

Acessibilidade ao Emprego e Produtividade na Região Metropolitana de São Paulo

Leandro Batista Duarte¹  | Andrews Augusto Diniz Barros² 

¹ Universidade Estadual de Feira de Santana. Email: lbduarte@uefs.br

² Universidade Federal de Pernambuco. Email: andrews.barros@ufpe.br

RESUMO

A hipótese do *mismatch* espacial entre trabalhadores e oportunidades de trabalho identifica vários mecanismos através dos quais o acesso físico restrito aos empregos pode resultar em resultados insatisfatórios no mercado de trabalho. Dessa forma, este estudo tem como objetivo analisar a relação entre acessibilidade ao emprego e salários no mercado de trabalho da Região Metropolitana de São Paulo. Como existem problemas de endogeneidade entre a localização residencial e a mensuração de oportunidades de trabalho, o estudo fez uso de uma subamostra (jovens adultos que moram com os pais) como forma de amenizar os problemas de endogeneidade. O artigo utilizou dois bancos de dados: o Censo e a RAIS. Os principais resultados encontrados mostram que há uma relação entre duas medidas diferentes de acessibilidade aos empregos e salários, ou seja, o aumento do acesso físico ao CBD e subcentros aumenta a produtividade do trabalhador na RMSP.

PALAVRAS-CHAVE

Mismatch Espacial, Acessibilidade, Salários.

Accessibility to Employment and Productivity in the Metropolitan Region of São Paulo

ABSTRACT

The hypothesis of spatial mismatch between workers and job opportunities identifies several mechanisms through which restricted physical access to jobs can result in unsatisfactory labor market outcomes. Thus, this study aims to analyze the relationship between accessibility to employment and wages in the labor market in the Metropolitan Region of São Paulo. As there are endogeneity problems between the residential location and the measurement of work opportunities, the study used a subsample (young adults who live with their parents) as a way of alleviating the endogeneity problems. The article used two databases: the Census and the RAIS. The main results found show that there is a relationship between two different measures of accessibility to jobs and wages, that is, the increase in physical access to the CBD and sub-centers increases worker productivity in the RMSP.

KEYWORDS

Spatial Mismatch, Accessibility, Wage

CLASSIFICAÇÃO JEL

R10, R30

1. Introdução

Com uma população de 20 milhões de habitantes em 2010, a Região Metropolitana de São Paulo (RMSP) é o centro urbano mais importante do Brasil contribuindo com aproximadamente 20% para o PIB nacional e concentrando 10% da população brasileira. Não muito diferente de algumas megacidades como Tóquio, Nova York, Seul e Cidade do México, a RMSP é conhecida por problemas urbanos crônicos associados a níveis muito altos de aglomeração, incluindo congestionamento de tráfego e muitas externalidades associadas (Wheeler, 2003).

No Brasil, há a tendência da população de renda mais elevada morar nas regiões centrais, enquanto os grupos com renda mais baixa ocupam o entorno das aglomerações urbanas, na periferia. Este cenário não é diferente para a RMSP que exibe uma divisão núcleo-periferia particularmente marcada (Ramos, 2016). Moreno-Monroy e Ramos (2015) argumentaram que a formação da RMSP se deu com um rápido crescimento populacional, tendo a população pobre e menos instruída ocupando um vasto cinturão periférico suburbano ao redor do centro formado por um processo de expansão centrífuga não planejada. Dessa forma, Nadalin e Iglori (2010) afirmaram que isso explica, pelo menos parcialmente, o padrão de localização das famílias na RMSP de acordo com a renda, com as mais ricas, menos distantes dos centros de emprego.

A localização das famílias mais pobres na periferia da RMSP tanto pode refletir escolhas em função da renda, como dificuldades de acesso aos empregos (*mismatch* espacial). Segundo a hipótese do *mismatch*, trabalhadores localizados em regiões socioeconomicamente vulneráveis, enfrentam maiores dificuldades para encontrar empregos nos centros de negócios das cidades (Kain, 1968; Ihlanfeldt, 2006). O *mismatch* espacial é a desconexão espacial entre trabalhadores e oportunidades de trabalho. Dessa forma, o *mismatch* espacial identifica vários mecanismos¹ através dos quais o acesso físico restrito aos empregos pode causar resultados insatisfatórios no mercado de trabalho.

Um achado consistente do ponto de vista teórico é o argumento de que os modelos de salário-eficiência urbanos indicam que a produtividade e a duração do deslocamento estão relacionadas, ou seja, os trabalhadores que passam por viagens de deslocamento mais longo fornecem níveis de esforço mais baixos do que aqueles que residem mais perto dos empregos (Zenou, 2002; Pierrard, 2008).

Em relação à literatura empírica, alguns autores estudaram a relação entre acessibilidade ao emprego e produtividade dos trabalhadores (Van Ommeren e Gutiérrez-i Puigarnau, 2011; Stone e Schneider, 2016; Vieira e Haddad, 2015; Haddad e Barufi, 2017). Estes estudos indicaram que deslocamentos mais longos diminuem a produtividade dos agentes e afetam negativamente uma ampla gama de indicadores de saúde, como níveis percebidos de estresse (Costa e Steckel, 1997; Hansson et al., 2011). Van Ommeren e Gutiérrez-i Puigarnau (2011) examinaram o impacto do des-

¹Ver Kain (1968) e Gobillon et al. (2011) para mais detalhes.

locamento sobre a produtividade dos trabalhadores, que se manifesta por níveis mais altos de absenteísmo para aqueles com viagens mais longas na Alemanha. Stone e Schneider (2016) mostraram para uma amostra de residentes dos Estados Unidos que os deslocamentos mais longos foram associados ao aumento do estresse e do cansaço, gerando baixo bem estar.

Embora exista uma reduzida literatura abordando evidências de acessibilidade ao emprego determinando os salários intraurbanos no Brasil, pode-se citar o estudo de Vieira e Haddad (2015) que avaliaram os impactos econômicos do metrô de São Paulo, considerando a relação entre mobilidade, acessibilidade e produtividade, e seus efeitos de equilíbrio geral; Haddad e Barufi (2017) que exploraram a relação entre tempo de deslocamento, acessibilidade para empregos e preços urbanos no contexto dos mercados urbanos na Região Metropolitana de São Paulo (RMSP), observando como os salários individuais são afetados pelo potencial acesso às oportunidades de trabalho; Bonomi Barufi et al. (2016) que investigaram a relação entre mismatch espacial, salários e desemprego para as grandes áreas urbanas do Brasil, que inclui a RMSP. E por fim, tem-se o trabalho de Duarte (2020) que analisou a relação entre acessibilidade ao emprego e a produtividade do mercado de trabalho na cidade do Recife.

Neste estudo, o objetivo é analisar a relação entre acessibilidade ao emprego e os resultados do mercado de trabalho na RMSP. Mais especificamente, pretende-se a partir do Censo 2010 e do georreferenciamento da localização dos empregos estimar uma relação empírica entre o nível de acesso aos empregos e salários, considerando covariáveis ou fatores que afetam ou estão associados simultaneamente a estas duas variáveis. A variável de interesse é o salário por hora do trabalhador que pode ser afetado conjuntamente, entre outras características, pela acessibilidade aos empregos na região.

