



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA – PIMES

LEANDRO BATISTA DUARTE

**ACESSIBILIDADE AO EMPREGO E RESULTADOS NO MERCADO DE
TRABALHO**

RECIFE – PE
2020

LEANDRO BATISTA DUARTE

**ACESSIBILIDADE AO EMPREGO E RESULTADOS NO MERCADO DE
TRABALHO**

*Trabalho apresentado ao Programa de Pós-graduação
em Economia–PIMES do Departamento de Economia
da Universidade Federal de Pernambuco como
requisito parcial para obtenção de grau de Doutor em
Ciências Econômicas. Área de Concentração: Teoria
Econômica.*

Orientador: Raul da Mota Silveira Neto

RECIFE – PE
2020

Catálogo na Fonte
Bibliotecária Maria Betânia de Santana da Silva, CRB4-1747

D812a	Duarte, Leandro Batista Acessibilidade ao emprego e resultados no mercado de trabalho / Leandro Batista Duarte. - 2020. 160 folhas: il. 30 cm. Orientador: Prof. Dr. Raul da Mota Silveira Neto. Tese (Doutorado em Ciências Econômicas) – Universidade Federal de Pernambuco, CCSA, 2020. Inclui referências. 1. Economia do trabalho. 2. Oportunidade de emprego. 3. Mercado de trabalho. I. Silveira Neto, Raul da Mota (Orientador). II. Título. 331 CDD (22. ed.) UFPE (CSA 2020 – 121)
-------	--

LEANDRO BATISTA DUARTE

**ACESSIBILIDADE AO EMPREGO E RESULTADOS NO MERCADO DE
TRABALHO**

*Trabalho apresentado ao Programa de Pós-graduação
em Economia–PIMES do Departamento de Economia
da Universidade Federal de Pernambuco como
requisito parcial para obtenção de grau de Doutor em
Ciências Econômicas. Área de Concentração: Teoria
Econômica.*

Aprovada em: 27/08/2020

BANCA EXAMINADORA

Prof. Dr. Raul da Mota Silveira Neto (Orientador)
Universidade Federal de Pernambuco

Prof.^a Dr.^a Tatiane Almeida de Menezes (Examinadora Interna)
Universidade Federal de Pernambuco

Prof. Dr. André Matos Magalhães (Examinador Interno)
Universidade Federal de Pernambuco

Prof.^a Dr.^a Roberta de Moraes Rocha (Examinadora Externa)
Universidade Federal de Pernambuco

Prof.^a Dr.^a Ana Barufi Bonomi Barufi (Examinadora Externa)
Universidade Federal de São Paulo

AGRADECIMENTOS

Para que essa conquista fosse alcançada, foram necessárias muitas renúncias e dedicação. Felizmente não estive sozinho nesta caminhada e gostaria de agradecer a todos os que colaboraram de alguma forma para que esse momento se concretizasse.

Primeiramente a Deus, pelos dons que me deu e por tudo o que proporciona em minha vida. Nem sei se sou tão merecedor dessas dádivas, contudo aceito de bom grado seus desígnios.

À minha família, pelos exemplos e apoio constantes. Em particular a minha esposa Nanda Belmiro que esteve ao meu lado o tempo todo. Aos meus pais, que, cada qual a seu modo, sempre se orgulharam de mim e confiaram em meu potencial. À minha sogra pela ajuda constante em momentos difíceis. Obrigado por estarem ao meu lado e acreditarem tanto em mim!

Aos professores do PIMES/UFPE, em especial ao Dr. Raul Silveira Neto, pela dedicação concedida durante a orientação, sempre disponível e disposto a ajudar. Agradeço também aos outros membros da Banca de Defesa pela disponibilidade: Prof.^a Dr.^a Tatiane Almeida de Menezes, Prof. Dr. André Matos Magalhães, Prof.^a Dr.^a Roberta de Moraes Rocha e Prof.^a Dr.^a Ana Barufi Bonomi Barufi, os quais proporcionaram discussões e sugestões que servirão para o meu crescimento pessoal e o engrandecimento da pesquisa.

À secretaria composta por Jaqueline e Maria Luíza, pelo auxílio e presteza com que me atenderam durante a realização do curso.

À CAPES, pelo auxílio financeiro, o qual possibilitou que eu me dedicasse exclusivamente à pesquisa e à escrita desta tese.

À turma de doutorado, pelos momentos divididos juntos e que tornaram mais leve meu trabalho. Agradeço de forma especial à Andrews Barros, ao qual me ajudou com a programação estatística em momentos que demonstrei desconhecimento na área.

Enfim, dedico a tese ao meu filho Luís Fernando e agradeço mais uma vez a todas as pessoas que, direta ou indiretamente, contribuíram para a execução desta tese e conclusão do Doutorado, meu muitíssimo obrigado!

RESUMO

Esta tese é constituída por dois ensaios científicos que procuraram analisar o efeito da acessibilidade ao emprego nos resultados do mercado de trabalho. Os estudos utilizaram como fundamento teórico a hipótese do *mismatch* espacial entre trabalhadores e oportunidades de trabalho, o qual identifica vários mecanismos através dos quais o acesso físico restrito aos empregos pode resultar em resultados insatisfatórios no mercado de trabalho. Este fenômeno é especialmente observado em grandes áreas urbanas, nas quais muitos trabalhadores tendem a residir longe dos empregos e enfrentam barreiras geográficas para obter resultados mais satisfatórios. Dessa forma, o primeiro ensaio teve como objetivo analisar a relação entre acessibilidade ao emprego e os resultados do mercado de trabalho na cidade do Recife. Mais especificamente, estimou uma relação empírica entre o nível de acesso aos empregos e salários. Como existem problemas de endogeneidade entre a localização residencial e a mensuração de oportunidades de trabalho, o estudo fez o uso de instrumentos históricos para a acessibilidade aos empregos. Para tanto, o estudo utiliza a pesquisa Origem-Destino de 2016 e do georreferenciamento da localização dos empregos. 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 reduz a probabilidade do trabalhador em Recife pertencer ao grupo de menor renda. No segundo ensaio, analisou-se a relação entre *mismatch* espacial e os resultados do mercado de trabalho em uma grande área metropolitana do Brasil, a RMSP. O desafio da pesquisa foi calcular o efeito da acessibilidade ao emprego na probabilidade de estar empregado, de estar desempregado, de pertencer ao setor informal e de chance de ser *overeducated*. O estudo utilizou dois bancos de dados: o Censo de População de 2010 e a Relação de Informações Anuais de Responsabilidade Social (RAIS). As equações estimadas foram com base nos modelos de probabilidade linear e *Probit*, com e sem variáveis instrumentais. Foram encontradas evidências importantes com disparidades espaciais em termos de acesso a oportunidades de emprego nas áreas de ponderação da RMSP, bem como efeitos de gênero heterogêneos na conectividade nos resultados trabalhistas. Os resultados mostraram que, em geral, o melhor acesso físico aos locais de emprego aumenta a participação no mercado, diminui o desemprego, a informalidade e a condição de *overeducated*.

Palavras-chave: *Mismatch* Espacial; Acessibilidade ao emprego; Variável Instrumental.

ABSTRACT

This thesis consists of two scientific essays that sought to analyze the effect of accessibility to employment on the results of the labor market. The studies used as a theoretical basis the hypothesis of spatial mismatch between workers and job opportunities, which identifies several mechanisms through which restricted physical access to jobs can result in unsatisfactory results in the labor market. This phenomenon is especially observed in large urban areas, where many workers tend to live far from their jobs and face geographical barriers to obtain more satisfactory results. Thus, the first essay aimed to analyze the relationship between accessibility to employment and the results of the labor market in the city of Recife. More specifically, it estimated an empirical relationship between the level of access to jobs and wages. As there are problems of endogeneity between the residential location and the measurement of job opportunities, the study made use of historical instruments for accessibility to jobs. To this end, the study uses the 2016 Source-Destination survey and the georeferencing of job locations. 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 reduces the likelihood of workers in Recife belonging to the lowest income group. In the second essay, the relationship between spatial mismatch and labor market results in a large metropolitan area in Brazil, the RMSP, was analyzed. The challenge of the research was to calculate the effect of accessibility to employment on the probability of being employed, being unemployed, belonging to the informal sector and the chance of being overeducated. The study used two databases: the 2010 Population Census and the Annual Social Responsibility Information List (RAIS). The estimated equations were based on the linear probability and Probit models, with and without instrumental variables. Important evidence was found with spatial disparities in terms of access to employment opportunities in the areas of consideration of the RMSP, as well as heterogeneous gender effects on connectivity in labor outcomes. The results showed that, in general, better physical access to places of employment increases market share, decreases unemployment, informality and being overeducated.

Keywords: Space *Mismatch*; Accessibility to employment; Instrumental variable.

LISTA DE FIGURAS

ENSAIO 1: ACESSIBILIDADE AO EMPREGO E PRODUTIVIDADE DOS TRABALHADORES NA CIDADE DO RECIFE

Figura 1.1 - Distribuição do Emprego e População para a Cidade do Recife, 2010.....	25
Figura 1.2 - Distribuição da população por faixa de renda.	25
Figura 1.3 - Centros de empregos e localização das residências dos trabalhadores no Recife.	35
Figura 1.4 - Trilhos antigos da cidade do Recife	37
Figura 1.5 - Mapa das linhas férreas do Estado de Pernambuco em 1886.....	43
Figura 1.6 - Acessibilidade ao emprego para CBD e Subcentros, 2016	45

ENSAIO 2: *MISMATCH* ESPACIAL E MERCADO DE TRABALHO: O EFEITO DA ACESSIBILIDADE AO EMPREGO SOBRE PARTICIPAÇÃO, DESEMPREGO, INFORMALIDADE E *OVEREDUCATION* NA REGIÃO METROPOLITANA DE SÃO PAULO

Figura 2.1 - Áreas de Ponderação da região metropolitana de São Paulo	79
Figura 2.2 - Subcentros das áreas de ponderação da região metropolitana de São Paulo	85
Figura 2.3 - Tipologia de vizinhança na RMSP	91
Figura 2.4 - Percorso dos rios e centros de empregos na RMSP	93
Figura 2.5 - Percentual de participação no mercado na RMSP	95
Figura 2.6 - Percentual de trabalhadores desempregados na RMSP	96
Figura 2.7 - Percentual de trabalhadores informais na RMSP	98
Figura 2.8 - Percentual de trabalhadores informais na RMSP	99

LISTA DE GRÁFICOS

ENSAIO 1: ACESSIBILIDADE AO EMPREGO E PRODUTIVIDADE DOS TRABALHADORES NA CIDADE DO RECIFE

Gráfico 1.1 - Salário dos moradores e média da distância ao CBD - Recife 26

Gráfico 1.2 - Distribuição do sistema de transporte por gênero em Recife 28

ENSAIO 2: *MISMATCH* ESPACIAL E MERCADO DE TRABALHO: O EFEITO DA ACESSIBILIDADE AO EMPREGO SOBRE PARTICIPAÇÃO, DESEMPREGO, INFORMALIDADE E *OVEREDUCATION* NA REGIÃO METROPOLITANA DE SÃO PAULO

Gráfico 2.1 - Relação entre a participação e a acessibilidade aos empregos para RMSP 95

Gráfico 2.2 - Relação entre a taxa de desemprego e a acessibilidade para RMSP 97

Gráfico 2.3 - Relação entre a taxa de informalidade e a acessibilidade para RMSP 99

Gráfico 2.4 - Relação entre a taxa de *overeducated* e a acessibilidade para RMSP 100

Gráfico 2.5 - Validação Cruzada “*leave-one-out*” para PEA - Subcentros 125

Gráfico 2.6 - Validação Cruzada “*leave-one-out*” para Desemprego - Subcentros 126

Gráfico 2.7 - Validação Cruzada “*leave-one-out*” para Informalidade - Subcentros 127

Gráfico 2.8 - Validação Cruzada “*leave-one-out*” para *overeducated* - Subcentros 127

LISTA DE QUADROS

ENSAIO 1: ACESSIBILIDADE AO EMPREGO E PRODUTIVIDADE DOS TRABALHADORES NA CIDADE DO RECIFE

Quadro 1.1 - Descrição das variáveis do modelo econométrico. 33

ENSAIO 2: *MISMATCH* ESPACIAL E MERCADO DE TRABALHO: O EFEITO DA ACESSIBILIDADE AO EMPREGO SOBRE PARTICIPAÇÃO, DESEMPREGO, INFORMALIDADE E *OVEREDUCATION* NA REGIÃO METROPOLITANA DE SÃO PAULO

Quadro 2.1 - Descrição das variáveis do modelo econométrico. 88

LISTA DE TABELAS

ENSAIO 1: ACESSIBILIDADE AO EMPREGO E PRODUTIVIDADE DOS TRABALHADORES NA CIDADE DO RECIFE

Tabela 1.1 - Distribuição dos trabalhadores entre diferentes modais por faixa de renda, cidade do Recife	27
Tabela 1.2 - Estatísticas descritivas das principais variáveis para o OD Survey, 2016.	44
Tabela 1.3 - Resultado da estimação do 1º Estágio.....	46
Tabela 1.4 - Efeitos marginais da estimação para a probabilidade de até 2 salários mínimos.	48
Tabela 1.5 - Efeitos marginais para a probabilidade de até 2 salários mínimos por modal.	49
Tabela 1.6 - Efeitos marginais para a probabilidade de até 2 salários mínimos por gênero ...	50
Tabela 1.7 - Efeitos marginais para a probabilidade de até 2 salários mínimos com a amenidades.....	52
Tabela 1.8 - Efeitos marginais para diferentes funções de decaimento da distância	52
Tabela 1.9 - Efeitos marginais para diferentes funções de decaimento da distância (sem população do CBD).....	53

ENSAIO 2: *MISMATCH* ESPACIAL E MERCADO DE TRABALHO: O EFEITO DA ACESSIBILIDADE AO EMPREGO SOBRE PARTICIPAÇÃO, DESEMPREGO, INFORMALIDADE E *OVEREDUCATION* NA REGIÃO METROPOLITANA DE SÃO PAULO

Tabela 2.1 - Variáveis usadas na análise de componentes principais e seus fatores.....	90
Tabela 2.2 - Média das características dos tipos de vizinhança	90
Tabela 2.3 - Estatísticas resumidas das variáveis incluídas no modelo de regressão	101
Tabela 2.4 - Resultado da estimação para a probabilidade de participação no mercado (regressões com todos os indivíduos) – 1º Estágio	102
Tabela 2.5 - Efeitos Marginais para a probabilidade de participação no mercado (regressões com todos os indivíduos).....	104
Tabela 2.6 - Efeitos marginais com e sem variáveis instrumentais para a probabilidade de participação no mercado (regressões para homens e mulheres)	106

Tabela 2.7 - Resultado da estimação para a probabilidade de desempregados (regressões com todos os indivíduos) – 1º Estágio	107
Tabela 2.8 - Efeitos marginais para a probabilidade de estar desempregado (regressões com todos os indivíduos).	109
Tabela 2.9 - Efeitos marginais com e sem variáveis instrumentais para a probabilidade de estar desempregado (regressões para homens e mulheres)	110
Tabela 2.10 - Resultado da estimação para a probabilidade de condição de trabalhador informal (regressões com todos os indivíduos) – 1º Estágio	111
Tabela 2.11 - Efeitos marginais para a probabilidade de trabalhadores informais (regressões com todos os indivíduos).....	112
Tabela 2.12 - Efeitos marginais com e sem variáveis instrumentais para a probabilidade de estar na informalidade (regressões para homens e mulheres)	114
Tabela 2.13 - Resultado da estimação para a probabilidade de <i>overeducated</i> (regressões com todos os indivíduos) – 1º Estágio	115
Tabela 2.14 - Efeitos marginais para a probabilidade de <i>overeducated</i> (regressões com todos os indivíduos).....	117
Tabela 2.15 - Efeitos marginais com e sem variáveis instrumentais para a probabilidade de <i>overeducated</i> (regressões para homens e mulheres)	118
Tabela 2.16 - Efeitos marginais da acessibilidade na probabilidade de participação, desempregados, na informalidade e condição de <i>overeducated</i> (regressões com jovens adultos que moram com os pais).....	120
Tabela 2.17 - Efeitos marginais da acessibilidade na participação, desemprego, informalidade e de <i>overeducated</i> (50% mais pobres da amostra)	121
Tabela 2.18 – Efeitos marginais da acessibilidade na probabilidade de participação, desempregados e informalidade (indivíduos que moram em casa própria)	122
Tabela 2.19 - Efeitos marginais da acessibilidade na probabilidade de participação, desemprego, informalidade e <i>overeducated</i> (regressões para diferentes funções de distância)	124

SUMÁRIO

INTRODUÇÃO	14
1 ACESSIBILIDADE AO EMPREGO E PRODUTIVIDADE DOS TRABALHADORES NA CIDADE DO RECIFE.....	17
1.1 Introdução.....	17
1.2 Acessibilidade ao emprego e mercado de trabalho	21
1.3 A cidade do Recife e distribuição das ocupações	23
1.4 Metodologia	28
1.4.1 Modelagem Econométrica	28
1.4.2 Banco de Dados.....	32
1.4.3 Identificação da localização dos empregos do Recife.....	34
1.4.4 Identificação econométrica com instrumentos históricos.....	36
1.4.5 Rotas e trilhos antigos do século XIX na cidade do Recife.....	41
1.5 Resultados	44
1.6 Robustez dos resultados.....	51
1.7 Considerações Finais	53
APÊNDICE	56
 2 MISMATCH ESPACIAL E MERCADO DE TRABALHO: O EFEITO DA ACESSIBILIDADE AO EMPREGO SOBRE PARTICIPAÇÃO, DESEMPREGO, INFORMALIDADE E OVEREDUCATION NA REGIÃO METROPOLITANA DE SÃO PAULO	 59
2.1 Introdução	59
2.2 <i>Mismatch</i> espacial e Mercado de Trabalho.....	65
2.3 Estudos Empíricos.....	70
2.4 Estratégia Empírica	78
2.4.1 Área de Estudo e Dados	78
2.4.2 Especificações Econométricas	80
2.4.3 Medidas de acessibilidade.....	83
2.4.4 Descrição das variáveis de indivíduos	87
2.4.5 Definição das características da vizinhança	89
2.4.6 A Endogeneidade da Localização Residencial e Variável Instrumental.....	91
2.5 A estrutura espacial da RMSP	93
2.6 Resultados	100
2.6.1 Estatísticas Descritivas.....	100
2.6.2 Participação no mercado	101
2.6.3 Desempregado.....	107
2.6.4 Informalidade	111
2.6.5 <i>Overeducation</i>	115
2.6.6 Robustez	118
2.7 Considerações Finais.....	128
APÊNDICE	131
 REFERÊNCIAS	 148

INTRODUÇÃO

Os espaços e as estruturas urbanas são vistos como locais ou polos que atraem aglomerações de indivíduos gerando oportunidades de bem estar. As decisões de localização de empresas e trabalhadores moldam a distribuição espacial da atividade econômica dentro das cidades e, assim, afetam as relações no mercado de trabalho. Esta elevada concentração de atividade econômica não somente define a distribuição das oportunidades, como também diversos problemas econômicos e socioambientais (PROQUE, 2014).

Com aproximadamente 85% de sua população vivendo em áreas urbanas em 2010, segundo o Censo Demográfico Brasileiro, o processo de urbanização no Brasil está altamente avançado. Essa grande concentração de pessoas nas áreas urbanas resulta em uma ampla gama de implicações e desafios para a sociedade, desde as possibilidades de ganho econômico das economias de aglomeração até a necessidade de planejamento urbano e a busca de soluções para as questões de mobilidade e poluição (SILVEIRA NETO & MOURA, 2019).

Os grandes centros urbanos do Brasil exibem dificuldades e necessidades que influenciam as condições de habitação, como a urbanização de áreas impróprias, adensamento urbano, congestionamentos, maior tempo gasto nos deslocamentos diários, maior consumo de energia, poluição dos rios e outros problemas ambientais em virtude do crescimento muitas vezes desordenado e sem planejamento urbano. Como consequência, muitas cidades apresentam deseconomias de aglomeração, menor eficiência energética e menor produtividade (PROQUE, 2014).

O aumento da intensidade do tráfego nas áreas urbanas, que é comumente apontado como um problema importante nas grandes cidades surge quando a infraestrutura de transporte não segue as mudanças dos indivíduos que deslocam a dinâmica. No Brasil, essa questão ganhou muita atenção recentemente, dado que a taxa de urbanização dobrou em apenas algumas décadas e não houve melhora significativa na infraestrutura de transporte (RENKOW & HOOVER, 2000; SOHN, 2005). Além disso, políticas econômicas anticíclicas mais recentes (ou seja, redução de impostos em produtos industrializados, como automóveis e motocicletas), favoreceram um aumento drástico da frota automotiva brasileira, levando a consequências para o tempo de deslocamento (OLIVEIRA *et al.*, 2015).

O deslocamento mais longo está associado ao maior custo monetário e de oportunidade incorporado no transporte. Estes custos são ainda mais altos em centros urbanos carentes de boas infraestruturas de transporte, como parece ser o caso de muitas cidades

brasileiras. O tempo adicional gasto na mudança de um lugar para outro poderia ser alocado para atividades mais produtivas (OLIVEIRA *et al.*, 2015).

Como dito, o tempo de descolamento para o trabalho é um fator considerável dos custos de transportes, variável também importante para a escolha do local de residência pelos indivíduos. Essa diferença dos custos de deslocamentos dos grupos distintos altera os padrões de segregação dos grupos, podendo causar a piora nas condições de mobilidade urbana, afetando o grau de segregação entre pobres e ricos.

Em grandes áreas urbanas, um elevado segmento da população de baixa renda enfrenta maiores custos de deslocamento, distâncias e tempos de deslocamento mais longos para a mesma distância percorrida. Esta situação resultou de uma história de *déficits* de infraestrutura de transporte, priorização de carros em relação a outros modos de transporte e, em alguns casos, falta de conscientização sobre os benefícios de uma provisão de transporte mais diversificada e equitativa (MORENO-MONROY, 2016). A partir daí, a literatura do *mismatch* espacial identificou vários mecanismos através dos quais o acesso físico restrito aos empregos pode resultar em mais desemprego (KAIN, 1968; ROGERS, 1997; IMMERGLUCK, 1998; ONG e MILLER, 2005; DURJANDIN *et al.*, 2008). Essa análise pode ser ampliada para explicar como estes mecanismos podem levar a uma menor produtividade, pior participação no mercado de trabalho e maior informalidade.

Apesar da desigualdade e da segregação espacial das cidades do Brasil, as evidências existentes com a utilização de instrumentos exógenos nos modelos econométricos para o país são limitadas em duas dimensões. Primeiro, somente se conhece resultados para a Região Metropolitana de São Paulo - RMSP (Haddad e Barufi, 2017), por exemplo, segundo, as evidências existentes são apenas para relação entre acesso ao emprego e salários. Haddad e Barufi (2017) exploram a relação entre tempo de deslocamento, acessibilidade para empregos e preços urbanos no contexto dos mercados urbanos na Região Metropolitana de São Paulo, observando como os salários individuais são afetados pelo potencial acesso às oportunidades de trabalho. Os autores utilizam o percurso dos rios como instrumento exógeno às estimativas dos salários.

Com base nesta perspectiva, a tese discorrerá sobre alguns desses problemas urbanos devido à urbanização, tomando como pano de fundo o mercado de trabalho, utilizando o instrumental teórico da Economia Urbana.

Como exemplo desse processo de urbanização e no interesse em analisar a acessibilidade ao emprego, escolheu-se estudar a cidade do Recife. A escolha geográfica deveu-se ao fato de que tal região metropolitana apresenta alto tempo de *commuting*, além de

elevada desigualdade de renda e pobreza. Além disto, o estudo conta com dados de uma pesquisa recente de mobilidade urbana realizada pela Prefeitura do Recife, Pesquisa Origem-Destino de 2016. O primeiro capítulo da tese se debruça sobre o tema da relação entre acessibilidade ao emprego e produtividade dos trabalhadores e tem como objetivo analisar como os agentes econômicos se beneficiam de tais interações dentro de uma área específica do mercado de trabalho urbano, ou seja, medir como os salários são afetados pelo potencial acesso para oportunidades de trabalho no Recife.

Os resultados da pesquisa indicam que existe um *mismatch* espacial no mercado de trabalho em Recife, 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. As evidências sugeriram que o aumento do acesso físico ao CBD e subcentros reduz a probabilidade do trabalhador em Recife pertencer ao grupo de menor renda. Observou-se também que os indivíduos que residem em localidades mais distantes, independentemente do tipo de transporte são menos produtivos.

O segundo capítulo fornece uma análise da relação entre *mismatch* espacial e os resultados do mercado de trabalho em uma grande área metropolitana do Brasil, a Região Metropolitana de São Paulo (RMSP), que é o maior centro urbano do país. Assim como muitas outras cidades, a RMSP é conhecida por problemas urbanos crônicos associados a níveis muito altos de aglomeração, incluindo congestionamento de tráfego e muitas externalidades associadas. Especificamente, o objetivo da pesquisa foi estimar o efeito da acessibilidade ao emprego na probabilidade de participar do mercado de trabalho, de estar desempregado, de pertencer ao setor informal e na chance de ser *overeducated*. O estudo, portanto, contribui para a literatura analisando o efeito da acessibilidade em diferentes variáveis dependentes, bem como o *mismatch* espacial em mercados de trabalho urbanos na região metropolitana de São Paulo.

A respeito dos principais resultados obtidos neste capítulo, pode-se dizer que a RMSP apresenta padrões importantes de *mismatch* espacial e segregação residencial que afetam a participação da força de trabalho. As oportunidades de emprego medidas pelo índice de acessibilidade são muito importantes para todos da amostra, porém mais especificamente para as mulheres. Os resultados mostraram que, em geral, o melhor acesso físico aos locais de emprego aumenta a participação no mercado, diminui o desemprego, a informalidade e a condição de *overeducated*.

1 ACESSIBILIDADE AO EMPREGO E PRODUTIVIDADE DOS TRABALHADORES NA CIDADE DO RECIFE

1.1 Introdução

Com uma população que representa 41,6% da Região Metropolitana (RM), a cidade do Recife, densamente povoada com 7.403,5 hab/km² em 2015, é a quarta maior densidade urbana do país (OLIVEIRA & SIVEIRA NETO, 2015). Recife, não muito diferente de algumas capitais do Brasil, é conhecida por problemas urbanos crônicos associados a níveis altos de aglomeração, incluindo congestionamento de tráfego e muitas externalidades associadas. Dentre todas as RMs do país, a Região metropolitana do Recife (RMR) foi a que obteve o maior crescimento do tempo de *commuting* no percurso de casa ao local de trabalho, entre os anos de 2003 e 2013 (BARBOSA, 2015). Além disso, é marcada por elevada desigualdade de renda e alto índice de pobreza. De fato, a RMR apresentou crescimento de 16,3% nesses indicadores no período entre 2011 e 2015. Os componentes de infraestrutura urbana (mobilidade), renda e trabalho foram os maiores responsáveis pelo mau desempenho da região (IPEA, 2017).

No Brasil, há um cenário que evidencia a segregação residencial em função da renda, ou seja, existe 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 (CÔRTEZ, 2008). Nas regiões centrais coexiste infraestrutura adequada com melhores condições de vida e bem-estar, enquanto, na periferia predomina condições precárias com famílias habitando em áreas impróprias. Nessas áreas, geralmente observa-se a ausência de serviços urbanos e sociais básicos, além de universidades, bibliotecas, parques e outros. Além disso, considerando o espaço urbano e os indivíduos que nele residem quanto maior a distância do local de trabalho, maiores serão os gastos de tempo e dinheiro, fazendo com que ricos escolham morar mais próximos aos centros de empregos (PROQUE, 2014). Além do mais, a locomoção da classe de renda mais elevada geralmente é feita por automóveis particulares, enquanto a de baixa renda utiliza o transporte público (PROQUE, 2014). Em particular, Oliveira e Silveira Neto (2015), em um estudo sobre mensuração a segregação residencial e a identificação da sua configuração espacial para a cidade do Recife, mostram evidências que apontam para padrões de macrosegregação na cidade.

Especificamente, os mais ricos concentram-se próximos aos centros e às amenidades, enquanto os mais pobres residem de forma mais dispersa no território.

Na literatura sobre a desconexão espacial entre trabalhadores e oportunidades de trabalho prevalece a hipótese do *mismatch* espacial, que identifica vários mecanismos¹ através dos quais o acesso físico restrito aos empregos pode levar a resultados insatisfatórios no mercado de trabalho. Segundo a teoria, trabalhadores localizados em regiões socioeconomicamente vulneráveis, enfrentam maiores dificuldades para encontrar empregos nos centros de negócios das cidades (KAIN, 1968; IHLANFELDT; SJOQUIST, 1998). Um achado consistente do ponto de vista teórico é o argumento de que os modelos de salário-eficiência urbano indicam que a produtividade e a duração do deslocamento estão relacionadas, ou seja, os trabalhadores que passam por viagens de deslocamentos mais longos fornecem níveis de esforço mais baixos do que aqueles que residem mais perto dos empregos, visto que eles poderiam se atrasar ou se cansariam devido ao deslocamento excessivo (ZENOU, 2002). Em outro estudo mais recente, Pierrard (2008) mostrou analiticamente através de um modelo simples para o caso de Luxemburgo que a decisão de comutar depende da produtividade dos trabalhadores e das perspectivas de emprego.

Há uma literatura empírica recente em economia urbana que estuda a relação entre acessibilidade ao emprego e produtividade dos trabalhadores (VAN OMMEREN & GUTIÉRREZ-I-PUIGARNAU, 2011; STONE & SCHNEIDER, 2016; VIEIRA & HADDAD, 2015; HADDAD & BARUFI, 2017). Estes estudos indicam 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 *et al.*, 1988; NOVACO & COLLIER, 1994; WENER *et al.*, 2003; KARLSTROM & ISACSSON, 2009; HANSSON *et al.*, 2011; NELEN, 2012). Van Ommeren e Gutiérrez-i-Puigarnau (2011) examinam o impacto do deslocamento sobre a produtividade dos trabalhadores na Alemanha, que se manifesta por níveis mais altos de absenteísmo para aqueles com viagens mais longas. Seu trabalho se baseia em pesquisas anteriores de Ross e Zenou (2008), que encontraram uma relação positiva entre o deslocamento, desemprego e salários usando uma Amostra de Microdados de Uso Público dos EUA do Censo Decenal de 2000. Estas conclusões são consistentes com seu modelo de salário-eficiência urbana. 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.

¹ Esses mecanismos foram abordados na seção 2.

Apesar da reduzida literatura abordando a relação entre acessibilidade e produtividade dos trabalhadores para o Brasil, ou seja, evidências de acessibilidade ao emprego determinando os salários intra-urbanos, pode-se citar o estudo de Vieira e Haddad (2015) que utilizam o método desenvolvido em Haddad *et al.* (2015) para a avaliação dos impactos econômicos do metrô de São Paulo, considerando a relação entre mobilidade, acessibilidade e produtividade, e seus efeitos de equilíbrio geral. Os autores utilizam a medida de produtividade dos trabalhadores como um indicador de desempenho econômico associado às fricções do sistema de transporte urbano. Eles mapeiam as mudanças de produtividade por local de residência e local de trabalho, e que, permite estimar os efeitos econômicos mais abrangentes (e.g. PIB e bem-estar) por meio de simulações com um modelo espacial de equilíbrio geral computável (EGC). Haddad e Barufi (2017) exploram a relação entre tempo de deslocamento, acessibilidade para empregos e preços urbanos no contexto dos mercados urbanos na Região Metropolitana de São Paulo, observando como os salários individuais são afetados pelo potencial acesso às oportunidades de trabalho.

Neste estudo, o objetivo foi analisar a relação entre acessibilidade ao emprego e os resultados do mercado de trabalho na cidade do Recife. Mais especificamente, pretende-se a partir da mais recente pesquisa Origem-Destino de 2016 e do georreferenciamento da localização dos empregos, primeiro, traçar um panorama descritivo da relação entre o grau de acessibilidade aos empregos e os resultados do mercado de trabalho. Além disto, busca-se 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 a renda do trabalhador que pode ser afetada conjuntamente, entre outras características, pela acessibilidade aos empregos na cidade.

O maior desafio da pesquisa é lidar com os principais problemas econométricos decorrentes do fato de que a localização residencial e a mensuração de oportunidades de trabalho são potencialmente endógenas. Essa endogeneidade pode surgir através da auto-seleção de trabalhadores mais ou menos produtivos para áreas específicas, pela potencial causalidade reversa de oportunidades de emprego, ou através de decisões simultâneas de localização de firmas e trabalhadores em um ambiente de equilíbrio geral. (IHLANFELDT, 2006). 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. Esta vertente da literatura propõe abordar essas questões com o uso de variáveis instrumentais. Para o presente trabalho optou-se por utilizar essa estratégia (BAUM-SNOW, 2007; MICHAELS,

2008; DURANTON E TURNER, 2012; VOLPE MARTINCUS *ET AL.*, 2013; REDDING E TURNER, 2014; HSU E ZHANG, 2014; BAUM-SNOW *ET AL.*, 2015; GARCIA-LOPEZ *ET AL.*, 2015; MAYER E TREVIEN, 2017), baseando-se na inclusão de uma variável histórica (acesso aos empregos por trilhos antigos) como instrumento para acessibilidade à empregos. A validade da estratégia adotada baseia-se na correção para possível seleção endógena.

Note-se que a disponibilidade da pesquisa OD realizada pela Prefeitura do Recife representa uma oportunidade para se conhecer as condições recentes de acesso ao emprego na cidade do Recife, visto que é a única pesquisa disponível que permite saber o local de residência e o de trabalho do empregado na RMR. Além disto, o esforço de georreferenciamento das ocupações no âmbito do NERU (Núcleo de Estudos Regionais e Urbanos) deste Departamento de Economia da UFPE gera da mesma forma, subsídios essenciais para a pesquisa, já que também não há outras informações disponíveis sobre a distribuição das ocupações no referido espaço metropolitano.

A escolha da área de estudo deve-se ao fato da inexistência de trabalhos sobre o tema para cidades do Nordeste brasileiro e, para o Brasil, apenas para a cidade de São Paulo e RM (que é bem diferente das demais cidades do Brasil). Além disso, Recife é cidade muito antiga no país e isto permite a construção de bom instrumento. Recife é o centro urbano mais desenvolvido do estado de Pernambuco, com forte fluxo de pessoas entre as áreas periféricas e o centro da cidade, tornando a questão da mobilidade urbana crucial para se discutir o mercado de trabalho e o desenvolvimento local. Ademais, a desorganização do sistema de transporte urbano e problemas presentes no cotidiano do Recife podem criar barreiras à inserção na atividade econômica dos cidadãos mais pobres. Assim, será discutido no artigo como o desajuste espacial entre locais residenciais e emprego também pode ser particularmente relevante para indivíduos mais pobres, que dependem mais do transporte público.

Mais especificamente, a principal contribuição para a literatura feita por este estudo consiste em analisar o efeito da acessibilidade do trabalho nos salários dos trabalhadores utilizando uma variável exógena. Além disto, a análise também é realizada para diferentes transportes e separada por gêneros. A importância de realizar a análise para diferentes modais está no fato de que o transporte público e privado representam formas substitutivas de deslocamento do local de residência para o local de trabalho (que podem afetar o emprego e a qualidade do emprego) e estimativas separadas para homens e mulheres, a fim de destacar as diferenças de gênero em termos do efeito da acessibilidade dos transportes públicos e privados no emprego e na incompatibilidade. Uma contribuição adicional deste estudo é que a

análise concentra-se em uma área urbana específica, o que limitaria implicitamente a importância da heterogeneidade territorial não observada.

Com esses objetivos, este capítulo se organiza da seguinte maneira. Na seção seguinte, contextualiza-se a relação entre a acessibilidade ao emprego e o mercado de trabalho. Na seção 1.3 prossegue com informações a respeito da distribuição das ocupações na cidade do Recife. Em seguida, a seção 1.4, descreve a metodologia adotada e a base de dados. Na seção 1.5 são analisadas evidências para o mercado de trabalho em Recife, seguidas pela apresentação dos testes de robustez dos resultados. Por fim, as considerações finais são apresentadas na seção 1.6 e no apêndice são apresentados outros resultados referentes a testes de robustez adicionais.

1.2 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 tempo de deslocamento de casa ao trabalho com fricções na procura por emprego, para em seguida apresentar as hipóteses do *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, MACHADO e PERO, 2016). A produtividade dos trabalhadores tende a diminuir com o maior tempo gasto nos deslocamentos intra-urbanos, 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 & FREY, 2008).

A literatura do *mismatch* espacial identifica vários mecanismos através dos quais o acesso físico restrito aos empregos pode resultar em mais desemprego (KAIN, 1968;

GOBILLON *et al.*, 2011; GOBILLON & SELOD, 2014). Principalmente em se tratando de países em desenvolvimento, essa análise pode ser ampliada para explicar como tais mecanismos podem levar a um emprego mais informal, uma vez que o acesso a empregos formais é muitas vezes difícil e oneroso para uma grande parte da população. Como resultado do *mismatch* espacial, os trabalhadores de baixa renda podem ser desencorajados a realizar deslocamentos para empregos formais, a não terem acesso a informações sobre oportunidades de emprego e enfrentam discriminação por empregadores com base no local de residência. (MORENO-MONROY, 2016).

Quando as comutações são longas, os trabalhadores podem decidir trabalhar de maneira informal em casa ou perto de casa ao invés de aceitar uma oferta para um trabalho formal mais distante. Além disso, os trabalhadores com baixos rendimentos em busca de oportunidades de trabalho formal podem ficar relutantes em viajar para áreas centrais nos quais os empregos formais são concentrados, limitando sua busca ao seu bairro local, onde o emprego informal pode ser mais prevalente. Esse mecanismo está associado aos custos de deslocamento, uma vez que quando um trabalhador recebe uma oferta para um emprego localizado longe de seu local de residência, ele prevê que terá que incorrer em custos diários de deslocamento se aceitar a oferta. Esse mecanismo é particularmente relevante para grupos que sem recursos suficientes para adquirir um carro e pagar por seu seguro e manutenção, não têm outra escolha senão confiar em transporte público (GOBILLON & SELOD, 2014).

Outro mecanismo enfatiza a falta de informação que está relacionada à separação geográfica entre oportunidades de emprego e indivíduos que residem longe desses centros, normalmente trabalhadores de baixa renda. Essa distância física, então, tem sido reconhecida como uma barreira de emprego causada, muitas vezes, pela fraca rede de informação. A segregação residencial diminui as interações com indivíduos, organizações e empregadores fora dos bairros, ou seja, o número mais baixo de pessoas pode ser uma barreira adicional às oportunidades econômicas (ONG & MILLER, 2005). Essa distância aos empregos é prejudicial para os trabalhadores, porque diminui a eficiência da procura de emprego. Ao procurar um emprego, um indivíduo pode ter pouca informação sobre quais locais têm ofertas de trabalho adequadas o que limita a probabilidade de obter melhor ocupação. Para empregos de serviços pouco qualificados, em particular, os métodos de recrutamento de empregadores continuam a ser locais, o que pode reduzir ainda mais as informações que os candidatos têm sobre ofertas de emprego distantes (GOBILLON & SELOD, 2014). Além disso, quando a maioria das pessoas em um local é prejudicada pela distância, é mais provável que muitos residentes estejam desempregados, levando a rede social local a ser de má qualidade, o que

implica que os vizinhos não podem ser usados como referências a potenciais empregadores (GOBILLON *et al.*, 2011). Ou seja, quando a taxa de desemprego é elevada em uma determinada área, é menos provável que os trabalhadores conheçam vizinhos empregados que possam informá-los sobre vagas existentes (CALVÓ-ARMENGOL & JACKSON, 2004; SELOD & ZENOU, 2006).

Os indivíduos que residem em áreas distantes de centros de emprego nos quais a habitação é mais acessível podem ter menos incentivo (esforço) para procurar ativamente por um emprego. Como as despesas com moradia são mais baixas, elas podem permanecer desempregadas por um período de tempo mais longo do que as famílias que vivem em áreas mais próximas dos empregos. Em contrapartida, os desempregados que moram em áreas nas quais os aluguéis são caros podem se sentir mais pressionados a procurar intensamente por um emprego para evitar a necessidade de se mudar (SMITH & ZENOU, 2003; PATTACHINI & ZENOU, 2006; GOBILLON & SELOD, 2014).

Por fim, os indivíduos com baixa renda que residem em áreas com pouca acessibilidade ao transporte podem ser discriminados pelos empregadores (ver Wilson, 1996, para histórias de empresas que não contratam indivíduos localizados em bairros "ruins"). Os empregadores formais podem estar preocupados com as consequências de longos deslocamentos dos trabalhadores, mas ainda são relutantes em oferecer uma compensação pelos custos desse deslocamento (MORENO-MONROY, 2016). As firmas frequentemente evitam o recrutamento nessas áreas, e os candidatos às vezes são estigmatizados por estereótipos atribuídos a seus bairros (ONG & MILLER, 2005). Para os empregadores, existe o preconceito associado à localização residencial dos candidatos, pois consideram que, em média, os indivíduos de áreas estigmatizadas têm maus hábitos de trabalho ou são mais propensos a serem criminosos (GOBILLON *et al.*, 2011). Em setores nos quais os trabalhadores estão em contato com os clientes, os empregadores podem discriminar indivíduos segregados residencialmente a fim de satisfazer os preconceitos percebidos de seus clientes (HOLZER & IHLANFELDT, 1998).

De fato, cidades, cujo transporte funciona bem, reduzem a distância entre pessoas e oportunidades econômicas (AVNER & LALL, 2016). A má conectividade entre os indivíduos e as oportunidades de emprego são, portanto, obstáculo para o crescimento econômico. Do lado da demanda, limitam o tamanho efetivo do mercado de trabalho, evitando a correspondência bem-sucedida entre trabalhadores e empresas e levando a uma má alocação de recursos produtivos e ineficiências econômicas. Do lado da oferta de mão de obra, a acessibilidade ao emprego disfuncional podem dissuadir a participação no mercado de

trabalho e do emprego formal, pois pode aumentar os salários de reserva e fricções de informação (GOBILLON & SELOD, 2014).

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 & GUTIÉRREZ-I-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 que os trabalhadores cuja residência é mais distante do local de trabalho têm maior probabilidade de estar atrasados ou cansados (GOBILLON & SELOD, 2014). Por outro lado, os trabalhadores que realizam deslocamentos mais curtos podem apresentar maiores níveis de esforço no trabalho, sendo mais produtivos (ZENOU, 2002).

