salarial minceriana. Embora o retorno à educação não tenha se alterado, após o controle da dependência espacial salarial, as estimativas encontradas para a defasagem espacial salarial mostraram que os efeitos de vizinhança importam para a determinação salarial individual. As evidências deste estudo são compatíveis com a literatura de *Peer Effects*, des-

de que sugerem a existência de efeitos endógenos para os indivíduos pertencentes ao mesmo grupo ou espaço. Nesse sentido, os resultados encontrados neste estudo apontam para a necessidade de considerar a presença de *Peer Effects* na determinação salarial nos estudos da literatura empírica.

5 Conclusão

O objetivo deste estudo foi identificar possíveis efeitos espaciais existentes na estimação econométrica do retorno à educação dos indivíduos residentes na Região Metropolitana de Salvador, supondo a existência de algum tipo de dependência espacial na equação salarial minceriana. Foi utilizada a amostra de microdados do Censo Populacional para as áreas de ponderação da Região Metropolitana de Salvador (IBGE, 2010). Uma equação minceriana foi estimada, com base no *Spatial Autoregressive Model-SAR*, utilizando os métodos de econometria espacial. Foi incluída uma defasagem espacial da variável dependente na equação salarial, para corrigir a possível autocorrelação espacial existente. A defasagem foi definida para a mesma área de ponderação do indivíduo.

A revisão dos trabalhos teóricos e empíricos da literatura que integra Economia Espacial e Economia do Trabalho sugere a existência de ganhos de aglomeração sobre os salários em áreas urbanas densas. Existem evidências de que a educação atua ampliando os ganhos salariais associados à localização. Nesse sentido, as interações sociais entre os indivíduos no espaço podem gerar retornos salariais diferenciados espacialmente para cada ano de educação. Embora alguns estudos mais recentes na literatura empírica nacional tenham considerado a importância da localização para a determinação salarial e dos retornos à educação, não foram encontrados estudos que utilizassem uma abordagem de econometria espacial para estimar os efeitos de interação social.

A literatura recente sobre interações sociais na teoria econômica tem mostrado que o comportamento do indivíduo pode ser influenciado pelo grupo de referência no qual está inserido. O trabalho seminal de Manski (1993) definiu como Efeitos Sociais Endógenos os canais através dos quais a sociedade pode afetar o comportamento dos indivíduos. Essa literatura busca entender se alguma forma de interação pode explicar porque os agen-

tes pertencentes ao mesmo grupo se comportam similarmente. Nesse sentido, a literatura de *Peer Effects* assume a existência de efeitos dos grupos em que os indivíduos estejam inseridos sobre os resultados desses indivíduos.

Os resultados estimados, através dos métodos de MQO e GMM em dois estágios, mostraram um retorno à educação na região metropolitana de Salvador de 5,6%, sugerindo o ganho em média que um trabalhador deve obter para cada ano adicional de escolaridade, nessa área. Embora o retorno à educação não tenha se alterado, após o controle da dependência espacial salarial, as estimativas encontradas para a defasagem espacial salarial mostraram que os efeitos de vizinhança importam para a determinação salarial individual. Os indivíduos residentes na região metropolitana de Salvador apresentaram um ganho salarial de 1,24% em decorrência dos salários dos seus vizinhos mais próximos, pertencentes à mesma área de ponderação. Tais efeitos sociais seriam compatíveis com a literatura de *Peer Effects* e mostram que a distribuição salarial de determinado grupo ou espaço pode atuar como um determinante direto do salário de um indivíduo inserido neste grupo ou espaço.

As evidências empíricas encontradas sugerem que não somente a localização do indivíduo pode afetar suas escolhas e seus resultados econômicos, como também a de seus vizinhos mais próximos. Interações sociais de indivíduos pertencentes a determinados grupos (espaços) podem explicar comportamentos individuais e resultados decorrentes desses. Nesse sentido, os efeitos de interações sociais no espaço podem ser relevantes para a determinação dos salários individuais e devem ser considerados nos estudos que estimam uma equação salarial minceriana.