A aplicação empírica desta pesquisa está ligada a um grande número da literatura que analisa o efeito da acessibilidade ao emprego nos resultados do mercado de trabalho. Para o Brasil, são poucas evidências disponíveis sobre esta relação. O estudo, portanto, contribui para a literatura investigando o *mismatch* espacial em mercados de trabalho urbanos na RMSP. Além disto, compreender os padrões de acessibilidade ao emprego nas cidades dos países em desenvolvimento é de fundamental importância, bem como mostrar empiricamente se o *mismatch* espacial é relevante para a produtividade.

Apesar do maior interesse recente sobre o tema, a maior parte dos estudos empíricos ainda não chegou a um consenso sobre o papel dos fatores espaciais na explicação dos resultados individuais no mercado de trabalho. Como este estudo utiliza como estrutura conceitual um modelo espacial de equilíbrio geral de uso da terra com locais de trabalho endógenos em uma cidade circular (Fujita e Ogawa, 1982; Lucas e Rossi-Hansberg, 2002) e propriedades de eficiência da alocação de terras nessa mesma estrutura de cidade (Rossi-Hansberg, 2004) é necessário estar ciente do potencial viés de simultaneidade nas estimações empíricas para resultados no mercado

de trabalho com base em medidas de acessibilidade. É possível lidar com a questão da simultaneidade incluindo instrumentos históricos ou geográficos que influenciaram a localização da infraestrutura de transporte dentro de uma cidade sem determinar diretamente a localização de trabalhadores e empresas.

Entre os poucos estudos empíricos que tentam lidar com a questão da endogeneidade, a maioria prefere restringir a amostra a jovens adultos que residem com os pais, argumentando que a escolha do local de residência dos pais pode ser considerada exógena ao status de emprego de jovens adultos (ver, por exemplo, (O'Regan e Quigley, 1998; Dujardin et al., 2008). No entanto, tem sido argumentado que esta solução não elimina completamente a endogeneidade, pois podem existir características familiares não observadas que influenciam tanto a escolha residencial quanto o status de emprego de adultos jovens (Glaeser et al., 1996).

O presente artigo se concentrou em utilizar uma subamostra de jovens adultos (18 a 25 anos de idade) que vivem com seus pais com o argumento de que a escolha do local de residência dos pais pode ser considerada exógena a busca por melhores salários próximos aos centros de empregos. A literatura mostra que o número de vagas de emprego e o nível salarial tendem a diminuir à medida que a distância ao centro urbano aumenta. Empregos com melhores salários podem estar concentrados perto do centro, pois se beneficiam mais de transbordamentos de conhecimento que geram externalidades de aglomeração (Partridge et al., 2009). Além disso, essa relação é considerada mais forte para áreas maiores e mais densas, porque os custos de congestionamento e o tamanho da expansão urbana levam a um custo de vida mais alto para as áreas centrais.

O artigo está estruturado da seguinte forma. Além desta introdução, a Seção 2 apresenta discussões acerca da relação da acessibilidade ao emprego e mercado de trabalho, a Seção 3 descreve a área de estudo, banco de dados e a estratégia econométrica, enquanto na Seção 4 serão analisados os resultados, seguida da robustez. Considerações finais seguem na Seção 6.

2. Acessibilidade ao emprego e mercado de trabalho

Como a questão foco do trabalho é a análise da relação entre acessibilidade ao emprego e resultados no mercado de trabalho, cabe esclarecer alguns conceitos e hipóteses voltados para esses argumentos. A seção dará início mencionando a relação do deslocamento de casa ao trabalho com fricções na produtividade, para em seguida apresentar a relação com o *mismatch* espacial.

Para a maior parte da população assalariada dos grandes centros urbanos do país, a viagem de casa ao trabalho ocorre em várias etapas, que envolvem o uso de transportes coletivos de massa (ônibus, metrô, trem etc). São estes trabalhadores, de classes mais baixas, por habitarem em locais mais afastados do centro, que incorrem em um ônus crescente com os congestionamentos, atrasos, baldeações e desorganização dos

transportes. Para eles, o tempo de deslocamento casa/trabalho é perdido, podendo ter impactos na sua qualidade de vida bem como na sua produtividade e predisposição para o trabalho (Mihessen et al., 2014). A produtividade dos trabalhadores tende a diminuir com o maior tempo gasto nos deslocamentos intraurbanos, uma vez que o nível de estresse e de atrasos é elevado (Proque, 2014). De acordo com a teoria microeconômica, os indivíduos não escolheriam ter um deslocamento mais longo, a menos que fossem compensados de alguma forma, seja na forma de melhores características de trabalho (incluindo remuneração) ou melhores perspectivas de moradia (Stutzer e Frey, 2008).

A literatura empírica sobre *mismatch* espacial pode ser subdividida em duas vertentes principais: a primeira visa compreender as causas, enquanto a segunda discute as consequências de um *mismatch* espacial (Gobillon e Selod, 2021). Conforme (Houston, 2005), as consequências de um *mismatch* espacial são geralmente avaliadas através de uma análise de segregação residencial, comparações de tempos de deslocamento, comparações de ganhos e medidas de proximidade ao trabalho. (Ihlanfeldt, 2006) destacaram o fato de que os efeitos do *mismatch* espacial têm sido investigados em salários mais baixos e viagens mais longas, especialmente no caso de trabalhadores negros nos Estados Unidos. Normalmente, as equações de emprego e renda incluem medidas de oportunidades de emprego locais, com uma estratégia baseada em um modelo de gravidade com função de decaimento da distância para levar em conta a distância das oportunidades de emprego.

Neste estudo, espera-se que deslocamentos mais longos diminuam a produtividade dos agentes, pois induzem o trabalhador a chegar mais tarde ao trabalho ou a sair mais cedo, além de aumentar o número de faltas (Van Ommeren e Gutiérrez-Puigarnau, 2011). Essa diminuição da produtividade pode ser observada por um empregador, que pode levar isso em conta ao contratar (ou demitir) determinado trabalhador (Zenou, 2002)). A razão pela qual a produtividade pode ser deteriorada pela distância é o fato de que quando a residência do trabalhador é mais distante do local de trabalho, este tem maior probabilidade de estar atrasado ou cansado (Gobillon e Selod, 2021). Por outro lado, os trabalhadores que realizam deslocamentos mais curtos podem apresentar maiores níveis de esforço no trabalho, sendo mais produtivos (Zenou, 2002).

3. Área de estudo e dados

A área de estudo é a Região Metropolitana de São Paulo (RMSP), a maior metrópole brasileira com cerca de 19,5 milhões de habitantes, onde cerca de, 30% da força de trabalho foi empregada informalmente em 2010. A unidade de análise é o indivíduo dentro da Área Espacial de Ponderação² (AEP), que é definida como uma área com-

²Define-se área de ponderação como sendo uma unidade geográfica, formada por um agrupamento de setores censitários, para a aplicação dos procedimentos de calibração das estimativas com as informações conhecidas para a população como um todo (IBGE, 2010).

posta por um conjunto de zonas mutuamente exclusiva, projetado para dar a robustez estatística necessária à estratégia de amostragem do Censo de População. De acordo com Censo Demográfico de 2010, a RMSP é particionada em 633 AEPs, com uma população média de 31.096 habitantes e uma área média de 12,55 km².