1.3 A cidade do Recife e distribuição das ocupações

A partir da revisão acima a respeito dos argumentos econômicos para o *mismatch* espacial e da motivação do tema dada na introdução, é de fundamental importância elencar os elementos que compõem a distribuição das ocupações na cidade do Recife.

Com base no georreferenciamento² das ocupações levada a efeito por esta investigação (base Origem-Destino) e de informações sobre as populações residentes dos bairros a partir do Censo Demográfico de 2010, a Figura 1.1, apresenta, a seguir, o padrão espacial da distribuição das densidades do emprego (base RAIS) e da população na cidade do Recife para o ano de 2010.

Com respeito à densidade do emprego, Figura 1.1 (a), percebe-se o maior adensamento das ocupações nos bairros situados imediatamente próximos ao centro antigo da cidade (Bairro do Recife, São José e Santo Antônio), área mais escura no centro leste do mapa, e uma elevada densidade de ocupações também no bairro litorâneo de Boa Viagem, área mais escura no sudeste do mapa. Por sua vez, como se percebe a partir da Figura 1.1 (b), à exceção do bairro de Boa Viagem, as mais altas densidades populacionais são identificadas fora das regiões de mais alta densidade do emprego, embora também não muito distantes das mesmas. Ou seja, as áreas ou bairros de maior adensamento populacional da cidade situam-se

² Foi utilizado o programa *BatchGeo* para o georreferenciamento das localidades residenciais dos indivíduos com base na Pesquisa Origem-Destino 2016.

em posições intermediárias entre as áreas de maior adensamento do emprego e as localidades periféricas da cidade, à exceção de algumas localidades periféricas mais ao norte da cidade que se apresentam densamente povoadas.

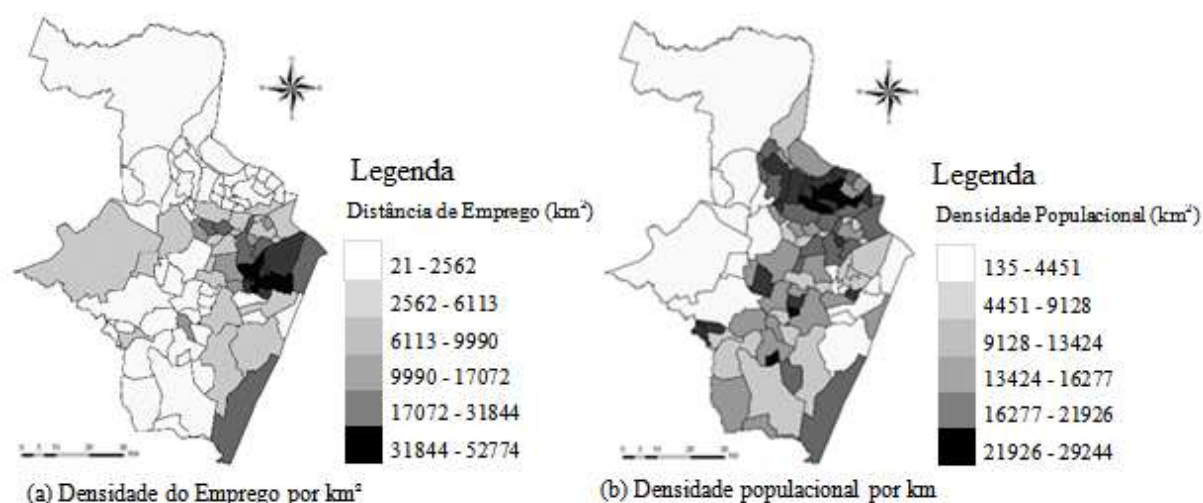


Figura 1.1 - Distribuição do Emprego e População para a Cidade do Recife, 2010.

Fonte: Elaboração própria com base nos dados da RAIS e IBGE.

Como pode ser visto no mapa da Figura 1.2, alguns bairros do Recife são considerados de alta renda, alguns de renda média e outros de renda baixa. Recife tem dois pontos bem característicos para população com renda mais elevada: o bairro de Boa Viagem e a Zona Norte da Cidade, que inclui os bairros de Casa Forte, Parnamirim, Jaqueira, Tamarineira, entre outros.

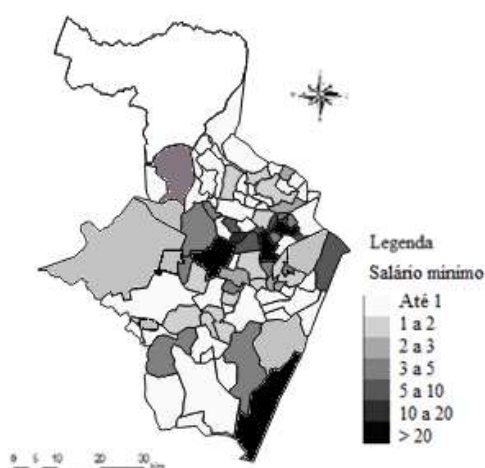
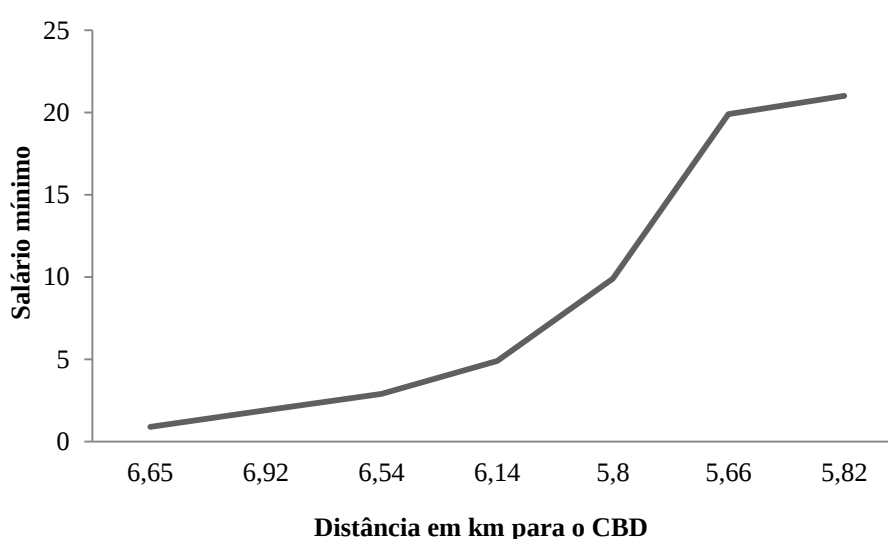


Figura 1.2 – Distribuição da população por faixa de renda.

Fonte: Elaboração própria com base nos dados da Pesquisa OD.

Na pesquisa foi utilizada a metodologia desenvolvida por McMillen (2001) para identificar os centros de empregos³. Com base nisto, o Gráfico 1.1 que apresenta a relação entre os grupos de renda e a distância ao “Central Business District” (CBD) da cidade, indica que a renda dos indivíduos é mais elevada perto do centro, ou seja, os salários são mais baixos para os indivíduos que moram mais longe do principal centro de emprego. O CBD da cidade do Recife além de apresentar maior percentual dos empregos formais é, essencialmente hoje, um centro de ocupações de serviços não comerciais, com forte atividade administrativa pública e privada, serviços de saúde, financeiros e de serviços de informação.

Gráfico 1.1 – Salário dos moradores e média da distância⁴ ao CBD - Recife



Fonte: Elaboração própria com base na Pesquisa OD, 2016.

Obviamente o resultado acima não se trata de relação causal, especialmente porque os indivíduos são analisados com referência à sua localização residencial. Assim, pode haver causalidade reversa nesse 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. Além do mais, não se pode claramente concluir que indivíduos que moram nas áreas mais distantes teriam maior tempo de *commuting*, já que essa base não analisou o indivíduo isolado, podendo o mesmo trabalhar em seu próprio bairro.

Analisando os dados da Pesquisa OD⁵ 2016 que representam os movimentos pendulares dos trabalhadores, ou seja, o fluxo diário entre o local de moradia e de trabalho das

³ O CBD e subcentros de empregos foram fornecidos por Rodrigues (2016) calculados em sua dissertação de mestrado. Na seção 1.4.3 será abordado com mais detalhes este procedimento de identificação.

⁴ Foi considerada a média da distância percorrida pelos indivíduos para o centro de emprego por faixa de renda. A distância foi calculada pelo *software* QGIS após o georreferenciamento da localização dos moradores com base nos dados da Pesquisa OD.

peças que desempenham atividades econômicas, foi possível observar informações fundamentais para a formulação do modelo empírico adotado neste estudo, pois a partir desses dados é possível mapear a estrutura da renda do trabalho, associando precisamente não somente a renda total dos moradores de cada região, mas também o local de origem da produção.

A partir da Tabela 1.1, que mostra os dados da Pesquisa de Mobilidade Urbana para a cidade do Recife em 2016, percebe-se que os resultados revelam a utilização do modo de transporte de acordo com a faixa de renda, com destaque na utilização do ônibus e do carro. A utilização do carro cresce juntamente com a renda informada, enquanto o uso do ônibus decresce de acordo com o aumento da renda. O modo a pé também foi mais representativo entre os de menor renda, sendo 11,51% dos que declararam até 1 salário mínimo vão a pé ao trabalho.

Tabela 1.1 – Distribuição dos trabalhadores entre diferentes modais por faixa de Renda, cidade do Recife

Renda	Até 1SM*	1-2 SM*	2-3 SM*	3-5 SM*	5-10 SM*	10-20 SM*	>20 SM*
Principal Modo de Transporte para o Trabalho (%)							
A pé	11,51	6,06	6,05	4,74	4,10	2,97	1,94
Bicicleta	3,64	3,60	1,97	1,35	0,90	1,11	0,39
Ônibus	74,37	67,90	40,67	15,16	5,42	2,28	0,90
Metrô	0,68	0,94	0,77	0,94	0,63	0,00	0,13
Carro	5,00	11,27	43,95	74,50	86,5	92,0	96,2
Moto	4,80	10,02	6,01	2,66	1,39	0,74	0,13
Táxi	0,00	0,03	0,23	0,53	0,63	0,64	0,13
Fretado	0,03	0,06	0,15	0,12	0,43	0,16	0,13
Distribuição do Sistema de Transporte (%)							
Público	84,84	73,44	17,02	17,02	6,38	2,37	1,05
Privado	15,16	26,56	82,98	82,98	93,62	97,63	98,95

Fonte: Elaboração própria com base em ICPS (2016).

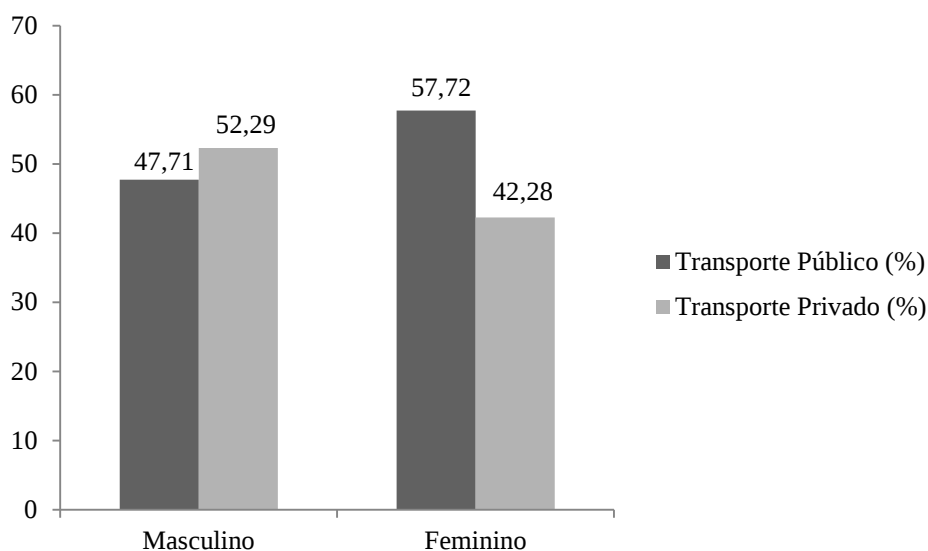
Quando analisado o modo de transporte na cidade, de acordo com a faixa de renda, a pesquisa OD indicou um claro decréscimo na utilização do transporte privado a cada diminuição de renda, da mesma forma que este decréscimo foi também constante na utilização do transporte público a cada aumento da faixa de renda das pessoas.

Nas viagens ao trabalho, o transporte público é também mais utilizado pelas mulheres do que pelos homens, enquanto 57,72% da população do sexo feminino utiliza a categoria, 47,71% do sexo masculino afirmaram utilizá-lo. O modo mais utilizado pelos homens é o

⁵ A Pesquisa de Origem e Destino (OD) é um levantamento amostral de residências realizado pela Prefeitura de Recife com o objetivo de coletar informações sobre as viagens realizadas pela população da RMR em típico dia útil.

transporte privado, 52,29%, quase 10% a mais que o percentual de mulheres que o utilizam 42,28% (Gráfico 1.2).

Gráfico 1.2 – Distribuição do sistema de transporte por gênero em Recife.



Fonte: Elaboração própria com base em ICPS (2016).

1.4 Metodologia

A metodologia da pesquisa explora a relação entre acessibilidade para empregos e resultados no mercado de trabalho. Os objetivos do trabalho exigem, assim, estratégias de identificação da localização dos empregos do Recife (para descrição e caracterização espacial dos empregos da cidade) e de consideração da endogeneidade existente na relação entre salários e acessibilidade ao emprego (para obter a correta relação entre grau de acessibilidade e salários).

1.4.1 Modelagem Econométrica

A estratégia empírica desenvolvida baseia-se na estimativa da relação entre diferentes medidas de acessibilidade e resultados do mercado de trabalho (produtividade do trabalhador). As estimativas são realizadas para indivíduos que residem em uma área específica, a fim de capturar o efeito de cada variável dentro de uma estrutura urbana específica.

Partindo-se de uma equação salarial minceriana modificada, a equação que terá os parâmetros estimados nos modelos empíricos relacionados à determinação de salários é dada pela seguinte expressão:

$$Prob[w_{ij} = 1] = F[\alpha_0 + X_{ij}\beta + DP_j + \theta AC_{ij} + \varepsilon_i] \quad (1)$$

Dada a disponibilidade da informação sobre renda apenas para categorias, será estimado o efeito da acessibilidade sobre a probabilidade do trabalhador pertencer aos menores grupos de renda. Neste caso, o salário individual (w_{ij}) será igual a 1 quando o indivíduo pertencer a faixa de renda que ganha até 2 salários mínimos (s.m.). Ressalta-se que o estudo também estimou a equação para indivíduos que ganham 1 e até 3 salários mínimos como testes de robustez. Contudo, optou-se por considerar a faixa de renda de até 2 s.m. para a variável dependente para ter mais confiança nos resultados obtidos, uma vez que, os trabalhadores selecionados possam representar maior parte da amostra sendo ao mesmo tempo considerada de baixa renda. Em seguida, o salário é explicado por um conjunto de características individuais (X_i), atributos da região de trabalho (DP_j), o índice de acessibilidade (AC_{ij}) e o termo de erro ε_i .

O salário, que é a variável de interesse do estudo depende da oferta e da demanda local de emprego. Tais salários podem influenciar o número de postos de trabalho por meio do impacto sobre os custos de produção enfrentados pelas empresas. Assim, a acessibilidade aos empregos, que é a variável explicativa de interesse na análise, depende de alguns dos fatores que afetam os salários e também podem ser indiretamente afetados. Neste caso, o viés de simultaneidade surge à medida que os residentes e os locais das empresas são determinados ao mesmo tempo (HADDAD & BARUFI, 2017).

O estudo também fez o uso de uma especificação projetada para controlar as exigências de modelos com variáveis dependentes binárias. Dessa forma, a probabilidade de se observar $w_i = 1$ é dada pelo modelo Probit (função de distribuição cumulativa da distribuição normal padrão), ou seja,

$$Prob[w_i = 1|x] = \int_{-\infty}^{\beta x} \Phi(t) d_i = \Phi(\beta' x) \quad (2)$$

onde a função $\Phi(\cdot)$ é a notação usual para a distribuição normal padrão cumulativa e β é o vetor de parâmetros das variáveis explicativas consideradas. Posteriormente, foram estimados os efeitos marginais, que serão considerados nas análises a seguir:

$$\frac{\partial E[y|x]}{\partial x} = \Phi(x'\beta)\beta \quad (3)$$

Os efeitos marginais representam, por exemplo, o quanto uma mudança na variável de acessibilidade pode afetar a probabilidade de um indivíduo pertencer ao grupo de baixa renda. Os efeitos marginais (a variação da probabilidade de o evento ocorrer quando uma variável independente é modificada) são calculados por meio dos coeficientes estimados β 's.

No que concerne à acessibilidade aos empregos, pode-se dizer que o termo “acessibilidade” em estudos de Economia Urbana expressa a relação entre a atividade econômica de uma região qualquer e a infraestrutura de transporte (CASCETTA *et al.*, 2013). Há uma relação positiva entre acessibilidade e desenvolvimento, ou seja, quanto mais acessível for uma área em relação às diversas atividades de produção, maior será sua capacidade de crescimento e sua intensidade de desenvolvimento, dessa forma, a acessibilidade também pode ser definida como a oportunidade física que o residente de algum lugar possui para ingressar em determinado conjunto de ocupações (HANSEN, 1959). Conforme Fujita e Ogawa (1982), a ideia de acessibilidade surge através do potencial locacional, uma vez que fornece uma medida de vantagem geográfica referente à produção, influenciando positivamente o preço da terra.

A acessibilidade pode ser mensurada de várias maneiras, isto é, não existe um consenso quanto à metodologia específica para mensurá-la (Ingram, 1971; Geurs e van Wee, 2004), o importante é utilizar a medida que mais determina e a adequada à forma e à função urbana. Assim, o procedimento utilizado para o cálculo do índice de acessibilidade neste estudo é baseado na formulação de Hansen (1959), que define a acessibilidade como o potencial das oportunidades de interação. Mede a distância de viagem de cada local para os vários locais de trabalho na cidade do Recife, em uma configuração geográfica.

O modelo básico sugerido para a estimação da acessibilidade em oportunidades cumulativas é composto das seguintes equações:

$$A_i = \sum_{j=1}^n \frac{E_j}{d_{ij}(t)} \quad (4)$$

Na equação (4), A_i é a acessibilidade da região i (local de residência), E_j corresponde à quantidade de empregos disponíveis em qualquer região j (centros de empregos), enquanto d_{ij} é o impedimento (distância) existente para viajar entre i e j . As oportunidades E_j foram definidas como o emprego total E em cada região j .

$$d_{ij}(t) = e^{\alpha t} \quad (5)$$

Os valores de E_j foram extraídos da base de dados do Núcleo de Estudos Regionais e Urbanos (NERU) utilizando-se das informações disponibilizadas pela RAIS. Logo, E_j corresponde aos empregos formais dos principais centros de emprego do Recife, contabilizando 5 subcentros e 1 (centro principal), conforme discutido mais adiante. Como é possível observar em (5), esta é a função de impedimento $d_{ij}(t)$, ou seja, fator de repulsão que é representado por uma função exponencial composta do parâmetro de calibração α e da distância total do percurso da origem i para o destino j .

Os maiores desafios da pesquisa dizem respeito à dificuldade em identificar uma relação causal entre o grau de acessibilidade ao emprego e o salário do trabalhador. Neste sentido, a perspectiva adotada na pesquisa é de obter a relação mais confiável possível do ponto de vista empírico, tornando os resultados difíceis de serem explicados por potenciais endogeneidades não consideradas.

De acordo com a Teoria Econômica da Economia Urbana há possíveis fontes de variáveis omitidas que afetam simultaneamente a variável dependente e o indicador de acessibilidade, como a possibilidade de causalidade reversa. Neste último, a localização residencial está em função da renda, ou seja, a população de renda mais elevada escolhe onde residir, normalmente em áreas com melhor infraestrutura que correspondem, muitas vezes, a regiões centrais, enquanto os grupos de baixa renda ocupam o entorno das aglomerações urbanas, regiões mais precárias onde o preço do terreno é menor. Embora a população de renda mais elevada geralmente resida nos centros, onde prevalecem os empregos mais formais, também é possível verificar a procura de maiores áreas de habitação em regiões com baixa densidade e menos tráfego, por esta classe de renda mais elevada. A presença de amenidades urbanas também explica a concentração dos mais ricos nessas áreas.

Estas questões fazem as estimativas *Ordinary Least Squares* (OLS) provavelmente tendenciosas, podendo ser potencialmente atenuadas com uma estratégia de variáveis instrumentais (IV). Essa estratégia pode ser especificada tomando o índice de acessibilidade como uma função linear de uma covariável exógena (instrumento) z e um componente aleatório μ_i .

$$AC_{ij} = \gamma z_{ij} + X_{ij}\gamma + DP_j + \mu_j \quad (6)$$

Para alcançar a consistência, o instrumento z deve atender a duas condições: explicar o indicador de acessibilidade (ou seja, ter poder em uma regressão do primeiro estágio com variável dependente AC_{ij}) e satisfazer a restrição de afetar o resultado exclusivamente através do indicador no segundo estágio (ANGRISTE & PISCHKE, 2008)..

Para as equações acima é utilizado o logaritmo do índice de acessibilidade nas regressões. Esta estratégia é comum em estudos empíricos da Economia Urbana, que é geralmente sensível a observações desiguais em função da grandeza dos valores dessa variável. Todavia, o período em análise é restrito para o ano de 2016 em decorrência da Pesquisa Origem-Destino fornecido pela Prefeitura do Recife. O banco de dados, com informações sobre as variáveis empregadas em regressões econométricas, é detalhado na próxima seção.

1.4.2 Banco de Dados

O banco de dados construído permite contemplar diferentes aspectos sobre o município do Recife. As variáveis que constituem esse banco de dados (microdados) são apresentadas no Quadro 1.1, a seguir. As variáveis de acessibilidade ao emprego e distâncias às amenidades foram obtidas a partir de matrizes origem-destino da Pesquisa de Mobilidade da Região de Metropolitana do Recife para o ano de 2016. Foi criada uma "Matriz de custos Origem-Destino" pelo *software* QGIS, sendo que o parâmetro utilizado para construir a Matriz de Custo OD foi o comprimento da rota (em quilômetros) considerando distâncias euclidianas e assumindo que os trabalhadores podem caminhar em ambas as direções ao longo das rotas. As demais variáveis (*dummies*), também foram criadas com base na Pesquisa OD.

Mais especificamente, as variáveis pessoais 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 (LEE & MACDONALD, 2003; CRAINE, 2007). As variáveis quanto à faixa etária e de escolaridade traduzem

possíveis influências da disposição à locomoção e preferências quanto à vizinhança por parte dos indivíduos. Já a inclusão de variáveis associadas às características de ocupação reconhece as distintas distribuições das ocupações dos ramos de atividade. Por fim, as distâncias as amenidades (praia e parques) representam as preferências dos indivíduos quanto à escolha residencial.

Quadro 1.1 – Descrição das variáveis do modelo econométrico.

Variáveis	Descrição
W ((<i>dummy</i> para o salário)	Variável categórica referente ao salário de baixa renda
LnACi	Log da Medida de acessibilidade para CBD e Subcentros
Instrumento LnACi_IV	Log da acessibilidade por trilhos para o CBD e Subcentros
Variáveis de Controle	Descrição
Escolaridade (<i>dummies</i>)	1 – Baixa (Fundamental) 2 – Intermediária (Médio) 3 – Alta (Graduação) 4 – Elevada (Pós-graduação)
Masculino (<i>dummy</i>)	Uma variável binária que assume valor 0 no caso de mulheres e 1 para homens.
Faixa Etária (<i>dummies</i>)	1 - Jovens (16 a 24 anos) 2 – Adultos (25 a 39 anos) 3 – Adultos2 (40 a 59 anos) 4 – Idosos (Acima dos 60 anos)
Filhos (<i>dummy</i>)	Uma variável binária que assume valor 0 no caso de zero filho e valor 1 para pelo menos um filho.
Setor de Atividade (<i>dummies</i>)	1 – Construção 2 - Comércio 3 – Atividade Financeira 4 – Serviços de Saúde 5 – Serviços de Educação 6 – Serviços Especializados 7 – Serviços Públicos 8 - Outros
D_Praia	Distância euclidiana da residência à praia
D_Parques	Distância euclidiana da residência aos parques

Fonte: Elaboração própria.

A cidade do Recife está situada na Região Nordeste do Brasil e possui uma população, conforme o Censo Demográfico de 2010, da ordem de 1.537.000 habitantes. A Pesquisa Origem-Destino Metropolitana 2017/2018⁶ foi realizada por meio de uma parceria entre o Instituto da Cidade Pelópidas Silveira e o Grande Recife Consórcio de Transportes Metropolitanos e abrange toda a população que reside, trabalha, estuda ou busca serviços nos municípios da Região Metropolitana do Recife. A pesquisa OD contribui com o planejamento da mobilidade na RMR, além de ajudar a mapear as necessidades de deslocamento da

⁶ A pesquisa não utilizou a base de dados mais recente por não está disponível no início do estudo.

população, embasando as atualizações das linhas de ônibus, BRT e metrô, além de subsidiar as políticas municipais para pedestres e ciclistas (RECIFE, 2017).

O desenvolvimento do Plano de Mobilidade Urbana do Recife tem como objetivo descrever os comportamentos e a situação atual da mobilidade urbana no Recife. Os levantamentos origem-destino (OD) visam reunir informações sobre viagens realizadas pela população de uma determinada área metropolitana em um dia útil típico. A análise considera as informações sobre os trabalhadores com salários (renda) e suas viagens de casa para o trabalho. Uma distinção importante é feita entre as pessoas que viajam pelo transporte público e com veículos particulares.

Os dados do estudo referem-se a 2016, com 58.645 observações obtidas de famílias sobre as viagens feitas naquele ano na RMR, sendo desse total, 41.629 trabalhadores. A amostra total para a cidade do Recife após o georreferenciamento dos endereços com base do *BatchGeo* foi de 25.687 observações, sendo que 8.493 dos indivíduos utilizaram o carro apenas como meio de transporte principal; 8.362 para ônibus e metrô, representando o transporte público e 8.832 demais modais (a pé, bicicleta, motocicleta, táxi e veículo fretado).

A Região Metropolitana do Recife foi dividida em 249 Zonas de Tráfego delimitadas na Pesquisa OD⁷ desenvolvida em 1997, considerando, prioritariamente, os setores censitários do IBGE. Esse zoneamento foi mantido para que fosse possível realizar uma comparação entre os resultados obtidos na pesquisa atual e nas pesquisas anteriores (1972 e 1997). Na cidade do Recife são localizadas 108 zonas de tráfego. A amostra necessária é calculada para cada zona de tráfego, considerando estratos de faixa etária da população de cada zona e a faixa de renda em que ela se insere (RECIFE, 2017).

1.4.3 Identificação da localização dos empregos do Recife

Os subcentros de emprego para a cidade do Recife foram extraídos da base de dados do Núcleo de Estudos Regionais e Urbanos (NERU) da UFPE, utilizando-se das informações da RAIS 2015 disponibilizadas pelo IPEA. Os subcentros correspondem aos empregos

⁷As pesquisas origem-destino costumam ocorrer em intervalos de 10 anos no formato de pesquisa domiciliar. Isso gera longos períodos em que os dados são estimados e a implantação de grandes empreendimentos ou alterações no uso do solo não são considerados nessas estimativas. Devido ao elevado custo, poucas cidades conseguem manter a periodicidade planejada. A última pesquisa desse tipo, realizada no Recife, ocorreu em 1997. Hoje, esses dados servem apenas como referência histórica. Depois dessa pesquisa, houve, por exemplo, a implantação do Porto de Suape e do Parque Automotivo da Fiat em Goiana, alterando radicalmente as características dos comportamentos pesquisados à época. Desde 2015, o Recife seguiu pioneiro no planejamento urbano, adotando uma metodologia de pesquisa integrada às funções da cidade, que se repete anualmente, abrindo um novo panorama para o desenvolvimento da mobilidade urbana (RECIFE, 2017).

formais dos principais centros de emprego do Recife, contabilizando 5 subcentros e 1 centro principal (CBD).

Para identificar os subcentros, o trabalho faz uso de uma base de dados única que inclui o geoferrenciamento de 94% dos empregos formais da região, algo apenas possível a partir de um minucioso e intensivo trabalho feito com base nas informações da RAIS. Foi utilizada a localização de cada empresa para obter informações sobre o local de trabalho. A base de dados da RAIS permite identificar as firmas aos quais os empregos estão vinculados e os setores aos quais os empregos pertencem, favorecendo informação mais ampla do centro de emprego da cidade.

Com base no procedimento de McMillen (2001), foi possível identificar o centro (CBD) e subcentros de empregos na cidade do Recife, como se nota a partir da Figura 1.3, a seguir. Como já dito anteriormente, os subcentros de empregos foram fornecidos por Rodrigues (2016) calculados em sua dissertação de mestrado. Há diferentes formas de identificação de centros e subcentros de emprego urbano (Fernandez-Maldonado *et al.*, 2015), como as metodologias adotadas por Bourne (1989); Giuliano e Small (1991); McDonald & McMillen (1990); Craig & Ng (2001). Conforme Rodrigues *et al.*, (2019), as vantagens do procedimento de McMillen (2001) é que elimina a necessidade de usar *cut-offs* para a densidade e quantidade de emprego para identificar subcentros e utiliza estimação não-paramétrica para suavizar variações locais do emprego na cidade. Já na Figura 1.3 (b), tem-se a localização da residência dos trabalhadores de acordo com os dados da Pesquisa OD. Os agregados familiares estão espalhados pelo território, principalmente localizados nos arredores do centro, com densidade populacional decadente nos limites do território da metrópole.

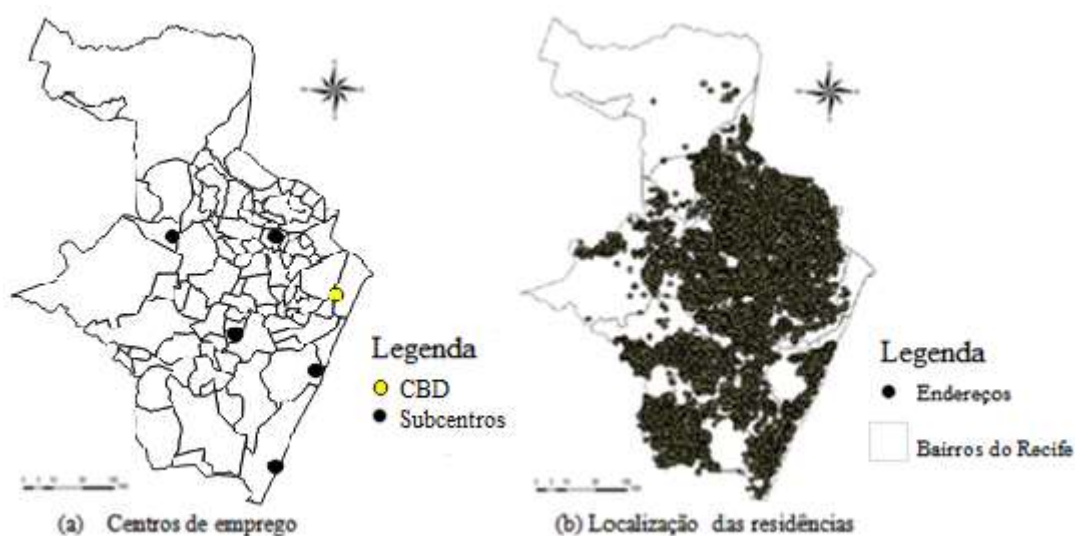


Figura 1.3 – Centros de empregos e Localização das residências dos trabalhadores no Recife.

Fonte: Elaboração própria com base nos dados da RAIS e da Pesquisa OD.

1.4.4 Identificação econométrica com instrumentos históricos

Em termos gerais, o procedimento de identificação econométrica do efeito causal pelo método de variáveis instrumentais serve para múltiplas aplicações em economia, onde o aproveitamento de estimadores de mínimos quadrados, ordinários ou generalizados, é comprometido em decorrência da correlação entre variáveis explicativas e o termo de erro. Um estimador de variáveis instrumentais tem a vantagem de fornecer estimativas consistentes e não-viesadas para o(s) parâmetro(s) de interesse, conforme Angrist e Pischke (2008) e Cameron e Trivedi (2005).

Como bem estabelecido pelos modelos de Economia Urbana (Fujita e Ogawa, 1982; Lucas e Rossi-Hansbeng, 2002; Fujita e Thisse, 2013), a determinação do local de residência e do local de trabalho (localização das firmas) ocorrem de forma simultânea: o trabalhador escolhe onde morar a partir dos possíveis salários (e aluguéis), dada uma determinada distribuição e firmas e, por sua vez, a decisão das firmas de onde produzir considera a localização das demais firmas (há ganho com a interação entre as mesmas) e os salários que terão de pagar, o que depende da localização do trabalhador (dados os custos de *commuting* envolvidos). Assim, o nível de acesso ao emprego (a localização das firmas) e o salário dos trabalhadores são determinados simultaneamente.

Além disso, é possível que a relação entre salários e acessibilidade ao emprego apresente outras fontes de endogeneidade: trabalhadores com maiores níveis de renda podem preferir locais mais amenos e afastados do emprego (Bruckener *et al.*, 2009; Melo e Grahán, 2009) ou preferir jornadas de *commuting* mais longas para obter salários mais altos.

A fim de obter uma estimativa da relação entre acessibilidade aos empregos nos salários, como adiantado acima, é necessário considerar o desafio colocado por dois problemas econométricos tradicionais: a possibilidade de causalidade reversa (*sorting*) e de simultaneidade na determinação de salários e acessibilidade. Na causalidade reversa, os trabalhadores com alta renda tendem a escolher morar próximos aos empregos, gerando uma associação espúria e fraca entre salários e acessibilidade. Além disso, os trabalhadores com mais habilidades “*skill*”, tendem a procurar locais com melhores salários e acesso a empregos, e a omissão de variáveis de controle gera associações espúrias e (agora) mais fortes entre os salários e a acessibilidade.

Três serão as estratégias utilizadas para fazer face a tais potenciais dificuldades. Primeiro, a pesquisa faz uso de um grande número de controles associados ao trabalhador e sua localização residencial, o que inclui características pessoais. Segundo, será utilizada uma variável instrumental para o indicador de acesso ao emprego, em uma estratégia similar àquela aplicada por Haddad e Barufi (2017). Em terceiro lugar, de forma a atenuar a endogeneidade associada ao possível *sorting* dos trabalhadores, estimações também serão efetivadas para amostras específicas, para os quais se acredita que o problema é bastante reduzindo. Ao menos, no que diz respeito à simultaneidade na determinação dos salários e grau de acesso ao emprego, espera-se que a endogeneidade possa ser no mínimo atenuada. De toda a forma, acredita-se que as relações estimadas representam ganho importante para o entendimento e caracterização da relação entre acesso a empregos e resultados do mercado de trabalho.

O vetor de características individuais é composto por nível de escolaridade, faixa etária e *dummy* para gênero e para filhos. Além disso, existem outros controles no modelo que incluem o setor de atividade. Diferentemente de Haddad e Barufi (2017) que buscam instrumentos em fatores geográficos⁸ que precedem à distribuição da população corrente e da própria malha viária dentro do município, a estratégia proposta na presente pesquisa de variáveis instrumentais se baseia no uso das rotas e trilhos antigos da cidade como instrumentos para o indicador de acessibilidade aos empregos, a saber, à distância física das residências aos trilhos⁹ antigos e, conseqüentemente, para os locais considerados centros de emprego da cidade do Recife. O instrumento foi construído a partir de uma base de dados GIS das rotas antigas dentro do Recife (Figura 1.4).

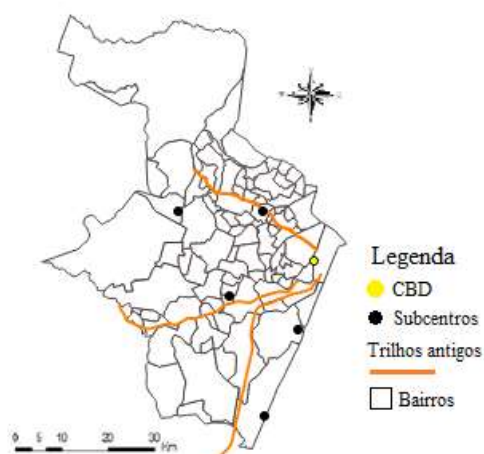


Figura 1.4 – Trilhos antigos da cidade do Recife.

Fonte: Elaboração própria com base nos dados da pesquisa.

⁸ Fazem o uso da distância física para o rio que dá acesso ao núcleo primitivo de São Paulo.

⁹ O *shapefile* das ferrovias foi obtido pelo Ministério da Infraestrutura.

Dentre os numerosos inventos que surgiram com a Revolução Industrial, o maquinário a vapor foi fundamental para os avanços tecnológicos da época. Com a produção em larga escala, houve a necessidade de transportar os produtos com rapidez, para chegar o quanto antes aos consumidores, sendo necessário facilitar o acesso entre fábrica e consumidores. Maiores atenções foram dadas às estradas para o escoamento dos produtos, contudo os meios de transportes ainda eram lentos e precários. Conforme Summerhill (2018), até 1870, apenas quatro rotas saíam do porto do Recife às localidades e estas eram de difícil navegação: trazia dificuldades para o transporte do açúcar e algodão. Além disso, existiam outras dificuldades do transporte terrestre antes das ferrovias como o alto custo das viagens em função de suas péssimas condições (estradas ruins e de baixa velocidade) limitava comércio e transporte a produtos de mais alto valor/peso, como café, algodão e açúcar. Após 1865, por exemplo, a queda do preço do algodão, dados os custos de transporte, inviabilizou o alcance do algodão brasileiro (SUMMERHILL, 2018). Então foi necessário revolucionar os transportes, utilizando o principal material daquela época do início do século XIX. Assim, foram desenvolvidas as Estradas de Ferro e as Estações Ferroviárias (Borges, 2014), mais importante inovação na mecanização da fase agrícola da produção (EISENBERG, 1977).

O estado de Pernambuco foi o primeiro do Nordeste e segundo do Brasil a possuir uma ferrovia. O primeiro trecho ferroviário foi inaugurado em 1858, a Estrada de Ferro do Recife ao São Francisco que tinha o propósito inicial de ligar a capital pernambucana à cachoeira de Paulo Afonso do rio São Francisco (FREITAS *et al.*, 2019). A partir daí surgiram outros caminhos de ferro que facilitaram muito a ligação entre o interior e o litoral do estado. Enquanto a Estrada de Ferro do Recife ao São Francisco era prolongada, atingindo Una (atual Palmares) em 1862, Catende (1882) e Garanhuns (1887), era criada na Inglaterra no ano de 1872 a *The Great Western of Brazil Railway Company* que ganhara do Império Brasileiro a concessão de explorar uma ferrovia no estado de Pernambuco, ligando Recife a Limoeiro. Em 1879 foi lançada a pedra fundamental desta ferrovia, sendo inaugurada em 1882. Neste mesmo ano, a Estrada de Ferro Central de Pernambuco, que tinha a proposta de ligar a capital pernambucana ao Agreste, iniciara suas obras de construção, inaugurando seu primeiro trecho entre a estação Central do Recife e Jaboatão em 1885 (BORGES, 2014).

As três ferrovias mais antigas ficaram organizadas da seguinte forma: a linha Sul, antiga Estrada de Ferro do Recife ao São Francisco, partindo da estação de Cinco Pontas, cruzando a Zona da Mata Sul e se ligando com o estado de Alagoas, possuindo ramais para Garanhuns, Barreiros e Cortês; na linha Norte, a Estrada de Ferro Recife-Limoeiro, partindo

da estação do Brum, no Recife, cortando a Zona da Mata Norte do estado e se ligando com o estado da Paraíba, com ramal de Carpina a Limoeiro; e por fim, a linha Centro, antiga Estrada de Ferro Central de Pernambuco, partindo da estação Central com direção ao Sertão, passando pela região central do estado (Figura 1.4).

Estas primeiras companhias ferroviárias que surgiram ligavam as regiões produtoras dos principais produtos agrícolas (açúcar e algodão) aos portos para exportação (MELO, 2000). O ciclo evolutivo da cultura da cana-de-açúcar no Brasil apresentou ao longo dos séculos uma importância crucial em face ao desenvolvimento econômico nacional. Embora o café ultrapassasse o açúcar na pauta de exportações em 1830, neste ano o volume absoluto das exportações do açúcar era 500% maior que aquele do apogeu do período açucareiro (EISENBERG, 1977). Segundo Summerhill (2018), em 1850, o açúcar representava 23% das receitas das exportações, em 1887, tal percentual era de 11%. Em ambos os anos, o açúcar era o segundo produto mais importante do país (atrás apenas do café).

O Recife do século XIX se firmava como centro financeiro e comercial da região norte brasileira. Nessa época, usufruía grande importância no país, uma vez que o seu potencial agroexportador, concentrado na cana-de-açúcar e no algodão, era reconhecidamente uma referência em termos de exportação para o mercado internacional (DUARTE, 2005). A economia canavieira começou em terrenos de várzeas dos rios (Capibaribe e Beberibe, principalmente) e só deixa estas áreas já no século XX, expulsa pelas atividades urbana (Melo, 1975), já em relação ao algodão, a região Nordeste foi pioneira na produção algodoeira no século XVIII. O *boom* do algodão nordestino ocorreu nos anos 60 do século XIX, sendo que, em meados do século até ao final da primeira metade do século XX o setor têxtil progrediu, inúmeras fabricas surgiram na Bahia, seguido do Maranhão, Alagoas, Pernambuco e Sergipe (DA COSTA OLIVEIRA, 2018).

Conforme observado por Redding e Turner (2014), a maioria dos trabalhos existentes¹⁰ que estimam o efeito das mudanças nas redes rodoviárias e nas estradas de ferro na distribuição da atividade econômica construiu esses instrumentos com base em infraestrutura planejada passada (DURANTON e TURNER, 2012; MICHAELS, 2008; BAUM-SNOW, 2007; BAUM-SNOW *et al.*, 2015; HSU e ZHANG, 2014; MAYER e TREVIEN, 2017) e mapas de rotas históricas (GARCIA-LOPEZ *et al.*, 2015; DURANTON e TURNER 2012, VOLPE MARTINCUS *et al.*, 2013; GIBBONS *et al.*, 2019).