Referências

- ABEL, J. R.; DEY, I.; GABE, T. M. Productivity and the density of human capital. New York: **Federal Reserve Bank Staff Reports**, n. 440, Sept. 2011 (Revised).
- BARROS, R. P. de; HENRIQUES, R.; MENDONÇA, R. Pelo fim das décadas perdidas: educação e desenvolvimento sustentado no Brasil. In: HENRIQUES, R. (Org.). **Desigualdade e pobreza no Brasil**. Rio de Janeiro: IPEA, 2000. 740 p.
- CAMPOS, F. M.; SILVEIRA NETO, R. M. A importância da dimensão do mercado de trabalho para os diferenciais de participação e salários entre gêneros: uma análise empírica para os centros urbanos brasileiros. In: ENCONTRO NACIONAL DE ECONOMIA - ANPEC, 37, 2009, Foz do Iguaçu. **Anais...** Foz do Iguaçu: ANPEC, 2009. Disponível em: <<http://www.anpec.org.br/novosite/br/encontro-2009#trabalhos>>. Acesso em: 20 dez. 2016.
- DANTAS, A. R. M.; FIGUEIREDO, E. A. Retorno da educação nos estados nordestinos: Piauí, Rio Grande do Norte e Bahia. **Revista Economia e Desenvolvimento**, v. 9, n. 1, 2010.
- FALCÃO, N.; SILVEIRA NETO, R. Concentração espacial de capital humano e externalidades: o caso das cidades brasileiras. In: ENCONTRO NACIONAL DE ECONOMIA - ANPEC, 35., 2007, Recife. **Anais...** Recife: ANPEC, 2007. Disponível em: <<http://www.anpec.org.br/novosite/br/encontro-2007#>>. Acesso em: 20 dez. 2016.
- FREITAS, U. R. P.; FREITAS, L. F. S. Estimativa dos determinantes do rendimento na Região Metropolitana de Salvador: uma avaliação à luz da teoria do capital humano. **Revista Desenbahia**, n. 8, Mar. 2008.
- GLAESER, E.; MARÉ. Cities and skills. **Journal of Labor Economics**, v. 19, n. 2, p. 316-342, 2001.
- GLAESER, E.; MARÉ. Cities and skills. **National Bureau of Economic Research-NBER Working Papers**, n. 4728, 1994.
- GLAESER, E.; RESSEGER, M. G. The complementarity between cities and skills. **Journal of Regional Science**, v. 50, n. 1, p. 221-244, 2010.
- HALFDANARSON, B.; HEUERMAN, D. F.; SÜDEKUM, J. Human capital externalities and the urban wage premium: two literatures and their interrelations. **Discussion Paper The Institute for the Study of Labor (IZA)**, n. 3493, May 2008.
- IBGE. INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo demográfico**. 2010. Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/resultados_gerais_amostra-re>

- sultados_gerais_amostra_tab_uf_microdados.shtm>. Acesso em: 27 out. 2013.
- KELEJIAN, H. H.; PRUCHA, I. R. A Generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. **Journal of Real Estate Finance and Economics**, v. 17, n. 1, p. 99-121, 1998.
- LACERDA, F. C. C. Desigualdade de rendimentos na Bahia: estimação de equações de rendimento com base nos microdados da PNAD 2005. In: ENCONTRO DE ECONOMIA BAIANA-EEB, 4, 2008, Salvador. **Anais...** Salvador: EEB, 2008.
- MANSKI, C. F. Economic Analysis of Social Interactions. **The Journal of Economic Perspectives**, v. 14, n. 3, 2000, p. 115-136.
- MANSKI, C. F. Identification of Endogenous Social Effects: The Reflection Problem. **The Review of Economic Studies**, vol. 60, n. 3. Jul., 1993, p. 531-542.
- MINCER, J. **Schooling, experience, and earnings**. New York: Columbia University Press. 1974. 167 p.
- RAUCH, J. Productivity gains from geographic concentration of human capital: evidence from Cities. **Journal of Urban Economics**, n.34, p. 380-400, 1993.
- ROCHA, R. de M. ; SILVEIRA NETO, R. da M. ; GOMES, S. M. F. P. O. Maiores cidades, maiores habilidades produtivas: ganhos de aglomeração ou atração de habilitados? Uma análise para as cidades brasileiras. In: FÓRUM BNB DE DESENVOLVIMENTO, 17. ENCONTRO REGIONAL DE ECONOMIA – ANPEC NORDESTE, 16, 2011, Fortaleza. **Anais...** Fortaleza: Fórum BNB de Desenvolvimento, 2011. Disponível em: <<http://www.bnb.gov.br/edicoes-antiores2>>. Acesso em: 20 dez. 2016.
- SERVO, L.; AZZONI, C. Education, cost of living and regional wage inequality in Brazil. **Papers in Regional Science**, v. 81, n. 2, p.157-175, 2002.
- SILVA, D. L. G. **Distribuição espacial dos efeitos dos ganhos de aglomeração sobre os retornos à educação no Brasil**. 2012. 157 f. Dissertação (Mestrado em Economia) – Faculdade de Economia, UFBA, Salvador, 2012.
- SULIANO, D. L.; SIQUEIRA, M. L. Um estudo do retorno da educação na região nordeste: análise dos estados da Bahia, Ceará e Pernambuco a partir da recente queda da desigualdade. Fortaleza: **Texto para Discussão do Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE)**, n. 72, Fev. 2010.
- WRIGLEY, N.; HOLT, T.; STEEL, D.; TRANMÉR, M. Analysing, modelling, and resolving the ecological fallacy. In: LONGLEY, P.; BATTY, M. (Eds.). **Spatial analysis: Modelling in a GIS environment**. Cambridge: GeoInformation International, p. 23–40, 1996.
- YANKOW, J. J. Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. **Journal of Urban Economics**, v. 60, p. 139-161, 2006.