Os micro dados para as características individuais são obtidos no Censo de População de 2010 (IBGE, 2010). Para o presente trabalho foi utilizado áreas de ponderação como definição de vizinhança, uma vez que, microdados ao nível individual também estão disponíveis para esta amostra. Também utilizou como fontes de recursos a Relação de Informações Anuais de Responsabilidade Social (RAIS) para o ano de 2010. Este é um dado administrativo mantido pelo Ministério do Trabalho do Brasil. A RAIS consiste em um painel de alta qualidade com informações sobre as características dos contratos entre empresas e funcionários, para todos os indivíduos formalmente empregados, desagregados no nível do município. Com isso, para a pesquisa foi utilizado os microdados georreferenciados da RAIS para identificação dos centros de empregos (subcentros), que contêm informações sobre o endereço de cada estabelecimento registrado. O conjunto de dados georreferenciado foi obtido no Instituto de Pesquisa Econômica Aplicada (IPEA).

3.1 Especificações Econométricas

A estratégia empírica desenvolvida baseia-se na estimativa da relação entre diferentes medidas de acessibilidade e resultados do mercado de trabalho (salários). As estimativas são realizadas para indivíduos que residem em uma área metropolitana específica, a fim de capturar o efeito de cada variável em termos relativos dentro de uma estrutura urbana específica. Tal estratégia visa explorar diferentes dimensões da hipótese do *mismatch* espacial em uma área metropolitana. Portanto, a equação salarial a ser estimada é com base em um modelo de OLS (*Ordinary Least Squares*), da seguinte maneira:

$$W_i = \beta_1 + \beta_2 AC_j + \beta_3 X_{ij} + \beta_4 H_{ij} + \beta_5 K_j \quad (1)$$

Nesta especificação, W_i refere-se ao logaritmo do salário³ por hora do trabalhador. Aqui, AC_j é o índice de acessibilidade por área de ponderação j , X_{ij} é um conjunto de características observadas para o indivíduo (sexo, idade, idade ao quadrado, cor ou raça, nível de escolaridade), H_j é um vetor de características domiciliares; K_j é um vetor de características de vizinhança (composição social); e β são vetores de parâmetros que serão estimados.

A acessibilidade aos empregos, que é a variável explicativa de interesse na análise, depende de alguns dos fatores que afetam as variáveis dependentes e podem ser indiretamente afetados. Neste caso, o viés de simultaneidade surge à medida que os

³Rendimento dividido pelo número de horas que a pessoa habitualmente dedicava no trabalho principal.

residentes e os locais das empresas são determinados ao mesmo tempo (Haddad e Barufi, 2017). Outras origens de endogeneidade podem surgir do comportamento dos trabalhadores. Por um lado, os trabalhadores são mais propensos a percorrer longas distâncias em troca de salários mais altos, o que influenciaria na maior produtividade. Por outro lado, indivíduos com rendimentos mais elevados podem se localizar mais longe de seus empregos para ter acesso a amenidades geralmente deslocadas nos arredores da cidade (Graham e Melo, 2009).

Tanto a presença de fatores omitidos que afetam simultaneamente as variáveis dependentes e o indicador de acessibilidade, como a possibilidade de causalidade reversa fazem as estimativas de OLS provavelmente tendenciosas. Contudo, o objetivo do estudo é encontrar uma associação entre a acessibilidade e produtividade dos trabalhadores. Essas preocupações podem ser potencialmente amenizadas com a utilização de subamostras.

Uma questão fundamental nos estudos de *mismatch* espacial é como medir a acessibilidade residencial às oportunidades de emprego. Esta variável leva em conta a distribuição espacial dos empregos e a distância ou custo de acesso para alcançá-los. As medidas de acessibilidade são geradas ao nível de AEP, com base em uma medida de oportunidade cumulativa. Este tipo de medida conta o número de trabalhos que podem ser alcançados a partir de um AEP e é calculado da seguinte forma:

$$AC_j = \sum_{k=1}^N \frac{E_k}{d_{jk}^\alpha(t)} \quad (2)$$

Na equação 2, C_j é a acessibilidade da região j (centroide da AEP), E_k corresponde à quantidade de empregos existentes em qualquer região j (centros de empregos), enquanto d_{jk} é o impedimento (por exemplo, distância euclidiana) existente para viajar entre j e k . O número total de empregos disponíveis em cada região da cidade e a impedância para alcançar essas regiões pode ser usado para definir a acessibilidade a empregos em um local específico. A medida de impedância é geralmente definida pela distância euclidiana ou pelo tempo de deslocamento entre o local residencial e os empregos, que pode ser derivado da disponibilidade de transporte em cada área da cidade.

Os valores de E_k foram obtidos da base de dados do Núcleo de Estudos Regionais e Urbanos (NERU) da UFPE, que utilizam informações da RAIS 2010 disponibilizadas pelo IPEA para o georreferenciamento das firmas e empregos. Logo, E_k corresponde aos empregos formais dos principais centros de emprego da RMSP, contabilizando 94 subcentros e 1 centro principal (CBD). Para identificar os subcentros presentes na RMSP, o trabalho faz uso de uma base de dados única que inclui o georreferenciamento de quase todos os empregos formais da região, algo apenas possível a partir de um minucioso e intensivo trabalho feito com base nas informações da RAIS (Figura 1).

Figura 1. Subcentros das áreas de ponderação da RMSP. Fonte: Elaboração própria com base nos dados da RAIS 2010.



Há diferentes formas de identificação de centros e subcentros de emprego urbano (Fernández-Maldonado et al., 2014). Quatro estratégias, porém, se destacam. A primeira, e mais utilizada, é baseada na proposta apresentada por Giuliano e Small (1991) que estabelece *thresholds* ou *cut-off points* para as densidades de emprego das localidades. Uma segunda estratégia utiliza picos da densidade das ocupações ou da razão ocupações/população (McDonald e McMillen, 1990; Craig e Ng, 2001). Informações sobre o fluxo de pessoas entre as localidades também são, numa terceira alternativa, utilizados para identificar centros de emprego de acordo com a proposta de Bourne (1989). Finalmente, também são utilizadas regressões localmente ponderadas e técnicas semi-paramétricas na tarefa (McMillen, 2001).

Em face da possibilidade de comparação com trabalhos recentes para cidades latino-americanas e da ausência ainda de consenso quanto à maior propriedade dos diferentes métodos, nesta pesquisa foi utilizado à metodologia desenvolvida por McMillen (2001). Como argumentado pelo autor, esse procedimento permite estudar diferentes centros urbanos sem necessariamente ter um conhecimento prévio completo da região de estudo em análise.

Este procedimento é realizado em duas etapas. O primeiro estágio apenas identifica os potenciais subcentros porque, apesar de detectar aumentos locais na função de densidade de emprego por meio dos resíduos, ele não determina se o local tem um efeito estatisticamente significativo na forma geral da função de densidade de emprego da região. A equação estimada no primeiro estágio é uma regressão geograficamente ponderada que pode ser expressa da seguinte forma:

$$y_i = g(DCBD_i) \quad (3)$$

onde os pesos são baseados na distância geográfica entre as observações. O y_i é o logaritmo natural da densidade de emprego na grade i e $DCBD_i$, a distância de cada grade i do CBD. Para realizar a estimativa da equação 3, precisa-se definir uma função do núcleo κ_i , que determina o peso dado à observação i com base na distância geográfica. Podem ser usadas diferentes funções, contudo o estudo seguiu McMillen (2001) e usou um Kernel tricúbico⁴

Por outro lado, o segundo estágio é baseado em um procedimento semi-paramétrico para verificar a relevância de cada candidato ao subcentro encontrado no estágio anterior. Assim, estimou-se os parâmetros da seguinte regressão semi-paramétrica:

$$y_i = g(DCBD_i) + \sum_{j=1}^S (\delta_{1j}D_{ij}^{-1} + \delta_{2j}D_{ij}) + u_i \quad (4)$$

em que D_{ij} denota a distância entre a grade i e ao candidato subcentro j ; $DCBD_i$ é a distância da grade i ao CBD; S é o número de potenciais subcentros, onde $j = 0, 1, \dots, S$. Os parâmetros δ_{1j} e δ_{2j} capturam a possível influência da distância dos subcentros em potencial na função geral da densidade de emprego.