¹⁰ Para maiores detalhes sobre o método, ver Baum-Snow (2007), Duranton e Turner (2011) e Garcia-Lopez (2012) que buscam instrumentos em projetos antigos ou velhas infraestruturas de transporte.

Baum-Snow (2007) avalia em que medida a construção de novas rodovias de acesso limitado contribuiu para o declínio populacional das cidades centrais dos EUA. Para dar conta da endogeneidade potencial da atribuição de rodovias, o autor instrumenta o número total de rodovias construídas com o número de rodovias em um Plano Interestadual de 1947. Garcia-Lopez (2012) usa instrumentos baseados em antigas infraestruturas de transporte na Espanha: estradas romanas, estradas principais e linhas ferroviárias construídas antes do final do século XIX. Duranton e Turner (2011) também utilizam instrumentos históricos, como as ferrovias do século XIX e as principais rotas de expedição do século XIX. O problema de identificação é a determinação simultânea do crescimento urbano e da infraestrutura de transporte das áreas metropolitanas dos EUA.

Para usar o instrumento, ele precisa ser exógeno e relevante. *A priori*, o instrumento parece ser exógeno devido ao período em que essas antigas infraestruturas foram construídas e as mudanças históricas significativas na sociedade e na economia (BAUM-SNOW, 2007; GARCIA-LOPEZ, 2012). Conforme Garcia-Lopez (2012) a exogeneidade de instrumentos depende do controle da geografia física e da população histórica. Para Duranton e Turner (2012), a infraestrutura de transporte atual de uma cidade depende de seu nível inicial dessa infraestrutura e a adequação de sua geografia para a construção de estradas. Conceitualmente, a relevância desses instrumentos reside no fato de que novas infraestruturas de transporte não são construídas isoladamente de infraestruturas previsíveis e antigas (GARCIA-LOPEZ, 2012).

A ideia adotada aqui, é que os planos de rede iniciais podem prever futuros desenvolvimentos de rede, mas são exógenos às mudanças nos salários. O argumento da exogeneidade baseia-se nas mudanças experimentadas no tamanho e estrutura da cidade antiga e atualmente, e que não poderiam ser previstas pelos planejadores no passado (MORENO-MONROY e RAMOS, 2015; DURANTON & TURNER, 2012). O que é necessário em particular é que as vias antigas não tenham sido projetadas para antecipar a mudança nos salários atualmente. As vias de transporte foram feitas em um contexto de forte crescimento econômico e grande planejamento, principalmente para satisfazer a requisitos imediatos de transporte de produtos coloniais. Parece plausível, que os planejadores de transporte urbano na época não pudessem prever o padrão de ocupação urbana nas décadas seguintes, nem o surgimento e a evolução de um mercado de trabalho segmentado.

A justificativa para a escolha das antigas ferrovias como instrumentos se deve aos primeiros trilhos que atuam como determinantes da localização da infraestrutura de transporte na região, de modo que a rede viária da cidade apresenta correlação espacial com antigas

ferrovias. Tais ferrovias foram criadas para ligar o porto da cidade do Recife aos municípios vizinhos ou locais produtores de açúcar e algodão, produtos mais importantes da economia local até o início do século XX, contudo a economia atual da cidade é quase que totalmente baseada no setor de serviços (sem vínculos com a produção colonial). Segundo números oficiais da Pesquisa Nacional por Amostra de Domicílios (PNAD) realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), em 2011, cerca de 80% do total de ocupações estavam relacionadas à atividade de serviços, incluindo comércio e educação, com 21% e 11% do total, respectivamente.

Em um estudo sobre os níveis de concentração das atividades econômicas, Silva *et al.*, (2019) identificaram maior concentração para o setor de serviços na Região Metropolitana do Recife. Entre as indústrias mais concentradas estavam atividades relacionadas a sistemas de informação, desenvolvimento de programas de computador, serviços de consultoria e suporte a negócios, serviços de concessão de crédito, serviços jurídicos e contabilidade. Outro setor predominante foi o comércio varejista com produtos em lojas especializadas que inclui farmácias, perfumaria e cosméticos, dispositivos de uso doméstico e pessoal, informática, instrumentos musicais, materiais de construção, livros e revistas, entre outros.

Além da justificativa acima, dadas às condições geográficas da cidade, como áreas de terras firmes circundadas por áreas de alagamentos e mangues, tais ferrovias seguiram sempre um curso determinado pela geografia mais favorável, evitando mangues e áreas alagadas. Os terrenos mais firmes possibilitaram a abertura de ruas, e, conseqüentemente, a ampliação do espaço sólido na área urbana (MELO, 1975). Sendo assim, o instrumento adotado satisfaz a condição de relevância, uma vez que, está relacionado à formação da cidade do Recife.

Conforme Duranton e Turner (2012), construir tanto ferrovias como estradas de automóveis requer nivelamento. Assim, uma antiga linha férrea provavelmente se tornará uma estrada moderna, porque as ferrovias antigas podem ser convertidas em rodovias de automóveis sem a despesa de nivelamento e classificação. Em segundo lugar, os construtores de estradas de ferro e os de estradas asfaltadas estão interessados em encontrar rotas de nível direto de um lugar para outro.

1.4.5 Rotas e trilhos antigos do século XIX na cidade do Recife.

A ferrovia contribuiu para o desenvolvimento econômico do país, sendo um dos alicerces para a integração regional e nacional, auxiliando os processos de urbanização e industrialização no Brasil, decisiva nas transformações urbanas das cidades ferroviárias e

influenciando a sociedade brasileira (MONASTIRSKY, 2006). Na segunda metade do século XIX, a Zona da Mata de Pernambuco passou por significativas mudanças nas relações sociais de produção fruto dos investimentos do capital industrial e financeiro em seu território. Dois aspectos foram relevantes neste processo: a implantação das ferrovias na região agrícola canavieira e o surgimento dos engenhos centrais e das usinas de açúcar que, pouco a pouco, substituindo os tradicionais engenhos banguês, tornaram-se os principais espaços de produção da agroindústria açucareira (FREITAS *et al.*, 2019). As primeiras companhias ferroviárias que surgiram, geralmente eram privadas e ligavam as regiões produtoras dos principais produtos agrícolas do Brasil aos portos para exportação (Figura 1.5).

A primeira estrada de ferro, no Brasil, foi inaugurada em 1854 pelo Imperador D. Pedro II e apenas um ano depois, no dia 8 de setembro de 1855, foi inaugurada a Estrada de Ferro Recife ao São Francisco, a segunda estrada férrea brasileira e a primeira no Nordeste e em Pernambuco. Sua expansão, em terras pernambucanas, durou praticamente um século, avançou em direção ao interior e alcançou o sertão pernambucano, totalizando 1322 km de extensão¹¹. A ideia inicial do projeto era sair do Recife, atravessar o leito do Sirinhaém, passando pelas povoações de Água Preta e Garanhuns e iria terminar em um dos pontos de extensa navegação do rio São Francisco (BORGES, 2014).

O primeiro trecho ferroviário foi inaugurado em 1858, mais precisamente a 8 de Fevereiro de 1858, entre as estações de Cinco Pontas (no Recife) e a estação ferroviária da Vila do Cabo (esta até hoje existente e em uso desde sua inauguração). Era este o trecho inicial da Estrada de Ferro do Recife ao São Francisco que tinha o propósito inicial de ligar a capital pernambucana à cachoeira de Paulo Afonso do rio São Francisco.

O desenvolvimento das ferrovias ia favorecendo os núcleos no interior pernambucano, visto que a concorrência terrestre ainda era muito acirrada com o tráfego naval. Muitos produtos vinham através de barcos para o armazém do porto de Recife. Quando a linha alcançava o Una (atual Palmares), os donos de terras cediam à ferrovia. Alguns números refletem o desenvolvimento e o impacto que a ferrovia trouxe para algumas regiões. Assim que Escada (distrito açucareiro mais rico de 1860) recebeu caminhos de ferro, o tráfego de mercadorias cresceu em 200% (MELO, 2000).

Duas outras linhas foram construídas em Pernambuco, sempre tendo como ponto de partida a cidade do Recife, se desenvolvendo para o interior e outros estados do Nordeste – a Estrada de Ferro do Recife a Limoeiro (iniciada em 1879) e a Estrada de Ferro Central de

¹¹ Inventário do Patrimônio Ferroviário em Pernambuco, pág. 11, v. 1. Recife: Iphan – PE, 2009.

Pernambuco (iniciada em 1881) –, além de vários ramais como o Ramal Garanhuns (inaugurado em 1887), o Ramal Cortês (provavelmente iniciado em 1890, sendo adquirido pela Great Western em 1907), o Ramal Barreiros (iniciado em 1912) e o Ramal Nazaré (inaugurado em 1882). Estas construções foram financiadas pelos proprietários dos engenhos centrais e das usinas de açúcar para o transporte da cana, da plantação para a fábrica, e do açúcar, da fábrica para as linhas, ramais e estações ferroviárias (FREITAS *et al.*, 2019).



Figura 1.5 – Mapa das linhas férreas do Estado de Pernambuco em 1886.
Fonte: Apresentado por Rodrigo Silva em seu "plano" de 1886.

Dessa forma, a condição para o uso do nosso instrumento exige que o mesmo afete os salários na cidade apenas por meio de seu efeito sobre a acessibilidade dos trabalhadores aos locais de emprego. As rotas e estradas do século XIX foram primeiramente desenhadas para satisfazer o transporte do açúcar e do algodão, não prevendo influenciar principalmente nos transportes de passageiros aos locais de trabalho. Quase totalidade da população¹² do Recife era concentrada nos bairros do Recife, Santo Antonio, São José, Boa Vista e Santo Amaro no período da construção das ferrovias e até final do século XIX, ou seja, não havia motivação

¹² Até 1844, ocupação urbana somente nos bairros do Recife, Santo Antônio, São José e parte da Boa Vista, ainda não existia Santo Amaro. A ocupação plena da Boa Vista somente no final do século XIX e início do século XX (MELO, 1975).

econômica para construir ferrovia para passageiros inicialmente (MELO, 1975). Dessa forma, as linhas férreas utilizadas como instrumento foram às três mais antigas do Recife: a Estrada de Ferro do Recife ao São Francisco, a Estrada de Ferro do Recife a Limoeiro e a Estrada de Ferro Central de Pernambuco.

A próxima seção discute os principais resultados das diferentes estimativas. O foco está na relevância do instrumento, à distância pelos trilhos antigos do Recife aos principais centros de empregos da cidade, como uma medida exógena da centralidade que moldou a infraestrutura urbana e os mercados urbanos.

1.5 Resultados

A análise estatística das variáveis selecionadas do banco de dados permite a obtenção de resultados preliminares. A tabela 1.2, a seguir, apresenta estatísticas descritivas das variáveis utilizadas nas estimações.

Tabela 1.2 - Estatísticas descritivas das principais variáveis para o OD Survey, 2016.

Variáveis	Média	Desvio Padrão	Mínimo	Máximo
W (até 2 s.m.)	0,44	0,49	0	1
LnACi_CBD	11,15	0,56	10,19	16,48
LnACi_Sub	11,72	0,42	10,69	14,41
LnACi_CBD_IV	10,58	0,55	9,61	12,47
LnACi_Sub_IV	10,36	1,10	7,21	14,40
Masculino	0,50	0,49	0	1
Filho	0,35	0,47	0	1
Baixa	0,83	0,37	0	1
Intermediária	0,15	0,12	0	1
Alta	0,11	0,31	0	1
Bem elevada	0,04	0,19	0	1
Jovens	0,05	0,22	0	1
Adultos	0,29	0,45	0	1
Adultos2	0,31	0,46	0	1
Idosos	0,20	0,40	0	1
Construção	0,04	0,18	0	1
Comércio	0,05	0,21	0	1
Ativ. Financeira	0,09	0,28	0	1
Serv. Saúde	0,12	0,32	0	1
Sev. Educação	0,05	0,22	0	1
Serv. Espec.	0,08	0,28	0	1
Serv. Público	0,05	0,21	0	1
Outros	0,51	0,49	0	1

Fonte: Elaboração própria com base na Pesquisa OD, 2016.

A princípio, é possível observar que foram atribuídos logaritmos para variáveis de acessibilidade ao emprego. Esta estratégia é comum em estudos empíricos (Quintanar, 2013; Haddad e Barufi, 2017; Carvalho, 2016; Boisjoly *et al.*, 2017) da Economia Urbana para

estimar a acessibilidade por oportunidades cumulativas, que é geralmente sensível a observações desiguais em função da grandeza dos valores dessa variável. Em particular, as variáveis *dummies* apresentam valores que descrevem a cidade do Recife. Em relação às variáveis educacionais, observa-se que o nível mais baixo é o que apresenta maior percentual do total dos trabalhadores na região. As *dummies* de faixa etária apontam que há concentração de adultos na faixa de 29 a 59 anos. Quanto às variáveis de setores econômicos, percebe-se certo equilíbrio na divisão, com maior percentual para os serviços de saúde e outros.

As figuras 1.6 mostram a distribuição espacial de dois índices de acessibilidade calculados para o CBD e subcentros. O “*Central Business District*” (CBD) considerado corresponde a maior concentração do total dos empregos da cidade, responsável pelas atividades de comércio, indústria e serviços públicos, já os subcentros foram potenciais centros de empregos, apresentando maior densidade de empregos e influencia local quanto mais próximo ao CBD e menor densidade de emprego e influencia local nas regiões mais periféricas. Ambos os mapas apresentam cores de acordo com os bairros do Recife; quanto mais escuras forem as cores, maior será a acessibilidade do grupo. Percebe-se que quanto mais próximo dos centros de empregos, maior é o índice de acessibilidade para os indivíduos.

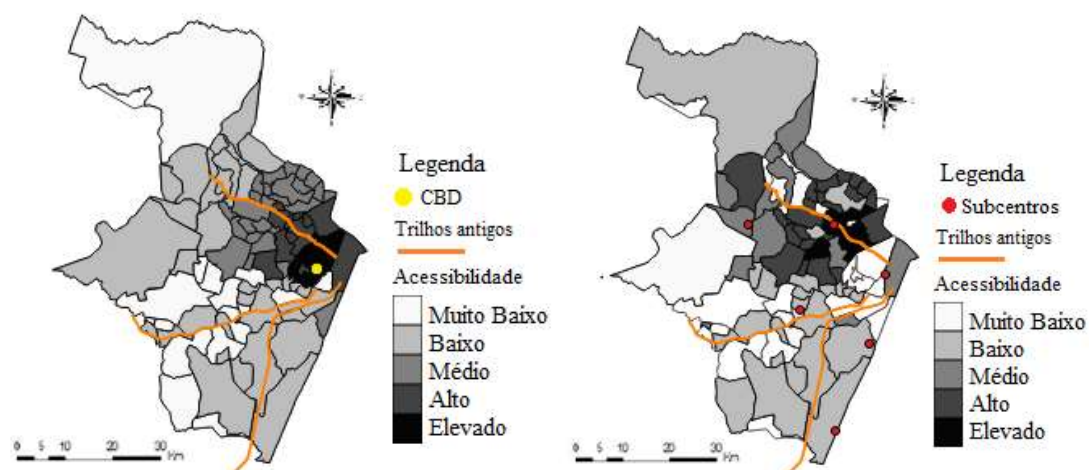


Figura 1.6 – Acessibilidade ao emprego para CBD e Subcentros, 2016.

Fonte: Elaboração própria com base nos dados da pesquisa.

A estratégia empírica discutida anteriormente exige a estimativa de diferentes modelos para avaliar a relevância da acessibilidade aos empregos para explicar os salários. No entanto, a acessibilidade é potencialmente endógena, o que implica a necessidade de se usar um instrumento exógeno para evitar ou atenuar a presença de estimativas enviesadas e inconsistentes que resultariam da aplicação do estimador de Mínimos Quadrados Ordinários.

Nessa perspectiva, como já dito anteriormente, foi assumida a exogeneidade do seguinte instrumento histórico: acesso físico aos centros de empregos utilizando os trilhos antigos como rotas. A regressão do primeiro estágio, entre instrumentos e variáveis potencialmente endógenas é evidenciada estatisticamente na Tabela 1.3, a seguir.

Tabela 1.3 - Resultado da estimação do 1º Estágio.

Variáveis	CBD		Subcentros	
	LnACi	MPL_IV	D.P.	MPL_IV
LnACi_IV				
Masculino	0,617***	0,005	0,132***	0,002
Filho	-0,011**	0,005	-0,022***	0,005
Baixa	-0,020***	0,005	0,029***	0,005
Intermediária	Ref.	Ref.	Ref.	Ref.
Alta	0,097***	0,022	0,056***	0,020
Bem elevada	0,019**	0,008	0,010	0,008
Jovens	0,078***	(0,014)	0,077***	0,014
Adultos	Ref.	Ref.	Ref.	Ref.
Adultos2	-0,081***	0,008	-0,199***	0,008
Idosos	-0,055***	0,008	-0,138***	0,008
Construção	0,042***	0,009	0,038***	0,009
Comércio	Ref.	Ref.	Ref.	Ref.
Ativ. Finan.	-0,042**	0,019	-0,108**	0,017
Serv.Saúde	0,040**	0,017	-0,007	0,015
Sev. Educaçã	-0,009	0,016	-0,100***	0,015
Serv.Espec.	0,020	0,018	-0,016	0,017
Serv.Público	0,021	0,017	-0,006	0,015
Outros	0,021	0,019	0,001	0,017
Constante	-0,031**	0,015	-0,138**	0,013
	4,665***	0,057	10,533***	0,033
R ² / Pseudo R ²			0,105	
Estat. F			2060,5***	
Hausman (Chi)			0,080	
Durbin (score, Chi)			0,777	
Wu-Hausman (F)			0,777	

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Inicialmente, testes foram executados para capturar a relevância dos instrumentos, qualificando-os como fortes ou fracos. No primeiro momento foi estimada a regressão simples¹³ da acessibilidade e, em seguida, foi integrado à equação o conjunto das seguintes covariadas. Em geral, a inclusão de novas variáveis em regressões múltiplas eleva o grau de ajuste da modelagem, mas pouco afeta o coeficiente da variável instrumental, que continua consideravelmente forte. Dessa forma, é possível empregar qualquer combinação de instrumento e variável explicativa de acessibilidade.

O instrumento utilizado, tanto nas estimativas para o CBD quanto para os subcentros, apresenta o coeficiente de interesse estatisticamente significativo com probabilidade de rejeição em 1% para cada tipo de regressão: com e sem covariadas. Observando-se as

¹³ Não foram apresentadas no trabalho.

estimativas do primeiro estágio, nota-se que, embora o coeficiente de determinação da regressão (R^2) não seja muito elevado, a estimativa para o coeficiente apresenta valor elevado e com sinal esperado. A estatística F para o primeiro estágio é estatisticamente significativa e assume um valor que facilmente excede 10, o valor de corte sugerido por Stock, Wright e Yogo (2002). As estatísticas Durbin e Wu-Hausman indicaram apenas para as estimativas do CBD a rejeição da hipótese nula de exogeneidade da variável representada. Assim, o estimador IV é preferível ao estimador¹⁴ OLS. O teste de Durbin-Wu-Hausman indica que os resíduos de uma regressão do índice de acessibilidade sobre as demais variáveis são estatisticamente significantes quando colocados como regressores de uma regressão dos salários sobre todas as variáveis ditas explicativas. Além disso, os resultados da primeira etapa fornecem confiança para o uso do instrumento proposto; utilizando a variável de acessibilidade por trilhos como regressor para o índice de acessibilidade, obteve-se uma influência positiva e estatisticamente significativa. De fato, o conjunto de estatísticas apresentadas para avaliar o instrumento fornece forte confiança para usá-lo.

A Tabela 1.4 exibe os resultados com base nas equações (1) e (2) usando índices de acessibilidade. As colunas (1)-(4) exibem os resultados para o CBD, enquanto as colunas (5)-(8) são estimativas para subcentros. A tabela considera a média dos efeitos marginais para a probabilidade dos trabalhadores pertencerem ao grupo de baixa renda. As estimativas dos regressores dos salários (até 2 salários mínimos) são obtidas a partir de modelos de Probabilidade Linear e Probit (efeitos marginais) com (colunas (3)-(4) e (7)-(8)) e sem (colunas (1)-(2) e (5)-(6)) variáveis instrumentais. O mesmo exercício é realizado considerando a variável dependente até 1 e 3 salários mínimos. Estes resultados encontram-se na Tabela A1.1 no apêndice do trabalho. No primeiro momento, percebe-se que a estatística F é alta para qualquer modelo empregado, o que garante a significância estatística conjunta das variáveis explicativas. Essas variáveis explicam de 19,2% a 24,5% da variação referente à variável dependente, conforme o coeficiente de determinação R^2 .

De acordo com as estimativas, a principal variável de interesse neste estudo, acessibilidade ao emprego, tem um impacto significativo na probabilidade do indivíduo pertencer ao grupo de renda que recebe até 2 salários mínimos. Tanto para o CBD quanto para os subcentros, um maior acesso aos empregos diminui a chance de o indivíduo ser de baixa renda. Percebe-se que os coeficientes são estáveis, tendo o índice de acessibilidade significativo a 1% e com efeito esperado em todas as estimativas.

¹⁴ Ambos os estimadores são tendenciosos nessa situação, mas somente IV pode gerar estimativas consistentes.

Tabela 1.4 – Efeitos Marginais para a probabilidade de até 2 salários mínimos.

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
W (Até 2 s.m)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	-0,090*** (0,004)	-0,089*** (0,004)	-0,125*** (0,008)	-0,068*** (0,007)	-0,172*** (0,006)	-0,166*** (0,006)	-0,165*** (0,023)	-0,166*** (0,022)
Masculino	-0,044*** (0,005)	-0,046*** (0,005)	-0,045*** (0,005)	-0,046*** (0,005)	-0,047*** (0,005)	-0,049*** (0,005)	-0,046*** (0,005)	-0,049*** (0,005)
Filho	-0,095*** (0,005)	-0,096*** (0,005)	-0,095*** (0,005)	-0,095*** (0,005)	-0,090*** (0,005)	-0,092*** (0,005)	-0,090*** (0,005)	-0,092*** (0,005)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,220*** (0,022)	0,217*** (0,023)	0,221*** (0,022)	0,213*** (0,023)	0,220*** (0,022)	0,216*** (0,023)	0,219*** (0,022)	0,216*** (0,023)
Alta	0,089*** (0,008)	0,084*** (0,008)	0,090*** (0,008)	0,084*** (0,008)	0,088*** (0,008)	0,082*** (0,008)	0,088*** (0,008)	0,082*** (0,008)
Bem elevada	-0,227*** (0,014)	-0,263*** (0,016)	-0,223*** (0,014)	-0,263*** (0,016)	-0,224*** (0,014)	-0,258*** (0,016)	-0,225*** (0,014)	-0,258*** (0,016)
Jovens	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Adultos	0,029*** (0,008)	0,026*** (0,008)	0,028*** (0,008)	0,030*** (0,008)	0,026*** (0,008)	0,022*** (0,008)	0,026*** (0,008)	0,023*** (0,008)
Adultos2	-0,001 (0,008)	-0,007 (0,007)	-0,002 (0,008)	-0,004 (0,008)	-0,001 (0,008)	-0,007 (0,007)	-0,001 (0,008)	-0,006 (0,007)
Idosos	-0,003 (0,008)	-0,002 (0,008)	-0,003 (0,008)	-0,004 (0,008)	-0,006 (0,008)	-0,004 (0,008)	-0,006 (0,008)	-0,005 (0,008)
Construção	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Comércio	0,365*** (0,018)	0,391*** (0,021)	0,360*** (0,018)	0,388*** (0,021)	0,355*** (0,018)	0,382*** (0,021)	0,356*** (0,018)	0,382*** (0,021)
Ativ. Finan.	0,069*** (0,016)	0,115*** (0,020)	0,068*** (0,016)	0,112*** (0,020)	0,064*** (0,016)	0,112*** (0,020)	0,064*** (0,016)	0,112*** (0,020)
S.Saúde	0,280*** (0,016)	0,319*** (0,019)	0,277*** (0,016)	0,316*** (0,019)	0,265*** (0,016)	0,306*** (0,019)	0,266*** (0,016)	0,306*** (0,019)
S. Educação	0,015 (0,018)	0,032 (0,023)	0,015 (0,018)	0,031 (0,023)	0,007 (0,018)	0,027 (0,023)	0,008 (0,018)	0,027 (0,023)
S. Espec.	0,026 (0,016)	0,056*** (0,021)	0,025 (0,016)	0,054** (0,021)	0,021 (0,016)	0,052** (0,021)	0,021 (0,016)	0,052** (0,021)
S. Público	-0,013 (0,018)	-0,011 (0,024)	-0,012 (0,018)	-0,013 (0,024)	-0,016 (0,018)	-0,014 (0,024)	-0,016 (0,018)	-0,014 (0,024)
Outros	0,490*** (0,014)	0,493*** (0,017)	0,484*** (0,014)	0,489*** (0,017)	0,471*** (0,014)	0,475*** (0,017)	0,472*** (0,014)	0,475*** (0,019)
Constante	1,182*** (0,057)	–	1,583*** (0,092)	–	2,209*** (0,078)	–	2,133*** (0,282)	–
Observações	25.687	25.687	25.687	25.687	25.687	25.687	25.687	25.687
Estat. F/Wald	492.48	6798.67	7766.00	5645.15	521.70	7140.93	7695.53	5476.69
R ² / Pseudo R ²	0.234	0.192	0.233	–	0.245	0.202	0.245	–

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1.

Fonte: Estimativas com base na média dos efeitos marginais.

Os modelos IV visam lidar com a questão da endogeneidade levantada anteriormente. A acessibilidade aos empregos é instrumentalizada pela distância aos principais centros, seguindo um caminho formado pelas rotas e trilhos antigos da cidade do Recife. Concentrando-se nos modelos IV, pode-se dizer que o aumento do acesso físico ao CBD e subcentros em 1%, reduz em 0,125 e 0,166 pontos percentuais a probabilidade do trabalhador em Recife pertencer ao grupo de menor renda. Estes valores são bastante significativos quando comparados com a estatística descritiva para a probabilidade do indivíduo ganhar até 2 salários mínimos. Ao dividir 0,125 e 0,166 pela média de 0,44 obtem valores como 0,28% e 0,37%, respectivamente, que correspondem a redução em termos percentuais da probabilidade

dos trabalhadores pertencerem a essa faixa de renda com o aumento de 1% no índice de acessibilidade.

Como esperado, o maior nível educacional diminui a chance de ser de baixa renda, como mostra nas estimativas da Tabela 1.4. Observa-se que os coeficientes para as *dummies* de gênero e filhos são significativos e com efeitos esperados. Quanto aos setores econômicos, a probabilidade aumenta para os trabalhadores do comércio, atividade financeira e serviços de saúde.

Seguindo uma estratégia de estimativa da equação Minceriana, a Tabela 1.5 apresenta os principais resultados para trabalhadores que utilizam veículos particulares (indivíduos que utilizam apenas o carro como veículo privado), transporte público (indivíduos usuários apenas do ônibus e metrô) e outros modais (indivíduos que vão ao trabalho a pé, de bicicleta, motocicleta, táxi e veículo fretado) de transporte para viagens diárias de casa para o trabalho. A razão fundamental para investigar o efeito do indicador de acessibilidade nos salários por diferentes modais 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. Resultados adicionais considerando até 3 salários mínimos por diferentes modais são representados na Tabela A1.2 no apêndice.

Tabela 1.5 – Efeitos marginais para a probabilidade de até 2 salários mínimos por modal

Variáveis	CBD					
	Transporte Próprio		Transporte Público		Outros Modais	
W (Até 2 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,032*** (0,007)	-0,015** (0,007)	0,012 (0,013)	-0,046*** (0,013)	-0,136*** (0,014)	-0,064*** (0,014)
Estat. F/Wald	1038.36	646.08	1185,80	855.80	2585.08	1933.77
R ² / Pseudo R ²	0.109	-	0,122	-	0,212	-
Observações	8,493	8,493	8.362	8.362	8.832	8.832

Variáveis	Subcentros					
	Transporte Próprio		Transporte Público		Outros Modais	
W (Até 2 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,040*** (0,013)	-0,078*** (0,014)	0,121 (0,024)	-0,076*** (0,022)	-0,311*** (0,038)	-0,108*** (0,034)
Estat. F/Wald	1031.65	626.88	1175.90	923.79	2499.79	2151.18
R ² / Pseudo R ²	0.113	-	0,097	-	0,191	-
Observações	8,493	8,493	8.362	8.362	8.832	8.832

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. Obs.: Controles adicionais: idade, gênero, nível de escolaridade, setor de atividade, *dummy* para filhos, termo constante.

Fonte: Estimativas com base na média dos efeitos marginais.

Analisando os resultados obtidos, percebem-se dois resultados importantes, primeiro para subcentros, a estimativa probit IV é equivalente para os modais, ou seja, o indicador de acessibilidade tem uma influência negativa sobre os salários dos indivíduos considerados de baixa renda. Segundo, observam-se para o CBD valores mais elevados dos coeficientes para

"outros moldais". Contudo, em geral, as estimativas confirmam suas evidências iniciais, indicando que os diferentes moldais não afetam os resultados principais antes obtidos.

De fato, a dependência do transporte público em Recife pode limitar a quantidade de tempo efetivo no trabalho, resultando em um salário mais baixo, além do mais, a distância mais longa pelo trânsito em massa também pode está associado a residências periféricas, o que também pode reduzir tanto a eficiência da busca por empregos quanto sua intensidade devido à menor acessibilidade.

A relação entre acessibilidade ao emprego e resultados do mercado de trabalho também mostra padrões mais distintos quando permite a heterogeneidade entre os gêneros, como pode perceber a partir da Tabela 1.6, a seguir. Os resultados mostram que, em geral, o conjunto de evidências não indica claramente que o acesso é mais decisivo para algum dos gêneros.

Tabela 1.6 – Efeitos marginais para a probabilidade de até 2 salários mínimos por gênero.

Variáveis	HOMENS							
	CBD				Subcentros			
	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
W (Até 2 s.m)								
LnACi	-0,075*** (0,006)	-0,076*** (0,006)	-0,088*** (0,011)	-0,068*** (0,011)	-0,141*** (0,009)	-0,137*** (0,008)	-0,176*** (0,035)	-0,134*** (0,033)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	275.90	3573.19	271.44	2867.77	286.36	3688.16	4077.93	2910.67
R ² / Pseudo R ²	0.241	0.200	0.241	-	0.248	0,206	0,247	-
Observações	13.014	13.014	13.014	13.014	13.014	13.014	13.014	13.014

Variáveis	MULHERES							
	CBD				Subcentros			
	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
W (Até 2 s.m)								
LnACi	-0,101*** (0,006)	-0,099*** (0,006)	-0,159*** (0,011)	-0,064*** (0,010)	-0,199*** (0,009)	-0,192*** (0,008)	-0,158*** (0,032)	-0,195*** (0,030)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	258.53	3343.28	256.05	2828.63	279.58	3571.35	3753.27	2607.55
R ² / Pseudo R ²	0.234	0,191	0,230	-	0,248	0,204	0,247	-
Observações	12.673	12.673	12.673	12.673	12.673	12.673	12.673	12.673

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1.

Obs.: Controles adicionais: idade, nível de escolaridade, setor de atividade, *dummy* para filhos, termo constante.

Fonte: Estimativas com base na média dos efeitos marginais.

Não há na literatura resultados claros que os homens são mais produtivos que as mulheres, contudo os estudos mostram que as mulheres tendem a ter deslocamentos mais curtos medidos tanto em termos de distância e tempo. Além do mais, as mulheres geralmente têm maior responsabilidade em manter a família e cuidar de dependentes (por exemplo, filhos) do que os homens, sendo assim, as demandas competitivas por tempo resultam em mulheres menos móveis - portanto, seu comprimento de deslocamento reduzido (CRANE, 2007; GIULIANO, 1998; LEE E MCDONALD, 2003; MADDEN, 1981; TURNER E

NIEMEIER, 1997). No entanto, as evidências nesta tabela de que o melhor acesso aos empregos é mais decisivo para as mulheres no mercado é fraca, não havendo muita diferença entre os gêneros. Os resultados para gêneros por diferentes modais tanto para o CBD quanto para os subcentros, encontram-se nas Tabelas A1.3 e A1.4 no apêndice, respectivamente.

Os resultados encontrados dos parâmetros corroboram com a literatura teórica de que distâncias mais longas para os centros de empregos diminuem a produtividade dos agentes, indo na mesma linha de diversas evidências empíricas sobre o tema, tais como os trabalhos de Van Ommeren e Gutiérrez-i-Puigarnau, (2011), Stone e Schneider (2016), Vieira e Haddad (2015) e Haddad e Barufi (2017), apesar dessas pesquisas apresentarem níveis de análise distintos e outras metodologias. Além disto, as evidências obtidas indicam que tal efeito estimado é importante do ponto de vista quantitativo, ou seja, possui relevância social.

1.6 Robustez dos Resultados

Com o fito de verificar a robustez dos resultados, alguns exercícios adicionais foram realizados. No primeiro, foi realizado o mesmo exercício inserindo controles adicionais como a distância física para algumas amenidades. No segundo, diferentes medidas de acessibilidade são formadas, utilizando-se diferentes funções de decaimento com a distância. Por fim, foram retirados da amostra os indivíduos que moram na região central do Recife.

Percebe-se na Tabela 1.7, a seguir, que os sinais e a significância dos coeficientes, em geral, permanecem inalterados com a adição de novos controles (distância euclidiana das residências para praias e parques). Estes controles permitem aprender potencialmente diferenças de preferências locacionais entre as famílias nas suas trocas entre acesso ao emprego e amenidades. Seabra *et al.*, (2016) mostrou que tais amenidades importam e são valoradas pelo mercado imobiliário da cidade do Recife. Para os autores, todas as amenidades incluídas no trabalho se mostraram importantes no processo de formação de preços dos imóveis residenciais. Especificamente, por um lado, os valores dos imóveis da referida cidade tendem a apresentar apreciações se os mesmos apresentam vista para o mar, estão próximos do mar, do rio e dos parques da cidade.

Os resultados parecem fazer sentido com a estrutura da cidade do Recife. Maiores distâncias às praias e aos parques aumentam a chance do indivíduo ser de baixa renda. O estudo de Oliveira e Silveira Neto (2015) para Recife mostra que indivíduos mais ricos moram em lugares com mais amenidades. Essas áreas são próximas ao centro e com ampla

oferta de serviços públicos. Os resultados por diferentes modais tanto para o CBD quanto para os subcentros, encontram-se na Tabela A1.5 do apêndice.

Tabela 1.7- Efeitos marginais para a probabilidade de até 2 salários mínimos com amenidades

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
W (Até 2 s.m)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	-0,034*** (0,005)	-0,029*** (0,005)	-0,107*** (0,010)	0,006 (0,009)	-0,084*** (0,008)	-0,079*** (0,007)	-0,182*** (0,024)	-0,067*** (0,023)
D_Praia	0,027*** (0,001)	0,026*** (0,001)	0,022*** (0,001)	0,028*** (0,001)	0,023*** (0,001)	0,022*** (0,001)	0,015*** (0,002)	0,023*** (0,002)
D_Parques	0,083*** (0,010)	0,078*** (0,010)	0,030*** (0,012)	0,081*** (0,012)	0,064*** (0,010)	0,058*** (0,010)	0,013 (0,016)	0,060*** (0,016)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	486,72	7420,07	488,45	6171,62	492,01	7490,51	486,14	6152,86
R² / Pseudo R²	0,254	0,210	0,249	-	0,256	0,212	0,252	-
Observações	25.687	25.687	25.687	25.687	25.687	25.687	25.687	25.687

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. Obs.: Controles adicionais: idade, gênero, nível de escolaridade, setor de atividade, *dummy* para filhos, termo constante.

Fonte: Estimativas com base na média dos efeitos marginais.

Em uma segunda etapa, complementou-se a análise de robustez com resultados para diferentes medidas de acessibilidade, considerando todos os indivíduos da amostra. A função de impedância¹⁵ pode assumir diferentes formas funcionais. Na Tabela 1.8 foram obtidos resultados utilizando duas funções de potência com diferentes parâmetros de decaimento. Este parâmetro mede a relação entre padrões de interação observados (comutação) e distância quando outros determinantes de interação são constantes. Percebe-se a partir da Tabela, a seguir, que os resultados obtidos inicialmente não parecem sensíveis ao grau de decaimento da função associada à distância ao emprego, ou seja, não alteram de forma importante os resultados já obtidos para o grau de acessibilidade ao emprego nos resultados do mercado de trabalho quanto aos sinais e magnitude. Novamente, resultados por diferentes modais tanto para o CBD quanto para os subcentros, encontram-se na Tabela A1.6 do apêndice.

Tabela 1.8 – Efeitos marginais para diferentes funções de decaimento da distância.

Função de Distância	Variáveis	CBD				Subcentros			
	W (Até 2 s.m)	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
$d_{ij}^{-0,5}$	LnACi	-0,090*** (0,004)	-0,089*** (0,004)	-0,125*** (0,008)	-0,068*** (0,007)	-0,172*** (0,006)	- 0,166*** (0,006)	-0,165*** (0,023)	-0,166*** (0,022)
$d_{ij}^{-1,5}$	LnACi	-0,074*** (0,001)	-0,068*** (0,001)	-0,047*** (0,002)	-0,078*** (0,002)	-0,037*** (0,001)	-0,036*** (0,001)	-0,044*** (0,004)	-0,035*** (0,004)
Observações		25.687	25.687	25.687	25.687	25.687	25.687	25.687	25.687

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. Obs.: Controles adicionais: idade, gênero, nível de escolaridade, setor de atividade, *dummy* para filhos, termo constante.

Fonte: Estimativas com base na média dos efeitos marginais.

¹⁵ Autores como Johnson (2006), Quintanar *et al.* (2013), Carvalho (2016) utilizaram diferentes medidas para a função de impedância como forma de buscar robustez em seus resultados.

Adicionalmente, foi considerada a mesma estimação apenas para os subcentros de empregos, ou seja, retirando-se da amostra os empregos do CBD da cidade. Este terceiro exercício também foi levado a efeito sem indivíduos que moram na região central do Recife (locais mais próximos aos empregos). A ideia de excluir os empregos do CBD está associada ao fato de que tal localização de empregos teria ditado à trajetória inicial das ferrovias uma vez que nele também se localiza o porto da cidade (onde era exportada a produção de açúcar e algodão). Embora uma parte muito pouco significativa de empregos associada à atividade de exportação do açúcar ainda exista neste CBD¹⁶, o exercício permite afastar completamente tal remota possibilidade de alguma endogeneidade do instrumento, já que os demais subcentros de emprego da cidade não guardam qualquer relação original com as ferrovias.

Tabela 1.9 – Efeitos marginais para diferentes funções de decaimento da distância (sem a população do CBD).

Função de Distância	Variáveis	CBD				Subcentros			
	W (Até 2 s.m)	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
d_{ij}^{-1}	LnACi	-0,117*** (0,007)	-0,117*** (0,007)	-0,118*** (0,010)	-0,106*** (0,009)	-0,170*** (0,007)	-0,166*** (0,006)	-0,113*** (0,026)	-0,170*** (0,025)
$d_{ij}^{-0,5}$	LnACi	-0,117*** (0,007)	-0,117*** (0,007)	-0,118*** (0,010)	-0,106*** (0,009)	-0,170*** (0,007)	-0,166*** (0,006)	-0,113*** (0,026)	-0,170*** (0,025)
$d_{ij}^{-1,5}$	LnACi	-0,111*** (0,002)	-0,101*** (0,001)	-0,044*** (0,003)	-0,124*** (0,003)	-0,035*** (0,001)	-0,034*** (0,001)	-0,027*** (0,004)	-0,035*** (0,004)
Observações		23.023	23.023	23.023	23.023	23.023	23.023	23.023	23.023

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. Obs.: Controles adicionais: idade, gênero, nível de escolaridade, setor de atividade, *dummy* para filhos, termo constante.

Fonte: Estimativas com base na média dos efeitos marginais.

Como se pode perceber através da Tabela 1.9, a exclusão das ocupações do CBD da cidade não altera de forma importante os resultados já obtidos para a importância do grau de acessibilidade ao emprego para os resultados salariais dos trabalhadores da cidade. Além disto, este resultado é válido mesmo quando também se excluem da amostra os indivíduos que moram na região do CBD. Resultados para diferentes modais são encontrados no apêndice na Tabela A1.7.

1.7 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

¹⁶Como mostraram recentemente, Belmiro *et al.* (2018), as ocupações do CBD tradicional da Cidade do Recife hoje majoritariamente são caracterizadas por serviços que utilizam trabalho qualificado sem pouca relação com exportação da produção local.

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.

Este estudo destina-se a preencher uma importante lacuna em relação ao mercado de trabalho da Cidade do Recife. Nesse sentido, a pesquisa explorou correlações entre resultados do mercado de trabalho e medidas de acessibilidade a empregos. Os resultados dependem de fortes hipóteses de identificação para evitar viés relacionado a decisões simultâneas de localização de trabalhadores e empresas dentro da cidade. A análise empírica, apesar de ser motivada por uma estrutura teórica de equilíbrio geral, foi baseada em estratégias de modelagem de equilíbrio parcial. A inclusão de um instrumento exógeno permitiu controlar pelo menos uma parcela do viés gerado nesse contexto. Os trilhos antigos do século XIX, hoje escondidos pelas estradas, parecem ter ajudado a moldar os mercados urbanos em Recife.

Para a cidade do Recife, existem desigualdades espaciais significativas, e, dependendo do modo de transporte, os trabalhadores podem ver sua produtividade ou suas oportunidades para encontrar uma melhor correspondência no mercado de trabalho um pouco mais dificultada. Dessa forma, o estudo tentou investigar se essa relação negativa entre acessibilidade aos empregos e resultados do mercado de trabalho é válida para a cidade do Recife após a utilização da Pesquisa Origem-Destino e o controle de características individuais. As conclusões indicam que há uma relação entre duas medidas diferentes de acessibilidade aos empregos e salários. Os resultados principais encontrados mostram que o aumento do acesso físico ao CBD e subcentros em 1%, reduz em 0,28% e 0,37% a probabilidade do trabalhador em Recife pertencer ao grupo de menor renda, significativo ao nível de 1%.

Os resultados da pesquisa indicam que existe um *mismatch* espacial no mercado de trabalho em Recife, 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, independentemente do tipo de transporte são menos produtivos.