APÊNDICES

APÊNDICE A – MODELO LINEAR: INFERÊNCIA DOS EFEITOS SOCIAIS ENDÓGENOS

Este apêndice apresenta um modelo linear de inferência dos efeitos de interação social, descrito em Manski (1993). Assuma que cada membro da população é caracterizado por um valor para $(y, x, z, u) \in \mathbb{R}^1 \times \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^1$, onde y é um escalar representando uma variável de resultado (dependente), x são os atributos caracterizando o grupo de referência de um indivíduo e (z, u) são os atributos que afetam y diretamente. As variáveis aleatórias (y, x, z) são diretamente observadas pelo pesquisador, enquanto u é não observado, tal que:

$$y = \alpha + \beta E(y|x) + E(z|x)' \gamma + z' \eta + u, E(u|x, z) = x' \delta \quad (3)$$

onde $(\alpha, \beta, \gamma, \delta, \eta)$ é um vetor de parâmetros. A regressão de y sobre (x, z) , em média, tem a seguinte forma linear:

$$E(y|x, z) = \alpha + \beta E(y|x) + E(z|x)' \gamma + x' \delta + z' \eta. \quad (4)$$

Se $\beta \neq 0$, a equação (4) apresenta um efeito endógeno: o resultado y de um indivíduo varia com $E(y|x)$, a média de y entre os membros no grupo de referência definido como x^8 . As evidências empíricas nos estudos existentes fornecem pouco suporte para a especificação dos grupos de referência. Geralmente é assumido que os indivíduos são influenciados em $E(y|x)$ e $E(z|x)$ por algum x específico. Se $\gamma \neq 0$, o modelo apresenta um efeito exógeno: y varia com $E(z|x)$, a média das variáveis exógenas entre os membros do grupo de referência. Se $\delta \neq 0$, o modelo apresenta efeitos correlacionados: indivíduos no grupo de referência x tendem a se comportar de maneira similar porque eles possuem características não observadas, u , similares ou estão inseridos em ambientes institucionais similares. O parâmetro η representa o efeito direto de z em y .

O “problema de reflexão” definido por Manski (1993) surge da presença de $E(y|x)$ como um regressor em (4). Após integrar ambos os lados de

⁸ Generalizações do modelo podem incluir múltiplos grupos de referência influenciando um indivíduo e o indivíduo dará mais peso ao comportamento de alguns grupos do que a outros.

(4), em relação à z , a equação de equilíbrio social será definida por:

$$E(y|x) = \alpha + \beta E(y|x) + E(z|x)' \gamma + x' \delta + E(z|x)' \eta. \quad (5)$$

Para $\beta \neq 1$, a equação (5) tem uma única solução, dada por:

$$E(y|x) - \beta E(y|x) = \alpha + E(z|x)' \gamma + x' \delta + E(z|x)' \eta E(y|x) = \frac{\alpha}{1-\beta} + E(z|x)' \frac{(\gamma + \eta)}{1-\beta} + x' \frac{\delta}{1-\beta} \quad (6)$$

Então, $E(y|x)$ é uma função linear de $[1, E(z|x), x]$, de modo que os parâmetros não são identificados. Inserindo (6) em (4) é obtido o modelo em sua forma reduzida:

$$E(y|x, z) = \alpha + \beta \left[\frac{\alpha}{1-\beta} + E(z|x)' \frac{(\gamma + \eta)}{1-\beta} + \frac{x' \delta}{1-\beta} \right] + E(z|x)' \gamma + x' \delta + z \eta \quad (7)$$

$$E(y|x, z) = \frac{\alpha}{1-\beta} + E(z|x)' \frac{(\gamma + \beta \eta)}{1-\beta} + \frac{x' \delta}{1-\beta} + z \eta$$