Seja x a variável de distância, que é separada em intervalos iguais e uma função cúbica que é aplicada a cada região. Defina como nós os limites entre os intervalos. Seja x_0 o valor mínimo, x_1 , x_2 e x_3 os nós e x_4 o valor máximo. Logo, a distância entre cada nó é definida como $\frac{x_4-x_0}{4}$. Neste caso, pode-se representar a função cúbica dos splines como:

$$g(DCBD_i) = \alpha + \beta_1(x_i - x_0) + \beta_2(x_i - x_0)^2 + \beta_3(x_i - x_0)^3 + \gamma_1 D_1(x_i - x_1)^3 + \gamma_2 D_2(x_i - x_2)^3 \quad (5)$$

onde D_k são variáveis dummy que são iguais a um quando $x_i > x_k$ para $k = 1, 2, 3$. Após a obtenção de $g(DCBD_i)$ da equação 4, a equação 5 é estimada por Mínimos Quadrados Ordinários. Como argumentam Rodrigues et al. (2019), para evitar problemas de multicolinearidade devido a muitos candidatos a subcentros, a lista final de subcentros é obtida a partir de um procedimento de regressão passo a passo reversa. Inicialmente, a equação 4 é estimada com todos os subcentros potenciais e a variável cujo coeficiente tem o menor valor t é eliminada. A equação reduzida é estimada novamente e essa rotina é repetida até que todas as variáveis de distância sejam positivas

⁴Deixe $d_i(x)$ definir a distância entre uma grade i e um ponto alvo x . Ordenando as observações de modo que $d_1(x) < d_2(x) < \dots < d_n(x)$, pode-se representar o Kernel tricúbico na equação, em que $I(d_i(x) < d_q(x))$ é um indicador de função para os iguais a 1. Segue equação: $\kappa_i = \left(1 - \left(\frac{d_i(x)}{d_q(x)}\right)^3\right)^3 I(d_i(x) < d_q(x))$.

e significativas no nível de 20%. Portanto, a lista final de subcentros inclui os locais com coeficientes positivos em δ_{1j} ou δ_{2j} no final do procedimento.

3.2 Descrição das variáveis de indivíduos

O banco de dados construído permitiu contemplar diferentes aspectos sobre a RMSP. As variáveis que constituem esse banco de dados e suas descrições são apresentadas na Tabela 1, a seguir:

Tabela 1. Descrição das variáveis dos modelos econométricos.

Variável	Descrição	Fonte
W	Logaritmo do salário por hora	Censo 2010
$LnAC_i$	Log do Índice de Acessibilidade (distância euclidiana)	RAIS 2010
Masculino	Variável <i>dummy</i> se for do sexo masculino	Censo 2010
Idade	Em anos	Censo 2010
Idade ²	<i>Proxy</i> para experiência	Censo 2010
Educação		
i) Baixa	i) Sem ensino fundamental completo	
ii) Intermediária	ii) Com ensino fundamental completo, mas sem médio completo	Censo 2010
iii) Alta	iii) Com ensino médio completo, mas sem superior	
iv) Elevada	iv) Com superior completo ou mais	
Raça	Branco Padro Negro Amarelo	Censo 2010
Tamanho da família	Número de pessoas por domicílio	Censo 2010
Tipo de vizinhança	Muito carente Carente Média Desenvolvida Muito Desenvolvida	Censo 2010

Fonte: Censo, RAIS 2010.

Entre as variáveis contidas no quadro, a maioria foi criada em nível de indivíduos por área de ponderação com base no Censo Demográfico 2010. Primeiramente, tem-se a variável dependente da estimativa (logaritmo do salário) referente ao salário por hora no mercado de trabalho. Em seguida, tem-se o índice de acessibilidade que foi construída a partir de uma base de dados GIS das rotas (distâncias euclidianas) dentro da RMSP. A tarefa principal foi calcular a distância dos centroides das AEP's para os centros de empregos para todas as unidades de observação no banco de dados. Assim, foi criada uma "Matriz Origem-Destino" pelo *software* QGIS, sendo que o parâmetro utilizado para construir a matriz foi o comprimento da rota (em quilômetros).

Foram consideradas nas estimativas um conjunto de características pessoais (sexo, idade, raça e escolaridade), que podem refletir diferenças nas preferências sobre a localização na cidade entre os indivíduos. Mais especificamente, as variáveis individuais absorvem as diferentes preferências com relação ao espaço urbano de acordo com certas características. Por exemplo, quanto à *dummy* para gênero, regularmente, observa-se nos trabalhos empíricos que as mulheres tendem a se localizar mais próximos aos locais de trabalho, o que é comumente atribuído à noção de maior responsabilidade para com as tarefas do lar atribuído às mulheres (Crane, 2007). As variáveis quanto à idade e escolaridade traduzem possíveis influências da disposição à locomoção e preferências quanto à vizinhança por parte dos indivíduos. O tamanho

da família capta a influência do poder aquisitivo da família e das preferências quanto a espaço da residência.

3.3 Definição das características de vizinhança

Em relação às variáveis que caracterizam a composição social da vizinhança, foi utilizada a técnica de análise fatorial para produzir indicadores compostos. A vantagem de usar a análise fatorial para produzir indicadores compostos das características da vizinhança em vez de variáveis individuais é que esses procedimentos removem o impacto potencial da colinearidade ao usar esses indicadores como variáveis preditoras nas análises de regressão, já que muitos indicadores de composição de vizinhança são altamente inter-relacionados (O'Regan e Quigley, 1998). Para contornar esse problema, foi utilizados métodos de análise fatorial padrão (Dujardin et al., 2008) para resumir essas múltiplas características em diferentes tipos de vizinhança nas AEP da RMSP.

O método de análise fatorial foi usado para identificar áreas socialmente homogêneas dentro da RMSP, que serão subsequentemente usadas nas análises de regressão. Primeiro foi executado uma análise de componentes principais que define um número limitado de fatores não correlacionados resumindo um conjunto de informações da vizinhança como percentual de mulheres chefes, nível educacional, percentual de pessoas desocupadas, percentual de domicílios sem acesso à rede de saneamento e coleta de lixo (Tabela 2).

Tabela 2. Variáveis usadas na análise de componentes principais e seus fatores.