Para futuras extensões da pesquisa, será importante investigar a relação do indicador de acessibilidade com os salários considerando uma variável contínua para a renda, bem como analisar o impacto da acessibilidade sobre participação no mercado de trabalho, sobre o

desemprego e sobre a informalidade. Além disto, será de fundamental importância um estudo que busque identificar as causas para a existência do *mismatch* espacial na cidade do Recife.

APÊNDICE

Tabela A1.1- Efeitos marginais para a probabilidade de 1 e até 3 s.m.

Tabela 11-12 - Efeitos marginais para a proximidade de 10 até 5 km								
Variáveis	CBD				Subcentros			
W (Até 1 s.m)	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
LnACi	-0,006* (0,003)	-0,007** (0,003)	-0,009 (0,006)	-0,005 (0,006)	-0,023*** (0,004)	-0,024*** (0,004)	-0,047*** (0,017)	-0,027* (0,016)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	164,14	2627,73	164,11	1933,87	165,60	2650,20	163,19	1947,17
R ² / Pseudo R ²	0.092	0,131	0,092	-	0,094	0,132	0,085	-
Observações	25.687	25.687	25.687	25.687	25.687	25.687	25.687	25.687
Variáveis	CBD				Subcentros			
W (Até 3 s.m)	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
LnACi	-0,102*** (0,004)	-0,123*** (0,005)	-0,162*** (0,007)	-0,065*** (0,007)	-0,201*** (0,006)	-0,253*** (0,008)	-0,309*** (0,023)	-0,189*** (0,021)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	532,41	6878,25	526,39	6107,87	577,58	7413,52	519,49	6231,88
R ² / Pseudo R ²	0,249	0,197	0,244	-	0,269	0,212	0,256	-
Observações	25.687	25.687	25.687	25.687	25.687	25.687	25.687	25.687

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1.

Fonte: Estimativas com base na média dos efeitos marginais.

Tabela A1.2 – Efeitos marginais para a probabilidade de até 3 s. m. por modal

Variáveis	CBD					
	Transporte Próprio		Transporte Público		Outros Modais	
W (Até 3 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,069***	-0,027**	-0,016*	-0,033***	-0,172***	-0,068***
	(0,010)	(0,010)	(0,009)	(0,008)	(0,014)	(0,013)
Estat. F/Wald	1221,40	934,79	1134,28	707,23	2249,12	1763,25
R ² / Pseudo R ²	0,125	-	0,122	-	0,186	-
Observações	8.493	8.493	8.362	8.362	8.832	8.832

Variáveis	Subcentros					
	Transporte Próprio		Transporte Público		Outros Modais	
W (Até 3 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,100***	-0,156***	0,055***	-0,058***	-0,429***	-0,133***
	(0,019)	(0,019)	(0,016)	(0,013)	(0,037)	(0,030)
Estat. F/Wald	1219,38	905,76	1114,86	753,51	2107,83	2166,68
R ² / Pseudo R ²	0,135	-	0,100	-	0,137	-
Observações	8.493	8.493	8.362	8.362	8.832	8.832

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1.

Fonte: Estimativas com base na média dos efeitos marginais.

Tabela A1.3 – Efeito marginais para a probabilidade de até 2 s.m. por modal – CBD

Variáveis	HOMENS					
	Transporte Privado		Transporte Público		Outros Modais	
W (Até 2 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,035***	-0,030**	0,091***	-0,063***	-0,105***	-0,068***
	(0,011)	(0,011)	(0,020)	(0,019)	(0,020)	(0,020)
Controles	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	578,58	377,76	677,58	498,34	1499,75	1071,66
R ² / Pseudo R ²	0,122	-	0,131	-	0,223	-
Observações	4.189	4.189	3.859	3.859	4.966	4.966

Variáveis	MULHERES					
	Transporte Privado		Transporte Público		Outros Modais	
W (Até 2 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,027***	0,001	-0,052***	-0,033**	-0,170***	-0,055***
	(0,009)	(0,009)	(0,018)	(0,018)	(0,020)	(0,019)
Controles	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	503,76	269,45	576,57	415,32	1164,06	897,67
R ² / Pseudo R ²	0,103	-	0,115	-	0,206	-
Observações	4.303	4.303	4.503	4.503	3.867	3.867

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1.

Fonte: Estimativas com base na média dos efeitos marginais.

Tabela A1.4 – Efeitos marginais para a probabilidade de até 2 s.m. por modal - Subcentros.

Variáveis	HOMENS					
	Transporte Privado		Transporte Público		Outros Modais	
W (Até 2 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,034*	-0,109***	0,281***	-0,060**	-0,253***	-0,102**
	(0,020)	(0,020)	(0,039)	(0,031)	(0,058)	(0,054)
Controles	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	573,25	366,85	659,72	608,91	1470,53	1158,61
R ² / Pseudo R ²	0,126	-	0,063	-	0,211	-
Observações	4.189	4.189	3.859	3.859	4.966	4.966

Variáveis	MULHERES					
	Transporte Privado		Transporte Público		Outros Modais	
W (Até 2 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,041**	-0,041**	-0,006	-0,084***	-0,363***	-0,108**
	(0,018)	(0,018)	(0,031)	(0,030)	(0,051)	(0,044)
Controles	Sim	Sim	Sim	Sim	Sim	Sim
Estat. F/Wald	502,50	264,73	567,71	409,52	1105,33	1017,09
R ² / Pseudo R ²	0,106	-	0,112	-	0,176	-
Observações	4.303	4.303	4.503	4.503	3.867	3.867

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1.

Fonte: Estimativas com base na média dos efeitos marginais.

Tabela A1.5 – Efeitos marginais para a probabilidade de até 2 s.m.

Variáveis	CBD					
	Transporte Próprio		Transporte Público		Outros Modais	
W (Até 2 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,042***	-0,004	0,042**	-0,022	-0,129***	-0,018
	(0,008)	(0,008)	(0,017)	(0,017)	(0,017)	(0,016)
Controles	Sim	Sim	Sim	Sim	Sim	Sim
D_Praia	0,007***	0,008***	0,013***	0,008***	0,010***	0,018***
	(0,001)	(0,001)	(0,002)	(0,002)	(0,002)	(0,002)
D_Parques	-0,007	-0,001	0,028*	0,011	-0,015	0,047**
	(0,012)	(0,012)	(0,017)	(0,017)	(0,020)	(0,021)
Estat. F/Wald	61,67	698,15	1229,53	890,77	153,90	2033,93
R ² / Pseudo R ²	0,114	-	0,125	-	0,216	-
Observações	8.493	8.493	8.362	8.362	8.832	8.832

Variáveis	Subcentros					
	Transporte Próprio		Transporte Público		Outros Modais	
W (Até 2 s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,065***	-0,052***	0,137***	-0,058**	-0,341***	-0,040
	(0,014)	(0,014)	(0,026)	(0,024)	(0,045)	(0,041)
Controles	Sim	Sim	Sim	Sim	Sim	Sim
D_Praia	0,004***	0,005***	0,022***	0,003	-0,007*	0,017***
	(0,001)	(0,001)	(0,002)	(0,002)	(0,004)	(0,004)
D_Parques	-0,014	-0,149	0,050***	0,004	-0,097***	0,040
	(0,014)	(0,013)	(0,017)	(0,017)	(0,028)	(0,027)
Estat. F/Wald	61,63	687,40	1235,65	933,96	2652,55	2248,09
R ² / Pseudo R ²	0,117	-	0,114	-	0,187	-
Observações	8.493	8.493	8.362	8.362	8.832	8.832

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1.

Fonte: Estimativas com base na média dos efeitos marginais.

Tabela A1.6 – Efeitos marginais para diferentes funções de decaimento da distância.

Função de Distância	CBD						
	Variáveis	Transporte Próprio		Transporte Público		Outros Modais	
	W(Até 2s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
$d_{ij}^{-0,5}$	LnACi	-0,032*** (0,007)	-0,015** (0,007)	0,012 (0,013)	-0,046*** (0,013)	-0,136*** (0,014)	-0,064*** (0,014)
$d_{ij}^{-1,5}$	LnACi	-0,021*** (0,004)	-0,010** (0,005)	0,004 (0,004)	-0,015*** (0,004)	-0,045*** (0,004)	-0,021*** (0,004)
Observações		8.493	8.493	8.362	8.362	8.832	8.832
Função de Distância	Subcentros						
	Variáveis	Transporte Próprio		Transporte Público		Outros Modais	
	W(Até 2s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
$d_{ij}^{-0,5}$	LnACi	-0,040*** (0,013)	-0,078*** (0,014)	0,121*** (0,024)	-0,076*** (0,022)	-0,311*** (0,038)	-0,108*** (0,034)
$d_{ij}^{-1,5}$	LnACi	-0,021*** (0,008)	-0,049*** (0,008)	0,021*** (0,006)	-0,014*** (0,005)	-0,085*** (0,008)	-0,018*** (0,007)
Observações		8.493	8.493	8.362	8.362	8.832	8.832

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. Obs.: Controles adicionais: idade, gênero, nível de escolaridade, setor de atividade, *dummy* para filhos, termo constante. Os pesos de amostragem são aplicados à estimativa.

Fonte: Estimativas com base na média dos efeitos marginais.

Tabela A1.7 – Efeito marginais para diferentes funções de decaimento da distância (sem população do CBD)

Função de Distância	CBD						
	Variáveis	Transporte Próprio		Transporte Público		Outros Modais	
	W(Até 2s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
d_{ij}^{-1}	LnACi	-0,021** (0,008)	-0,039*** (0,009)	0,026 (0,018)	-0,079*** (0,017)	-0,137*** (0,021)	-0,055*** (0,021)
$d_{ij}^{-0,5}$	LnACi	-0,021** (0,008)	-0,039*** (0,009)	0,026 (0,018)	-0,079*** (0,017)	-0,138*** (0,021)	-0,056*** (0,021)
$d_{ij}^{-1,5}$	LnACi	-0,014** (0,005)	-0,026*** (0,006)	0,008 (0,006)	-0,026*** (0,005)	-0,046*** (0,007)	-0,018*** (0,007)
Observações		7.618	7.618	7.867	7.867	7.538	7.538
Função de Distância	Subcentros						
	Variáveis	Transporte Próprio		Transporte Público		Outros Modais	
	W(Até 2s.m)	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
d_{ij}^{-1}	LnACi	-0,017 (0,015)	-0,090*** (0,015)	0,151*** (0,027)	-0,081*** (0,024)	-0,250*** (0,046)	-0,094** (0,043)
$d_{ij}^{-0,5}$	LnACi	-0,017 (0,015)	-0,090*** (0,015)	0,151*** (0,027)	-0,081*** (0,024)	-0,250*** (0,046)	-0,094** (0,043)
$d_{ij}^{-1,5}$	LnACi	-0,004 (0,004)	-0,018*** (0,004)	0,026*** (0,006)	-0,014*** (0,006)	-0,063*** (0,008)	-0,017** (0,008)
Observações		7.618	7.618	7.867	7.867	7.538	7.538

Desvio-padrão entre parênteses. ***p < 0.01, **p < 0.05, *p < 0.1. Obs.: Controles adicionais: idade, gênero, nível de escolaridade, setor de atividade, *dummy* para filhos, termo constante.

Fonte: Estimativas com base na média dos efeitos marginais.

2. MISMATCH ESPACIAL E MERCADO DE TRABALHO: O EFEITO DA ACESSIBILIDADE AO EMPREGO SOBRE PARTICIPAÇÃO, DESEMPREGO, INFORMALIDADE E OVEREDUCATION NA REGIÃO METROPOLITANA DE SÃO PAULO

2.1 Introdução

Com uma população de mais 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; ZHENG, 1998). De acordo com Instituto Brasileiro de Geografia e Estatística (IBGE), 2010, aproximadamente 55,3% dos passageiros da RMSP passaram mais de 30 minutos viajando, e cerca de 23% deles (aproximadamente 1,8 milhão de pessoas) passaram mais de uma hora indo de casa para o trabalho todos os dias (SILVEIRA NETO *et al.*, 2015).

Algumas cidades de países desenvolvidos são caracterizadas por fortes disparidades em oposição aos centros, sendo pobres e catalisadoras com muitos problemas sociais, enquanto os subúrbios são mais prósperos (DURJANDIN *et al.*, 2008). No Brasil, país em desenvolvimento, a maioria das cidades apresenta altos níveis de segregação socioeconômica (Oliveira e Silveira Neto, 2015; Boisjoly *et al.*, 2017), ou seja, as áreas centrais com boa acessibilidade concentram a maior parte dos empregos formais. Por outro lado, as áreas periféricas com população de baixa renda apresentam acessibilidade limitada, alta incidência de emprego informal e desemprego (MORENO-MOROY e RAMOS, 2015). Este cenário evidencia a segregação residencial em função da renda, ou seja, existe 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 (CÔRTEZ, 2008).

Esse cenário não é diferente para a RMSP que exhibe uma divisão núcleo-periferia particularmente marcada (RAMOS, 2014). Semelhante à maioria das grandes regiões metropolitanas brasileiras, o rápido desenvolvimento urbano da RMSP ocorreu sem expansão adequada da infraestrutura básica de transporte doméstico (MEYER *et al.*, 2004; NADALIN e IGLORI, 2010). Moreno-Monroy e Ramos (2015) argumentam 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. Slovic *et al.*, (2019), em um estudo para o município de São Paulo examinou as desigualdades da acessibilidade ao emprego e como ela se sobrepõe aos indicadores socioeconômicos, expectativa de vida e condições de infraestrutura. Os resultados mostraram que taxas mais baixa de acessibilidade ao emprego estão associadas a áreas de pior condição socioeconômica, com a expectativa de vida mais baixa e a infraestrutura proporcionalmente precária. Dessa forma, Nadalin e Iglioni (2010) argumentam que isso explica, pelo menos parcialmente, o padrão de localização das famílias da 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 resultar 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; SJOQUIST, 1998). Dessa forma, não se sabe ao certo sobre a importância relativa destas duas explicações.

Sabe-se que uma das explicações mais importantes para os resultados insatisfatórios do emprego em locais desfavorecidos é a hipótese do *Mismatch* espacial, primeiramente citada por John Kain no final da década de 1960 (KAIN 1968, 1992; KASARDA 1980). O *mismatch* espacial é a desconexão espacial entre trabalhadores e oportunidades de trabalho. Este *mismatch* pode ser causado por diversas formas dentre as quais, pode-se citar a segregação residencial que se refere à aglomeração territorial de famílias que pertencem ao mesmo grupo social definido em termos de *status* socioeconômico e etnicidade, entre outros classificadores. A segregação residencial pode gerar externalidades negativas no bairro que reduzam as oportunidades de trabalho para seus residentes (QUINTANAR, 2013). Dessa forma, o *mismatch* espacial identifica vários mecanismos¹⁷ através dos quais o acesso físico restrito aos empregos pode resultar em mais desemprego. Esta análise pode ser ampliada para explicar como esses mecanismos podem levar a um emprego mais informal. Com base nisto, pesquisadores têm observado uma crescente separação geográfica entre oportunidades de emprego e localização dos indivíduos, mostrando que a distância física tem sido reconhecida como uma barreira a empregos (SCOTT & URRY, 1994; STORPER & WALKER, 1983; CERVERO & WU, 1997; SUAREZ & DELGADO, 2007; MATAS, 2010; QUINTANAR, 2013).

¹⁷ Esses mecanismos foram abordados na seção 2.

A despeito do forte desenvolvimento da Economia Urbana (Bruckener, 2011) e largo conjunto de trabalhos aplicados, é de interesse de diversos estudos buscarem entender as forças econômicas que são responsáveis por resultados no mercado de trabalho urbano, o que inclui tanto a distribuição espacial das famílias e das ocupações, como a acessibilidade aos empregos. Neste sentido, Durjandin *et al.*, (2008) procuram entender como a acessibilidade aos empregos afeta o desemprego urbano na cidade de Bruxelas. Ong e Miller (2005) comparam os impactos do *mismatch* espacial (a separação geográfica dos trabalhadores e dos empregos) e o *mismatch* do transporte (a falta de acesso a um automóvel privado) nos índices de ocupação e taxas de desemprego em Los Angeles. No mesmo sentido, porém tentando explicar a relação do *mismatch* espacial para a informalidade, têm-se os trabalhos de Abane (1993); Miltra (2005); Suárez *et al.*, (2016); e Chen *et al.*, (2017).

Para o Brasil, Campos (2015) buscou identificar o efeito da decisão de localização do trabalhador na busca por emprego na cidade de Salvador. Motte *et al.*, (2016) testaram a influência de variáveis socioeconômicas e espaciais na distância e duração das viagens de deslocamento dos habitantes que trabalham nos setores formais e informais da região metropolitana do Rio de Janeiro (RMRJ). Especificamente para a região metropolitana de São Paulo (RMSP), Moreno-Monroy e Ramos (2015) estimaram o impacto das expansões do sistema de transporte público sobre as taxas de informalidade local entre 2000 e 2010. Barufi e Haddad (2017) investigaram a relação entre *mismatch* espacial, salários e desemprego para as grandes áreas urbanas do Brasil, que inclui a RMSP. Boisjoly *et al.*, (2017) exploraram a relação entre a informalidade do trabalho e a acessibilidade ao emprego nos transportes públicos. Por fim, Haddad e Barufi (2017) exploraram a relação entre tempo de deslocamento, acessibilidade para empregos e preços urbanos no contexto dos mercados urbanos.

Além da análise empírica acima, a hipótese do *mismatch* espacial permite estudar a relação com a literatura de *Overeducation*, ou seja, quando os indivíduos possuem níveis educacionais diferentes daqueles requeridos para suas ocupações (FREEDMAN, 1976; HARTOG, 2000, McGUINNESS, 2006). Esta literatura também conhecida como *mismatch* educacional no mercado de trabalho é caracterizada a partir da existência de indivíduos com escolaridade diferente daquela requerida pelo seu posto de trabalho, nesse sentido os trabalhadores podem ser classificados como *Overeducated* (sobre-educado) quando o trabalhador possuir o nível educacional superior ao que é requerido em sua ocupação, ocasionando em retorno salarial menor, se comparado a indivíduos que trabalham em atividades que requerem o seu nível de formação.

A extensa literatura sobre *mismatch* educacional (Leuven & Oosterbeek 2011) destaca especialmente que ter um emprego abaixo do nível de qualificação atingido representa uma situação de subemprego e gera um desperdício de capital humano. De fato, a supereducação está associada a menor produtividade, menor, elevado risco de abandonar os empregos e insatisfação (DI PAOLO *et al.*, 2014). A hipótese do *mismatch* espacial acima mencionada sugere que a falta de conexões com as oportunidades de emprego provavelmente será prejudicial para o emprego e para a qualidade do emprego. Logo, com base nesta previsão empírica, argumenta-se que o *mismatch* entre educação e trabalho, especialmente na forma de *overeducation*, pode ser o resultado de pesquisas de emprego com restrições espaciais. Portanto, além do desemprego e informalidade, a incidência de supereducação deve ser maior entre os trabalhadores com restrições espaciais. Em outras palavras, o aumento da acessibilidade ao emprego deve reduzir o risco de excesso de educação nos mercados de trabalho urbanos.

Desde o estudo seminal de Büchel e van Ham (2003), um número crescente de artigos preocupa-se com o papel desempenhado pela flexibilidade espacial na explicação das probabilidades de emprego e na propensão a ser *overeducated*, o que representa um caso específico de ter um emprego de baixa qualidade, a saber, Hensen *et al.*, (2009) para o caso da Holanda, Jauhainen (2011) na Finlândia, Huber (2012) em vários países da União Europeia, Iammarino & Martinelli (2012), Devillanova (2013) e Croce & Ghignoni (2013) na Itália e Ramos & Sanromà (2013) na Espanha. No Brasil, essa literatura tem-se desenvolvido mais recentemente e apresenta algumas contribuições importantes para o debate, sobretudo em relação à discussão das evidências empíricas no tema (SANTOS, 2002; SCHWARTZMAN, 2004; MACHADO, OLIVEIRA e CARVALHO, 2004; DIAZ e MACHADO, 2008; ESTEVES, 2009; VIANNA E OLIVEIRA, 2010), contudo não fazem a relação com o indicador de acessibilidade ao emprego. A alta incidência da incompatibilidade educacional no mercado de trabalho brasileiro e o consequente impacto sobre a produtividade que esse fenômeno parece implicar, como evidenciado nestes trabalhos, mostram a necessidade de acompanhar a evolução ao longo do tempo.

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 (IHLANFELDT, 2006; ANDERSSON *et al.*, 2014). Para o Brasil, são poucas evidências disponíveis sobre esta relação. Alguns autores estudaram o impacto do indicador de acessibilidade sobre os salários (Barufi e Haddad, 2017; Haddad e Barufi, 2017), na participação do trabalho (Campos, 2015), sobre o desemprego (Campos, 2015; Barufi e

Haddad, 2017) e sobre a informalidade (Moreno-Monroy e Ramos, 2015; Motte *et al.*, 2016; Boisjoly *et al.*, 2017), contudo não analisaram o efeito da acessibilidade aos empregos para diferentes variáveis dependentes ao mesmo tempo em um mercado urbano específico e nem separado por gênero.

Com base nessa perspectiva, este estudo fornece uma análise da relação entre *mismatch* espacial e os resultados do mercado de trabalho em uma grande área metropolitana do Brasil, a RMSP. O desafio da pesquisa foi calcular o efeito da acessibilidade ao emprego na probabilidade de participação no mercado, de estar desempregado, de pertencer ao setor informal e de chance de ser *overeducated*. 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 os *status* de ocupação considerados. Mais especificamente, a principal contribuição para a literatura feita por este estudo consiste em analisar o efeito da acessibilidade do trabalho aos centros de emprego para diferentes variáveis dependentes, inclusive na incidência e extensão da incompatibilidade na educação no emprego. (sob a forma de supereducação).

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. Fujita e Ogawa (1982) supõe a existência de uma cidade unidimensional cuja configuração resulta de interações econômicas entre famílias e empresas. Esta teoria determina a distribuição de terras comerciais e residenciais juntamente com as densidades de emprego e residenciais em todos os locais da cidade, bem como os preços urbanos, ou seja, salários e aluguéis de terras (ROSSI-HANSBERG, 2004).

Apesar da natureza reconhecidamente muito estilizada deste modelo, ele pode ser usado como uma estrutura conceitual para entender a interação de forças que são importantes na determinação da estrutura urbana (HADDAD e BARUFI, 2017). Como tal, fornece o contexto estrutural para obter implicações quantitativas em estudos empíricos mais

detalhados. Lucas e Rossi-Hansberg (2002) mostram que o modelo determina simultaneamente a localização de empregos e trabalhadores, bem como salários e aluguéis de terras. Em equilíbrio, os salários dentro da cidade aumentam com a densidade do emprego e os aluguéis de terras na cidade também aumentam com a densidade de trabalhadores. Como consequência, a acessibilidade aos empregos também é determinada simultaneamente, para qualquer função de custo de deslocamento. Como apontam Duranton e Puga (2015) os padrões de acessibilidade também são afetados pelas escolhas de localização de firmas e trabalhadores, que são determinadas pelos preços.

Entre poucos estudos empíricos que tentam lidar com a questão da endogeneidade, muitos preferem restringir a amostra (O'Regan e Quigley, 1996, 1998; Durjandin *et al.*, 2008; Matas *et al.*, 2010), no entanto esta solução não elimina completamente a endogeneidade, pois podem existir características não observadas que influenciam tanto a escolha residencial quanto o *status* de emprego (GLAESER, 1996). Segundo Ihlanfeldt (2006), entre os principais problemas econométricos decorrentes está o fato de que a localização residencial e a mensuração de oportunidades de trabalho são potencialmente endógenas. Essa endogeneidade pode surgir através da autoseleção de trabalhadores mais ou menos produtivos para áreas específicas, pela potencial causalidade reversa (*sorting*) de oportunidades de emprego e a probabilidade de estar desempregado, ou através de decisões simultâneas de localização de firmas e trabalhadores em um ambiente de equilíbrio geral.

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. Esta vertente da literatura propõe abordar essas questões com o uso de variáveis instrumentais. Para o presente trabalho obteve-se por utilizar essa estratégia (Ong e Miller, 2005; Moreno-Monroy e Ramos, 2015, Haddad e Barufi, 2017) baseando-se na inclusão de uma variável geográfica (acesso à margem fluvial) como instrumento para acessibilidade a empregos. A estratégia adotada permite lidar com a natureza simultânea dos locais comerciais e habitacionais e com a determinação dos resultados no mercado de trabalho com base no indicador de acessibilidade em um cenário geográfico específico. A validade da estratégia adotada baseia-se na correção para possível seleção endógena.

Para alcançar os resultados obtidos, o estudo utilizou principalmente dois bancos de dados: o Censo de População de 2010 e a Relação de Informações Anuais de Responsabilidade Social (RAIS). Os microdados para as características individuais são obtidos no Censo 2010 fornecido pelo IBGE, sendo a unidade de análise o indivíduo dentro

da Área Espacial de Ponderação (AEP). Já a base da RAIS permitiu o georreferenciamento da localização dos empregos para identificação dos centros de empregos (subcentros) para o ano de 2010. As equações estimadas foram com base nos modelos de probabilidade linear e *Probit*, com e sem variáveis instrumentais.

Os resultados deste estudo enfatizam a necessidade de analisar mais sobre o mercado de trabalho na Economia Urbana. Foram encontradas evidências importantes com disparidades espaciais em termos de acesso a oportunidades de emprego nas áreas de ponderação da RMSP, bem como efeitos de gênero heterogêneos na conectividade nos resultados trabalhistas. As evidências apontaram que, em geral, a acessibilidade ao emprego aumenta a participação da força de trabalho, diminui o desemprego, a informalidade e a chance do indivíduo ser *overeducated*.

O ensaio está estruturado da seguinte forma. A seção 2 fornece um breve referencial teórico sobre o *mismatch* espacial e os mercados de trabalho locais, enfocando as interações sociais dentro da cidade. Em seguida, será apresentada uma revisão de literatura sobre os principais resultados encontrados na literatura existente dos mercados urbanos. Na Seção 4, descreve a estratégia econométrica e o banco de dados, enquanto na Seção 5 serão analisados os resultados. Considerações finais seguem na Seção 6.

2.2 *Mismatch* espacial e Mercado de Trabalho

Na teoria da Economia Urbana, tem-se o Modelo Monocêntrico de localização residencial, que foi proposto por Alonso (1964) e desenvolvido por Muth (1969) e Mills (1967), apresentando uma cidade hipotética, onde todos os empregos localizam-se em um ponto central da cidade (*Central Business District – CBD*), para o qual todos os residentes devem se deslocar, ou seja, a teoria da localização afirma que as decisões relativas são tomadas em referência à atividade econômica no centro. A hipótese básica do modelo é a racionalidade econômica dos indivíduos/famílias e que estes buscam maximizar suas funções utilidades sujeita a sua restrição orçamentária. Os agregados familiares se localizam tão perto do centro como o seu orçamento permite.

À medida que a distância ao centro aumenta, o aluguel de terras diminui inversamente para os custos de transporte. Para este modelo, as famílias pobres localizam-se no centro da cidade, porque os lotes de moradia são menores e as estruturas mais antigas e, portanto, mais acessíveis, diferentemente dos ricos que residem mais afastados (SUÁREZ, MURATA &

CAMPOS, 2016). Glaeser, Kahn e Rappaport (2008) indicaram, para o caso das cidades americanas, que a localização dos indivíduos de baixa renda nas regiões centrais poderia ser explicada pela maior disponibilidade de transporte público nestas áreas.

Essa discussão foi inicialmente levantada por Kain (1968), que observou o mercado de trabalho para negros americanos residentes em guetos. Segundo o autor, existe segregação econômica e racial no mercado de trabalho e as regiões mais pobres tem o efeito de prejudicar a formação do trabalhador. Além disso, a distância desses locais em relação aos centros econômicos dificulta o acesso do trabalhador a novos empregos, elevando o tempo que ele passa desempregado. Essa abordagem teórica levou o nome de Hipótese do *Mismatch* Espacial. Ihlanfeldt (2006) aponta que as oportunidades de emprego dos negros foram corroídas por sua incapacidade de seguir os empregos da cidade central aos subúrbios à medida que a suburbanização dos empregos se acelerava.

Ihlanfeldt (2006) também aponta que as regulamentações de uso da terra dificultam a todos os trabalhadores menos qualificados, independentemente de raça ou etnia, encontrar moradias acessíveis perto de empregos suburbanos. Portanto, o *mismatch* espacial deve ser pensado como um problema de trabalhador menos qualificado, em vez de um problema de negros ou minorias. No entanto, os negros são desproporcionalmente afetados pelo *mismatch* espacial, porque seus locais residenciais são mais severamente restringidos do que os de brancos menos qualificados devido à discriminação racial nos mercados imobiliários.

As restrições nas localizações residenciais dos negros e dos trabalhadores pouco qualificados em geral são apenas uma das três condições que devem ser obtidas para que o *mismatch* espacial surja no mercado de trabalho urbano. Além disso, as empresas devem enfrentar custos mais altos (custos de instalação/produção) em áreas onde os trabalhadores menos qualificados são obrigados a viver, e os custos de busca ou deslocamento desses trabalhadores devem ser não triviais. Se a primeira condição não for satisfeita e as empresas puderem ser localizadas facilmente na cidade central, espera-se que os empregadores as localizem nessas áreas para capitalizar o excesso de oferta de mão-de-obra. Em longo prazo, a migração de empresas expandiria a demanda por mão-de-obra menos qualificada na cidade central e eliminaria o *mismatch* espacial. Se a segunda condição não for satisfeita e os trabalhadores das cidades centrais puderem comutar facilmente para empregos nos subúrbios, isso também serviria para eliminar um problema de *mismatch* no longo prazo (IHLANFELDT, 2006).

De fato as empresas encontram altos custos se estiverem localizadas na cidade central, o que inclui a dificuldade da montagem de terras, a alta criminalidade, os altos impostos sobre

a propriedade e o alto custo de contratar trabalhadores qualificados que preferem não trabalhar na cidade central porque preferem não ter que sair de suas casas nos subúrbios (IHLANFELDT, 2006). Entre esses fatores, verificou-se que o crime é um repelente particularmente forte ao emprego (Bollinger & Ihlanfeldt 2003), uma vez que aumenta os custos das empresas elevando seus prêmios de seguro, aumentando suas medidas de autoproteção e exigindo que paguem um prêmio a seus trabalhadores por trabalharem em áreas inseguras.

Com a mudança das cidades monocêntricas para cidades policêntricas, a abordagem teórica da estrutura urbana foi alterada, já que o modelo monocêntrico parece insuficiente para explicar a organização espacial das cidades de hoje (SMALL & SONG, 1992). De acordo com Fujita e Ogawa (1982), tornava-se necessária a introdução do conceito da não centralidade em estudos econômicos a partir de um modelo que concebesse a escolha de localização como endógena, de forma que não se assumissem *a priori* os espaços preenchidos por famílias e empresas. Este tipo de preocupação ganhava força ainda nos anos 1970, quando as investigações empíricas mostravam uma tendência generalizada à descentralização e consequente declínio da importância econômica do CBD (CARVALHO, 2016).

Segundo Brueckner, Thisse e Zenou (1999), a presença de amenidades urbanas presentes nas regiões centrais da cidade pode explicar um padrão de concentração dos mais ricos nas proximidades do centro e dos pobres nas periferias. Kain (1968) reconheceu que a suburbanização do trabalho pode criar um "*mismatch* espacial", ou seja, empregos com baixa remuneração estão presentes nos subúrbios, sendo assim, trabalhadores de baixa renda teriam mais dificuldade de acesso no centro da cidade, pois são obrigados a fazer comutações mais longas.

Nesse sentido, os candidatos a emprego que residem em áreas com poucas vagas de emprego locais ou em áreas localizadas longe de centros de emprego estão expostos apenas a um pequeno grupo de vagas. Residindo em mercados de trabalho locais de mais difícil acesso a emprego e serviços, eles devem passar mais tempo pesquisando antes de entrar em contato com um potencial empregador. Dessa forma, como argumentam Gobillon *et al.*, (2011), pelo menos três razões podem justificar a existência do *mismatch* espacial no mercado de trabalho. Primeiro, por causa de fricções informacionais, os candidatos a emprego podem não procurar muito longe de suas residências. Por exemplo, os trabalhadores que residem longe das oportunidades de emprego podem não ouvir falar de ofertas de emprego quando os empregadores recorrem a métodos de recrutamento que favorecem a força de trabalho local (ou seja, colocando cartazes "desejados" em lojas de varejo ou optando por não divulgar

ofertas de emprego além de certa distância). Alternativamente, os candidatos a emprego podem obter apenas informações parciais sobre a localização de empregos distantes ou podem ter apenas uma vaga ideia sobre os tipos de empregos oferecidos em partes na área com as quais não estão familiarizados (GOBILLON *et al*, 2011).

Em segundo lugar, como a pesquisa é dispendiosa, os trabalhadores podem restringir seu horizonte de pesquisa à vizinhança. Eles podem pesquisar com menos frequência para reduzir o número de viagens de busca ao emprego ou não procurar em todos os empregos localizados em lugares distantes. Nesse contexto, o acesso ao transporte público ou à posse de carros pode reduzir os custos da procura de emprego e expandir o horizonte da mesma (STOLL, 1999). Em terceiro lugar, o esforço na busca individual pode depender do custo de vida local, de modo que os trabalhadores que residem em áreas desconectadas das oportunidades de emprego não buscam intensamente. Smith e Zenou, (2003); Pattachini e Zenou, (2006) argumentam que os trabalhadores que residem em áreas com baixo custo de moradia podem sentir relativamente pouca pressão para procurar ativamente por um emprego para pagar seu aluguel. Como as despesas com moradia mais distantes dos centros de empregos são mais baixas, estes trabalhadores podem permanecer desempregados por um período de tempo mais longo do que as famílias que vivem em áreas mais próximas dos empregos. Por outro lado, os desempregados que moram em áreas nas quais os aluguéis são caros podem se sentir mais pressionados a procurar intensamente por um emprego para evitar a necessidade de se mudar (SMITH & ZENOU, 2003; PATTACHINI & ZENOU, 2006; GOBILLON & SELOD, 2014).

O *mismatch* espacial também pode explicar algo sobre a redução da frequência das propostas de emprego. Os empregadores podem então relutar em propor empregos para trabalhadores distantes porque o deslocamento em longas distâncias tornaria esses trabalhadores menos produtivos (eles apareceriam tarde ou se cansariam devido ao deslocamento excessivo, como coloca Zenou, 2002). À distância às oportunidades de emprego também pode reduzir a probabilidade de uma aceitação de emprego. De fato, os trabalhadores podem rejeitar uma oferta que envolva viagens longas demais se o deslocamento para esse trabalho fosse demasiadamente dispendiosa, tendo em conta o salário proposto (ZAX & KAIN, 1996).

Finalmente, o efeito da segregação residencial sobre a aquisição de empregos também pode ser prejudicial na medida em que os contatos profissionais geralmente ocorrem através de amigos e parentes (MORTENSEN & VISHWANATH, 1994). Como as redes sociais são pelo menos parcialmente localizadas, quando a taxa de desemprego é alta em uma

determinada área, é menos provável que os trabalhadores conheçam vizinhos empregados que possam informá-los sobre vagas existentes (CALVÓ-ARMENGOL & JACKSON, 2004; SELOD & ZENOU, 2006). Essa segregação social pode deteriorar a qualidade das redes sociais (Gobillon *et al.*, 2010) e, assim, a concentração espacial dos desempregados pode gerar uma externalidade negativa que diminui a probabilidade de ser empregada.

Entre alguns dos modelos teóricos dedicados a descrever os desequilíbrios espaciais no ambiente urbano, Zenou (2000) desenvolve um modelo com mecanismos de formação de cidades endógenas que resultam em empregos concentrados no CBD, indivíduos empregados que residem nas proximidades do centro da cidade, e desempregados estando mais longe dos empregos. O desemprego urbano é então reforçado na periferia da cidade, porque quanto mais longe um indivíduo está dos empregos, mais difícil é para o mesmo encontrá-los.

Na verdade, os estudos sobre acessibilidade ao emprego forneceram evidências de que as populações de baixa renda se beneficiam menos com a estrutura urbana (SUAREZ & DELGADO, 2007) e sugerem que um equilíbrio planejado entre empregos e habitação poderia permitir uma localização eficiente, reduzindo assim os tempos de *commuting* e o congestionamento do tráfego, tendo um efeito positivo no desempenho econômico (CERVERO & WU, 1997).

Enfim, a escolha da localização residencial é influenciada por diversos fatores, como acesso a amenidades, escolas, serviços e lojas. O gênero também tem sido relatado como fator importante nas decisões de viagem de trabalho. As mulheres geralmente têm viagens mais curtas (Hanson e Pratt, 1995; Madden, 1981; Turner e Niemeier, 1997; White, 1986) e, em muitos casos, especialmente as mães solteiras de baixa renda (Chapple, 2001) buscam emprego informal próximo ao local de residência. Silveira Neto *et al.*, (2015) estudaram o papel do estado civil e a presença de dependentes nos diferenciais de gênero no tempo de deslocamento. Os resultados sugeriram que o estado civil exerceu influência mais forte no tempo de deslocamento das mulheres trabalhadoras.

Portanto, todos esses mecanismos econômicos afetam os resultados no mercado de trabalho para os indivíduos, como por exemplo, na força de trabalho, na condição de desemprego, na taxa de informalidade ou na condição de overeducated que depende tanto das características individuais quanto das características locais. Na próxima seção, serão apresentados alguns estudos empíricos referentes a esta abordagem.

2.3 Estudos empíricos

A presente pesquisa relaciona-se com uma vertente da literatura que estuda a relação entre acessibilidade aos empregos e resultados no mercado de trabalho. Estudos mais antigos encontraram essa relação com base nos mecanismos do *mismatch* espacial (IHLANFELDT, 1993; HOLZER *et al.*, 1994). Ellwood (1986) para Chicago e Zax e Kain (1996) para Detroit, encontraram uma relação positiva entre as perspectivas de emprego da população negra e a acessibilidade ao emprego. Ihlanfeldt e Sjoquist (1990) mostraram que a proximidade com o trabalho aumenta a probabilidade de ser empregada para jovens nos Estados Unidos. Blumenberg e Ong (2001) descobriram que os tempos de deslocamento longos impactaram negativamente a capacidade das populações de baixa renda para encontrar trabalho em Los Angeles. Kawabata (2003) descobre que um melhor acesso ao transporte público aumenta a probabilidade de trabalho e horas de trabalho de indivíduos que não possuem um carro na área metropolitana dos EUA. No caso da Inglaterra, Patacchini e Zenou (2005) descobrem que os indivíduos que vivem longe ou que têm uma menor acessibilidade ao trabalho procuram menos emprego, enquanto os indivíduos que têm acesso ao carro aumentam a intensidade da pesquisa. Baum (2009) mostra que o acesso ao carro aumenta as oportunidades de trabalho e as horas de trabalho. Mata *et al.*, (2010) mostram que a baixa acessibilidade do emprego no transporte público afeta negativamente a probabilidade de emprego para o mercado feminino na região metropolitana de Barcelona e Madri. Sanchez (1999) e Tyndall (2015) descobriram que o acesso ao transporte público resulta em taxas de emprego mais altas no contexto norte-americano.

Dentre alguns estudos que analisaram os determinantes do desemprego ao nível de trabalhadores, têm-se o trabalho mais antigo de Rogers (1997) que descobriu que a maior acessibilidade ou proximidade ao trabalho resulta em menor duração do desemprego na área metropolitana de Pittsburgh. Ong e Miller (2005) comparam os impactos do *mismatch* espacial (a separação geográfica dos trabalhadores e dos empregos) e o *mismatch* do transporte (a falta de acesso a um automóvel privado) nos índices de ocupação e população do bairro e taxas de desemprego. A pesquisa estuda a área metropolitana de Los Angeles e utiliza uma abordagem de variável instrumental para corrigir a simultaneidade do emprego e da propriedade do carro. O instrumento baseia-se na construção de uma nova taxa de carro prevista concebida como uma função de três fatores exógenos como o custo de auto propriedade, o número de atividades que podem ser conduzidas dentro do bairro e a disponibilidade de transporte alternativo. Os resultados indicam que o *mismatch* do transporte

é o fator mais importante na geração de resultados do mercado de trabalho, particularmente para os bairros desfavorecidos. Áreas com relativamente mais empregos aumentam as taxas de emprego feminino, mas não as taxas de emprego masculinas. Por outro lado, menores taxas de propriedade do carro diminuem significativamente a proporção de emprego e aumentam a taxa de desemprego para ambos os sexos.

Grande parte dos estudos citados anteriormente encontraram evidências da existência da relação entre acessibilidade, segregação e emprego. Contudo, outros não tiveram resultados conclusivos ou não significativos (JENCKS E MAYER, 1990; HOLZER, 1991; IHLANFELDT e SJOQUIST, 1998). Isto se deveu principalmente a problemas metodológicos, tais como medidas inadequadas de acessibilidade ao trabalho, problemas de endogeneidade, pequenas amostras e dados agregados. Estudos posteriores introduziram melhores medidas de acessibilidade ao emprego (Rogers, 1997; Shen, 1998, Immergluck 1998; Johnson, 2006) e medidas de acesso ao transporte público ou privado (KAWABATA, 2003; ONG e MILLER, 2005; BAUM, 2009). Eles abordaram o problema de endogeneidade da localização residencial (WEINBERG, 2000 e 2004; GURMU *et al.*, 2008; ASLUND *et al.*, 2010).

Alguns estudos mostraram que o efeito da acessibilidade ao emprego desapareceu quando eles abordaram o problema da endogeneidade ou melhoraram as medidas de acessibilidade do trabalho. Sanchez *et al.*, (2004), Gurmu *et al.*, (2008) e Bania *et al.*, (2010) não encontraram relação entre a acessibilidade ao trabalho e o emprego para famílias pobres que receberam assistência temporária para famílias carentes (TANF) nos Estados Unidos. Os autores usaram várias medidas de acessibilidade ao emprego e buscaram atenuar o problema da endogeneidade entre o emprego e a localização residencial usando uma subamostra de indivíduos que receberam habitação pública. Durjandin *et al.*, (2008), investigaram os efeitos causais da organização espacial de Bruxelas sobre as propensões do desemprego. Usando dados censitários ao nível individual, a probabilidade de desemprego de jovens adultos foi estimada levando-se em conta as características pessoais, domésticas e de vizinhança. A endogeneidade dos locais residenciais foi resolvida pela restrição da amostra a jovens adultos que residem com seus pais, sendo o viés restante avaliado por meio de uma análise de sensibilidade. Os resultados sugeriram que o bairro de residência afeta significativamente a probabilidade de um jovem estar desempregado, porém a acessibilidade ao emprego não teve efeito sobre a probabilidade de desemprego.