Então, no modelo linear em (4), com $\beta \neq 1$, os parâmetros compostos $\alpha/(1-\beta)$, $(\gamma + \beta \eta)/(1-\beta)$, $\delta/(1-\beta)$ e η são identificados se os regressores $[1, E(z|x), x, z]$ são linearmente independentes na população de interesse. O modelo não permite distinguir os efeitos endógenos dos efeitos exógenos ou dos efeitos correlacionados, mas permite determinar se algum efeito social existe desde que os seus parâmetros sejam diferentes de zero. Entretanto, com $\beta \neq 1$, no modelo (4), os parâmetros de efeitos sociais não serão identificados se z for uma função linear de x , não variar com x e $E(z|x)$ for uma função linear de x . A inferência só será possível se $E(z|x)$ variar não-linearmente com x e $V(z|x) > 0$.

Estudos empíricos de efeitos sociais endógenos geralmente assumem que $\gamma = \delta = 0$, ou seja, nem o efeito exógeno e nem o efeito correlacionado estão presentes. Estes estudos utilizam métodos de estimação em dois estágios para estimar (β, η) sob o formato de um **modelo de correlação espacial**, dado por:

$$y_i = \beta W_{iN} Y + z_i' \eta + u_i, i=1, \dots, N \quad (8)$$

onde $Y = (y_i, i=1, \dots, N)$ é um vetor $N \times 1$ de realizações amostrais de y e W_{iN} é um vetor de pesos $1 \times N$, cujos componentes são não negativos e devem somar um. Geralmente é assumido que o termo de erro u apresenta uma distribuição normal, independente de x , e o modelo é estimado por máxima verossimilhança. A equação (8) mostra que o comportamento de cada indivíduo varia com uma média ponderada de comportamentos de outros membros do grupo.

O modelo de correlação espacial faz sentido em estudos de interações entre grupos grandes, se interpretado como um método de dois estágios para a estimação de efeitos endógenos puros. No primeiro estágio, os dados amostrais são utilizados para estimar $E(y | x)$ não parametricamente e no segundo estágio são estimados os parâmetros de interesse (β ,

η), através de ajuste de mínimos quadrados de y sobre $[1, E_N(y | x), z]$, onde $E_N(y | x)$ é a estimativa de primeiro estágio de $E(y | x)$. Estimativas não paramétricas de $E(y | x)$ podem ser médias ponderadas na forma de $E_N(y | x_i) = W_{iN} Y$. Deste modo, estimativas de (β , η) definidas na literatura de correlação espacial podem ser interpretadas como estimativas dos efeitos endógenos sociais puros. Por outro lado, estimativas de modelos não identificados podem ser obtidas. Mesmo se $E(y | x)$ é uma função linear de $[1, z]$, a estimativa de $E_N(y | x)$ é geralmente linearmente independente de $[1, z]$, de modo que os procedimentos de estimação em dois estágios fornecem estimativas de β mesmo quando esse parâmetro não é identificado.

APÊNDICE B – CLASSIFICAÇÃO DOS SETORES DE ATIVIDADE NO CENSO (IBGE, 2010)

Tabela 1 – Setor de Atividade Econômica

Setor	Código
Agricultura, pecuária, produção florestal, pesca e aquicultura	1
Indústrias extrativas	2
Indústrias de transformação	3
Eletricidade e gás	4
Água, esgoto, atividades de gestão de resíduos e descontaminação	5
Construção	6
Comércio; reparação de veículos automotores e motocicletas	7
Transporte, armazenagem e correio	8
Alojamento e alimentação	9
Informação e comunicação	10
Atividades financeiras, de seguros e serviços relacionados	11
Atividades imobiliárias	12
Atividades profissionais, científicas e técnicas	13
Atividades administrativas e serviços complementares	14
Administração pública, defesa e seguridade social	15
Educação	16
Saúde humana e serviços sociais	17
Artes, cultura, esporte e recreação	18
Outras atividades de serviços	19
Serviços domésticos	20
Organismos internacionais e outras instituições extraterritoriais	21
Atividades mal definidas	22

Fonte: Elaborado pela autora com base nos dados do Censo do IBGE (2010).

APÊNDICE C – MAPA DA REGIÃO METROPOLITANA DE SALVADOR POR ÁREA DE PONDERAÇÃO (107 APS)

Figura C.1 – Distribuição das áreas de ponderação da Região Metropolitana de Salvador (2010)



Fonte: Elaborado pela autora com base nos dados do Censo do IBGE (2010) e do software GEODA.