Variáveis	Fator 1	Fator 2
Autovalores	4,410	1,313
Porcentagem de variação explicada	55,14	16,42
% de Mulheres chefes	0,362	-0,226
% de pessoas com baixo nível educacional	-0,456	0,062
% de pessoas com ensino intermediário	0,229	0,581
% de pessoas com alto nível educacional	0,428	-0,277
% de pessoas desocupadas	-0,339	0,274
% de domicílios sem acesso à rede de saneamento	0,291	0,389
% de pessoas sem acesso a coleta de lixo	0,136	0,550
% de pessoas com menos de 15 anos	-0,453	-0,001

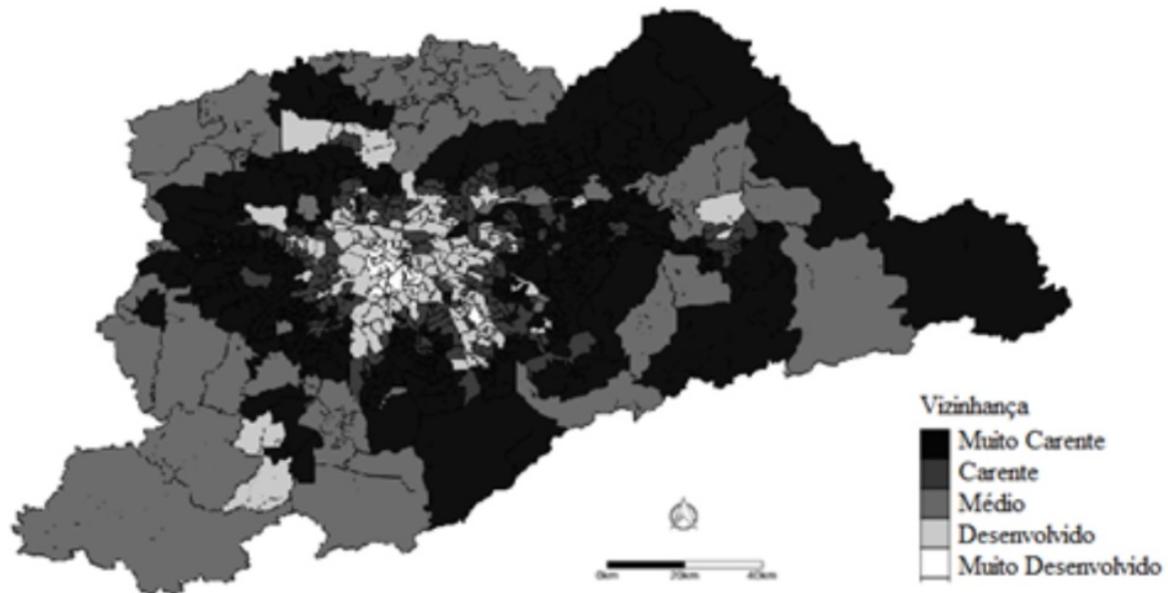
Notas: Apenas fatores com valores próprios superiores ou iguais a 1 foram selecionados.

Fonte: Elaboração própria.

Especificamente, para as variáveis que representam o nível educacional, foram considerados “baixos” os indivíduos sem ensino fundamental completo e com ensino fundamental completo, mas sem médio completo; “intermediário” indivíduos com ensino médio completo, mas sem superior; por fim, “alto” indivíduos com superior completo ou mais. Em seguida, as áreas são agrupadas de acordo com suas coordenadas nos eixos fatoriais, usando uma classificação hierárquica ascendente que minimiza a variância intragrupo (Ward Jr, 1963). Como no trabalho de Dujardin et al. (2008), foram obtidos cinco tipos de vizinhança. Este foi o número ideal de clusters de acordo com vários critérios, incluindo o critério de agrupamento cúbico. Os tipos de vizi-

nhança obtidos são apresentados na Figura 2.

Figura 2. Tipologia de vizinhança na RMSP. Fonte: Elaboração própria com base nos dados do Censo 2010.



3.4 A Endogeneidade da Localização Residencial

A relação entre os resultados individuais do mercado de trabalho à localização residencial levanta a questão da endogeneidade das escolhas de localização. É comum que indivíduos com características socioeconômicas semelhantes no mercado de trabalho tendem a se localizar em certas áreas do espaço urbano. Por exemplo, indivíduos com empregos bem remunerados escolherão residir em bairros com um bom ambiente social. Há, portanto, um problema de causalidade, por um lado, a localização residencial influencia os resultados do mercado de trabalho individual; e, por outro lado, os resultados individuais influenciam a escolha de um local residencial. Em outras palavras, é possível que características individuais que influenciam os resultados do mercado de trabalho também possam influenciar as escolhas residenciais. É claro que modelos padronizados como a equação 1 tornam possível controlar algumas características individuais e domiciliares que podem influenciar tanto a escolha da vizinhança quanto os resultados individuais. No entanto, é provável que algumas características individuais e familiares que não são observadas pelo pesquisador influenciem tanto o resultado do interesse quanto a escolha da vizinhança.

Várias estratégias foram desenvolvidas para corrigir a endogeneidade da escolha da vizinhança. Por exemplo, a existência de programas de realocação subsidiados pelo governo nos Estados Unidos, como os programas *Gautreaux* ou *Moving To Opportunity*, onde permitem obter estimativas mais confiáveis dos efeitos da vizinhança, visto que na medida em que as famílias são transferidas de uma vizinhança para outra atra-

vés de uma intervenção exógena (Oreopoulos, 2003). No entanto, devido à escassez de tais experimentos, os pesquisadores muitas vezes são obrigados a recorrer a estratégias mais questionáveis. A maioria dos estudos restringe sua amostra a jovens adultos que residem com seus pais, argumentando que as escolhas de localização foram feitas anteriormente pelos pais e podem ser consideradas exógenas ao *status* de emprego de jovens adultos (Dujardin et al., 2008). Esta também é a estratégia adotada neste artigo, ou seja, a amostra estudada é restrita a jovens participantes da força de trabalho (não estudantes) entre 18 e 25 anos que residem com os pais.

No entanto, essa abordagem não elimina completamente o viés de endogeneidade. Na verdade, ainda podem existir características parentais não observadas que determinam sua escolha residencial e também influenciam os resultados de emprego de seus filhos adultos (Glaeser et al., 1996). Por exemplo, a falta de compromisso com o trabalho pode induzir os pais a se localizarem em bairros de alta pobreza, mas também pode influenciar a motivação de seus filhos para procurar um emprego, bem como a intensidade de seu esforço de procura de emprego (Dujardin et al., 2008). Neste contexto, não é possível distinguir os efeitos de vizinhança do efeito dessas características parentais não observadas ao estimar a produtividade de jovens adultos que vivem com seus pais.

4. Resultados

A análise estatística das variáveis selecionadas do banco de dados permite a obtenção de resultados preliminares para a RMSP, conforme a Tabela 3, a seguir. Para a média da variável dependente, percebe-se que o salário hora é em média de R\$ 5,94 para os trabalhadores da RMSP. Em relação às variáveis educacionais, observa-se que o nível alto é o que apresenta maior percentual na região. A variável de composição étnica aponta que a concentração de brancos é maior, seguidas por pardos. Quanto às variáveis de vizinhança, percebe-se que a variável “Muito Carente” prevalece na região metropolitana de São Paulo.

A Tabela 3 exhibe os resultados com base na equação 1 usando índices de acessibilidade, sendo as estimativas⁵ dos regressores obtidas por OLS. Ao analisar os resultados para o CBD e subcentros de emprego, percebe-se que os coeficientes são significativos para todos os modelos estimados e mostra que os salários têm uma relação positiva com a acessibilidade para o CBD e subcentros da área metropolitana, sendo para o segundo, um efeito mais forte. Estes resultados corroboram com o trabalho de Vieira e Haddad (2015) que mostraram que uma redução generalizada no tempo de viagem corresponderia a um aumento potencial da produtividade dos trabalhadores da RMSP; Haddad e Barufi (2017) também para a RMSP mostraram ganhos nos salários com melhor acessibilidade tanto para os trabalhadores que utilizavam transporte próprio quanto público e por fim, o estudo de Duarte (2020) para Recife

⁵Em todas as regressões foram atribuídos os pesos amostrais.