Outro foco em análise relacionada com a hipótese do *mismatch* espacial é a informalidade presente em países em desenvolvimento. O presente trabalho relaciona-se com

uma vertente da literatura que estuda os determinantes da informalidade ao nível de trabalhadores (Gong e Van Soest, 2002; Marcouiller *et al.*, 1997). Estudos para países desenvolvidos não abordaram explicitamente a informalidade devido à sua pequena percentagem no mercado de trabalho desses países. Contudo, para países em desenvolvimento onde o setor informal é mais segregativo, Abane (1993) mostrou que em Acra (Gana) muitos trabalhadores informais, e particularmente as mulheres comerciantes, preferem viver perto de seus negócios, a fim de reduzir os custos de transportes. Na Índia, Mitra (2005) destacou que as mulheres no setor informal são limitadas a escolher um emprego próximo ao seu local de residência (cuja localização é determinada pelo marido), primeiro por causa das responsabilidades domésticas e, em segundo lugar, porque eles só podem acessar os empregos para contatos informais em sua vizinhança. Suárez *et al.*, (2016) mostraram para a cidade do México que os tempos de viagens observados para os trabalhadores de baixa renda (informais) têm o menor trajeto possível devido a sua localização residencial.

Por fim, a hipótese do *mismatch* espacial e a acessibilidade ao emprego permite estudar o mercado de trabalho urbano com base no *mismatch* educacional. Como já dito, o *mismatch* educacional no mercado de trabalho é caracterizada a partir da existência de indivíduos com escolaridade maior do que a requerida pela ocupação, os sobre-educados (*overeducated*), e indivíduos com escolaridade menor do que a requerida pela ocupação, os subeducados (RUBB, 2005). A existência de sobreeducação levanta questões interessantes acerca das teorias existentes sobre a dinâmica do mercado de trabalho e sobre sua ligação com o sistema educacional. Desde então, diversos estudos para diferentes países e períodos confirmam a incidência do *mismatch* educação-ocupação (FREEMAN, 1976; DUNCAN E HOFFMAN, 1981; CLOGG E SCHOKEY, 1984; VERDUGO e VERDUGO, 1989; ALBA-RAMIREZ, 1993; HARTOG, 2000; BAUER, 2002; MACGUINNESS, 2006).

Büchel e van Ham (2003) analisam como as oportunidades para diferentes níveis (características do mercado regional) e as restrições (até que ponto os candidatos a emprego estão restritos ao mercado regional) podem ajudar a explicar o fenômeno da supereducação. Os autores utilizaram um procedimento de duas etapas para controlar o acesso seletivo ao emprego e os resultados mostraram que o tamanho do mercado de trabalho é um fator importante para evitar a supereducação.

Hensen *et al.*, (2009) investigaram a relação entre mobilidade geográfica e incompatibilidade entre educação e trabalho na Holanda. Os autores se concentraram no papel da mobilidade geográfica na redução da probabilidade de os graduados trabalharem (i) empregos abaixo do nível de escolaridade; (ii) empregos fora de seu campo de estudo; (iii)

empregos de meio período; (iv) empregos flexíveis; ou (v) empregos pagos abaixo do salário esperado no início da carreira. Para isto, utilizaram dados sobre graduados do ensino médio e superior no período de 1996-2001. Os resultados mostraram que os graduados que são móveis têm maior probabilidade de encontrar emprego no nível educacional adquirido do que aqueles que não são. Além disso, os graduados em mobilidade têm maior probabilidade de encontrar empregos em período integral ou permanente. Isso sugere que a mobilidade é procurada para evitar não apenas a necessidade de aceitar um emprego abaixo do nível educacional adquirido, mas também outras incompatibilidades entre o ensino e o emprego; os graduados são espacialmente flexíveis, particularmente para garantir empregos em período integral.

Jauhiainen (2011) examina a influência da região e da mobilidade no risco de supereducação para homens e mulheres na Finlândia. Indivíduos com excesso de instrução são identificados com um método de medição estatística. No total, 9,6% de todos os trabalhadores são supereducados. Um modelo probit que controla o viés de seleção da amostra foi usado na análise empírica. As evidências sugeriram que a probabilidade de supereducação parece depender da região, em outras palavras, morar em um grande mercado de trabalho regional diminui a probabilidade de ser supereducado.

Iammarino & Martinelli (2012) analisam os determinantes em nível micro das correspondências educação-trabalho de recém-formados na Itália. A metodologia leva em consideração tanto a relação endógena entre migração e emprego, quanto o viés de auto-seleção entre emprego e educação - trabalho (des) correspondência. Usando uma pesquisa sobre a entrada de graduados italianos no mercado de trabalho, verificou-se que, embora se confirme que a migração em nível nacional tem um papel positivo tanto na procura de um emprego quanto na diminuição da probabilidade de excesso de escolaridade, surgem diferenças robustas quando se olha para a dimensão subnacional. De fato, as regiões do Norte, ao receber ingresso de graduados do Sul que conseguem obter uma boa correspondência entre educação e emprego nos mercados de trabalho receptores, estão aparentemente colhendo parte do retorno ao investimento em educação universitária no sul.

Devillanova (2013) estudam a relação entre mobilidade interna e supereducação. Usando uma grande pesquisa sobre o mercado de trabalho italiano, ele estima o efeito da flexibilidade espacial dos trabalhadores sobre a probabilidade de serem supereducados. A análise tenta lidar com duas possíveis causas de erros de especificação, que podem influenciar a correlação entre migração e supereducação para baixo: a endogeneidade da migração e a omissão de características relevantes do trabalho. O autor também lida com a seleção para emprego e controles para área e características pessoais, incluindo vários representantes para a

capacidade do indivíduo. Os resultados mostraram que o deslocamento está correlacionado positivamente com a qualidade da correspondência educação-emprego. Pelo contrário, a sabedoria convencional de que a migração interna reduz inequivocamente a incidência de supereducação não recebe apoio empírico. A correlação negativa entre migração e supereducação desaparece quando as características do trabalho são incluídas na análise e se torna positiva quando a migração é instrumentalizada. Por fim, o autor sugere que a relação entre migração interna e supereducação permanece incerta e que mais pesquisas são necessárias para melhor fundamentar as prescrições políticas.

Ramos & Sanromà (2013) analisaram a influência de variáveis individuais e algumas características de mobilidade espacial dos mercados de trabalho regionais na supereducação na Espanha. Para atingir esse objetivo, utilizaram os microdados da Pesquisa sobre Orçamento Domiciliar da Espanha para estimar um modelo *probit* para a probabilidade de supereducação, enquanto consideraram a seleção da amostra e a presença de dados em diferentes níveis (indivíduos e território). Os resultados obtidos permitiram concluir que o tamanho dos mercados de trabalho locais e a possibilidade de estender a procura de emprego para outros mercados de trabalho pendulares são fatores relevantes para explicar a supereducação no mercado de trabalho espanhol. Apesar das diferenças nas instituições do mercado de trabalho, os resultados foram muito semelhantes aos obtidos em outros países, principalmente ao trabalhar com dados mais recentes.

Di Paolo *et al.*, (2014) estende o estudo anterior de Matas *et al.*, (2010), considerando não apenas o efeito da acessibilidade do emprego na probabilidade de estar empregado, mas também seu impacto potencial em um aspecto extremamente relevante da qualidade do emprego, a saber, a correspondência entre a educação completa de um indivíduo e a sua ocupação alcançada (ou seja, emprego) *mismatch* educacional. Os resultados indicaram que, em geral, a acessibilidade ao emprego é importante tanto para o *mismatch* no emprego quanto na educação. Melhorar a acessibilidade ao emprego pode ser visto como um aumento no tamanho do mercado de trabalho (local) e, conseqüentemente, melhoraria o desempenho do mercado de trabalho.

Para o Brasil, apesar da reduzida literatura relacionando a acessibilidade aos empregos com resultados no mercado de trabalho, podem-se citar alguns estudos. Analisando a participação no mercado e o desemprego, tem-se o estudo de Campos (2015) que buscou identificar o efeito da decisão de localização do trabalhador na busca por emprego na cidade de Salvador. A partir dessa teoria do *mismatch* espacial foram especificados dois modelos econométricos que analisaram (i) o efeito da distribuição espacial na probabilidade do

trabalhador estar ou não ocupado, e (ii) o efeito da decisão de localização no tempo em que o trabalhador passa desempregado. As estimações utilizaram a base de microdados da Pesquisa de Emprego e Desemprego (PED) para a Região Metropolitana de Salvador no período de 1997 a 2014. Os autores chamam a atenção para o problema de endogeneidade existente quanto à escolha do local de residência poder estar relacionada à quantidade de ofertas de emprego existentes na região, o que resultaria em mais rápida saída do desemprego. Dessa forma, seguindo Andersson *et al.*, (2014), os autores filtraram os dados com trabalhadores que buscam emprego porque foram demitidos do último trabalho. Considerou-se que a decisão prévia do local da residência era exógena à duração do emprego. Os resultados alcançados mostraram que existem efeitos da distância sobre os resultados na busca por emprego, onde trabalhadores residentes nos bairros pertencentes à região do Miolo e do Subúrbio Ferroviário tiveram maiores dificuldades para encontrar um emprego do que os demais habitantes da cidade. Estas regiões do Miolo e Subúrbio Ferroviário apresentavam níveis educacionais e renda domiciliar mais baixos do que nas demais áreas da cidade, apontando para a existência de um ambiente de vulnerabilidade social.

Estudo posterior ampliou a análise para várias regiões como o estudo de Barufi e Haddad (2017) que investigaram a relação entre *mismatch* espacial, salários e desemprego em grandes áreas do Brasil. Este trabalho teve como objetivo investigar se a relação negativa entre *mismatch* espacial e resultados do mercado de trabalho é válida no Brasil após o controle de características individuais. A estratégia de identificação foi baseada em hipóteses mais restritivas, presumindo que no curto prazo os preços no mercado de trabalho estivessem próximos do nível de equilíbrio, e trabalhadores e empresas são relativamente imóveis. Além disso, os autores apontaram que a medição das oportunidades de emprego locais poderia ser indireta (usando a suposição de que existe um centro geográfico na cidade ou considerando o tempo de deslocamento como uma possível medida da distância aos empregos). O local específico das oportunidades de emprego não foi incluído na análise, o que significa que esse problema de endogeneidade pode ser menos relevante. As conclusões indicaram que não houve relação clara entre diferentes medidas de acessibilidade a empregos e a probabilidade de estar desempregado. No entanto, para os salários, houve uma correlação clara, que é mais forte nas grandes áreas metropolitanas do país.

Analisando a informalidade, tem-se o trabalho de Motte *et al.*, (2016) que descobriu para a Região Metropolitana do Rio de Janeiro que os trabalhadores informais viajam com menos frequência do que os trabalhadores formais e fazem viagens mais curtas, e isso está relacionado ao fato de que os empregos informais são mais descentralizados do que os

empregos formais. Moreno-Monroy e Ramos (2015) estimaram o impacto das expansões do sistema de transporte público sobre as taxas de informalidade para a Região Metropolitana de São Paulo (RMSP) entre 2000 e 2010. Os autores comparam as mudanças médias na informalidade em áreas que receberam novas infraestruturas de transporte público com as mudanças médias em áreas que deveria receber infraestrutura de acordo com os planos oficiais, mas não receberam por atrasos. Após controlar a seleção endógena através do método de diferença em diferença utilizando como instrumento exógeno um plano de rede histórico de 1968 para mudanças no acesso ao transporte, acham que a informalidade diminuiu em média 16% mais rápido em áreas que recebem novas infraestruturas de transporte público em comparação com áreas que enfrentaram atrasos no projeto.

Boisjoly *et al.*, (2017) exploraram a relação entre a informalidade do trabalho e a acessibilidade ao emprego nos transportes públicos na RMSP, no Brasil. Os autores consideraram um importante preditor de ser informal que não foi incluído na regressão, a renda. Incluir a variável resultaria em estimativas tendenciosas devido à endogeneidade, visto que os indivíduos poderiam ser informais porque têm baixos rendimentos, mas também poderiam ter baixos rendimentos porque são informais. Para resolver esta questão, os autores dividiram a amostra em duas categorias com base no salário mínimo oficial no Brasil em 2010 (ou seja, se a renda mensal nominal do trabalhador foi inferior ao salário mínimo ou mais) e correu o mesmo modelo de regressão para cada categoria. Os resultados da regressão mostraram que para os trabalhadores que ganham menos do que o salário mínimo, um maior nível de acessibilidade aos empregos pelos transportes públicos está associado a uma menor probabilidade de ser um trabalhador no setor informal.

Também analisando a Região Metropolitana de São Paulo, contudo fazendo uma análise por gênero, Silveira Neto *et al.*, (2015) relatam pesquisas sobre os determinantes do tempo de deslocamento para o trabalho. Os resultados obtidos indicaram que o estado civil é um preditor significativo do tempo de deslocamento para as mulheres. Em relação aos homens, as mulheres têm maior probabilidade de ter trajetos mais curtos, independentemente da estrutura da família, o efeito é mais pronunciado para as mulheres casadas, especialmente em famílias com dupla renda. Como dito na seção anterior, o gênero tem sido relatado como fator importante nas decisões de viagem ao trabalho. As mulheres geralmente têm viagens mais curtas (Hanson e Pratt, 1995; Madden, 1981; Turner e Niemeier, 1997; White, 1986) e, em muitos casos, especialmente as mães solteiras de baixa renda (Chapple, 2001) buscam emprego informal próximo ao local de residência.

Finalmente para a literatura de sobre-educação, pode-se dizer que a mesma está bem documentada nos países desenvolvidos. Groot e Maassen van den Brink (2000) e Rubb (2003) fornecem uma ampla metanálise dos resultados sobre a incidência e o impacto do *mismatch* educacional nesses países. Contudo, apesar de sua importância para compreensão de fenômenos sociais em andamento no país, bem como para a condução de políticas educacionais eficientes, a sobre-escolarização foi pouco estudada no Brasil. Machado *et al.*, (2003), utilizando dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) para o período de 1981 a 2001, corroboram a literatura apontando que há uma tendência no Brasil de que pessoas mais bem preparadas ocupem posições que antes eram exercidas por trabalhadores menos qualificados. Ademais, postulam que há uma relação inversa entre a idade e *mismatch* educacional e que o descasamento é maior nas categorias sócio-ocupacionais manual e doméstica. Os dados da PNAD também são empregados por Santos (2002) em uma especificação do tipo Verdugo e Verdugo (1989). Os anos avaliados são 1992, 1995, 1997 e 1999. Os resultados encontrados pelo autor condizem com os modelos de discriminação salarial da mulher no mercado de trabalho, com o prêmio salarial feminino no caso da sobre-escolarização sendo inferior aos dos homens. Além disso, é constatado que o efeito da incompatibilidade é mais evidente para as ocupações de nível médio, não obstante também estar presente em ocupações que demandam nível superior de ensino.

Diaz e Machado (2008) avaliaram a incidência e os retornos da sobreeducação e da subeducação no Brasil, nas Grandes Regiões e no Estado de São Paulo. Também foram analisadas as diferenças por gênero e por grande grupo ocupacional. Foram utilizados dados do Censo de 2000 e da Classificação Brasileira das Ocupações de 2002. Os autores utilizaram a perspectiva do *Job Analysis* e a regressão proposta por Duncan e Hoffman (1981). Os resultados mostraram que os retornos à escolaridade requerida são maiores do que os da educação que o trabalhador possui, e, para o Brasil e suas macro-regiões, os retornos da escolaridade requerida são maiores do que os da sobre-escolarização. Uma conclusão interessante apresentada no estudo é a de que, no Brasil, o retorno da sobre-educação, para homens e mulheres, é cerca de três quartos do retorno da escolaridade requerida. Assim, há um incentivo maior a ser sobre-escolarizado no país relativamente aos outros países, nos quais essa razão é de metade a dois terços conforme citado anteriormente.

Viana e Oliveira (2010) utilizando dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) procuraram mensurar o *mismatch* educacional e estimar o papel dos efeitos de ciclo de vida, de conjuntura do mercado e de geração através de um modelo de idade-período-coorte (IPC). Os resultados encontrados apontaram para uma predominância

dos efeitos de conjuntura para os homens, enquanto as trabalhadoras mulheres sofreram mais pela influência dos efeitos ciclo de vida e de geração.

Machado e Oliveira (2013) analisaram os determinantes da mobilidade sócio-ocupacional no mercado de trabalho metropolitano brasileiro, enfocando o efeito de condições cíclicas sobre os movimentos ascendente e descendente dos indivíduos. A análise é feita com base no *status* de incompatibilidade educacional dos indivíduos nas ocupações, verificado a partir da comparação entre a escolaridade observada dos indivíduos e a escolaridade requerida numa dada ocupação. A estratégia de identificação dos modelos baseou-se na construção de um pseudopainel no nível de coortes de nascimento, utilizando os dados da Pesquisa Mensal do Emprego (PME), do Instituto Brasileiro de Geografia e Estatística (IBGE), no período 2002-2008. A principal evidência obtida foi que a ascensão ocupacional dos sobre-educados apresenta um comportamento pró-cíclico em relação às flutuações na taxa de desemprego no mercado de trabalho metropolitano. Esse resultado contribui para o entendimento dos determinantes da mobilidade socioeconômica e da incompatibilidade educacional no país.

2.4 Estratégia Empírica

A finalidade desta seção é apresentar a metodologia empírica e o banco de dados utilizados na pesquisa, além de explorar o possível problema de endogeneidade decorrente da determinação simultânea das variáveis de interesse: indicador de acessibilidade, participação no mercado, condição de desempregado, na informalidade e *overeducated*. O objetivo do estudo exige, assim, uma estratégia de identificação para a possível endogeneidade existente entre a localização residencial e a mensuração das oportunidades de trabalho.

2.4.1 Área de estudo e dados

A área de estudo é a Região Metropolitana de São Paulo (RMSP), a maior metrópole brasileira com uma população estimada em 21,5 milhões em 2018 pelo IBGE. A RMSP é geograficamente dividida pelos rios Tietê e Tamanduatei, correndo do Noroeste para o Sudeste da região. A unidade de análise é o indivíduo dentro da Área Espacial de

Ponderação¹⁸ (AEP), que é definida como uma área composta por um conjunto de zonas mutuamente exclusiva, projetado para dar a robustez estatística necessária à estratégia de amostragem do Censo de População (ver Figura 2.1). 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.

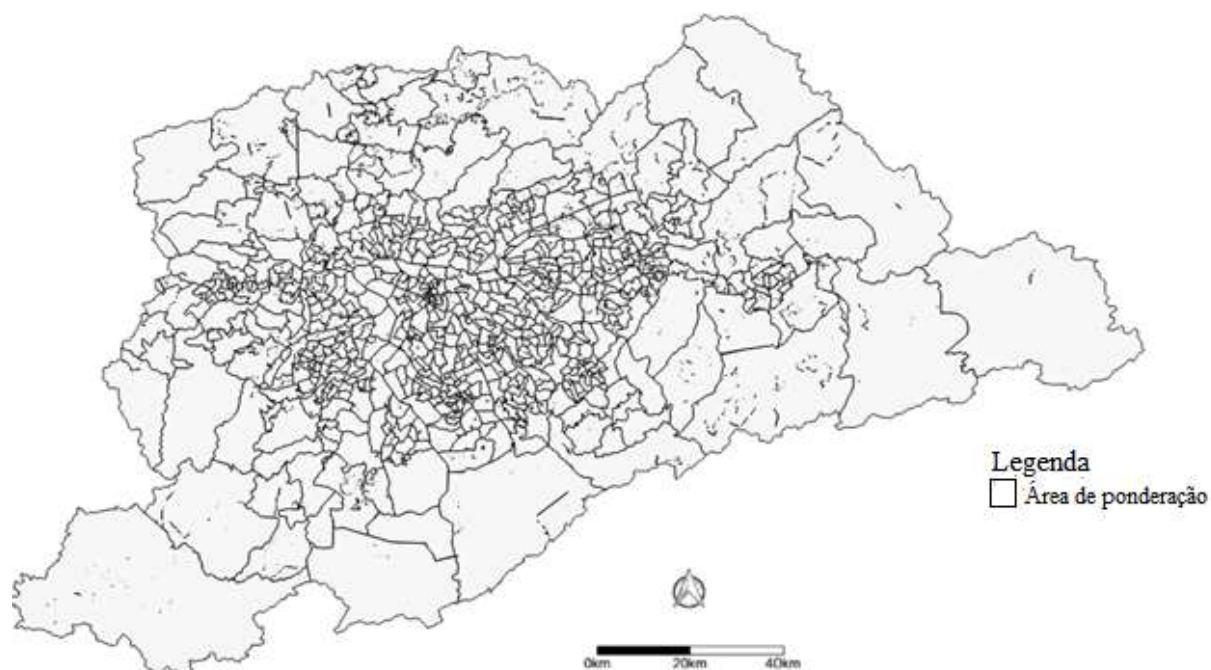


Figura 2.1 – Áreas de Ponderação da região metropolitana de São Paulo.
Fonte: Elaboração própria com base nos dados do Censo 2010.

Os microdados para as características individuais são obtidos no Censo de População de 2010 (Instituto Brasileiro de Geografia e Estatística - IBGE, 2010), juntamente com a Classificação Brasileira das Ocupações domiciliar e as descrições da Classificação Brasileira das Ocupações de 2002. O IBGE realiza, a cada dez anos, um Censo Demográfico, com desagregação regional ao nível municipal (ou ao nível da área censitária dos municípios maiores). O Censo Demográfico coleta informações sobre as principais características dos indivíduos e famílias, fornecendo detalhes sobre as condições de vida da população em cada município e servindo como um instrumento de política muito importante em um país. Um questionário mais curto aplica-se a toda a população ao nível do setor censitário, enquanto as características individuais específicas são investigadas em um conjunto mais longo de questões que são dadas a uma amostra e são representativas ao nível das áreas ponderadas

18 Define-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).

(conglomerados de setores censitários com pelo menos 400 domicílios). 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 usamos 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).

2.4.2 Especificações Econométricas

A estratégia empírica desenvolvida baseia-se na estimativa da relação entre diferentes medidas de acessibilidade e resultados do mercado de trabalho (probabilidade de participação no mercado, de estar desempregado, de estar na informalidade e na condição de *overeducated*). 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. Portanto, a equação a ser estimada é com base em um modelo de probabilidade, da seguinte maneira:

$$Prob[y_{ij} = 1] = F[\beta_1 + \beta_2 AC_j + \beta_3 X_{ij} + \beta_4 H_{ij} + \beta_5 K_j] \quad (1)$$

Nesta especificação, y_i refere-se ao *status* de emprego (é igual a 1 quando uma pessoa está na condição de interesse, PEA, desempregada, na informalidade ou na situação de *overeducated*) e F é uma função de probabilidade. 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, se o indivíduo é casado, e se a pessoa é ou não a chefe da família), 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.

Uma regressão linear simples que representa um Modelo de Probabilidade Linear estimado por OLS pode não ser apropriada, pois os termos de erros são heterocedásticos (Hair *et al.*, 2009), podendo ocorrer que as probabilidades condicionais calculadas não pertençam ao intervalo fechado 0 e 1 (DIAS FILHO e CORRAR, 2012). Por outro lado, a menor exigência com respeito à distribuição dos resíduos representa uma vantagem em relação a modelos mais tradicionais com variáveis dependentes binárias. Portanto, o estudo também considerou uma especificação projetada para controlar as exigências de modelos com variáveis dependentes binárias. Dessa forma, a probabilidade de se observar $y_i = 1$ é dada pelo modelo Probit (função de distribuição cumulativa da distribuição normal padrão), ou seja,

$$Prob[y_i = 1] = \int_{-\infty}^{\beta'x} \Phi(t) d_{i=1} = \Phi(\beta'x) \quad (2)$$

onde a função $\Phi(\cdot)$ é a notação usual para a distribuição normal padrão cumulativa e β é o vetor de parâmetros das variáveis explicativas consideradas. No modelo *Probit*, é postulado a distribuição normal reduzida e $F(\cdot)$ designa, então, a função de distribuição normal,

$$\Phi(x) = \int_{-\infty}^{\beta'x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \quad (3)$$

a função de densidade de probabilidade associada é dada, como é sabido, por:

$$\Phi(x) = \frac{d\Phi(x)}{dx} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (4)$$

A equação geral do Modelo em função das variáveis explicativas adotadas é representada por:

$$\Phi(\beta'x) = \Phi(\beta_1 + \beta_2 AC_j + \beta_3 X_{ij} + \beta_4 H_{ij} + \beta_5 K_j) \quad (5)$$

De acordo com Stock e Watson (2004), a expressão $\beta'x$ no *Probit*, desempenha o papel de “z”, na tabela de distribuição acumulada normal padrão. Posteriormente, serão estimados os efeitos marginais, que serão considerados nas análises, a seguir. Os efeitos marginais representam, por exemplo, o quanto uma mudança na variável de acessibilidade pode afetar a probabilidade de um indivíduo participar no mercado de trabalho. Os efeitos marginais (a variação da probabilidade de o evento ocorrer quando uma variável independente é modificada) são calculados por meio dos coeficientes estimados β 's.

$$\frac{\partial E[y|x]}{\partial x} = \Phi(x'\beta)\beta \quad (6)$$

É claro que modelos padronizados como a equação (5) tornam possível controlar algumas características individuais e domiciliares que podem influenciar tanto a escolha da vizinhança quanto os resultados individuais no mercado de trabalho. No entanto, é provável que algumas características individuais e familiares que não são observadas pelo pesquisador (e, portanto, não incluídas em X_i ou H_i), influenciem tanto o resultado do interesse quanto a escolha da vizinhança. Por exemplo, indivíduos com baixa adesão ao mercado de trabalho (que influencia diretamente a probabilidade de desemprego, informalidade, menor participação no mercado de trabalho e maior chance de *overeducated*) possam optar por residir em bairros pobres por alguma razão econômica ou social. Como consequência, o que o pesquisador percebe como um efeito de vizinhança através do parâmetro estimado β_5 pode simplesmente resultar de um efeito correlacionado que reflete a escolha residencial.

Como o estudo baseia-se em um modelo em que a estrutura espacial da cidade passa a ser determinada endogenamente como resultado de interações entre firmas e famílias (Fujita & Ogawa, 1982), as equações acima passam a estar sujeitas a um possível viés de simultaneidade nas estimações empíricas com medidas de acessibilidade. Nesse modelo, as firmas e famílias são livres quanto à escolha locacional dentro da cidade. No entanto, esta escolha por residência é dependente da localização das firmas, e estas, por sua vez, fazem suas opções com base na localização de outras firmas. Sob as firmas, os ganhos de aglomeração são opostos aos crescentes custos de produção. Sob as famílias, o custo de transporte e o tempo despendido são compensados pelo decréscimo dos aluguéis da terra (FUJITA & OGAWA, 1982; LUCAS & ROSSI-HANSBERG, 2002).

Assim, a acessibilidade aos empregos, que é a variável explicativa de interesse na análise, depende de alguns dos fatores que afetam as variáveis dependentes e também podem ser indiretamente afetados. Neste caso, o viés de simultaneidade surge à medida que os residentes e os locais das empresas são determinados ao mesmo tempo (HADDAD & 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 participação da força de trabalho, menor a chance de estar desempregado, na informalidade ou na condição de *overeducated*. 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 (MELO & GRAHAM, 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. Essas preocupações podem ser potencialmente eliminadas com uma estratégia de variáveis instrumentais (IV). Essa estratégia pode ser especificada como uma função linear de uma covariável exógena (instrumento) z e um componente aleatório μ_i .

$$AC_j = \beta_0 + \beta_1 z_j + \beta_2 X_{ij} + \beta_3 H_{ij} + \beta_4 K_j + \mu_j \quad (7)$$

Para alcançar a consistência, o instrumento z deve atender a duas condições: estar associada ao indicador de acessibilidade (ou seja, ter poder de explicação em uma regressão do primeiro estágio com variável dependente AC_j) e satisfazer a restrição de afetar o resultado exclusivamente através do indicador no segundo estágio. Na próxima seção será apresentada a medida de acessibilidade e, em seguida, o instrumento utilizado no primeiro estágio das regressões.

2.4.3 Medidas de acessibilidade

Uma questão fundamental nos estudos de *mismatch* espacial é como medir a acessibilidade residencial às oportunidades de emprego. Seguindo Rogers (1997), essa variável deve levar 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. Essas medidas são utilizadas na pesquisa, pois fornecem uma medida de acessibilidade fácil de interpretar (GEURS & VAN WEE, 2004). As medidas de oportunidade cumulativa são frequentemente usadas para medir a distribuição espacial da acessibilidade, conforme feito, por exemplo, por Wang e Chen (2015). Este tipo de medida conta o número de trabalhos que podem ser alcançados a partir de um AEP e é calculado da seguinte forma:

$$AC_j = \sum_{k=1}^n \frac{E_k}{d_{jk}^\alpha(t)} \quad (8)$$

Na equação (4), AC_j é a acessibilidade da região j (centróide da AEP), E_k corresponde à quantidade de empregos existentes em qualquer região k (centros de empregos), enquanto d_{jk} é o impedimento (por exemplo, distância) existente para viajar entre j e k .

A função de impedância pode assumir diferentes formas funcionais. Na literatura empírica (Quintanar, 2013; Carvalho, 2016; Haddad e Barufi, 2017), as funções mais utilizadas são a função de potência e a função exponencial, sendo α o parâmetro de decaimento. Este parâmetro mede a relação entre padrões de interação observados (comutação) e distância (tempo) quando outros determinantes de interação são constantes. Um valor de α muito elevado indica, assim, que a distância ou o tempo desencorajam mais a interação. Neste estudo, foram considerados diferentes medidas para d^α como forma de robustez.

Os valores de E_j foram obtidos da base de dados do Núcleo de Estudos Regionais e Urbanos (NERU) da UFPE, que utiliza 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 2.2). Foram utilizados a localização de cada empresa para obter informações sobre o local de trabalho. Como em Rodrigues *et al.*, (2019), ressalta-se uma limitação da RAIS que é a falta de informações sobre trabalhadores que não estão formalmente empregados. Essa é uma característica importante no contexto brasileiro, onde as taxas de informalidade ultrapassam 40% de todos os trabalhadores. A informalidade é uma característica importante no Brasil, abrangendo 41,1% dos trabalhadores em 2016. Além disto, é a principal ocupação da população de 11 estados brasileiros e sendo a principal ocupação de mais de 40% da população em 21 estados (IBGE, 2019). Atualmente, o setor de trabalho informal em países como o Brasil pode atingir até 50% do mercado de trabalho em algumas cidades (BOISJOLY *et al.*, 2017).

Contudo, como bem argumentam Rodrigues *et al.*, (2019) a parcela de empregos formais nas regiões metropolitanas é bem superior àquela da média nacional. Segundo dados do Censo 2010, o percentual de empregos formais da RMSP corresponde a 63,3% do total de empregos. Além disto, a distribuição espacial dos trabalhadores informais é semelhante aos trabalhadores formais (Rodrigues *et al.*, 2019); portanto, não há perda considerável de informações ao identificar os subcentros de emprego.

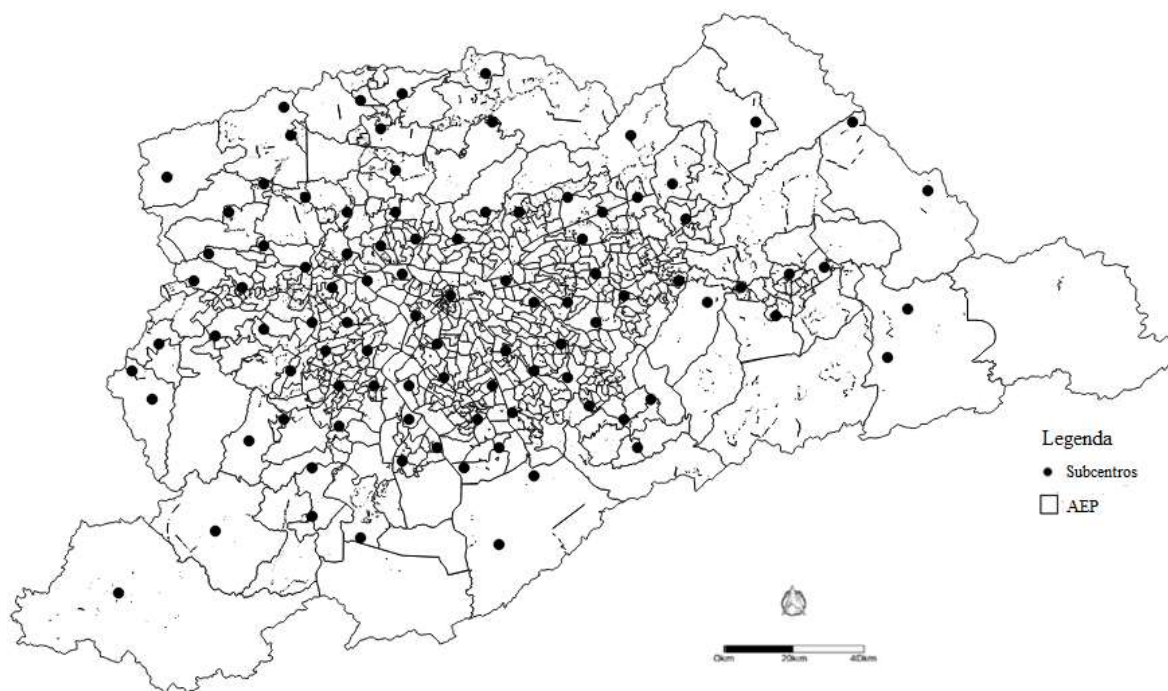


Figura 2.2 – Subcentros das áreas de Ponderação da região metropolitana de São Paulo.
Fonte: Elaboração própria com base nos dados da RAIS 2010.

Há diferentes formas de identificação de centros e subcentros de emprego urbano (FERNANDEZ-MALDONADO *et al.*, 2015). 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 & MCMILLEN, 1990; CRAIG & 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 a 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:

$$y_i = g(DCBD_i) \quad (9)$$

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 9, 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 tricúbico¹⁹.

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:

$$y_i = g(DCBD_i) + \sum_{j=1}^S (\delta_{1j} D_{ij}^{-1} + \delta_{2j} D_{ij}) + u_i \quad (10)$$

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

¹⁹ Deixe $d_i(x)$ 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 1.

$$\kappa_i = \left(1 - \left(\frac{d_i(x)}{d_q(x)} \right)^3 \right)^3 I(d_i(x) < d_q(x))$$

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 $(x_4 - x_0)/4$. Neste caso, pode-se representar a função cúbica dos splines como:

$$g(DCBD_1) = \alpha + \beta_1(x_i - x_0) + \beta_2(x_i - x_0)^2 + \beta_3(x_i - x_0)^3 + \gamma_1 D_1(x_i - x_1)^3 + \gamma_2 D_2(x_i - x_2)^3 + \gamma_3 D_3(x_i - x_3)^3 + \varepsilon_i \quad (11)$$

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 10, a equação 11 é estimada por mínimos quadrados ordinários (OLS). 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 10 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 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.

2.4.4 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 no Quadro 2.1, a seguir. Entre as variáveis contidas no quadro, a maioria foi criada em nível de indivíduos por área de ponderação com base no Censo Demográfico 2010, produzido pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Conforme discutido na seção 2.4.3, as variáveis de distância para construir o índice de acessibilidade ao emprego foram obtidas a partir do *software* QGIS, tomando as coordenadas geográficas.

Primeiramente, têm-se as variáveis dependentes das estimativas (PEA, desemprego, informal e overeducated) referentes à condição do indivíduo no mercado de trabalho que, como se viu, por determinar um maior ou menor custo associado ao deslocamento para o trabalho, pode afetar a localização das famílias. São também consideradas nas estimativas um conjunto de características pessoais (sexo, idade, raça e escolaridade), além do estado civil, que podem refletir diferenças nas preferências sobre a localização na cidade entre os indivíduos.

Quadro 2.1 – Descrição das variáveis dos modelos econométricos.

Amostra (pessoas da RMSP)	
VARIÁVEIS	DESCRIÇÃO
<i>Variáveis dependentes</i>	
PEA	Se a pessoa está na PEA (15 a 64 anos)
Desemprego	Se a pessoa está na PEA e desempregada
Informal	Se a pessoa está ocupada e na informalidade (sem carteira de trabalho)
Overeducated	Se a pessoa tem educação maior que a exigida no emprego
<i>Variáveis explicativas</i>	
LnACi	Log da Acessibilidade (distância euclidiana) para CBD e Subcentros
LnACi_IV	Log da Acessibilidade (distância através dos rios) para CBD e Subcentros
Masculino	Variável <i>dummy</i> se for do sexo masculino
Idade	Em anos
Idade ²	<i>Proxy</i> para não linearidade da experiência
Educação	
i) Baixo	i) Sem ensino fundamental completo,
ii) Intermediário	ii) Com ensino fundamental completo, mas sem médio completo,
iii) Alto	iii) Com ensino médio completo, mas sem superior,
iv) Bem elevado	iv) Com superior completo ou mais
Casado	Variável <i>dummy</i> se for casado
Chefe	Variável <i>dummy</i> se for chefe da família
Raça	Branco Pardo Negro Amarelo
Tamanho da família	Numero de pessoas por domicilio
Tipo de Vizinhaça	Muito Carente Carente Médio Desenvolvida Muito Desenvolvida
Acesso à infraestrutura	Coleta de lixo Saneamento básico

Fonte: Elaboração própria.

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 (LEE & MACDONALD, 2003; CRAINE, 2007). De forma similar, a condição de casado pode atribuir um caráter mais permanente à localização dos domicílios e, assim, menor vinculação imediata ao local de trabalho. 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.

2.4.5 Definição das Características da vizinhança

Dois tipos de indicadores são usados para explicar as características da vizinhança que podem potencialmente influenciar as probabilidades individuais de se estar na condição de participar no mercado de trabalho, de estar desempregado, na informalidade e de se estar na condição de *overeducated*. O primeiro tipo relaciona-se com a teoria do *mismatch* espacial que caracteriza a desconexão de bairros a empregos, enquanto o segundo tipo relaciona-se com a composição social dos bairros.

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, visto que considerar todas essas características individuais juntas em uma única regressão podem causar problemas de colinearidade, já que muitos indicadores de composição de vizinhança são altamente inter-relacionados (O'REGAN & QUIGLEY, 1996; MORENOFF & SAMPSON, 1997; JOHNSTON *et al.*, 2004). Para contornar esse problema, foi utilizados métodos de análise fatorial padrão (Johnston, 1978; Durjandin *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 na seção 6. 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 2.1). Especificamente, para as variáveis que representam o nível educacional, foi considerado “baixo” 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.

Tabela 2.1 -Variáveis usadas na análise de componentes principais e seus fatores.

	Fator 1	Fator 2
Autovalores	4,410	1,313
Porcentagem de variação explicada	55,14	16,42
<i>Variáveis</i>		
% 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.

Em seguida, as áreas são agrupadas de acordo com suas coordenadas nos eixos fatoriais, usando uma classificação hierárquica ascendente (método Ward²⁰, que minimiza a variância intragrupo). Como no trabalho de Durjandin *et al.* (2008), foram obtidos cinco tipos de vizinhança (Tabela 2.2). Este foi o número ideal de *clusters* de acordo com vários critérios, incluindo o critério de agrupamento cúbico.

Tabela 2.2 – Média das características dos tipos de vizinhança

	Muito Carente	Carente	Média	Desenvolvido	Muito Desenvolvido	Total
% de Mulheres chefes	11,7	13,3	10,9	15,1	20,1	12,9
% de pessoas com baixo nível educacional	72,8	60,3	79,1	46,5	29,4	64,6
% de pessoas com ensino intermediário	22,9	26,9	17,8	26,4	22,6	23,9
% de pessoas com alto nível educacional	4,2	12,6	3,1	27,1	47,9	11,4
% de pessoas desocupadas	5,1	3,9	4,8	2,9	2,3	4,4
% de domicílios sem acesso à rede de saneamento	17,6	6,1	6,2	1,9	0,1	15,1
% de pessoas sem acesso a coleta de lixo	0,3	0,0	2,7	0,0	0,0	0,3
% de pessoas com menos de 15 anos	25,5	20,1	27,8	15,4	11,2	22,3

Nota: Os cinco tipos de vizinhança foram definidos por uma classificação hierárquica ascendente com o Critério de Ward.

²⁰ Segundo Hair *et al.*, (2005), o método de Ward consiste em um procedimento de agrupamento hierárquico no qual a medida de similaridade usada para juntar agrupamentos é calculada como a soma de quadrados entre os dois agrupamentos feita sobre todas as variáveis. Esse método tende a resultar em agrupamentos de tamanhos aproximadamente iguais devido a sua minimização de variação interna. Em cada estágio, combinam-se os dois agrupamentos que apresentarem menor aumento na soma global de quadrados dentro dos agrupamentos.

Seja $x_{l,j,k}$: valor para a variável p na observação j pertencente ao cluster l . A soma total dos erros quadrados (agrupando os clusters l e i), será $SS_{l,i} = \sum_{k=1}^{n_l} \sum_{j=1}^p (x_{l,k,j} - \bar{x}_j)^2 + \sum_{k=1}^{n_i} \sum_{j=1}^p (x_{i,k,j} - \bar{x}_j)^2$.

O primeiro tipo de vizinhança compreende áreas muito carentes com percentual alto da população com baixo nível educacional e sem acesso a saneamento básico. Estas áreas também possuem os níveis de desemprego mais alto de toda a aglomeração. Eles estão socialmente próximos a um grupo de áreas carentes que apresentam características semelhantes, mas com uma proporção menor de famílias sem acesso à saneamento e uma situação menos grave em termos de educação e desocupação. O terceiro grupo têm um nível socioeconômico mais baixo, em particular uma super-representação de trabalhadores e indivíduos com menor escolaridade. No entanto, a taxa de desocupação está próxima da média da região. Os dois grupos restantes são caracterizados em média por níveis mais altos de educação e percentuais baixos de domicílios sem acesso a saneamento básico e com pessoas com menos de 15 anos. Os tipos de vizinhança obtidos são apresentados na Figura 2.3, a seguir.