Tabela 3. Variáveis usadas na análise de componentes principais e seus fatores.

Variáveis	Média	Desvio Padrão	Mínimo	Máximo
<i>W</i>	5,94	25,08	0,02	3676,5
<i>LnAC_iCBD</i>	10,60	0,52	9,17	14,13
<i>LnAC_iSUB</i>	12,66	0,44	11,17	15,62
Masculino	0,57	0,49	0	1
Idade	21,73	2,19	18	25
Idade ²	477,14	95,17	324	625
Baixa	0,1	0,31	0	1
Intermediária	0,18	0,38	0	1
Alta	0,57	0,49	0	1
Elevada	0,13	0,33	0	1
Branco	0,55	0,49	0	1
Negro	0,06	0,24	0	1
Amarelo	0,01	0,10	0	1
Pardo	0,38	0,48	0	1
Muito Carente	0,6	0,48	0	1
Carente	0,21	0,4	0	1
Médio	0,07	0,26	0	1
Desenvolvida	0,09	0,29	0	1
Muito desenvolvida	0,01	0,11	0	1
T Família	4,63	1,7	2	25

Fonte: Elaboração própria.

mostrando que o melhor acesso físico aos centros de empregos diminui a probabilidade dos trabalhadores pertencerem ao grupo de baixa renda.

Os resultados encontrados na Tabela 4 costumam serem mais significativos para áreas metropolitanas maiores. Portanto, os salários são mais baixos para os indivíduos que moram mais longe dos centros de emprego. No entanto, este resultado demonstra mais uma correlação do que um efeito causal, especialmente porque os indivíduos são analisados com referência à sua localização residencial. Embora a restrição da amostra tente reduzir a endogeneidade presente, não elimina a causalidade reversa neste caso, já que a escolha do local pelo indivíduo pode ser afetada pelo salário recebido anteriormente, e isso também pode afetar as perspectivas atuais do mercado de trabalho e a produtividade.

Os coeficientes estimados provavelmente estarão subestimando o efeito real. Portanto, se essa questão da causalidade reversa for tratada corretamente, a acessibilidade aos empregos deve ser ainda mais relevante na determinação dos níveis salariais, já que seria possível reduzir o efeito da realocação observando as oportunidades de emprego na região.

Como esperado, o efeito é positivo com o aumento do nível educacional, como mostra nas estimativas da Tabela 4. Observa-se que os coeficientes das variáveis para gênero e raça são significativos e com efeitos esperados. Ou seja, olhando especificamente para a *dummy* de sexo, pode-se dizer que além do *mismatch* espacial existente há uma possível discriminação de gênero no mercado de trabalho, visto que os salários aumentam com o indivíduo sendo do gênero masculino. O mesmo acontece com as variáveis de raça, visto que o fato de ser pardo ou negro há uma redução do salário comparado à referência branca. Para as variáveis de vizinhança, percebe-se que além

Tabela 4. Regressão do logaritmo do salário por hora (todos os indivíduos).

Variáveis	CBD	Subcentros
LnW	Coefficiente	Coefficiente
$LnAC_i$	0,093*** (0,017)	0,122*** (0,017)
Masculino	0,099*** (0,007)	0,099*** (0,007)
Idade	0,051 (0,033)	0,051 (0,033)
Idade ²	0,001 (0,000)	-0,001 (0,000)
Baixa	Ref	Ref
Intermediária	0,005 (0,016)	0,005 (0,016)
Alta	0,037*** (0,014)	0,036*** (0,014)
Elevada	0,559*** (0,020)	0,555*** (0,020)
Branco	Ref	Ref
Negro	-0,121*** (0,015)	-0,121*** (0,015)
Amarelo	0,015 (0,039)	0,017 (0,039)
Pardo	-0,107*** (0,008)	-0,107*** (0,008)
Muito Carente	Ref	Ref
Carente	0,050*** (0,014)	0,051*** (0,014)
Média	-0,001 (0,021)	0,015 (0,021)
Desenvolvida	0,162*** (0,021)	0,168*** (0,021)
Muito desenvolvida	0,354*** (0,052)	0,376*** (0,052)
Tamanho da família	-0,013*** (0,002)	-0,013*** (0,002)
Constante	-0,626 (0,400)	-1,185 (0,400)
Est. F/Wald	248,50	246,38
Coefficiente R ²	0,179	0,180
Observações	39.379	39.379

Notas:*** Significativo ao nível de 1%; ** 5% e * 10%. Foram utilizados clusterização dos erros robustos à heterocedasticidade com os devidos pesos amostrais na estimação dos modelos econométricos.

Fonte: Elaboração própria.

da significância, a magnitude do efeito aumenta com o melhor nível de vizinhança.

A relação entre acessibilidades ao emprego e resultados do mercado de trabalho mostra padrões mais distintos quando permite a heterogeneidade entre os gêneros. Na Tabela 5 têm-se efeitos heterogêneos para homens e mulheres. Os coeficientes para o CBD e subcentros novamente apresentam sinais esperados, com efeito, um pouco mais forte para os subcentros e para o gênero feminino. Note que estes novos resultados estão em linha com resultados da literatura (Duarte, 2020), ou seja, acesso é muito mais importante para as mulheres.

Na literatura, existem diversas explicações sobre por que as mulheres são mais sensíveis ao mercado de trabalho local do que homens, resumido em MacDonald

Tabela 5. Regressão do logaritmo do salário por hora (regressões para homens e mulheres).

HOMENS		
Variáveis	CBD	Subcentros
<i>LnW</i>	Coefficiente	Coefficiente
<i>LnAC_i</i>	0,088*** (0,018)	0,115*** (0,016)
Controles	Sim	Sim
Est. F/Wald	174,03	174,10
Coefficiente R ²	0,165	0,166
Observações	22.730	22.730
MULHERES		
Variáveis	CBD	Subcentros
<i>LnW</i>	Coefficiente	Coefficiente
<i>LnAC_i</i>	0,100*** (0,020)	0,132*** (0,018)
Controles	Sim	Sim
Est. F/Wald	155,22	155,70
Coefficiente R ²	0,199	0,199
Observações	16.649	16.649

Notas:*** Significativo ao nível de 1%; ** 5% e * 10%. Foram utilizados clusterização dos erros robustos à heterocedasticidade com os devidos pesos amostrais na estimação dos modelos econométricos.

Fonte: Elaboração própria.

(1999). Uma explicação é que as mulheres ganham menos do que os homens e os salários médios das mulheres variam menos no território, visto que as mulheres recebem uma compensação menos assalariada por viagens mais longas do que os homens. Outra explicação é que, na maioria dos casos, as mulheres não são chefes de famílias, portanto, buscam empregos parciais, sazonais ou temporários menos remunerados do que empregos em tempo integral.

5. Robustez

A fim de verificar a robustez dos resultados, alguns exercícios adicionais foram realizados, de forma a atenuar a endogeneidade associada ao possível *sorting* dos trabalhadores; neste sentido, estimações também foram efetivadas para amostras específicas, para as quais se acredita que o problema é bastante reduzindo. Ao menos, no que diz respeito à simultaneidade na determinação dos resultados no mercado de trabalho e grau de acesso ao emprego, espera-se que a possível endogeneidade possa ser no mínimo atenuada.

No primeiro exercício, foi utilizada uma subamostra com uma população que não possui automóvel próprio como meio de locomoção, visto que estes indivíduos teriam maiores dificuldades para percorrerem longas distâncias. Segundo, considerou apenas indivíduos que não moram de aluguel, visto que, os mesmos quando comparados com aqueles que moram, tem menor flexibilidade em relação à escolha do local de emprego, ou seja, tal situação poderia os levar a uma condição de menor influência das oportunidades de trabalho sobre a escolha da moradia. Por fim, complementou-se a análise de robustez com resultados para diferentes medidas de acessibilidade utilizando diferentes medidas para a função de distância (Tabela 6). A expectativa é

que, assim, tal expediente reduziria potenciais endogeneidades associadas à relação entre acessibilidade ao emprego e localização residencial.

Tabela 6. Robustez para a regressão do logaritmo do salário por hora.

Variáveis	Sem automóvel	Casa Própria	Função de distância $d_{ij}^{-0,5}$	Função de distância $d_{ij}^{-1,5}$
CBD				
<i>LnW</i>	Coefficiente	Coefficiente	Coefficiente	Coefficiente
<i>LnAC_i</i>	0,084*** (0,017)	0,108*** (0,016)	0,028* (0,016)	0,009* (0,005)
Controles	Sim	Sim	Sim	Sim
Est. F/Wald	41,64	218,98	239,94	239,94
Coefficiente R ²	0,059	0,190	0,177	0,177
Observações	17.417	30.753	39.379	39.379
Subcentros				
<i>LnW</i>	Coefficiente	Coefficiente	Coefficiente	Coefficiente
<i>LnAC_i</i>	0,107*** (0,020)	0,131*** (0,016)	0,122* (0,015)	0,017* (0,003)
Controles	Sim	Sim	Sim	Sim
Est. F/Wald	42,54	220,14	246,38	238,98
Coefficiente R ²	0,059	0,190	0,180	0,178
Observações	17.417	30.753	39.379	39.379

Notas:*** Significativo ao nível de 1%; ** 5% e * 10%. Foram utilizados clusterização dos erros robustos à heterocedasticidade com os devidos pesos amostrais na estimação dos modelos econométricos.

Fonte: Elaboração própria.

Comparando estes resultados com aqueles obtidos para a amostra geral, percebe-se que os resultados para os indivíduos que não possuem automóvel como meio de locomoção mantêm-se positivo e significativo. A razão fundamental para investigar o efeito do indicador de acessibilidade nos salários para indivíduos sem automóveis é o fato de que podem existir variáveis omitidas não observáveis (exemplos preferências locacionais) que afetam a escolha do modal e a localização residencial e isto, pode afetar as estimativas (introduzindo viés). Outra razão é que tais variáveis também permitem algum controle para escolha locacional.

A acessibilidade afeta o emprego das diferentes maneiras. Alguns mecanismos são mais relevantes para alguns grupos populacionais do que outros (Gobillon et al., 2011). O acesso aos empregos depende da distribuição geográfica das oportunidades de trabalho e da flexibilidade espacial dos indivíduos. A flexibilidade espacial significa a possibilidade de que o indivíduo possa se deslocar ou se mudar de casa. Em outras palavras, é a possibilidade de ajustar a localização residencial (van Ham et al., 2001). Por exemplo, se os indivíduos moram de aluguel, é mais provável que eles sejam mais espacialmente flexíveis e, conseqüentemente, são menos sensíveis ao mercado de trabalho local. Dessa forma, na Tabela 6, consideraram-se também os indivíduos que moram em casa própria. Os resultados corroboram para a robustez das evidências já obtidas para a amostra em geral.

Em uma segunda etapa, complementou-se a análise de robustez com resultados para diferentes medidas de acessibilidade utilizando diferentes medidas para a função de distância. A função de impedância pode assumir diferentes formas funcionais. Aqui, utilizaram-se duas funções de potência com diferentes parâmetros de decaimento. Foram considerados na Tabela 6, valores de -0,5 e -1,5 para os parâmetros

de decaimento, ao invés de -1,0 (valor utilizado nos exercícios do trabalho). Estes parâmetros medem a relação entre padrões de interação observados (comutação) e distância quando outros determinantes de interação são constantes. Autores como Quintanar (2012) e Carvalho (2017), por exemplo, utilizaram diferentes medidas para a função de impedância como forma de buscar robustez em seus resultados.

Percebe-se que os resultados obtidos inicialmente são mais sensíveis ao grau de decaimento da função associada à distância ao emprego, embora continuem positivos e significativos a 10%. Contudo, a magnitude dos coeficientes é menor quando considerado o grau de decaimento de -0,5 e -1,5 comparados ao grau de -1,0, mostrando que o resultado do índice de acessibilidade na variável dependente tem efeito mais fraco. O melhor acesso físico as oportunidades de empregos aumentam as chances de os indivíduos serem mais produtivos.

6. Considerações Finais

A literatura sobre *mismatch* espacial sugere que a influência da localização espacial e da distância aos empregos nos resultados do mercado de trabalho é mais forte para as grandes áreas urbanas e que é mais relevante para minorias pouco qualificadas, para as quais os custos de congestionamento são relativamente mais importantes. Além disso, essas minorias (população de baixa renda) podem enfrentar mais limitações em suas interações sociais, com um impacto significativo em sua capacidade de encontrar uma melhor correspondência no mercado de trabalho.

Nesse sentido, a pesquisa explorou correlações entre resultados do mercado de trabalho e medidas de acessibilidade a empregos. A análise empírica, apesar de ser motivada por uma estrutura teórica de equilíbrio geral, foi baseada em estratégias de modelagem de equilíbrio parcial. Os resultados da pesquisa indicam que existe um *mismatch* espacial no mercado de trabalho na RMS, ou seja, há influência da localização espacial e da distância aos empregos nos resultados do mercado de trabalho, além do mais, os salários estão fortemente relacionados à distância aos empregos e à distância para os subcentros. Observou-se que os indivíduos no qual residem em localidades mais distantes são menos produtivos.

As implicações políticas do *mismatch* espacial entre trabalhadores e oportunidades de trabalho dependem do contexto e dos mecanismos que geram essa desconexão. Dentro das políticas que podem ser implementadas, estão a facilitação da mobilidade residencial, a disseminação espacial de informações sobre empregos disponíveis.

A acessibilidade às oportunidades de emprego é um dos muitos desafios. Famílias mais pobres localizados em áreas periféricas e de difíceis acessos, geralmente têm dificuldade de se proteger e se engajar em atividades econômicas significativas. Ao nível mais agregado da cidade, a má conectividade evita o crescimento de *clusters* e a produtividade urbana. O aumento da acessibilidade não só é importante para os habitantes da cidade, mas também é essencial para estimular economias de aglomeração

e crescimento econômico a longo prazo.

Outros estudos poderiam sugerir uma melhor solução para o problema de endogeneidade de acessibilidade, através da disponibilidade de bons instrumentos e da estimativa de equações estruturais. Embora na literatura ainda não haja consenso sobre a forma adequada de modelar essas equações estruturais, bem como as decisões de localização residencial e local de trabalho.