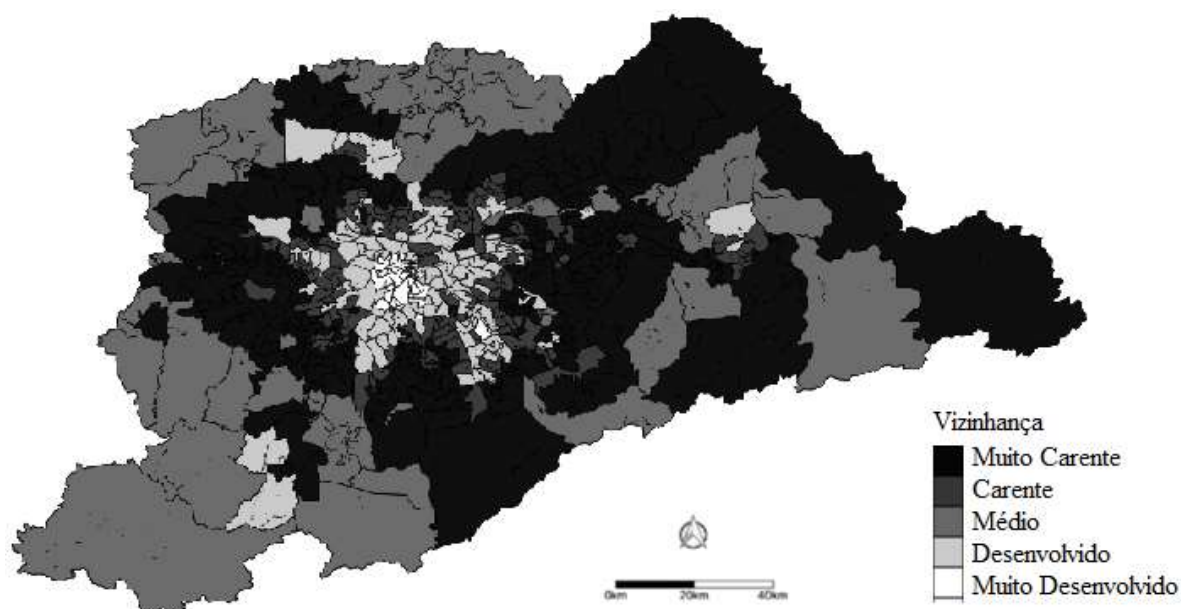


Figura 2.3 – Tipologia de vizinhança na RMSP.

Fonte: Elaboração própria com base nos dados do Censo 2010.

2.4.6 A Endogeneidade da Localização Residencial e Variável Instrumental

Analisar a relação entre os resultados individuais dos mercados de trabalhos urbanos e a localização residencial levanta a questão da endogeneidade, uma vez que envolve as escolhas de localização (GLAESER, 1996; DIETZ, 2002). Notadamente, é comum que indivíduos com características socioeconômicas semelhantes obtenham resultados parecidos no mercado de trabalho, visto que tendem a residir em certas áreas do espaço urbano. Por

exemplo, indivíduos com empregos bem remunerados escolherão residir em bairros com melhores condições sociais. Há, portanto, uma causalidade reversa, ou seja, 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 (DURJANDIN *et al.*, 2008). Alguns modelos tornam possíveis controlar características individuais e domiciliares que podem influenciar tanto a escolha da vizinhança quanto os resultados individuais (DURJANDIN *et al.*, 2008). No entanto, é provável que algumas características individuais e familiares que não são observadas pelo pesquisador influenciem tanto o resultado de interesse quanto a escolha da vizinhança.

Várias estratégias foram desenvolvidas para corrigir a endogeneidade da escolha da vizinhança (Chetty *et al.*, 2014; Chetty & Hendren, 2018) como por exemplo, programas de realocação subsidiados pelo governo dos Estados Unidos, onde permitiram obter estimativas mais confiáveis dos efeitos da vizinhança, uma vez que as famílias foram transferidas de uma área para outra através de uma intervenção exógena (OREOPOULOS, 2003). Contudo, devido à escassez de tais experimentos, os pesquisadores muitas vezes são obrigados a recorrer a outras estratégias. Alguns estudos restringem sua amostra a jovens²¹ 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 (O'REGAN & QUIGLEY, 1996, 1998; DURJANDIN *et al.*, 2008). No entanto, essa abordagem não elimina completamente o viés de endogeneidade, pois podem existir características parentais não observadas que determinam sua escolha residencial e também influenciam os resultados de emprego de seus filhos adultos (GLAESER, 1996).

O principal desafio da abordagem adotada nesta pesquisa é a forte preocupação teórica (Fujita e Ogawa, 1982; Lucas e Rossi-Hansbeng, 2002) de que, os resultados no mercado de trabalho são determinados simultaneamente com a acessibilidade aos empregos, onde este será potencialmente correlacionado com o termo de erro, podendo levar ao viés e inconsistência na estimativa de OLS. A solução para este problema exige encontrar um instrumento adequado para a medida de acessibilidade.

Desse modo, a estratégia de variáveis instrumentais baseia-se na inclusão de uma variável geográfica como o instrumento da acessibilidade à empregos, a saber, neste caso, a

²¹ Amostra estudada é restrita a jovens participantes da força de trabalho entre 19 e 25 anos que residem com os pais.

distância através do rio para as oportunidades de emprego (CBD e subcentros). Quanto à justificativa para a escolha desse instrumento, baseia-se na mesma adotada no estudo de Haddad e Barufi (2017). A geografia atuou como determinante da localização da infraestrutura de transporte na região, de modo que a rede rodoviária da cidade apresenta forte correlação espacial com "vias navegáveis" pré-urbanas. Atualmente, a grande maioria dos rios e riachos foi cobertos por asfalto e cimento e são praticamente desconhecidos pelos agentes econômicos (Figura 2.4). *A priori* o caso para a exogeneidade do instrumento parece, pois, plausível.

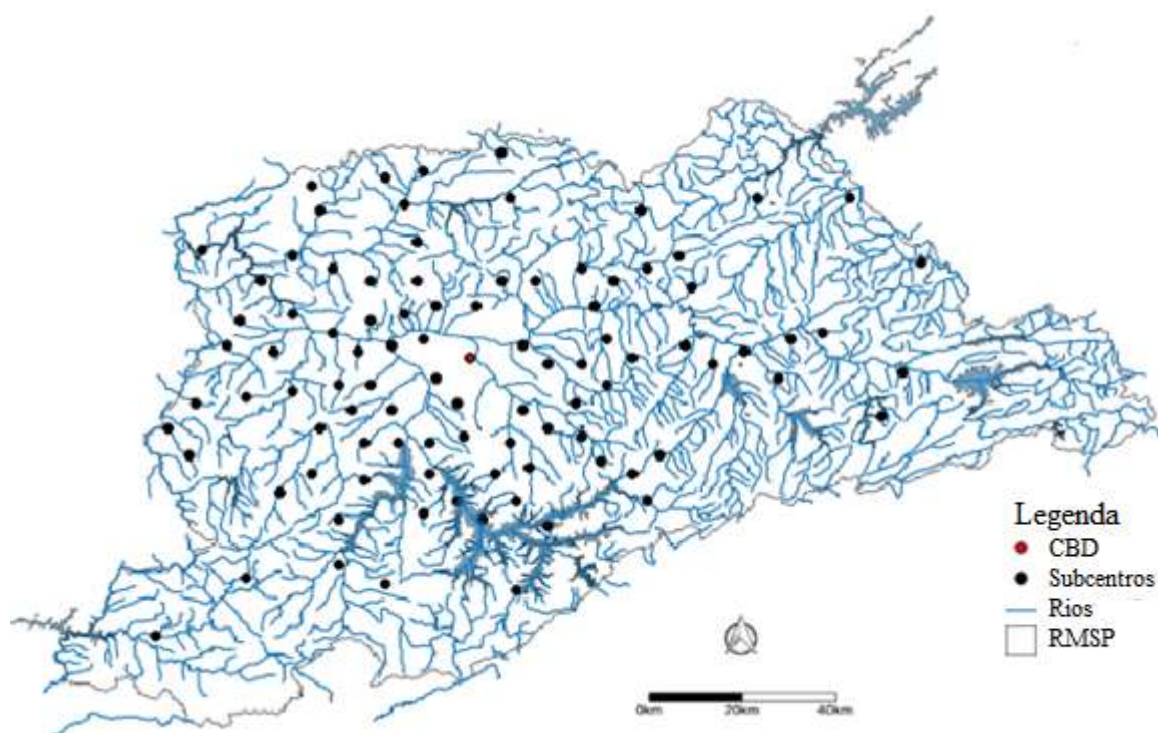


Figura 2.4 – Percurso dos rios e centros de empregos na RMSP.

Fonte: Elaboração própria com base nos dados do Censo 2010.

Haddad e Barufi (2017) com base numa perspectiva histórica, argumentam que em meados do século XIX, a cidade de São Paulo iniciou uma ocupação mais sistemática das várzeas dos principais cursos de água da cidade, tendo o motivo associado à implementação de infraestrutura regional e urbana (ferrovias e estradas). Em 1929, o Plano de Avenidas elaborado por Prestes Maia reforçou ainda mais a ocupação das planícies de inundação dos rios, concentrando-se no desenvolvimento de grandes avenidas nos rios da região (HADDAD e BARUFI, 2017). Seu projeto constituía-se pela implantação do modelo radial-perimetral, formado por três circuitos: 1) o do Perímetro de Irradiação; 2) o das vias férreas transferidas para as marginais, tendo seus antigos leitos transformados em avenidas; 3) o circuito de *parkways*, em que a avenida mais importante é a Tietê. São esses os principais elementos

estruturadores do Plano de Avenidas, sendo que os dois últimos dizem respeito diretamente ao uso das várzeas dos rios Tietê, Pinheiros e Tamanduateí (CUSTÓDIO, 2004). Essa concepção do uso de várzeas como espaço preferencial de circulação prevaleceu nos planos seguintes de desenvolvimento para a cidade, de modo que as avenidas de rodovias representam hoje as principais artérias de São Paulo (TRAVASSOS & GROSTEIN, 2003; MEYER, GROSTEIN & BIDERMAN, 2004).

Como em Haddad e Barufi (2017), o estudo buscou instrumentos em fatores geográficos que precedem à distribuição da população corrente e da própria malha viária dentro do município, baseando-se no uso das rotas dos rios da cidade como instrumentos de acessibilidade aos empregos. O instrumento foi construído a partir de uma base de dados GIS das rotas dos rios dentro da RMSP. A tarefa principal foi calcular a distância dos centróides das AEP's para os rios e depois 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), assumindo que os indivíduos podem caminhar em ambas as direções ao longo das rotas.

2.5 A estrutura espacial da RMSP

A Área Urbana da região metropolitana de São Paulo apresenta uma estrutura espacial caracterizada por disparidades sociais importantes, claramente evidentes quando se considera o centro da cidade e as regiões periféricas. A força de trabalho considerada corresponde aos indivíduos que se encontram na condição de População Economicamente Ativa (PEA) na faixa etária de 15 a 64 anos de idade. A Figura 2.5 mostra a distribuição espacial dessa participação do mercado de trabalho em porcentagem²² por área de ponderação. A RMSP exibe uma forte divisão núcleo-periferia: os empregos estão altamente concentrados na parte oeste central da região, onde a população economicamente ativa é mais presente.

²² A taxa de participação calculada é a PEA/PIA.

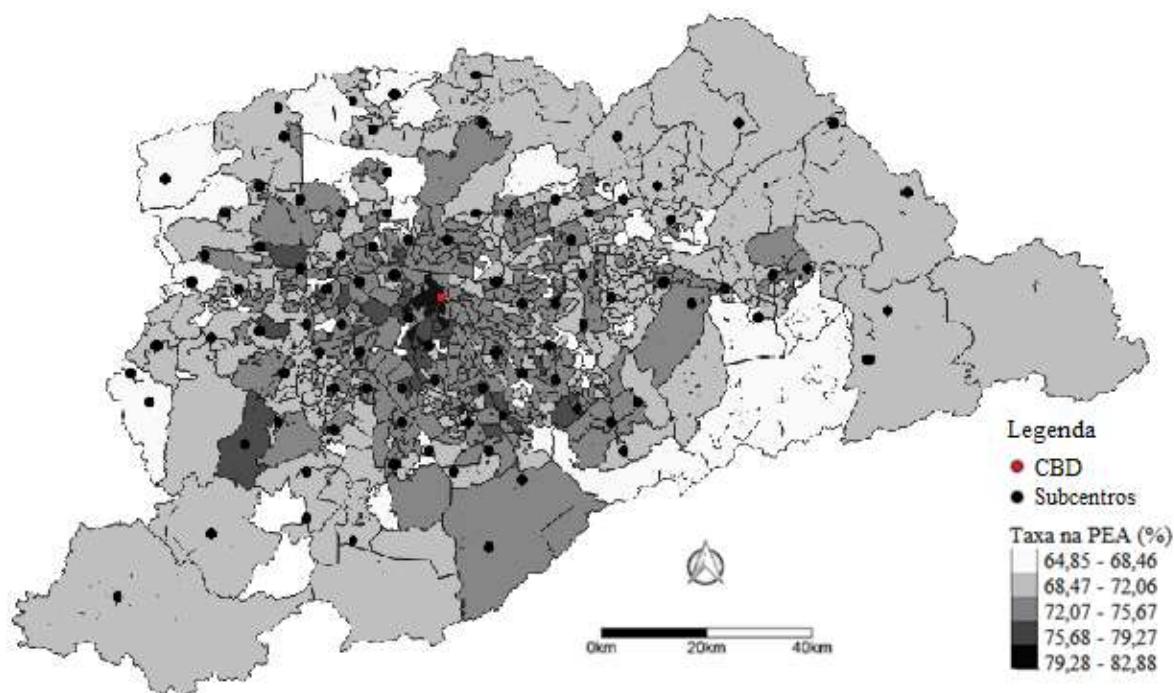
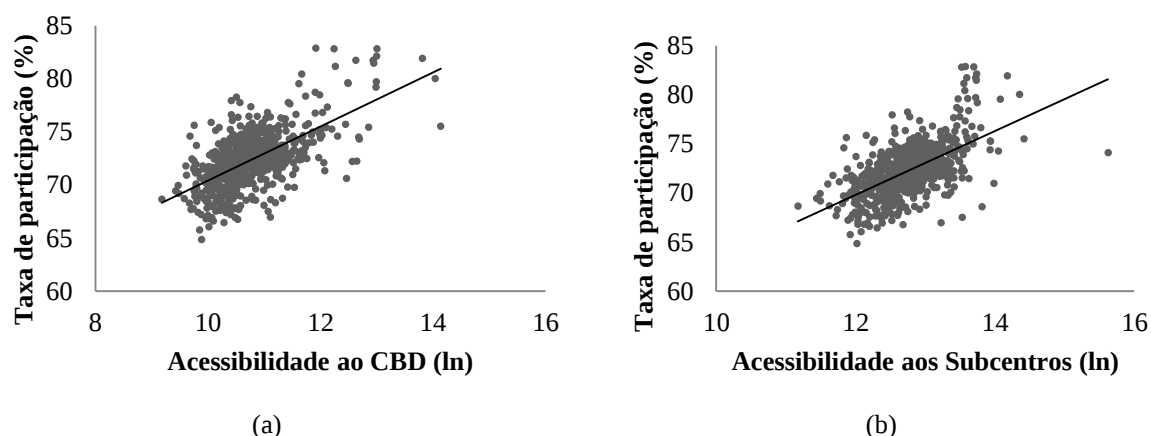


Figura 2.5 – Percentual de participação no mercado na RMSP.

Fonte: Elaboração própria com base nos dados do Censo 2010.

Percebe-se nos gráficos 2.1 a relação positiva entre a acessibilidade ao emprego com a participação dos indivíduos no mercado de trabalho. Esta relação mostra que o melhor acesso ao CBD e aos subcentros fornece maior chance de participação. Obviamente, não se trata de relação causal, contudo apresenta uma clara associação positiva entre as variáveis.

Gráfico 2.1 – Relação entre a participação e a acessibilidade aos empregos para RMSP.



(a) (b)
Fonte: Elaboração própria com base nos dados da pesquisa.

A Figura 2.6 mostra a percentagem de trabalhadores desempregados entre os participantes da força de trabalho, destacando uma zona de taxas de desemprego baixa na parte central da área urbana. Foram considerados desempregados, os indivíduos que se

encontram na PEA, porém na condição de desocupados por área de ponderação. A relação observada nos gráficos 2.2 (a) e (b) mostram uma associação negativa entre a taxa de desemprego e a acessibilidade aos empregos. Nesta relação, o melhor acesso aos centros de empregos diminui o percentual de desempregos.

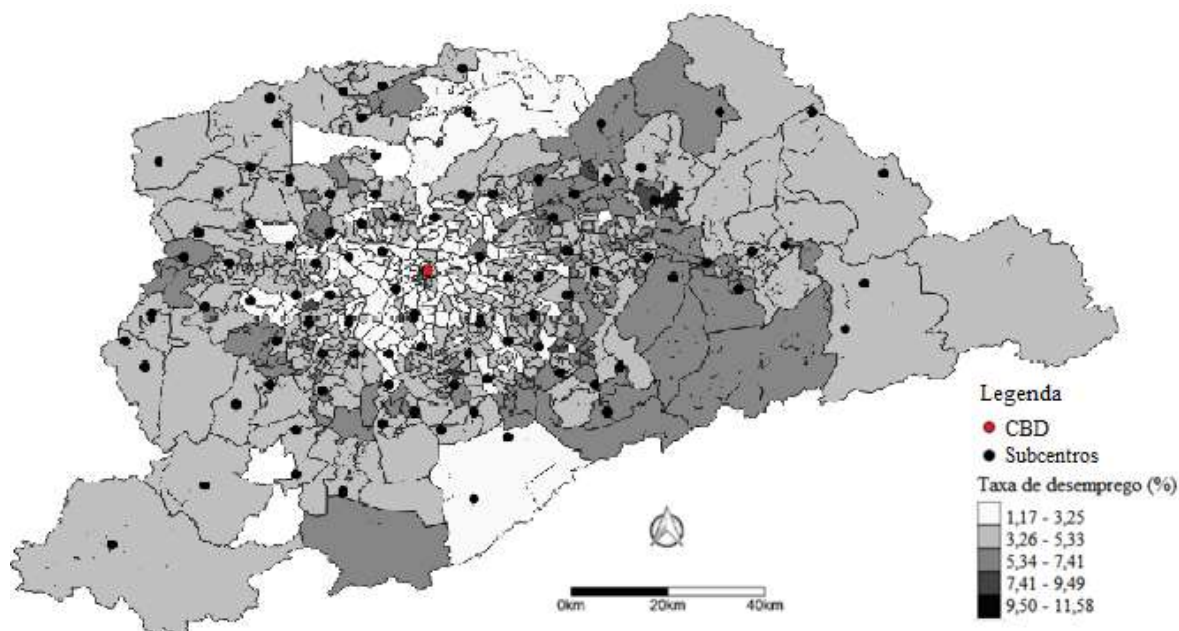
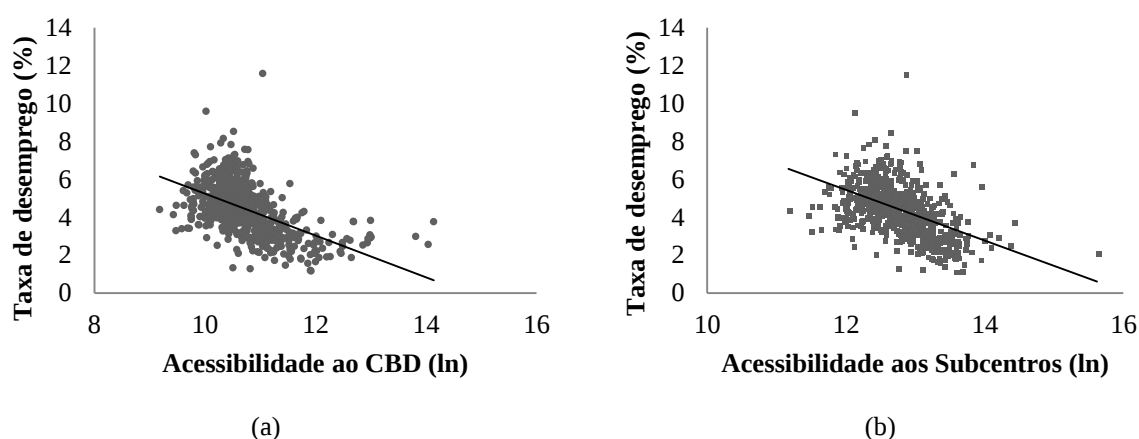


Figura 2.6 – Percentual de trabalhadores desempregados na RMSP.

Fonte: Elaboração própria com base nos dados do Censo 2010.

Uma implicação importante do padrão de distribuição espacial dos trabalhadores e moradores observada na RMSP, juntamente com o comportamento da demanda de viagens, é a dificuldade que alguns trabalhadores têm no acesso aos empregos. Em particular, a descentralização do trabalho, de alguma forma difusa, torna mais difícil para os trabalhadores sem carro chegar a potenciais empregadores, uma vez que estes estão localizados longe dos pontos focos de acessibilidade associados à infraestrutura de transporte público. Este possível *mismatch* espacial torna muito mais difícil o acesso para trabalhadores com baixas oportunidades de emprego, uma vez que tradicionalmente residem nas partes menos centrais das cidades e dependem fortemente da estrutura radial da rede de transporte público (BARUFI e HADDAD, 2017).

Gráfico 2.2 – Relação entre a taxa de desemprego e a acessibilidade para RMSP.



Fonte: Elaboração própria com base nos dados da pesquisa.

A Figura 2.7 mostra a taxa de informalidade de cada AEP da RMSP em 2010. No que diz respeito aos conceitos de informalidade, as diversas definições podem levar a resultados substancialmente diferentes, dificultando uma comparação entre estudos. Uma classificação tradicional na literatura brasileira enquadra como trabalhadores informais, àqueles que não possuem carteira de trabalho assinada, tal como o trabalho de Fernandes (1996), Carneiro e Henley (2001), Pianto e Pianto (2002) e Menezes-Filho, Mendes e Almeida (2004). Estudos recentes, como os de Dalberto (2014), Cirino e Dalberto (2014) estabelecem como conceito de trabalho informal aquele que não contribui para o Instituto Nacional do Seguro Social. Outros trabalhos como de Machado, Oliveira e Antigo (2008) e Duarte *et al.*, (2018) combina a definição de regulação do trabalho, ou seja, carteira assinada como característica de formalidade e sem carteira assinada como informalidade, e a definição de subordinação. Dessa forma, o trabalhador informal é o que se autodeclara conta própria, exceto as ocupações de profissionais liberais, e os empregados sem carteira assinada. Apesar de existirem diversas definições para a informalidade, no trabalho foi considerada a definição mais simples possível, o trabalhador apenas sem carteira assinada para representar a informalidade.

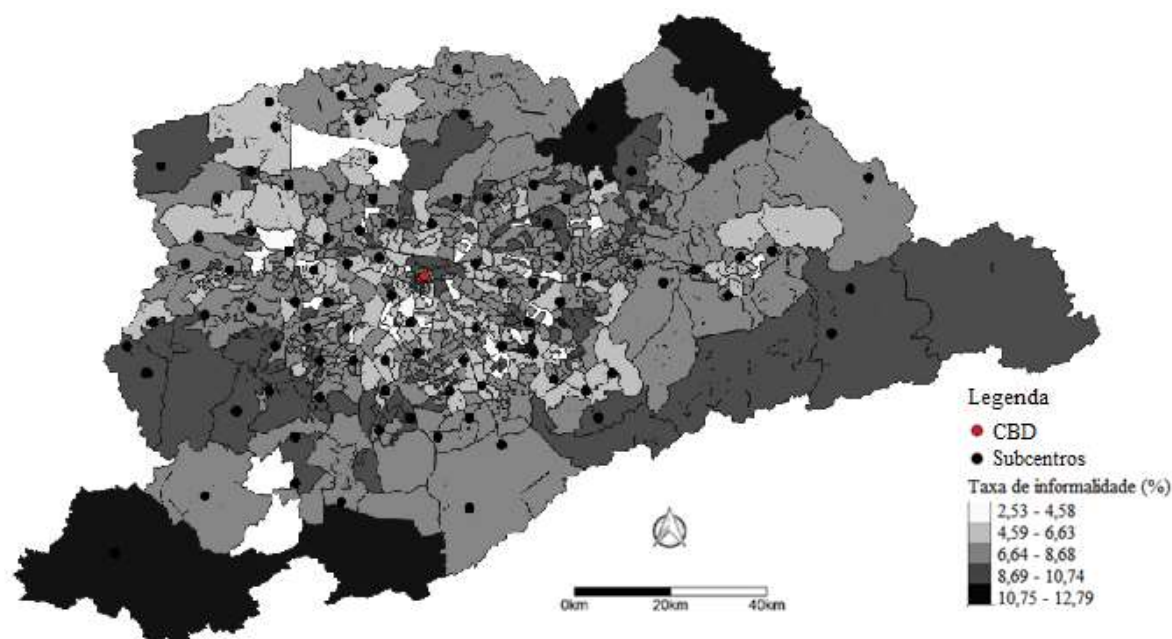


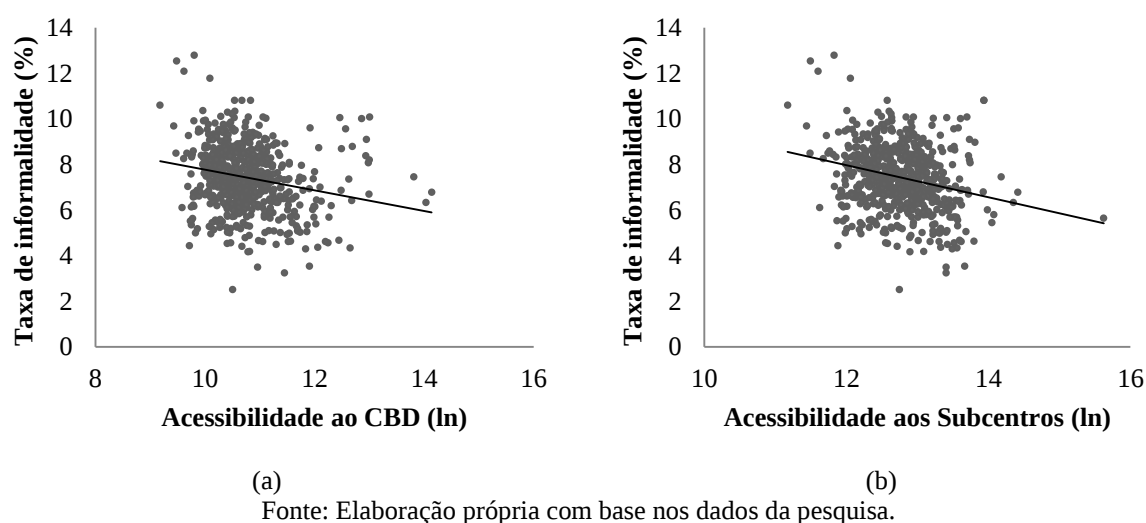
Figura 2.7 – Percentual de trabalhadores informais na RMSP.

Fonte: Elaboração própria com base nos dados do Censo 2010.

As taxas de informalidade são geralmente mais baixas no CBD e na região urbana sudoeste do CBD. Em contrapartida, a informalidade parece ser mais alta a nordeste do centro e nas áreas periféricas. Como coloca Boisjoly *et al.* (2017), este padrão geral sugere que os residentes que estão localizados centralmente têm acesso a um maior número de postos de trabalho formais do que aqueles situados na periferia.

Observando os gráficos 2.3 (a) e (b), percebe-se que a distribuição espacial da taxa de informalidade parece estar correlacionada com o padrão espacial de acessibilidade aos empregos, onde AEP com baixos níveis de acessibilidade parece caracterizar-se por maiores taxas de informalidade. Em outras palavras, áreas com altos níveis de acessibilidade aos empregos geralmente têm taxas mais baixas de informalidade. No gráfico 2.3, a taxa de informalidade de 2010 de cada AEP é plotada como uma função da acessibilidade aos empregos. Este número confirma que as áreas com menores níveis de acessibilidade aos postos de trabalho tendem a ter maiores taxas de informalidade.

Gráfico 2.3 – Relação entre a taxa de informalidade e a acessibilidade para RMSP.



A Figura 2.8 mostra a percentagem de trabalhadores na condição de *overeducated* na RMSP. Foram considerados *overeducated* os trabalhadores que possuem o nível educacional superior ao que é requerido em sua ocupação. Essa classificação foi realizada com base na Classificação Brasileira de Ocupações (CBO) tomando o nível educacional exigido em cada ocupação e comparando com a escolaridade de cada indivíduo.

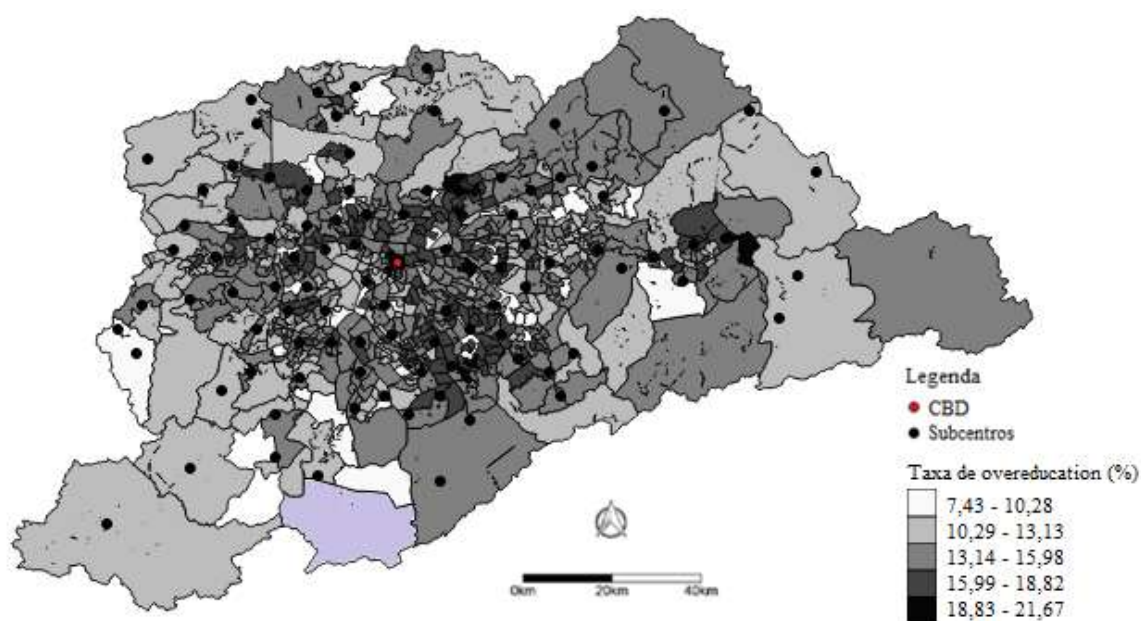


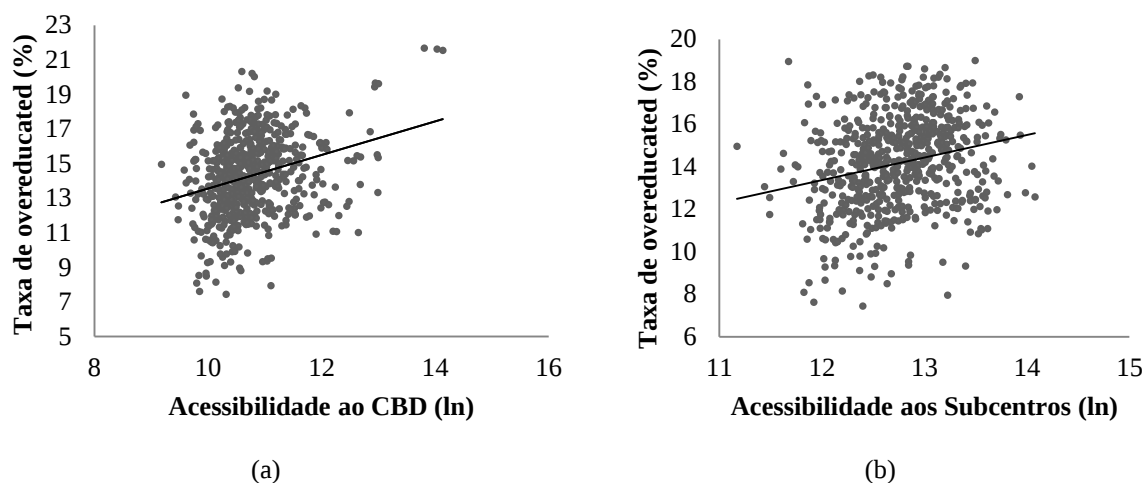
Figura 2.8 – Percentual de trabalhadores *overeducated* na RMSP.

Fonte: Elaboração própria com base nos dados do Censo 2010.

A relação observada nos gráficos 2.4 (a) e (b) mostram uma associação positiva entre a condição de *overeducated* e a acessibilidade aos empregos. Nesta relação, o melhor acesso aos centros de empregos aumentaria o percentual da condição de *overeducated*. Percebe-se que

para o CBD a associação positiva é mais forte do que quando considerado os subcentros. Esperava-se uma relação negativa entre as variáveis mostrando que o pior acesso aos empregos, aumentaria a chance de *overeducation*.

Gráfico 2.4 – Relação entre a taxa de *overeducated* e a acessibilidade para RMSP.



Fonte: Elaboração própria com base nos dados da pesquisa.

2.6 Resultados

Nesta seção serão apresentados os resultados estatísticos e econométricos desenvolvidos para investigar a probabilidade de participação no mercado, de desemprego individual, informalidade e da condição de *overeducated*, considerando o acesso físico para o CBD e subcentros. As estimações referem-se aos modelos de probabilidade linear (estimador OLS) e modelos Probit (estimador ML) com e sem variáveis instrumentais. Em cada modelo, considera-se o papel das características individuais, características de vizinhança e domésticas.

2.6.1 Estatísticas Descritivas

A análise estatística das variáveis selecionadas do banco de dados permite a obtenção de resultados preliminares para a RMSP, conforme a Tabela 2.3, a seguir. A princípio, é possível observar que foram atribuídos logaritmos para variáveis de acessibilidade ao emprego, como em Haddad e Barufi (2017). Esta estratégia é comum em estudos empíricos da Economia Urbana, que é geralmente sensível a observações desiguais em função da grandeza dos valores dessa variável. Em particular, as variáveis *dummies* apresentam

valores que descrevem a RMSP. Para a média das variáveis dependentes, percebe-se uma elevada taxa percentual para a População Economicamente Ativa (PEA), tendo as variáveis para o desemprego, informalidade e *overeducated* valores menores, porém significativos para uma região metropolitana. Quanto às variáveis de acessibilidade, os indivíduos tem uma menor acessibilidade média aos empregos para o CBD do que para os subcentros. Em relação às variáveis educacionais, observa-se que o nível mais baixo 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.

Tabela 2.3- Estatísticas resumidas das variáveis incluídas no modelo de regressão.

Variáveis	Média	Desvio Padrão	Mínimo	Máximo
PEA	0,71	0,45	0	1
Desemprego	0,08	0,28	0	1
Informalidade	0,15	0,36	0	1
<i>Overeducated</i>	0,14	0,34	0	1
LnACi_CBD	10,70	0,60	9,17	14,13
LnACi_sub	12,72	0,47	11,17	15,62
LnACi_CBD_IV	10,73	0,58	9,16	13,38
LnACi_sub_IV	12,76	0,45	11,15	13,95
Masculino	0,48	0,49	0	1
Idade	37,06	13,01	10	138
Idade ²	1543,40	1076,84	100	19044
Baixo	0,49	0,49	0	1
Intermediário	0,16	0,37	0	1
Alto	0,23	0,42	0	1
Bem Elevado	0,10	0,30	0	1
Casado	0,31	0,46	0	1
Chefe	0,30	0,45	0	1
Branco	0,57	0,49	0	1
Negro	0,06	0,24	0	1
Amarelo	0,01	0,12	0	1
Pardo	0,34	0,47	0	1
Muito Carente	0,55	0,49	0	1
Carente	0,21	0,41	0	1
Médio	0,07	0,25	0	1
Desenvolvida	0,12	0,33	0	1
Muito Desenvolvido	0,03	0,17	0	1
T_Família	4,08	1,93	1	42
Sanemamento	0,84	0,36	0	1
Coleta	0,94	0,23	0	1

Fonte: Elaboração própria.

2.6.2 Participação no mercado

Os modelos IV visam lidar com a questão da endogeneidade levantada anteriormente. A acessibilidade aos empregos é instrumentalizada pela distância ao CBD e aos subcentros,

seguindo um caminho formado por rios. A Tabela 2.4 mostra as regressões do 1º Estágio, bem como as estatísticas de Durbin e Wu-Hausman²³ que indicam a rejeição da hipótese nula de exogeneidade da variável representada. Assim, o estimador IV é preferível ao estimador²⁴ OLS. De fato, o conjunto de estatísticas apresentadas para avaliar o instrumento fornece forte confiança para usá-lo. Primeiro, observe que a estatística F para o primeiro estágio é estatisticamente significativa e assume um valor que facilmente excede 10, o valor de corte sugerido por Stock, Wright e Yogo (2002), segundo, os coeficientes de determinação R² apresentam valores altos.

Tabela 2.4 – Resultado da estimação para a probabilidade de participação no mercado (regressões com todos os indivíduos) – 1º Estágio

Variáveis	CBD		Subcentros	
	MPL_IV	D.P.	MPL_IV	D.P.
LnACi_IV	0,9122***	0,0004	0,8154***	0,0006
Masculino	-0,0001	0,0003	0,0003	0,0004
Idade	0,0002***	0,0001	0,0002***	0,0001
Idade ²	-0,0000***	0,0000	-0,0000***	0,0000
Baixa	Ref.	Ref.	Ref.	Ref.
Intermediária	-0,0001	0,0005	-0,0001	0,0006
Alta	0,0003	0,0004	-0,0036***	0,0005
Bem elevada	-0,0038***	0,0006	0,0022***	0,0008
Casado	-0,0070***	0,0004	-0,0048***	0,0005
Chefe	0,0036***	0,0004	0,0018***	0,0005
Branco	Ref.	Ref.	Ref.	Ref.
Negro	0,0004	0,0007	-0,0040***	0,0009
Amarelo	-0,0345***	0,0014	-0,0246***	0,0017
Pardo	0,0005	0,0004	-0,0014***	0,0005
Muito Carente	Ref.	Ref.	Ref.	Ref.
Carente	0,0066***	0,0005	0,0286***	0,0006
Média	-0,0181***	0,0007	0,0001	0,0009
Desenvolvida	0,0925***	0,0007	0,1726***	0,0008
Muito Desenvolvida	0,5100***	0,0012	0,4275***	0,0013
Tamanho da família	-0,0018***	0,0001	-0,0010***	0,0001
Saneamento	0,0013**	0,0007	0,0239***	0,0006
Coleta	0,0079***	0,0008	0,0015	0,0010
<i>Constante</i>	0,8593***	0,0047	2,2550***	0,0006
R ² / Pseudo R ²	0,9250		0,8136	
Estat. F	542387.43***		191964.84***	
Durbin (score, <i>Chi</i>)	11,17***		9,23***	
Wu-Hausman (F)	11,17***		9,23***	

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. D.P: desvio padrão.

Um aspecto importante de qualquer análise intra urbana é que o espaço desempenha um papel relevante na determinação dos resultados locais. No presente estudo essa questão foi trabalhada de várias maneiras. Em primeiro lugar, cada variável relacionada à acessibilidade

²³ O teste de Durbin-Wu-Hausman indica que os resíduos de uma regressão do índice de acessibilidade sobre as demais variáveis são estatisticamente significantes quando colocados como regressores de uma regressão da probabilidade de estar na PEA sobre todas as variáveis ditas explicativas.

²⁴ Ambos os estimadores são tendenciosos nessa situação, mas apenas IV pode gerar estimativas consistentes.

está lidando, pelo menos indiretamente, com a distância entre origem e destino. Além disso, algumas dessas variáveis serão constantes para os indivíduos que residem na mesma área de ponderação, e especialmente no caso da acessibilidade, a possível concentração espacial dos trabalhos é incorporada no cálculo dessa variável.

A Tabela 2.5 exibe os resultados com base nas equações 1 e 2 usando índices de acessibilidade. As colunas (1)-(4) exibem os resultados para o CBD, enquanto as colunas (5)-(8) são estimativas para centros (CBD e subcentros). A tabela considera a média dos efeitos marginais das variáveis para todos os indivíduos em idade de trabalhar. As estimativas dos regressores são obtidas a partir de modelos de Probabilidade Linear e Probit (efeitos marginais) com (colunas (3)-(4) e (7)-(8)) e sem (colunas (1)-(2) e (5)-(6)) variáveis instrumentais. Sobre os efeitos marginais, é importante destacar que a sua interpretação em modelos de variável dependente binária, é a seguinte: i) para variáveis contínuas, representa o impacto médio em termos de pontos percentuais na probabilidade de um indivíduo participar do mercado de trabalho para aumento unitário da variável explicativa contínua considerada; e ii) para as variáveis qualitativas, determina a mudança média na probabilidade de um indivíduo estar na PEA, em pontos percentuais, devido à presença da característica indicada pela *dummy* considerada.

De acordo com as estimativas, a principal variável de interesse neste estudo, acessibilidade ao trabalho, tem um impacto significativo na probabilidade de emprego. Tanto para o CBD quanto para os subcentros, um maior acesso à empregos aumenta a probabilidade de participação no mercado para todos os indivíduos da amostra. Percebe-se que os coeficientes são estáveis, mas a magnitude dos resultados para os estimadores ML (efeitos marginais) são maiores que OLS apenas nos modelos sem variável instrumental. O índice de acessibilidade é significativo a 1% e com efeito esperado em todas as estimativas. Concentrando-se nos modelos IV e tomando como exemplo a coluna 7, pode-se dizer que o aumento do acesso físico ao CBD e subcentros em 10%, aumenta em 0,14 ponto percentual a probabilidade de participação no mercado de trabalho. Este resultado é significativo comparado com a média de chance de participação que é de 0,71 com desvio padrão de 0,45. Ou seja, o aumento de 10% no índice de acessibilidade aumenta em aproximadamente 0,2% a probabilidade de se participar no mercado.

Como esperado, a educação aumenta a chance de se participar no mercado de trabalho, como mostra nas estimativas da Tabela 2.5. Observa-se que os coeficientes das variáveis socioeconômicas (idade, idade², educação e chefe de família) são significativos e com efeitos

esperados. Mais experiência e mais educação aumentam a probabilidade de trabalhar. Ser chefe do agregado familiar aumenta a probabilidade de se participar no mercado de trabalho.

Tabela 2.5 – Efeitos Marginais para a probabilidade de participação no mercado (regressões com todos os indivíduos).

Variáveis	CBD				Subcentros			
PEA	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
LnACi	0,0128*** (0,0010)	0,0120*** (0,0010)	0,0143*** (0,0011)	0,0053*** (0,0012)	0,0112*** (0,0013)	0,0098*** (0,0012)	0,0141*** (0,0016)	0,0054*** (0,0017)
Masculino	0,1611*** (0,0009)	0,1600*** (0,0008)	0,1611*** (0,0009)	0,1752*** (0,0009)	0,1611*** (0,0009)	0,1600*** (0,0008)	0,1611*** (0,0009)	0,1751*** (0,0009)
Idade	0,0505*** (0,0002)	0,0440*** (0,0001)	0,0505*** (0,0002)	0,0489*** (0,0002)	0,0505*** (0,0002)	0,0441*** (0,0001)	0,0505*** (0,0002)	0,0489*** (0,0002)
Idade²	-0,0006*** (0,0000)	-0,0005*** (0,0000)	-0,0006*** (0,0000)	-0,0006*** (0,0000)	-0,0006*** (0,0000)	-0,0006*** (0,0000)	-0,0006*** (0,0000)	-0,0006*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0567*** (0,0013)	0,0510*** (0,0012)	0,0567*** (0,0013)	0,0547*** (0,0012)	0,0566*** (0,0013)	0,0509*** (0,0012)	0,0566*** (0,0013)	0,0546*** (0,0012)
Alta	0,1663*** (0,0012)	0,1546*** (0,0011)	0,1663*** (0,0012)	0,1601*** (0,0011)	0,1662*** (0,0012)	0,1545*** (0,0011)	0,1662*** (0,0012)	0,1600*** (0,0011)
Bem elevada	0,2410*** (0,0016)	0,2522*** (0,0017)	0,2411*** (0,0016)	0,2161*** (0,0010)	0,2407*** (0,0016)	0,2519*** (0,0017)	0,2407*** (0,0016)	0,2159*** (0,0010)
Casado	-0,0311*** (0,0010)	-0,0250*** (0,0010)	-0,0310*** (0,0010)	-0,0280*** (0,0011)	-0,0313*** (0,0010)	-0,0253*** (0,0010)	-0,0313*** (0,0010)	-0,0282*** (0,0011)
Chefe	0,0650*** (0,0010)	0,0719*** (0,0010)	0,0650*** (0,0010)	0,0780*** (0,0011)	0,0651*** (0,0010)	0,0719*** (0,0010)	0,0651*** (0,0010)	0,0781*** (0,0011)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0429*** (0,0018)	0,0422*** (0,0018)	0,0428*** (0,0018)	0,0448*** (0,0018)	0,0430*** (0,0018)	0,0423*** (0,0018)	0,0429*** (0,0018)	0,0449*** (0,0018)
Amarelo	-0,0269 (0,0035)	-0,0290*** (0,0035)	-0,0269*** (0,0035)	-0,0335*** (0,0041)	-0,0265*** (0,0035)	-0,0286*** (0,0035)	-0,0265*** (0,0035)	-0,0329*** (0,0041)
Pardo	0,0228*** (0,0010)	0,0217*** (0,0010)	0,0228*** (0,0010)	0,0239*** (0,0011)	0,0228*** (0,0010)	0,0217*** (0,0010)	0,0228*** (0,0010)	0,0239*** (0,0011)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0088*** (0,0012)	-0,0075*** (0,0012)	0,0037* (0,0022)	-0,0089*** (0,0014)	-0,0068*** (0,0022)	-0,0054*** (0,0012)	-0,0077*** (0,0012)	-0,0067*** (0,0014)
Média	-0,0136*** (0,0019)	-0,0132*** (0,0018)	-0,0132*** (0,0019)	-0,0147*** (0,0021)	-0,0135*** (0,0019)	-0,0134*** (0,0018)	-0,0126*** (0,0019)	-0,0142*** (0,0021)
Desenvolvida	-0,0216*** (0,0017)	-0,0203*** (0,0017)	-0,0229*** (0,0018)	-0,0233*** (0,0021)	-0,0172*** (0,0017)	-0,0159*** (0,0017)	-0,0191*** (0,0018)	-0,0182*** (0,0020)
Muito Desenvolvida	-0,0531*** (0,0031)	-0,0569*** (0,0032)	-0,0556*** (0,0032)	-0,0641*** (0,0041)	-0,0435*** (0,0029)	-0,0471*** (0,0030)	-0,0464*** (0,0031)	-0,0541*** (0,0038)
Tamanho da família	-0,0030*** (0,0002)	-0,0031*** (0,0002)	-0,0030*** (0,0002)	-0,0035*** (0,0002)	-0,0030*** (0,0002)	-0,0031*** (0,0002)	-0,0030*** (0,0002)	-0,0035*** (0,0002)
Saneamento	-0,0003 (0,0013)	0,0006 (0,0013)	-0,0004 (0,0013)	0,0005 (0,0015)	-0,0002 (0,0013)	0,0007 (0,0013)	-0,0006 (0,0014)	0,0006 (0,0015)
Coleta	0,0158*** (0,0020)	0,0167*** (0,0019)	0,0160*** (0,0020)	0,0190*** (0,0022)	0,0156*** (0,0020)	0,0164*** (0,0019)	0,0159*** (0,0020)	0,0187*** (0,0022)
Constante	0,535*** (0,0120)	-	-0,4740*** (0,0125)	-	-0,4658*** (0,0165)	-	-0,5006*** (0,0201)	-
Estat. F/Wald	8466,47	141745,08	8466,88	123761,12	8461,79	141674,39	8461,87	123711,32
R²/ Pseudo R²	0,1614	0,1419	0,1614	-	0,1613	0,1418	0,1613	-
Observações	835.804	835.804	835.804	835.804	835.804	835.804	835.804	835.804

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Estimativas com base na média dos efeitos marginais.

Os resultados para os modelos de probabilidade linear com IV implicam que o coeficiente de acessibilidade é maior do que suas contrapartes OLS. Isso sugere que os

modelos de OLS estão subestimando os efeitos da acessibilidade aos empregos nos resultados do mercado de trabalho, devido a um viés de simultaneidade. Portanto, a acessibilidade pode ser positivamente correlacionada com o termo de erro, levando a um coeficiente subestimado tendencioso.

A relação entre acessibilidades do trabalho e resultados do mercado de trabalho mostra padrões mais distintos quando permite a heterogeneidade em todos os gêneros. Na Tabela 2.6 têm-se efeitos heterogêneos para homens e mulheres. Os primeiros estágios para estes resultados por gênero estão disponíveis na Tabela A2.1 no apêndice, bem como os resultados completos dos controles (Tabelas A2.2 e A2.3). Em geral, os resultados para mulheres tendem a seguir o padrão anteriormente observado, enquanto para os homens a acessibilidade não parece positivamente associada à participação no mercado de trabalho. Mais especificamente, analisando resultados apenas para o gênero masculino, observa-se a não significância quando considerado apenas o CBD como oportunidade de emprego. Porém, quando considerado os subcentros, incluindo o CBD, a relação entre acessibilidade a empregos e participação no mercado de trabalho apresenta sinal oposto ao esperado, ou seja, o melhor acesso físico aos centros de empregos diminui a probabilidade de participação no mercado de trabalho. Para as mulheres, é interessante observar o efeito marginal que é quase o dobro daquele obtido para amostra em geral.

A insignificância do coeficiente estimado para acessibilidade pode estar associado ao fato de que os homens, em geral, participam mais do mercado de trabalho (55% dos ocupados são do gênero masculino na amostra) que as mulheres, já que são majoritariamente os chefes da família e podem sofrer maior pressão social para inserção no mercado de trabalho. Assim, seus resultados sugerem que acesso mais difícil não é suficiente para inibir a participação destes. Este resultado também é consistente com os menores tempos de deslocamento da residência ao trabalho para as mulheres na RMSP apontados por Silveira Neto *et al.*, (2015). Outra explicação para este resultado pode ser o fato desta subamostra apresentar características não observadas na composição do grupo familiar.

Os resultados para os demais controles utilizados na regressão apresentados no apêndice parecem fazer sentido com a literatura (Quintanar, 2013; Campos, 2015), olhando especificamente para a educação, o fato de ser casado e chefe do agregado familiar. Como resultado interessante, tem-se que a probabilidade de trabalho aumenta quando o homem é casado e diminui quando a mulher é casada, uma vez que a mulher tem maior probabilidade de se dedicar ao trabalho doméstico quando casada. Se observarmos os efeitos marginais

médios, estes são mais significativos e maiores para as mulheres do que para os homens (Tabela A2.2 e A2.3).

Tabela 2.6 – Efeitos marginais com e sem variáveis instrumentais para a probabilidade de participação no mercado (regressões para homens e mulheres)

participação no mercado (regressões para homens e mulheres)								
Variáveis		CBD			Subcentros			
HOMENS								
PEA	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
LnACi	0,0018 (0,0013)	0,0003 (0,0013)	0,0022 (0,0014)	-0,0013 (0,0015)	-0,0036** (0,0016)	-0,0054*** (0,0016)	-0,0044** (0,0020)	-0,0030 (0,0084)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	401.311	401.311	401.311	401.311	401.311	401.311	401.311	401.311
Estat. F/Wald	3993,23	59633,15	3993,26	56038,89	3993,41	59643,54	3993,42	53043,93
R ² / Pseudo R ²	0,1519	0,1494	0,1519		0,1519	0,1494	0,1519	
MULHERES								
LnACi	0,0235*** (0,0015)	0,0236*** (0,0016)	0,0259*** (0,0017)	0,0133*** (0,0019)	0,0256*** (0,0019)	0,0248*** (0,0019)	0,0325*** (0,0024)	0,0115*** (0,0024)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	434.493	434.493	434.493	434.493	434.493	434.493	434.493	434.493
Estat. F/Wald	3821,22	62253,10	3821,46	56232,27	3818,11	62198,34	3818,60	56213,74
R ² / Pseudo R ²	0,1367	0,1091	0,1367		0,1366	0,1090	0,1366	

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Estimativas com base na média dos efeitos marginais.

Os resultados obtidos indicam que a acessibilidade às oportunidades de emprego representam barreiras mais importantes para a participação no mercado de trabalho para mulheres que para os homens. Embora o exercício levado a efeito não identifique as razões para tal na literatura, existem diversas explicações sobre por que as mulheres são mais sensíveis às condições de mercado de trabalho local do que homens, resumido em MacDonald (1999). Uma explicação é que as mulheres ganham menos do que os homens e os salários médios das mulheres variam menos no território. 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. Esse tipo de trabalho geralmente não justifica grandes mudanças. Além disso, o duplo papel de trabalhadora, ou seja, ser dona de casa/mãe restringe os deslocamentos. Os deslocamentos curtos são mais fáceis de combinar com as demandas de tempo às tarefas domésticas e/ou de assistência à criança (QUINTANAR, 2013).

Por fim, a distribuição do emprego através do território pode afetar as condições do mercado de trabalho local. Por exemplo, se os empregos dominados por mulheres forem uniformemente distribuídos em todo o território, é mais fácil que as mulheres encontrem um emprego perto de onde vivem. Se o mercado de trabalho local estiver segmentado espacialmente, as oportunidades de trabalho para as mulheres são menores. Nem todos os trabalhos podem ser alcançado por mulheres porque alguns deles podem estar longe.

Uma das políticas mais eficazes para aumentar a taxa de emprego feminino é melhorar o nível de educação das mulheres. A contribuição do estudo é confirmar duas possibilidades adicionais. Primeiro, reduzir a segregação residencial provavelmente levará a um aumento na taxa de emprego. Em segundo lugar, melhorar a acessibilidade do trabalho provavelmente também será uma política eficaz, principalmente para as mulheres com menos escolaridade.

2.6.3 Desempregado

Apresenta-se agora a estimativa IV para a influência da acessibilidade ao emprego na probabilidade de estar desempregado. A partir da Tabela 2.7, nota-se que existe uma associação positiva e estatisticamente significativa entre o instrumento (acesso físico ao emprego por rios) e a variável dependente (probabilidade de estar desempregado), algo bem vindo por ter confiança em usar a variável como instrumento. Além disso, como no caso da participação no mercado de trabalho, as estatísticas apresentadas na Tabela 2.7, a seguir, indicam que os resultados pelo modelo IV são mais adequados (variável de interesse endógena e instrumento parece adequado a partir das estatísticas de teste para sua força).

Tabela 2.7 – Resultado da estimação para a probabilidade de desempregados (regressões com todos os indivíduos) – 1º Estágio

Variáveis	CBD		Subcentros	
	MPL_IV	D.P.	MPL_IV	D.P.
LnACi_IV	0,9153***	0,0005	0,8121***	0,0007
Masculino	-0,0003	0,0004	-0,0006	0,0005
Idade	-0,0002***	0,0001	-0,0000	0,0001
Idade ²	0,0000***	0,0000	0,0001	0,0000
Baixa	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0003	0,0006	-0,0001	0,0007
Alta	0,0010*	0,0005	-0,0042***	0,0006
Bem elevada	-0,0034***	0,0007	0,0002	0,0009
Casado	-0,0081***	0,0004	-0,0056***	0,0005
Chefe	0,0005***	0,0058	0,0024***	0,0005
Branco	Ref.	Ref.	Ref.	Ref.
Negro	-0,0004	0,0008	-0,0047***	0,0010
Amarelo	-0,0343***	0,0016	-0,0246***	0,0019
Pardo	0,0008*	0,0005	-0,0012***	0,0005
Muito Carente	Ref.	Ref.	Ref.	Ref.
Carente	0,0062***	0,0006	0,0291***	0,0007
Média	-0,0163***	0,0009	0,0020*	0,0011
Desenvolvida	0,0948***	0,0008	0,1773***	0,0009
Muito Desenvolvida	0,5116***	0,0013	0,4320***	0,0015
Tamanho da família	-0,0019***	0,0001	-0,0009***	0,0001
Saneamento	0,0015**	0,0006	0,0266***	0,0008
Coleta	0,0102***	0,0010	0,0061***	0,0011
<i>Constante</i>	0,8337***	0,0005	2,2977***	0,0095
R ² / Pseudo R ²	0,9239		0,8142	
Estat. F	394036.75***		142216.58***	
Durbin (score, Chi)	6,21**		25,53***	

Wu-Hausman (F)	6,21**	25,53***
----------------	--------	----------

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. D.P: desvio padrão.

Analisando os resultados para os modelos da Tabela 2.8, tem-se que o melhor acesso à oportunidades de empregos diminui a probabilidade do indivíduo estar desempregado, sendo os efeitos do CBD semelhantes àqueles obtidos para os subcentros. Ainda nesta tabela, percebe-se que os resultados para os efeitos marginais são fortes quando comparado à média da chance de estar desempregado que é de 0,08 e o desvio padrão de 0,28 (Tabela 2.3). Tomando os maiores valores dos efeitos marginais que correspondem as colunas (3) e (7), tem-se que um aumento de 1% no índice de acessibilidade reduz a probabilidade de estar desempregado em 0,13% e 0,17% para o CBD e subcentros, respectivamente.

Analisando as características da vizinhança, observa-se que áreas mais carentes aumentam a probabilidade de desemprego. Como sugerido pela teoria, isso pode depender de redes sociais locais de menor qualidade: quanto maior a proporção de vizinhos desempregados, mais difícil é a inserção de um indivíduo no mercado de trabalho. Isso confirma que a probabilidade de desemprego é afetada por uma grande variedade de características de vizinhança, em vez de apenas pela acessibilidade ao emprego. Tudo o mais sendo igual, a probabilidade de desemprego é a mais baixa para os indivíduos que residem em vizinhanças mais desenvolvidas. De acordo com a teoria do *mismatch* espacial e de estudos empírico como o de Durjandin *et al.*, (2008), viver em uma área com um ambiente socioeconômico de baixa qualidade (muito carente e carente) aumenta significativamente a probabilidade de desemprego em comparação com os que vivem em áreas ricas.

Em todas as regressões, é menos provável que os homens ou trabalhadores instruídos estejam desempregados do que mulheres ou trabalhadores com menor escolaridade. A probabilidade de desemprego também diminui com a idade do indivíduo, quando casado e chefe da família. Os negros e pardos são mais desfavorecidos do que os brancos.

Em relação à infraestrutura, os resultados mostram que indivíduos que moram em locais com acesso a saneamento básico e coleta de lixo diminuem as chances de estar desempregado. Estes resultados estão ligados novamente ao ambiente social, ou seja, áreas que apresentam melhores condições quanto a infraestrutura reduzem a probabilidade de desemprego.

Tabela 2.8 – Efeitos marginais para a probabilidade de estar desempregado (regressões com todos os indivíduos).

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
Desemprego	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	-0,0101*** (0,0007)	-0,0106*** (0,0008)	-0,0109*** (0,0008)	-0,0059*** (0,0008)	-0,0098*** (0,0015)	-0,0097*** (0,0010)	-0,0135*** (0,0012)	-0,0027*** (0,0011)
Masculino	-0,0410*** (0,0007)	-0,0408*** (0,0007)	-0,0410*** (0,0007)	-0,0384*** (0,0006)	-0,0410*** (0,0007)	-0,0408*** (0,0007)	-0,0410*** (0,0007)	-0,0384*** (0,0006)
Idade	-0,0132*** (0,0001)	-0,0091*** (0,0001)	-0,0132*** (0,0001)	-0,0084*** (0,0001)	-0,0132*** (0,0001)	-0,0091*** (0,0001)	-0,0132*** (0,0001)	-0,0084*** (0,0001)
Idade ²	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0005*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0090*** (0,0010)	0,0031*** (0,0009)	0,0090*** (0,0006)	0,0029*** (0,0009)	0,0090*** (0,0010)	0,0032*** (0,0009)	0,0090*** (0,0010)	0,0030*** (0,0009)
Alta	-0,0211*** (0,0009)	-0,0191*** (0,0009)	-0,0212*** (0,0009)	-0,0172*** (0,0007)	-0,0210*** (0,0009)	-0,0190*** (0,0009)	-0,0211*** (0,0009)	-0,0170*** (0,0007)
Bem elevada	-0,0438*** (0,0012)	-0,0599*** (0,0014)	-0,0438*** (0,0012)	-0,0452*** (0,0008)	-0,0435*** (0,0012)	-0,0596*** (0,0014)	-0,0435*** (0,0012)	-0,0450*** (0,0008)
Casado	-0,0136*** (0,0007)	-0,0184*** (0,0008)	-0,0137*** (0,0007)	-0,0167*** (0,0007)	-0,0134*** (0,0007)	-0,0183*** (0,0008)	-0,0135*** (0,0007)	-0,0166*** (0,0007)
Chefe	-0,0093*** (0,0007)	-0,0154*** (0,0008)	-0,0093*** (0,0007)	-0,0141*** (0,0007)	-0,0094*** (0,0008)	-0,0155*** (0,0008)	-0,0094*** (0,0008)	-0,0141*** (0,0007)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0066*** (0,0013)	0,0071*** (0,0013)	0,0066*** (0,0013)	0,0067*** (0,0013)	0,0065*** (0,0013)	0,0070*** (0,0013)	0,0066*** (0,0013)	0,0067*** (0,0013)
Amarelo	-0,0025 (0,0026)	-0,0056* (0,0031)	-0,0025 (0,0026)	-0,0049* (0,0027)	-0,0028 (0,0026)	-0,0059** (0,0031)	-0,0028 (0,0026)	-0,0051** (0,0027)
Pardo	0,0072*** (0,0007)	0,0069*** (0,0007)	0,0072*** (0,0007)	0,0064*** (0,0007)	0,0072*** (0,0007)	0,0069*** (0,0007)	0,0072*** (0,0007)	0,0065*** (0,0007)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0091*** (0,0009)	-0,0085*** (0,0009)	-0,0088*** (0,0009)	-0,0074*** (0,0008)	-0,0104*** (0,0009)	-0,0100*** (0,0009)	-0,0092*** (0,0009)	-0,0083*** (0,0008)
Média	-0,0105*** (0,0015)	-0,0094*** (0,0014)	-0,0107*** (0,0015)	-0,0084*** (0,0012)	-0,0109*** (0,0015)	-0,0095*** (0,0014)	-0,0120*** (0,0015)	-0,0092*** (0,0012)
Desenvolvida	-0,0097*** (0,0013)	-0,0136*** (0,0014)	-0,0089*** (0,0013)	-0,0118*** (0,0012)	-0,0126*** (0,0013)	-0,0170*** (0,0013)	-0,0102*** (0,0013)	-0,0144*** (0,0011)
Muito Desenvolvida	-0,0013*** (0,0023)	-0,0076*** (0,0027)	0,0000 (0,0024)	-0,0079** (0,0024)	-0,0080*** (0,0022)	-0,0154** (0,0026)	-0,0042*** (0,0023)	-0,0140*** (0,0021)
Tamanho da família	0,0057*** (0,0002)	0,0038*** (0,0001)	0,0057*** (0,0002)	0,0035*** (0,0001)	0,0057*** (0,0002)	0,0038*** (0,0001)	0,0057*** (0,0002)	0,0035*** (0,0001)
Saneamento	-0,0061*** (0,0010)	-0,0046*** (0,0010)	-0,0060*** (0,0010)	-0,0042*** (0,0009)	-0,0060*** (0,0010)	-0,0045*** (0,0010)	-0,0055*** (0,0010)	-0,0041*** (0,0009)
Coleta	-0,0091*** (0,0015)	-0,0080*** (0,0014)	-0,0092*** (0,0015)	-0,0077*** (0,0014)	-0,0090*** (0,0015)	-0,0077*** (0,0014)	-0,0093*** (0,0015)	-0,0077*** (0,0014)
Constante	0,5108*** (0,0088)	-	0,5195*** (0,0095)	-	0,5282*** (0,0127)	-	0,5734*** (0,0155)	-
Estat. F/Wald	1854,09	31742,53	1854,03	28853,39	1850,58	31674,94	1851,73	23828,52
R ² / Pseudo R ²	0,0540	0,0875	0,0540		0,0539	0,0873	0,0539	
Observações	616.754	616.754	616.754	616.754	616.754	616.754	616.754	616.754

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Estimativas com base na média dos efeitos marginais.

Os resultados da Tabela 2.9, a seguir, referem-se à probabilidade de estar desempregado para homens e mulheres. Os resultados para os primeiros estágios estão no apêndice na Tabela A2.4, assim como os efeitos marginais do 2º estágio para as variáveis de controle (Tabelas A2.5 e A2.6). Observa-se claramente que o efeito do coeficiente é mais forte para as mulheres do que para os homens quando analisado o índice de acessibilidade.

Focando nos resultados para subcentros com a variável instrumentalizada e tomando a coluna (7) ao nível de comparação entre os gêneros, tem-se que um aumento de 1% no indicador de acessibilidade reduz em 0,08% e 0,21% para homens e mulheres a probabilidade de estar desempregado na RMSP, respectivamente.

Tabela 2.9 – Efeitos marginais com e sem variáveis instrumentais para a probabilidade de estar desempregado (regressões para homens e mulheres)

Variáveis	CBD				Subcentros			
	HOMENS							
Desemprego	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
LnACi	-0,0056*** (0,0009)	-0,0054*** (0,0010)	-0,0056*** (0,0010)	-0,0056*** (0,0010)	-0,0046*** (0,0012)	-0,0043*** (0,0012)	-0,0058*** (0,0014)	-0,0021*** (0,0014)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	333.347	333.347	333.347	333.347	333.347	333.347	333.347	333.347
Estat. F/Wald	926,50	14311,00	926,22	13297,58	925,37	14294,93	925,43	13289,09
R ² / Pseudo R ²	0,0476	0,0870	0,0476	-	0,0476	0,0869	0,0476	-
MULHERES								
LnACi	-0,0159*** (0,0013)	-0,0169*** (0,0013)	-0,0179*** (0,0014)	-0,0073*** (0,0014)	-0,0170*** (0,0016)	-0,0160*** (0,0015)	-0,0236*** (0,0020)	-0,0042*** (0,0020)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	283.407	283.407	283.407	283.407	283.407	283.407	283.407	283.407
Estat. F/Wald	915,98	15258,06	916,42	13852,05	913,44	15198,89	914,92	13837,23
R ² / Pseudo R ²	0,0550	0,0782	0,0550	-	0,0548	0,0779	0,0548	-

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Estimativas com base na média dos efeitos marginais.

Os resultados confirmam que o tipo da vizinhança tem um efeito negativo na chance de estar desempregado. Moradores de áreas muito carentes têm menor probabilidade de encontrar um emprego (Tabela A2.5 e A2.6). Como sugerido pela teoria do *mismatch* espacial, essa relação opera por meio de dois canais principais: a deterioração das redes sociais locais e as maiores dificuldades na aquisição de capital humano. Os resultados do presente estudo estão em linha com os resultados obtidos em outros trabalhos. Gobillon e Selod (2006), por exemplo, concluem que a segregação residencial impede que trabalhadores desempregados encontrem um emprego na área metropolitana de Paris. Por seu turno, Dujardin *et al.*, (2008) confirmam que residir em um bairro pobre aumenta a probabilidade de estar desempregado em Bruxelas.

Ressalta-se como resultado importante o fato de ser chefe do agregado familiar para os gêneros. Para os homens, ser chefe diminui a probabilidade de estar desempregado ao nível de significância a 1%, no entanto, para as mulheres essa mesma variável não foi significativa.

2.6.4 Informalidade

Como explicado anteriormente, a validade do instrumento proposto se baseia em sua relevância e exogeneidade em relação à variável de resultado. Para avaliar sua relevância, estimaram-se as regressões apresentadas na Tabela 2.10 com o objetivo de explicar no 2º Estágio a informalidade presente no mercado de trabalho da RMSP. Como pode ser visto nas duas primeiras linhas, o instrumento é estatisticamente significativo e positivo. A tabela também apresenta as estatísticas de Durbin e Wu-Hausman, mostrando que o estimador IV é preferível ao estimador OLS, bem como a estatística F e o coeficiente de determinação R² com valores altos e significativos. Assim, o instrumento parece ser relevante pelos testes apresentados.

Tabela 2.10 – Resultado da estimação para a probabilidade de condição de trabalhador informal (regressões com todos os indivíduos) – 1º Estágio

Variáveis	CBD		Subcentros	
	MPL_IV	D.P.	MPL_IV	D.P.
LnACi_IV	0,9163***	0,0005	0,8112***	0,0007
Masculino	-0,0003	0,0004	-0,0006	0,0005
Idade	-0,0003***	0,0001	-0,0001	0,0001
Idade ²	0,0000***	0,0000	0,0001***	0,0000
Baixa	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0004	0,0007	-0,0001	0,0008
Alta	0,0011*	0,0006	-0,0041***	0,0007
Bem elevada	-0,0036***	0,0008	-0,0000	0,0009
Casado	-0,0087***	0,0005	-0,0060***	0,0006
Chefe	0,0039***	0,0005	0,0021***	0,0006
Branco	Ref.	Ref.	Ref.	Ref.
Negro	-0,0004	0,0008	-0,0052***	0,0010
Amarelo	-0,0347***	0,0017	-0,0242***	0,0020
Pardo	0,0011**	0,0005	-0,0010*	0,0006
Muito Carente	Ref.	Ref.	Ref.	Ref.
Carente	0,0063***	0,0006	0,0300***	0,0007
Média	-0,0165***	0,0010	0,0017	0,0012
Desenvolvida	0,0945***	0,0008	0,1786***	(0,0009
Muito Desenvolvida	0,5108***	0,0014	0,4328***	(0,0016
Tamanho da família	-0,0019***	0,0001	-0,0010***	(0,0001
Saneamento	0,0012*	0,0007	0,0266***	0,0008
Coleta	0,0099***	0,0010	0,0053***	0,0007
<i>Constante</i>	0,8246***	0,0055	2,3106***	0,0101
R ² / Pseudo R ²	0,9231		0,8146	
Estat. F	355840,82***		130290,52***	
Durbin (score, <i>Chi</i>)	0,15		27,88***	
Wu-Hausman (F)	0,15		27,89***	

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. D.P: desvio padrão.

A acessibilidade ao emprego é negativamente relacionada com a probabilidade de se trabalhar no setor informal, sugerindo que a informalidade pode ser uma escolha relacionada à

falta de acesso. A escolha de trabalhar no setor informal é muitas vezes vista como uma resposta à falta de acessibilidade nas grandes cidades dos países em desenvolvimento (AVNER & LALL, 2016). A Tabela 2.11 aponta para o acesso desigual aos empregos em termos espaciais como determinante da informalidade, e sugere que a baixa acessibilidade aos empregos formais, devido às características da estrutura espacial, determina quem vai acabar trabalhando no setor informal.

Tabela 2.11 – Efeitos marginais para a probabilidade de trabalhadores informais (regressões com todos os indivíduos).

Variáveis	CBD				Subcentros			
	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
Informalidade								
LnACi	-0,0048*** (0,0010)	-0,0042*** (0,0010)	-0,0048*** (0,0011)	-0,0034*** (0,0011)	-0,0080*** (0,0013)	-0,0078*** (0,0013)	-0,0133*** (0,0016)	0,0021 (0,0016)
Masculino	-0,0741*** (0,0009)	-0,0715*** (0,0009)	-0,0741*** (0,0009)	-0,0724*** (0,0009)	-0,0741*** (0,0009)	-0,0716*** (0,0009)	-0,0741*** (0,0009)	-0,0724*** (0,0009)
Idade	-0,0093*** (0,0001)	-0,0080*** (0,0001)	-0,0093*** (0,0001)	-0,0079*** (0,0001)	-0,0093*** (0,0001)	-0,0080*** (0,0001)	-0,0093*** (0,0001)	-0,0079*** (0,0001)
Idade ²	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	-0,0443*** (0,0014)	-0,0376*** (0,0011)	-0,0443*** (0,0014)	-0,0353*** (0,0011)	-0,0443*** (0,0014)	-0,0376*** (0,0011)	-0,0443*** (0,0014)	-0,0353*** (0,0011)
Alta	-0,1114*** (0,0012)	-0,1016*** (0,0010)	-0,1114*** (0,0012)	-0,0944*** (0,0010)	-0,1114*** (0,0012)	-0,1017*** (0,0010)	-0,1116*** (0,0012)	-0,0945*** (0,0010)
Bem elevada	-0,1521*** (0,0016)	-0,1612*** (0,0010)	-0,1521*** (0,0016)	-0,1241*** (0,0010)	-0,1519*** (0,0016)	-0,1610*** (0,0010)	-0,1519*** (0,0016)	-0,1240*** (0,0010)
Casado	-0,0309*** (0,0010)	-0,0332*** (0,0010)	-0,0309*** (0,0010)	-0,0326*** (0,0010)	-0,0309*** (0,0010)	-0,0332*** (0,0046)	-0,0310*** (0,0010)	-0,0327*** (0,0010)
Chefe	0,0018*** (0,0010)	-0,0020** (0,0010)	0,0018* (0,0010)	-0,0020** (0,0010)	0,0018* (0,0010)	-0,0020** (0,0010)	0,0018* (0,0010)	-0,0021** (0,0010)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0147*** (0,0018)	0,0137*** (0,0019)	0,0147*** (0,0018)	0,0140*** (0,0019)	0,0147*** (0,0018)	0,0141*** (0,0019)	0,0148*** (0,0018)	0,0142*** (0,0019)
Amarelo	-0,0015 (0,0035)	-0,0032 (0,0037)	-0,0015 (0,0035)	-0,0031 (0,0037)	-0,0017 (0,0035)	-0,0033 (0,0037)	-0,0017 (0,0035)	-0,0030 (0,0037)
Pardo	0,0101*** (0,0010)	0,0095*** (0,0010)	0,0101*** (0,0010)	0,0095*** (0,0010)	0,0102*** (0,0010)	0,0096*** (0,0010)	0,0102*** (0,0010)	0,0096*** (0,0010)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	0,0021* (0,0013)	0,0011 (0,0013)	0,0021* (0,0013)	0,0012 (0,0013)	0,0026** (0,0013)	0,0018 (0,0013)	0,0043*** (0,0013)	0,0031*** (0,0013)
Média	0,0068*** (0,0020)	0,0063*** (0,0020)	0,0068*** (0,0020)	0,0063*** (0,0020)	0,0056*** (0,0021)	0,0050*** (0,0020)	0,0039* (0,0021)	0,0034* (0,0020)
Desenvolvida	0,0065*** (0,0018)	0,0042** (0,0018)	0,0065*** (0,0018)	0,0043** (0,0019)	0,0073*** (0,0017)	0,0054*** (0,0018)	0,0107** (0,0018)	0,0061*** (0,0019)
Muito Desenvolvida	0,0099*** (0,0031)	0,0082** (0,0035)	0,0100*** (0,0032)	0,0079** (0,0035)	0,0102*** (0,0029)	0,0091*** (0,0033)	0,0155*** (0,0031)	0,0076*** (0,0034)
Tamanho da família	0,0038*** (0,0002)	0,0031*** (0,0002)	0,0038*** (0,0002)	0,0031*** (0,0002)	0,0038*** (0,0002)	0,0031*** (0,0002)	0,0038*** (0,0002)	0,0031*** (0,0002)
Saneamento	-0,0190*** (0,0014)	-0,0164*** (0,0014)	-0,0190*** (0,0014)	-0,0167*** (0,0014)	-0,0185*** (0,0014)	-0,0162*** (0,0062)	-0,0177*** (0,0015)	-0,0159*** (0,0015)
Coleta	-0,0131*** (0,0021)	-0,0108*** (0,0021)	-0,0131*** (0,0021)	-0,0110*** (0,0021)	-0,0133*** (0,0021)	-0,0113*** (0,0021)	-0,0137*** (0,0021)	-0,0117*** (0,0021)
Constante	0,535*** (0,0120)	-	0,5364*** (0,0129)	-	0,5854*** (0,0173)	-	0,6499*** (0,0212)	-
Estat. F/Wald	1411,85	26042,45	1411,70	24430,34	1412,68	26060,38	1414,08	24475,41
R ² / Pseudo R ²	0,0454	0,0533	0,0454	-	0,0455	0,0533	0,0454	-
Observações	563.463	563.463	563.463	563.463	563.463	563.463	563.463	563.463

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Estimativas com base na média dos efeitos marginais.

O tamanho da informalidade econômica é uma resposta à incapacidade da economia formal para acomodar todos os trabalhadores, então a estrutura econômica urbana formal determinará quais indivíduos se dedicam a atividades econômicas informais e onde moram (não onde eles decidem viver). A localização do trabalho informal pode, portanto, depender, em parte, da distância ao local de residência, para que tais empregos estejam localizados em um ponto em que a renda seja maximizada e o custo de transporte minimizado. Nesse caso, quando observados os modelos com variáveis instrumentais, os resultados novamente corroboram com a literatura do *mismatch* espacial, ou seja, o melhor acesso à oportunidades de empregos diminui a probabilidade de um indivíduo estar no setor informal. Observando os efeitos marginais das variáveis de acesso aos empregos e tomando o valor da coluna 7 da tabela, temos: um aumento de 10% do índice de acessibilidade está associado a uma redução de quase 1% na chance de se estar na informalidade (Tabela 2.11). Os efeitos marginais parecem assim grandes quando observado na amostra a chance de se estar na informalidade que é de 0,15 com desvio padrão de 0,36 (Tabela 2.3). Os resultados encontrados seguem na mesma linha de diversas evidências empíricas sobre o tema, tais como os trabalhos de Miltra (2005), Suárez et al. (2016), Chen *et al.* (2017), Moreno-Monroy e Ramos (2015), Motte *et al.*, (2016) e Boisjoly *et al.*, (2017), apesar dessas pesquisas apresentarem níveis de análise distintos e outras metodologias.

Compreender melhor o nível de capital humano e a acessibilidade relacionados às cidades dos países em desenvolvimento é importante. Como a literatura do *mismatch* espacial mostra, famílias mais pobres provavelmente serão mais penalizadas com viagens dispendiosas e longas. Estas também são susceptíveis de serem menos escolarizadas e dependem principalmente do setor informal. Observando agora todos os demais efeitos marginais, percebe-se que os sinais dos controles parecem corroborar com a literatura, por exemplo, trabalhadores mais escolarizados tem menor probabilidade de trabalhar no setor informal.

A Tabela 2.12, a seguir, apresenta a relação entre acessibilidades ao emprego e resultados do mercado de trabalho, mostrando padrões mais distintos quando se permite a heterogeneidade entre os gêneros. Os resultados apresentam efeitos heterogêneos para homens e mulheres, com estimativas OLS e ML. Os resultados para os primeiros estágios estão no apêndice na Tabela A2.7, bem como os efeitos marginais do 2º estágio para as variáveis de controle (Tabelas A2.8 e A2.9). Dessa forma, observa-se que a relação entre o melhor acesso físico ao CBD para homens não foi significativo, enquanto para as mulheres se obteve uma

relação inversa e significativa. Quando analisado os subcentros como oportunidades de empregos, a relação foi significativa para ambos os gêneros, sendo bem mais forte para o sexo feminino.

Além disto, note que os efeitos marginais obtidos para as mulheres são maiores que aqueles obtidos para amostra total. Estes resultados sugerem que existe uma segmentação de gênero nos mercados de trabalho, com predominância da informalidade para as mulheres. De acordo com a Teoria da Responsabilidade Familiar (Johnston-Anumonwo, 1992), as mulheres geralmente têm uma maior responsabilidade em manter a família e cuidar de dependentes (por exemplo, filhos) do que os homens. As demandas competitivas por tempo resultam em mulheres menos móveis - portanto, seu comprimento de deslocamento reduzido (Crane, 2007; Giuliano, 1998; Lee e McDonald, 2003; Madden, 1981; Turner e Niemeier, 1997) o que diminui o seu esforço na busca por emprego nos centros (CBD e subcentros) onde prevalece a formalidade e acabam recorrendo aos empregos informais, próximos de casa.

Tabela 2.12 – Efeitos marginais com e sem variáveis instrumentais para a probabilidade de estar na informalidade (regressões para homens e mulheres)

Variáveis	CBD				Subcentros			
	HOMENS							
Informalidade	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
LnACi	-0,0014 (0,0013)	-0,0007 (0,0013)	-0,0016 (0,0014)	-0,0046 (0,0073)	-0,0044*** (0,0016)	-0,0037** (0,0016)	-0,0083*** (0,0020)	-0,0378*** (0,0101)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	310.912	310.912	310.912	310.912	310.912	310.912	310.912	310.912
Estat. F/Wald	596,16	10012,90	596,17	9646,69	596,49	10017,63	597,02	9660,16
R ² / Pseudo R ²	0,0334	0,0422	0,0334		0,0334	0,0422	0,0334	
MULHERES								
LnACi	-0,0102*** (0,0017)	-0,0100*** (0,0017)	-0,0104*** (0,0018)	-0,0086*** (0,0019)	-0,0148*** (0,0002)	-0,0148*** (0,0022)	-0,0224*** (0,0027)	-0,0001*** (0,0027)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	252.551	252.551	252.551	252.551	252.551	252.551	252.551	252.551
Estat. F/Wald	823,34	14364,35	823,08	13467,59	823,92	14376,81	825,13	13506,34
R ² / Pseudo R ²	0,0554	0,0581	0,0554		0,0555	0,0581	0,0554	

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Estimativas com base na média dos efeitos marginais.

Ressalta-se que a relação entre o melhor acesso físico ao CBD para homens não foi significativo mesmo utilizando o instrumento como variável exógena, enquanto para as mulheres do mesmo modo obteve uma relação inversa e significativa. Neste sentido, entende-se que as barreiras espaciais não são iguais para toda a população no caso da RMSP. As mulheres são as mais sensíveis às condições locais do mercado de trabalho.

2.6.5 Overeducation

Nos países desenvolvidos, a literatura (Groot e Maassen van den Brink, 2000; Rubb, 2003; Macguiness, 2006) acerca da sobreeducação encontra-se em expansão. Há vários autores preocupados com a existência de diferenças entre os requisitos educacionais das ocupações e a escolaridade possuída pelos indivíduos e os seus efeitos. Contudo, o presente estudo visa lidar com os efeitos da acessibilidade ao emprego com a chance do indivíduo ser *overeducated*. A partir da Tabela 2.13, regressões para o 1º Estágio, nota-se que existe uma associação positiva e estatisticamente significativa entre o instrumento (acesso físico ao emprego por rios) e a variável dependente, algo bem vindo por ter confiança em usar a variável como instrumento.

Tabela 2.13 – Resultado da estimação para a probabilidade de *overeducated* (regressões com todos os indivíduos) – 1º Estágio

Variáveis	CBD		Subcentros	
	MPL_IV	D.P.	MPL_IV	D.P.
LnACi_IV	0,9163***	0,0005	0,8112***	0,0103
Masculino	-0,0003***	0,0004	-0,0006	0,0004
Idade	-0,0003***	0,0000	-0,0001	0,0000
Idade ²	0,0000***	0,0000	0,0000	0,0000
Baixa	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0008*	0,0005	0,0041***	0,0006
Alta	-0,0038***	0,0007	-0,0001	0,0008
Casado	-0,0087***	0,0005	-0,0060***	0,0006
Chefe	0,0039***	0,0005	0,0021***	0,0006
Branco	Ref.	Ref.	Ref.	Ref.
Negro	-0,0004	0,0009	-0,0052***	0,0010
Amarelo	-0,0347***	0,0017	-0,0242***	0,0020
Pardo	0,0011**	0,0005	0,0010*	0,0006
Muito Carente	Ref.	Ref.	Ref.	Ref.
Carente	0,0063***	0,0006	0,0300***	0,0007
Média	-0,0165***	0,0010	0,0017	0,0012
Desenvolvida	0,0946***	0,0008	0,1786***	0,0009
Muito Desenvolvida	0,5108***	0,0014	0,4328***	0,0016
Tamanho da família	-0,0019***	0,0001	-0,0010***	0,0001
Saneamento	0,0012*	0,0007	0,0266***	0,0008
Coleta	0,0099***	0,0010	0,0099***	0,0010
<i>Constante</i>	0,8415***	0,0060	2,3088***	0,0103
R ² / Pseudo R ²	0,9231		0,8146	
Estat. F	375610,09***		137529,13***	
Durbin (score, <i>Chi</i>)	21,07***		8,36***	
Wu-Hausman (F)	21,07***		8,36***	

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

A estatística F para o primeiro estágio é estatisticamente significativa e assume valores altos, bem como as estatísticas Durbin e Wu-Hausman indicaram significância nos resultados. Assim, o estimador IV é preferível ao estimador OLS. O teste de Durbin-Wu-Hausman indica

que os resíduos de uma regressão do índice de acessibilidade sobre as demais variáveis são estatisticamente significantes quando colocados como regressores de uma regressão para a condição de *overeducated* sobre todas as variáveis ditas explicativas. Além disto, os resultados da primeira etapa fornecem confiança para o uso do instrumento proposto; utilizando a variável de acessibilidade por rios como regressor para o índice de acessibilidade, obteve-se uma influência positiva e estatisticamente significativa. De fato, o conjunto de estatísticas apresentadas para avaliar o instrumento fornece forte confiança para usá-lo.