Referências

- Bonomi Barufi, A. M., Amaral Haddad, E., et al. (2016). Spatial mismatch, wages and unemployment in brazilian metropolitan areas. Texto para discussão, Núcleo de Economia Regional e Urbana da Universidade de São Paulo (NEREUS).
- Bourne, L. S. (1989). Are new urban forms emerging? empirical tests for canadian urban areas. *Canadian Geographer/Le Géographe canadien*, 33(4):312–328.
- Carvalho, D. S. A. D. (2017). Transporte e densidade populacional na área urbana de salvador. Dissertação de Mestrado, Universidade Federal da Bahia.
- Costa, D. e Steckel, R. H. (1997). Long-term trends in health, welfare, and economic growth in the united states. In: *Health and welfare during industrialization*, Página 47–90. University of Chicago Press.
- Craig, S. G. e Ng, P. T. (2001). Using quantile smoothing splines to identify employment subcenters in a multicentric urban area. *Journal of Urban Economics*, 49(1):100–120.
- Crane, R. (2007). Is there a quiet revolution in women's travel? revisiting the gender gap in commuting. *Journal of the American planning association*, 73(3):298–316.
- Duarte, L. B. (2020). *Acessibilidade ao emprego e resultados no mercado de trabalho*. Tese de Doutorado, Universidade Federal de Pernambuco.
- Dujardin, C., Selod, H., e Thomas, I. (2008). Residential segregation and unemployment: the case of brussels. *Urban Studies*, 45(1):89–113.
- Fernández-Maldonado, A. M., Romein, A., Verkoren, O., e Parente Paula Pessoa, R. (2014). Polycentric structures in latin american metropolitan areas: Identifying employment sub-centres. *Regional Studies*, 48(12):1954–1971.
- Fujita, M. e Ogawa, H. (1982). Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional science and urban economics*, 12(2):161–196.
- Giuliano, G. e Small, K. A. (1991). Subcenters in the los angeles region. *Regional science and urban economics*, 21(2):163–182.

- Glaeser, E. L., Sacerdote, B., e Scheinkman, J. A. (1996). Crime and social interactions. *The Quarterly journal of economics*, 111(2):507–548.
- Gobillon, L., Magnac, T., e Selod, H. (2011). The effect of location on finding a job in the paris region. *Journal of Applied Econometrics*, 26(7):1079–1112.
- Gobillon, L. e Selod, H. (2021). Spatial mismatch, poverty, and vulnerable populations. *Handbook of regional science*, Página 573–588.
- Graham, D. J. e Melo, P. C. (2009). Agglomeration economies and labour productivity: evidence from longitudinal worker data for gb's travel-to-work areas.
- Haddad, E. A. e Barufi, A. M. B. (2017). From rivers to roads: Spatial mismatch and inequality of opportunity in urban labor markets of a megacity. *Habitat International*, 68:3–14.
- Hansson, E., Mattisson, K., Björk, J., Östergren, P.-O., e Jakobsson, K. (2011). Relationship between commuting and health outcomes in a cross-sectional population survey in southern sweden. *BMC public health*, 11(1):1–14.
- Houston, D. S. (2005). Methods to test the spatial mismatch hypothesis. *Economic Geography*, 81(4):407–434.
- Ihlanfeldt, K. R. (2006). A primer on spatial mismatch within urban labor markets. *A companion to urban economics*, Página 404–417.
- Kain, J. F. (1968). Housing segregation, negro employment, and metropolitan decentralization. *The quarterly journal of economics*, 82(2):175–197.
- Lucas, R. E. e Rossi-Hansberg, E. (2002). On the internal structure of cities. *Econometrica*, 70(4):1445–1476.
- MacDonald, H. I. (1999). Women's employment and commuting: explaining the links. *Journal of Planning Literature*, 13(3):267–283.
- McDonald, J. F. e McMillen, D. P. (1990). Employment subcenters and land values in a polycentric urban area: the case of chicago. *Environment and Planning A*, 22(12):1561–1574.
- McMillen, D. P. (2001). Nonparametric employment subcenter identification. *Journal of Urban economics*, 50(3):448–473.
- Mihessen, V., Machado, D. C., e Pero, V. (2014). Mobilidade urbana e mercado de trabalho na região metropolitana do rio de janeiro. *Revista da ABET*, 14(2):310–327.
- Moreno-Monroy, A. I. e Ramos, F. R. (2015). The impact of public transport expansions on informality: the case of the são paulo metropolitan region.

- Nadalín, V. G. e Iglíori, D. C. (2010). Evolução urbana e espraiamento na região metropolitana de São Paulo.
- O'Regan, K. M. e Quigley, J. M. (1998). Spatial effects upon employment outcomes: The case of New Jersey teenagers.
- Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood. *The quarterly journal of economics*, 118(4):1533–1575.
- Partridge, M. D., Rickman, D. S., Ali, K., e Olfert, M. R. (2009). Agglomeration spillovers and wage and housing cost gradients across the urban hierarchy. *Journal of International Economics*, 78(1):126–140.
- Pierrard, O. (2008). Commuters, residents and job competition. *Regional Science and Urban Economics*, 38(6):565–577.
- Proque, A. L. (2014). O efeito da distribuição espacial do uso da terra e das políticas de transporte sobre a eficiência energética em áreas urbanas. Dissertação de Mestrado, Universidade Federal da Bahia.
- Quintanar, S. (2012). Job accessibility and probability of being employed by educational level and informality in metropolitan area of Mexico City. Texto para discussão, Working Paper, Autonomous University of Barcelona.
- Ramos, F. R. (2016). Identificando subcentros de emprego na região metropolitana de São Paulo: Potenciais locais e consequências para a estrutura urbana. *Anais*, Página 1–21.
- Rodrigues, F. A. d. C., SILVEIRA NETO, R., e Miranda, F. (2019). Identification of employment subcenters in Brazilian metropolitan regions. *ENCONTRO NACIONAL DE ECONOMIA*, 47.
- Rossi-Hansberg, E. (2004). Optimal urban land use and zoning. *Review of Economic Dynamics*, 7(1):69–106.
- Stone, A. A. e Schneider, S. (2016). Commuting episodes in the United States: Their correlates with experiential wellbeing from the American Time Use Survey. *Transportation Research Part F: Traffic Psychology and Behaviour*, 42:117–124.
- Stutzer, A. e Frey, B. S. (2008). Stress that doesn't pay: The commuting paradox. *Scandinavian Journal of Economics*, 110(2):339–366.
- van Ham, M., Mulder, C. H., e Hooimeijer, P. (2001). Spatial flexibility in job mobility: macrolevel opportunities and microlevel restrictions. *Environment and Planning A*, 33(5):921–940.
- Van Ommeren, J. N. e Gutiérrez-i Puigarnau, E. (2011). Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics*, 41(1):1–8.

- Vieira, R. S. e Haddad, E. A. (2015). An accessibility index for the metropolitan region of são paulo. In: *The Rise of the City*, Página 242–258. Edward Elgar Publishing.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244.
- Wheeler, C. H. (2003). Evidence on agglomeration economies, diseconomies, and growth. *Journal of Applied Econometrics*, 18(1):79–104.
- Zenou, Y. (2002). How do firms redline workers? *Journal of urban Economics*, 52(3):391–408.

 Este artigo está licenciado com uma *CC BY 4.0 license*.