Na Tabela 2.14, tem-se os resultados dos modelos de Probabilidade Linear e Probit com as variáveis de interesse, índices de acessibilidades e *overeducated* para o CBD e subcentros. Estes resultados são importantes, pois podem evidenciar a má alocação existente no mercado de trabalho da RMSP. Note-se que, os coeficientes do índice de acessibilidade para o CBD são positivos e significativos nos modelos sem variáveis instrumentais. Esses resultados não são esperados de acordo com a teoria do *mismatch* espacial e educacional. Ao analisar os subcentros de emprego, percebe-se que os coeficientes são negativos e significativos para todos os modelos estimados. Ou seja, o melhor acesso físico aos centros de emprego diminui a probabilidade do indivíduo ser *overeducated*, isto é, apresentar uma educação maior que a exigida no trabalho. Observando os efeitos marginais das variáveis de acesso aos empregos e tomando o valor da coluna 7 da tabela, temos: um aumento de 10% do índice de acessibilidade está associado a uma redução de 0,56% na chance de ser *overeducated*.

Como esperado, o efeito é positivo com o aumento do nível educacional, como mostra nas estimativas da Tabela 2.14. Observa-se que os coeficientes das variáveis para gênero, idade, casado e raça são significativos e com efeitos esperados. Ou seja, olhando especificamente para a *dummy* de sexo, pode-se dizer que além do *mismatch* educacional existente há uma possível discriminação de gênero no mercado de trabalho, visto que a probabilidade de chance de ser *overeducated* diminui com o indivíduo sendo do gênero masculino. O mesmo acontece com as variáveis de raça, visto que o fato de ser pardo ou negro aumenta a chance de apresentarem sobreeducação. Ser chefe do agregado familiar não apresenta coeficientes significativos e para as variáveis de vizinhança, percebe-se que além da significância, a magnitude do efeito aumenta com o melhor nível de vizinhança.

Tabela 2.14 – Efeitos marginais para a probabilidade de *overeducated* (regressões com todos os indivíduos).

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
Overeducated	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	0,0035*** (0,0012)	0,0042*** (0,0012)	0,0010 (0,0013)	0,0058 (0,0045)	-0,0045*** (0,0016)	-0,0032** (0,0016)	-0,0079*** (0,0020)	-0,0026** (0,0019)
Masculino	-0,0096*** (0,0011)	-0,0078*** (0,0011)	-0,0096*** (0,0011)	-0,0087*** (0,0012)	-0,0045*** (0,0016)	-0,0079*** (0,0011)	-0,0097*** (0,0011)	-0,0087*** (0,0012)
Idade	-0,0057*** (0,0002)	-0,0054*** (0,0002)	-0,0057*** (0,0002)	-0,0059*** (0,0002)	-0,0057*** (0,0002)	-0,0054*** (0,0002)	-0,0057*** (0,0002)	-0,0059*** (0,0002)
Idade ²	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0000*** (0,0000)	0,0001*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,3774*** (0,0013)	0,3460*** (0,0010)	0,3773*** (0,0013)	0,3995*** (0,0014)	0,3772*** (0,0013)	0,3458*** (0,0010)	0,3771*** (0,0013)	0,3993*** (0,0014)
Alta	0,2590*** (0,0017)	0,2597*** (0,0016)	0,2589*** (0,0017)	0,3185*** (0,0021)	0,2589*** (0,0017)	0,2595*** (0,0016)	0,2589*** (0,0017)	0,3185*** (0,0021)
Casado	-0,0047*** (0,0012)	-0,0030** (0,0012)	-0,0048*** (0,0012)	-0,0033** (0,0013)	-0,0049*** (0,0012)	-0,0032** (0,0012)	-0,0050*** (0,0012)	-0,0036** (0,0013)
Chefe	-0,0001 (0,0012)	-0,0030 (0,0012)	-0,0001 (0,0012)	-0,0010 (0,0014)	-0,0001 (0,0012)	-0,0009 (0,0012)	-0,0001 (0,0012)	-0,0010 (0,0014)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0063*** (0,0022)	0,0028 (0,0022)	0,0064*** (0,0022)	0,0031 (0,0025)	0,0064*** (0,0022)	0,0029 (0,0022)	0,0065*** (0,0022)	0,0032 (0,0025)
Amarelo	-0,0062 (0,0042)	-0,0029 (0,0040)	-0,0062 (0,0042)	-0,0031 (0,0045)	-0,0062 (0,0042)	-0,0027 (0,0040)	-0,0062 (0,0042)	-0,0030 (0,0045)
Pardo	0,0173*** (0,0012)	0,0141*** (0,0012)	0,0174*** (0,0012)	0,0157*** (0,0014)	0,0174*** (0,0012)	0,0142*** (0,0012)	0,0174*** (0,0012)	0,0157*** (0,0014)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0242*** (0,0015)	-0,0211*** (0,0015)	-0,0230*** (0,0015)	-0,0231*** (0,0016)	-0,0211*** (0,0015)	-0,0181*** (0,0015)	-0,0199*** (0,0015)	-0,0198*** (0,0016)
Média	-0,0057** (0,0024)	-0,0061** (0,0028)	-0,0064** (0,0024)	-0,0067** (0,0028)	-0,0080*** (0,0024)	-0,0082*** (0,0025)	-0,0091*** (0,0025)	-0,0090*** (0,0028)
Desenvolvida	-0,0731*** (0,0021)	-0,0640*** (0,0021)	-0,0707*** (0,0021)	-0,0673*** (0,0021)	-0,0669*** (0,0020)	-0,0580*** (0,0020)	-0,0647*** (0,0020)	-0,0613*** (0,0020)
Muito Desenvolvida	-0,1483*** (0,0037)	-0,1330*** (0,0037)	-0,1441*** (0,0038)	-0,1268*** (0,0029)	-0,1379*** (0,0035)	-0,1227*** (0,0034)	-0,1344*** (0,0035)	-0,1185*** (0,0028)
Tamanho da família	0,0021*** (0,0003)	0,0019*** (0,0003)	0,0021*** (0,0003)	0,0021*** (0,0003)	0,0021*** (0,0003)	0,0019*** (0,0003)	0,0021*** (0,0003)	0,0021*** (0,0003)
Saneamento	0,0089*** (0,0017)	0,0110*** (0,0018)	0,0092*** (0,0017)	0,0120*** (0,0019)	0,0100*** (0,0017)	0,0119*** (0,0018)	0,0104*** (0,0017)	0,0131*** (0,0019)
Coleta	0,0061** (0,0026)	0,0083** (0,0026)	0,0059** (0,0026)	0,0092** (0,0029)	0,0055** (0,0026)	0,0078*** (0,0026)	0,0053** (0,0026)	0,0086*** (0,0029)
Constante	0,2282*** (0,0144)	-	0,2550*** (0,0156)	-	0,3221*** (0,0210)	-	0,3652*** (0,0257)	-
Estat. F/Wald	5658,61	90187,15	101850,59	82875,69	5658,63	90179,85	101865,85	82887,97
R ² / Pseudo R ²	0,1531	0,1311	0,1531	-	0,1531	0,1311	0,1531	-
Observações	563.463	563.463	563.463	563.463	563.463	563.463	563.463	563.463

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Estimativas com base na média dos efeitos marginais.

A Tabela 2.15 mostra os resultados das regressões separadas por gêneros. Os resultados para os primeiros estágios estão no apêndice na Tabela A2.10, bem como os efeitos marginais do 2º estágio para as variáveis de controle (Tabelas A2.11 e A2.12). Os coeficientes para os subcentros novamente apresentam sinais esperados, com efeito, um pouco mais forte para o gênero feminino quando significativos. Para o CBD, os resultados para os homens apresentaram coeficientes positivos e significativos, enquanto para as mulheres não foram

significativos. Os homens, em geral, estão mais dispostos a percorrerem longas distâncias para adquirirem emprego e o CBD para muitos indivíduos da amostra é distante com relação à localização residencial.

Note que estes novos resultados estão em linha com resultados anteriores: acesso é muito mais importante para as mulheres. Para os homens, tomando os resultados da coluna 7, tem-se que um aumento de 10% no indicador de acessibilidade está associado a uma redução de 0,23% na chance de se estar na condição de *overeducated*. Para as mulheres este efeito marginal é mais forte. Tomando como exemplo a mesma coluna 7, tem-se que a redução na chance de se estar na condição de *overeducated* será de 0,33% para o mesmo aumento de 10% no índice de acessibilidade aos empregos.

Tabela 2.15 – Efeitos marginais com e sem variáveis instrumentais para a probabilidade de *overeducated* (regressões para homens e mulheres)

Variáveis	CBD				Subcentros			
	HOMENS							
<i>Overeducated</i>	MPL (1)	Probit (2)	MPL_IV (3)	Probit_IV (4)	MPL (5)	Probit (6)	MPL_IV (7)	Probit_IV (8)
LnACi	0,0049*** (0,0017)	0,0058*** (0,0016)	0,0023 (0,0018)	0,0187*** (0,0018)	-0,0044** (0,0021)	-0,0027 (0,0021)	-0,0067** (0,0026)	0,0007 (0,0026)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	310.912	310.912	310.912	310.912	310.912	310.912	310.912	310.912
Estat. F/Wald	3145,34	47148,84	53466,69	43812,38	3145,05	47138,17	53470,93	43814,64
R ² / Pseudo R ²	0,1468	0,1267	0,1467	-	0,1467	0,1267	0,1467	-
MULHERES								
LnACi	0,0013 (0,0019)	0,0018 (0,0019)	-0,0012 (0,0021)	-0,0025 (0,0068)	-0,0053** (0,0024)	-0,0045** (0,0024)	-0,0105*** (0,0030)	-0,0048 (0,0030)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	252.551	252.551	252.551	252.551	252.551	252.551	252.551	252.551
Estat. F/Wald	2842,27	42971,19	48321,58	38928,59	2842,57	42973,73	48333,37	38939,33
R ² / Pseudo R ²	0,1606	0,1364	0,1606	-	0,1606	0,1364	0,1606	-

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Estimativas com base na média dos efeitos marginais.

2.6.6 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 (Tabela 2.16), foi utilizada uma subamostra de jovens adultos de 19 a 25 anos que moram com os pais, visto que a escolha do local de residência dos pais pode ser considerada

exógena ao *status* de emprego de jovens adultos. Segundo, foi considerada uma subamostra com os 50% dos indivíduos mais pobres. A ideia é que tais amostras incluem indivíduos com mais dificuldades de "sorting", ou seja, indivíduos mais pobres em consequência da mais baixa renda terão menor flexibilidade da escolha da moradia em relação ao local de emprego (Tabela 2.17). Terceiro, 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 à uma condição de menor influência das oportunidades de trabalho sobre a escolha da moradia. (Tabela 2.18). Por fim, complementou-se a análise de robustez com resultados para diferentes medidas de acessibilidade utilizando diferentes medidas para a função de distância (Tabela 2.19). A expectativa é que, assim, tal expediente reduziria potenciais endogeneidades associadas à relação entre acessibilidade ao emprego e localização residencial.

Nas Tabelas, a seguir, foram calculados o efeito da acessibilidade na probabilidade de participação no mercado, de estar desempregado, de estar na informalidade e na condição de *overeducated* para amostras diferentes. Nesta seção serão analisados apenas os resultados para o 2º estágio das regressões, os resultados para o 1º estágio estão no apêndice. Percebe-se que o melhor acesso aos centros de empregos diminui a probabilidade de participação. Este resultado pode ser justificado por características parentais influenciando jovens a não procurarem emprego, além do mais, grande parte desses jovens ainda estudam. Essas subamostras não podem abordar o viés de seleção se pais e filhos compartilhassem a mesma heterogeneidade não observada, por exemplo, em termos de produtividade (ASLUND *et al.*, 2010).

Em relação às variáveis dependentes como o desemprego e a informalidade apresentam relação inversa com o índice de acessibilidade. O melhor acesso as oportunidades de emprego diminui a probabilidade de o indivíduo jovem estar desempregado e no setor informal. Tomando os coeficientes dos efeitos marginais da coluna para subcentros com estimativas MPL_IV, tem-se que o aumento de 1% no índice de acessibilidade está associado a uma redução de 0,13% na probabilidade do indivíduo estar desempregado e de 0,07% de estar no setor informal. Comparando estes resultados com aqueles obtidos para a amostra geral, percebe-se que os efeitos marginais para os jovens são um pouco mais fracos. Por outro lado, para a condição de *overeducated*, percebe-se que o efeito é mais forte para os indivíduos jovens do que para a amostra em geral. Ou seja, o melhor acesso ao emprego reduz a chance de se estar na condição de *overeducated* (escolaridade maior do que a exigida em seu trabalho).

Tabela 2.16 – Efeitos marginais da acessibilidade na probabilidade de participação, desempregados, na informalidade e condição de *overeducated* (regressões com jovens adultos que moram com os pais).

Variáveis	CBD		Subcentros	
PEA	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,0169*** (0,0038)	-0,0130*** (0,0037)	-0,0223*** (0,0050)	-0,0122*** (0,0050)
Controles	Sim	Sim	Sim	Sim
Observações	79.037	79.037	79.037	79.037
Estat. F/Wald	325,34	4813,54	325,32	4814,31
R ² / Pseudo R ²	0,0654	-	0,0654	-
Desemprego	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,0147*** (0,0037)	-0,0045 (0,0038)	-0,0186*** (0,0049)	-0,0104*** (0,0049)
Controles	Sim	Sim	Sim	Sim
Observações	61.480	61.480	61.480	61.480
Estat. F/Wald	84,30	1360,42	84,24	1359,61
R ² / Pseudo R ²	0,0228	-	0,0228	-
Informalidade	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,0077** (0,0041)	-0,0239*** (0,0041)	-0,0104** (0,0055)	-0,0081 (0,0054)
Controles	Sim	Sim	Sim	Sim
Observações	52.610	52.610	52.610	52.610
Estat. F/Wald	48,57	788,92	48,57	789,23
R ² / Pseudo R ²	0,0155	-	0,0155	-
Overeducated	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,0413*** (0,0055)	-0,0095* (0,0055)	-0,0718*** (0,0072)	-0,0140** (0,0072)
Controles	Sim	Sim	Sim	Sim
Observações	52.610	52.610	52.610	52.610
Estat. F/Wald	304,28	4671,52	306,80	4712,27
R ² / Pseudo R ²	0,0894	-	0,0898	-

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs.: Controles adicionais: Idade, Idade², nível de escolaridade, *dummy* para masculino, raça, tipos de vizinhança, *dummy* para infraestrutura, termo constante.

Fonte: Elaboração própria com base na média dos efeitos marginais.

A fim de obter uma estimativa da relação entre acessibilidade aos empregos nos resultados do mercado de trabalho, como adiantado acima, é necessário considerar o desafio colocado por problemas econométricos tradicionais: a possibilidade de causalidade reversa (*sorting*). Na causalidade reversa, os trabalhadores com alta renda tendem a escolher morar próximos aos empregos, gerando uma associação espúria e fraca entre a variável dependente e o índice de acessibilidade. Com tal preocupação, considerou-se na Tabela 2.17 os 50% dos indivíduos mais pobres da amostra. Novamente, os resultados encontrados para participação no mercado, condição de desemprego e informalidade corroboram com a literatura teórica. A participação aumenta quando há um aumento no indicador de acessibilidade, contudo observando os valores dos efeitos marginais, percebe-se que os mesmos são mais fracos para os mais pobres do que para a amostra em geral. Tomando o valor de (0,0059) na tabela, tem-

se que, o aumento em 10% no acesso aos empregos está associado a um aumento de 0,08% na probabilidade de participação.

Tabela 2.17 – Efeitos marginais da acessibilidade na participação, desemprego, informalidade e de *overeducated* (50% mais pobres da amostra)

Variáveis	CBD		Subcentros	
PEA	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	0,0097*** (0,0014)	-0,0012 (0,0014)	0,0059*** (0,0019)	0,0086*** (0,0019)
Controles	Sim	Sim	Sim	Sim
Observações	621.342	621.342	621.342	621.342
Estat. F/Wald	5720,69	82507,03	5718,60	82481,78
R ² / Pseudo R ²	0,1422	-	0,1421	-
Desemprego	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,0133*** (0,0012)	-0,0066*** (0,0012)	-0,0135*** (0,0016)	-0,0021*** (0,0016)
Controles	Sim	Sim	Sim	Sim
Observações	431.510	431.510	431.510	431.510
Estat. F/Wald	1194,98	18376,92	1192,16	18591,65
R ² / Pseudo R ²	0,0475	-	0,0473	-
Informalidade	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,0097*** (0,0016)	-0,0107*** (0,0016)	-0,0178*** (0,0021)	-0,0018* (0,0021)
Controles	Sim	Sim	Sim	Sim
Observações	382.892	382.892	382.892	382.892
Estat. F/Wald	898,01	15813,84	899,77	15848,61
R ² / Pseudo R ²	0,0427	-	0,0427	-
Overeducated	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	0,0058*** (0,0017)	0,0210*** (0,0017)	-0,0012 (0,0022)	0,0057*** (0,0022)
Controles	Sim	Sim	Sim	Sim
Observações	382.892	382.892	382.892	382.892
Estat. F/Wald	6536,12	81234,94	6535,08	81230,28
R ² / Pseudo R ²	0,2249	-	0,2249	-

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs.: Controles adicionais: Tipos de vizinhança, Idade, Idade², nível de escolaridade, *dummy* para masculino, *dummy* para chefe, raça, *dummy* para infraestrutura, termo constante.

Fonte: Elaboração própria com base na média dos efeitos marginais.

Analisando os resultados para o desemprego e tomando o coeficiente de 0,0135 da tabela, tem-se que um aumento de 1% no índice de acessibilidade está associado a uma redução de 0,12% na probabilidade de estar desempregado. Este resultado é menor do que a amostra em geral, o que sugere que o melhor acesso para os mais pobres não seja tão determinante para reduzir o desemprego. Para a informalidade, tomando o valor do coeficiente de 0,0178 na tabela, tem-se que um aumento em 10% no acesso reduz em quase 1% na chance de estar no setor informal. Esta evidência corrobora com os resultados obtidos na amostra em geral. A condição de informalidade é comum na população em geral, e esta condição está ainda mais presente para a população mais pobre, onde busca, muitas vezes, o setor informal como forma de sobrevivência. Por fim, observando os resultados para a condição de *overeducated*, percebe-se que os valores foram positivos e significativos tanto

para o CBD quanto para os subcentros, o que sugere que o melhor acesso para a população mais pobre não é suficiente para diminuir a probabilidade de estar na condição de sobreeducado (Tabela 2.17).

A acessibilidade afeta o emprego das diferentes maneiras. Alguns mecanismos são mais relevantes para alguns grupos populacionais do que outros (GOBILLON & SELOD, 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 2.18, a seguir, considerou-se os indivíduos que moram em casa própria. Os resultados corroboram para a robustez das evidências já obtidas para a amostra em geral.

Tabela 2.18 – Efeitos marginais da acessibilidade na probabilidade de participação, desempregados e informalidade (indivíduos que moram em casa própria)

Variáveis	CBD		Subcentros	
	MPL_IV	Probit_IV	MPL_IV	Probit_IV
PEA				
LnACi	0,0084*** (0,0014)	0,0023* (0,0014)	0,0085*** (0,0019)	0,0013*** (0,0019)
Controles	Sim	Sim	Sim	Sim
Observações	602.773	602.773	602.773	602.773
Estat. F/Wald	122461,90	93566,39	122443,98	93559,37
R ² / Pseudo R ²	0,1689	-	0,1688	--
Desemprego				
LnACi	-0,0081*** (0,0010)	-0,0031*** (0,0011)	-0,0107*** (0,0014)	-0,0004*** (0,0014)
Controles	Sim	Sim	Sim	Sim
Observações	438.308	438.308	438.308	438.308
Estat. F/Wald	1335,87	20612,77	1335,67	20614,37
R ² / Pseudo R ²	0,0547	-	0,0547	-
Informalidade				
LnACi	-0,0046*** (0,0014)	-0,0059*** (0,0014)	-0,0107*** (0,0014)	-0,0031* (0,0019)
Controles	Sim	Sim	Sim	Sim
Observações	401.342	401.342	401.342	401.342
Estat. F/Wald	18251,94	16606,90	18271,44	16627,23
R ² / Pseudo R ²	0,0435	-	0,0435	-
Overeducated				
LnACi	-0,0061*** (0,0017)	-0,0012** (0,0005)	-0,0131*** (0,0023)	-0,0026 (0,0023)
Controles	Sim	Sim	Sim	Sim
Observações	401.342	401.342	401.342	401.342
Estat. F/Wald	3710,91	55214,01	3712,09	55227,27
R ² / Pseudo R ²	0,1427	-	0,1427	-

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs .: Controles adicionais: Tipos de vizinhança, Idade, Idade², nível de escolaridade, *dummy* para masculino, *dummy* para chefe, raça, *dummy* para infraestrutura, termo constante.

Fonte: Elaboração própria com base na média dos efeitos marginais.

O melhor acesso as oportunidades de empregos aumentam a participação no mercado, diminuem a condição de se estar desempregado, na informalidade ou de *overeducated*. Comparando os coeficientes obtidos na tabela dos indivíduos que moram em casa própria com aqueles da amostra em geral, percebe-se que os efeitos marginais são mais fracos para esta subamostra, porém como já dito, significativos (Tabela 2.18).

Em uma segunda etapa, complementou-se a análise de robustez com resultados para diferentes medidas de acessibilidade utilizando-se 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 2.19, a seguir, 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 Johnson (2006), Quitanar *et al.* (2013) e Carvalho (2016), por exemplo, utilizaram diferentes medidas para a função de impedância como forma de buscar robustez em seus resultados.

Percebe-se através da Tabela 2.19, que os resultados obtidos inicialmente não parecem sensíveis ao grau de decaimento da função associada à distância ao emprego, ou seja, não alteram de forma importante os resultados já obtidos para o grau de acessibilidade ao emprego nos resultados do mercado de trabalho quanto aos sinais. Contudo, a magnitude dos coeficientes são maiores quando considerados o grau de decaimento de -0,5 e -1,5 comparados ao grau de -1,0, mostrando que o resultado do índice de acessibilidade nas variáveis dependentes tem efeito mais forte. Especificamente para a condição de *overeducated*, percebe-se que os coeficientes das regressões são semelhantes aos coeficientes dos resultados principais, foram significativos e com sinais esperados apenas para os subcentros de empregos. O melhor acesso físico as oportunidades de empregos reduz as chances dos indivíduos serem sobreeducados, ou seja, apresentarem grau de escolaridade maior que o trabalho de ocupação.

Tabela 2.19 – Efeitos marginais da acessibilidade na probabilidade de participação, desemprego, informalidade e *overeducated* (regressões para diferentes funções de distância)

Variáveis	Função de distância $d_{ij}^{-0,5}$				Função de distância $d_{ij}^{-1,5}$			
	CBD		Subcentros		CBD		Subcentros	
PEA	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	0,1181*** (0,0094)	0,0087* (0,0093)	0,0165*** (0,0015)	-0,0014 (0,0015)	0,0393*** (0,0031)	0,0029*** (0,0031)	0,0032*** (0,0005)	0,0007*** (0,0005)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	835.804	835.804	835.804	835.804	835.804	835.804	835.804	835.804
Estat. F/Wald	8406,37	125909,94	8463,62	123742,88	8406,37	125909,94	8458,45	123694,38
R ² / Pseudo R ²	0,1554	-	0,1613	-	0,1554	-	0,1612	-
Desemprego	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,0942*** (0,0074)	-0,0071*** (0,0074)	-0,0145*** (0,0011)	0,0017 (0,0011)	-0,0305*** (0,0024)	-0,0023*** (0,0024)	-0,0033*** (0,0004)	-0,0005*** (0,0004)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	616.754	616.754	616.754	616.754	616.754	616.754	616.754	616.754
Estat. F/Wald	1834,71	30842,77	1853,31	28855,17	1835,60	30718,77	1847,89	28802,77
R ² / Pseudo R ²	0,0440	-	0,0539	-	0,0446	-	0,0537	-
Informalidade	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	-0,0423*** (0,0100)	-0,0019** (0,0101)	-0,0125*** (0,0016)	0,0022 (0,0016)	-0,0141*** (0,0033)	-0,0006*** (0,0033)	-0,0245*** (0,0006)	-0,0011*** (0,0006)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	563.463	563.463	563.463	563.463	563.463	563.463	563.463	563.463
Estat. F/Wald	1409,83	24539,07	1413,90	24470,39	1409,83	24539,07	1413,53	24484,94
R ² / Pseudo R ²	0,0442	-	0,0455	-	0,0442	-	0,0454	-
Overeducated	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV	MPL_IV	Probit_IV
LnACi	0,0102 (0,0118)	0,0007 (0,0117)	-0,0067*** (0,0019)	-0,0012* (0,0019)	0,0034 (0,0039)	0,0002 (0,0039)	-0,0038*** (0,0007)	-0,0001* (0,0007)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	563.463	563.463	563.463	563.463	563.463	563.463	563.463	563.463
Estat. F/Wald	5657,89	82912,19	5658,84	82884,01	5657,89	82912,19	5659,08	82915,51
R ² / Pseudo R ²	0,1530	-	0,1531	-	0,1530	-	0,1530	-

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs.: Controles adicionais: Tipos de vizinhança, Idade, Idade², nível de escolaridade, *dummy* para masculino, *dummy* para chefe, raça, *dummy* para infraestrutura, termo constante. Função de distância $d_{ij}^{-0,5}$ e $d_{ij}^{-1,5}$

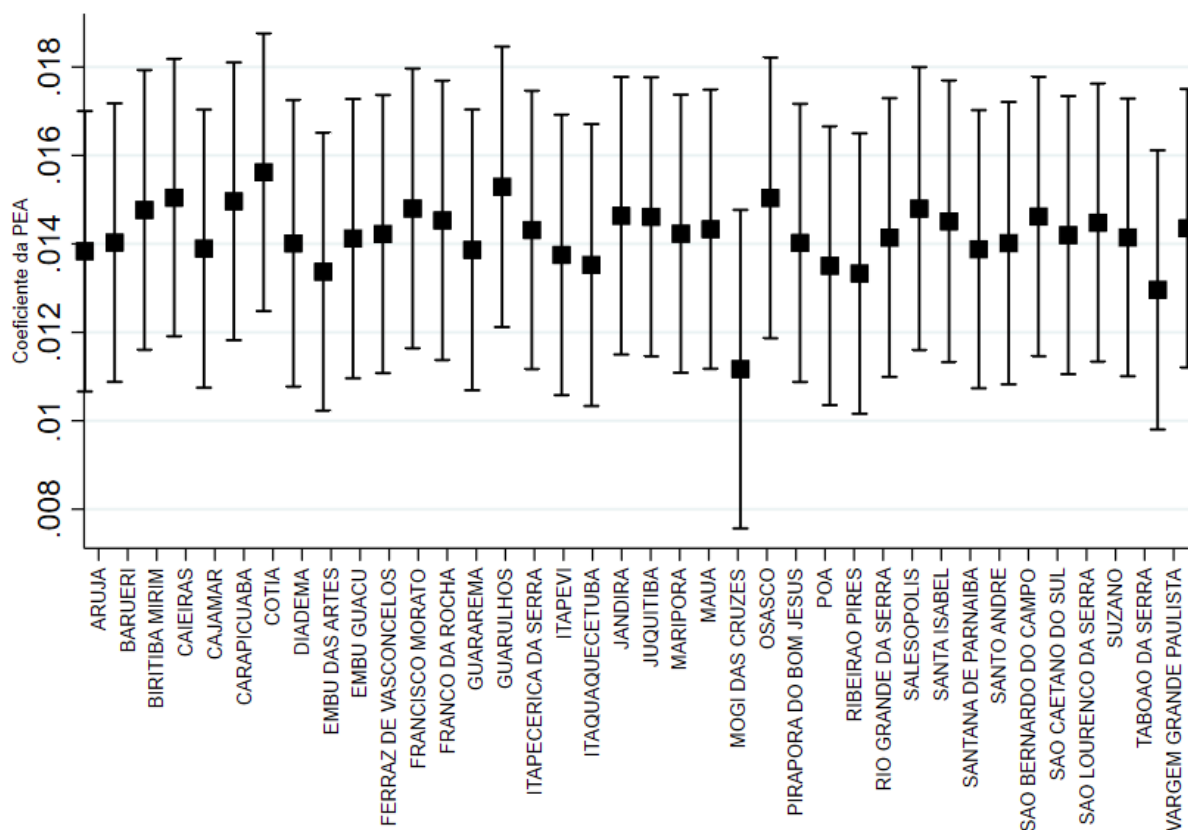
Fonte: Elaboração própria com base na média dos efeitos marginais.

Por fim, como terceira etapa para resultados de robustez, foram elaborados gráficos de validação cruzada²⁵ “*Leave One Out*” (LOOCV). O método foi realizado usando o conjunto de dados consistindo em 38 municípios. Este método de validação cruzada deixa de fora um conjunto de observações associadas a cada um dos municípios da amostra. Em seguida, o processo é repetido, apresentando outra observação única diferente para validação, e o treinamento é feito com os dados das outras observações. A vantagem em utilizar essa técnica é que ela investiga completamente à variação do modelo em relação aos dados utilizados, além disto, a estimativa do erro é não tendenciosa, ou seja, tende à taxa verdadeira. Ressalta-se que o exercício foi realizado com a exclusão do município de São Paulo, visto que o mesmo é o

²⁵ A validação cruzada é uma técnica para avaliar a capacidade de generalização de um modelo, a partir de um conjunto de dados. Esta técnica é amplamente empregada em problemas onde o objetivo da modelagem é a predição. Busca-se então estimar o quão preciso é este modelo na prática, ou seja, o seu desempenho para um novo conjunto de dados.

destino de muito dos ocupados de outros municípios. Os resultados usando o *leave-one-out*²⁶ para os subcentros são mostrados nos gráficos 2.5 a 2.8 a seguir, enquanto os resultados para o CBD encontram-se no apêndice.

Gráfico 2.5 – Validação Cruzada “*leave-one-out*” para PEA - Subcentros

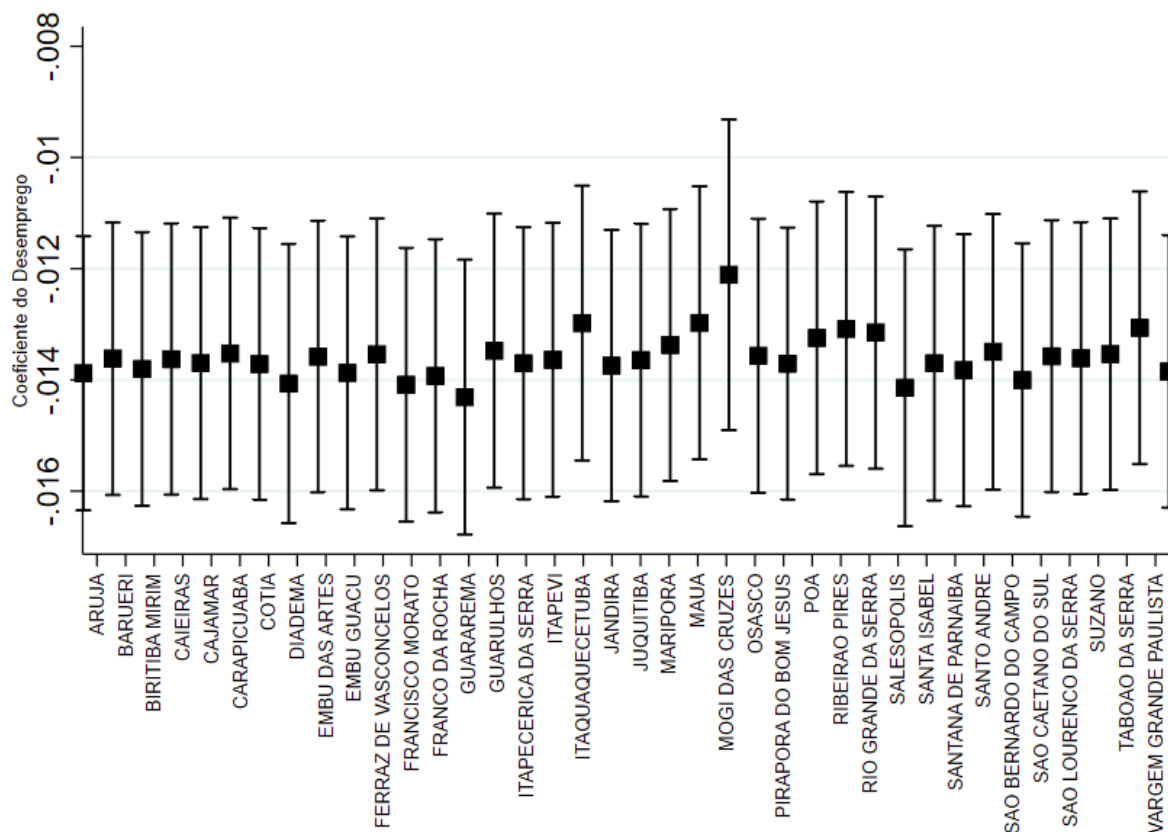


Fonte: Elaboração própria com base nas estimativas

Os gráficos apresentam a trajetória dos valores obtidos para o estimador *leave-one-out* considerando como variável dependente a participação no mercado, à condição de estar desempregado, no setor informal e de *overeducated*. Os pontos representam a média e as barras indicam o desvio padrão para cada valor de N . Vale destacar que os valores exibidos são previsões da base de teste utilizando para treinamento o conjunto de treino e validação.

²⁶ O método *leave-one-out* é um caso específico do *k-fold*, com k igual ao número total de dados N . Nesta abordagem são realizados N cálculos de erro, um para cada dado.

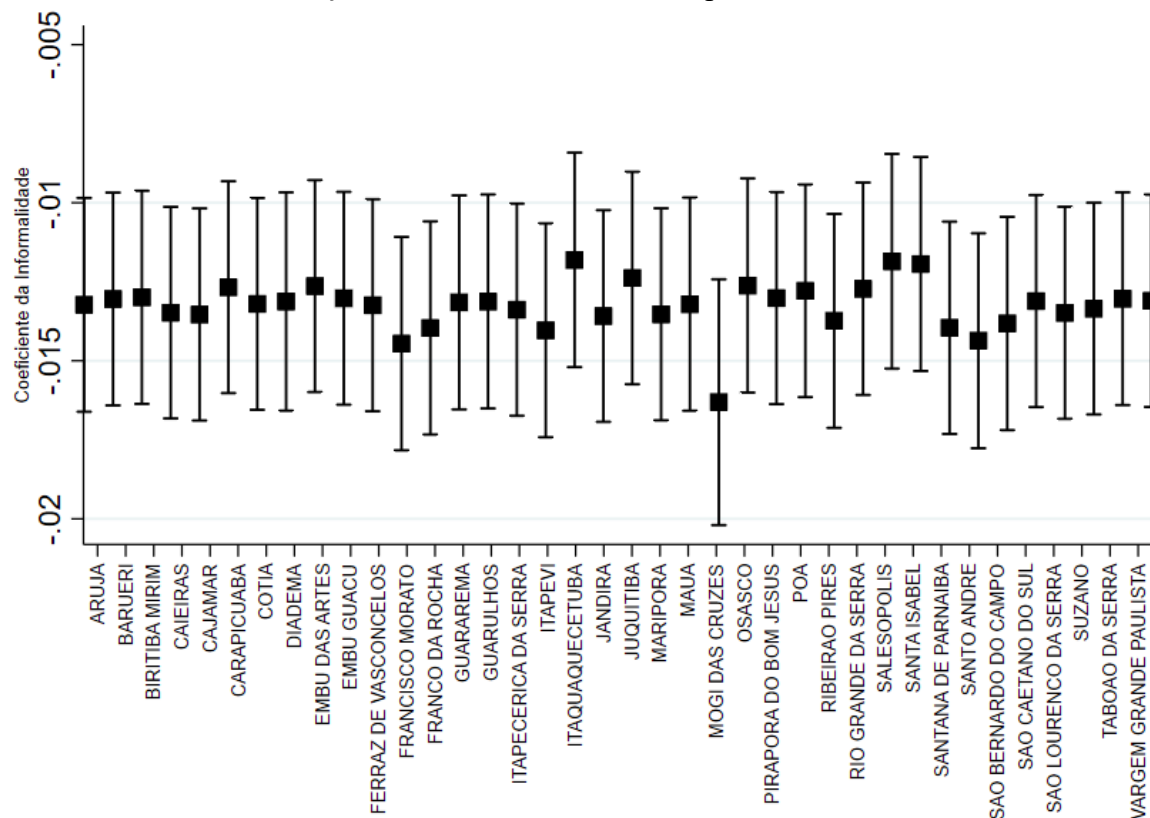
Gráfico 2.6 – Validação Cruzada “leave-one-out” para Desemprego - Subcentros



Fonte: Elaboração própria com base nas estimativas

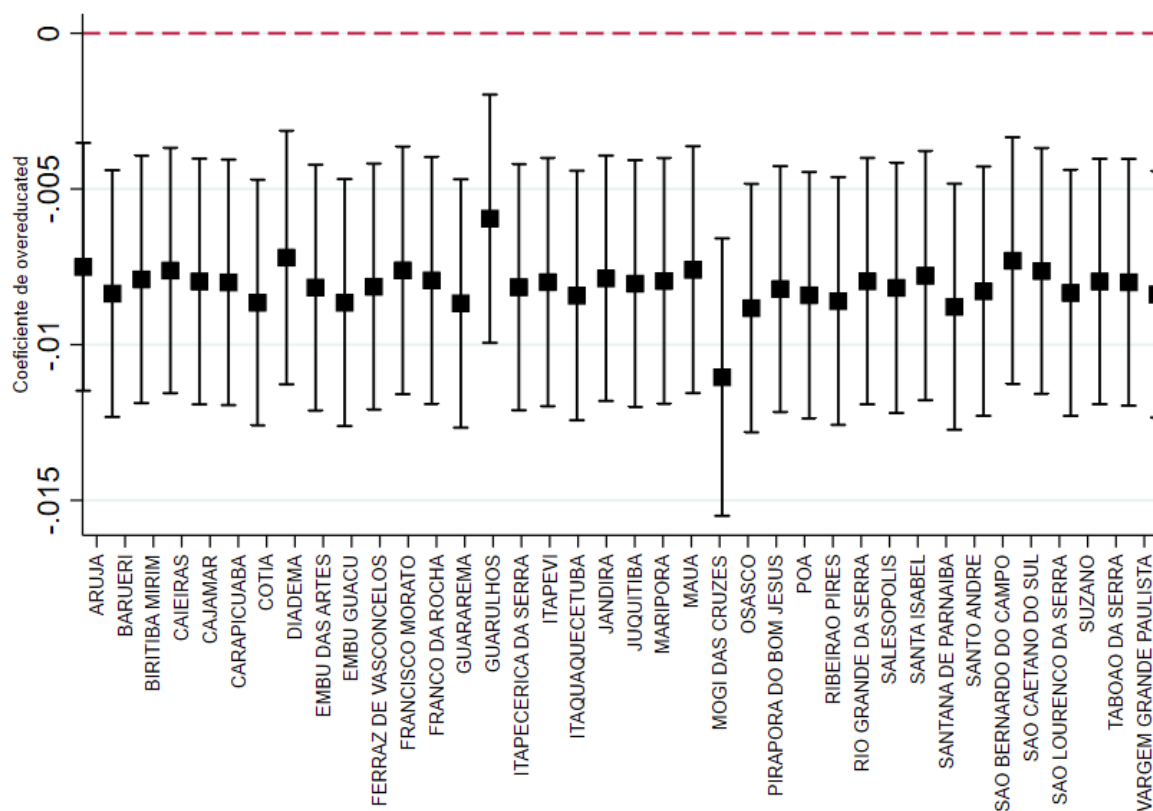
Note que os valores obtidos em todos os intervalos são diferentes de zero o que indica que os coeficientes são necessários. Percebe-se que apesar do modelo estimado está bem ajustado, apresentam algumas variações ao retirar um conjunto de observações associadas a cada município da amostra, ou seja, as estimações não apresentam grandes variabilidades. Como pode ser visto, o modelo pode ser considerado robusto, já que pequenas flutuações no valor são observadas com até 38 conjuntos de observações removidas.

Gráfico 2.7 – Validação Cruzada “leave-one-out” para Informalidade - Subcentros



Fonte: Elaboração própria com base nas estimativas

Gráfico 2.8 – Validação Cruzada “leave-one-out” para overeducated - Subcentros



Fonte: Elaboração própria com base nas estimativas

2.7 Considerações Finais

Este artigo estuda a acessibilidade ao emprego e sua relação com os resultados do mercado de trabalho da região metropolitana de São Paulo. Uma das principais contribuições foi fornecer estimativas consistentes da acessibilidade aos empregos usando um instrumento exógeno para analisar o impacto em diferentes variáveis dependentes, a saber, participação no mercado, condição de estar desempregado, na informalidade e de ser *overeducated*. Além disto, o estudo buscou analisar esta relação separada por gêneros.

A RMSP apresenta padrões importantes de *mismatch* espacial e segregação residencial que afetam a participação da força de trabalho. Com respeito às oportunidades de emprego medidas pelo índice de acessibilidade, estas são muito importantes para todos da amostra, porém mais especificamente para as mulheres. Os resultados mostraram que, em geral, o melhor acesso físico aos locais de emprego aumenta a participação no mercado, diminui o desemprego e a informalidade.

Especificamente para a participação no mercado, observou-se que o índice de acessibilidade foi significativo a 1% e com efeito esperado em todas as estimativas. Concentrando-se nos modelos IV percebeu-se que o aumento de 10% no índice de acessibilidade aumenta em aproximadamente 0,2% a probabilidade de se participar no mercado. A análise por gênero sugeriu que a acessibilidade às oportunidades de emprego representa barreiras mais importantes para a participação no mercado de trabalho para mulheres que para os homens. Considerando a variável dependente desemprego, teve-se que o melhor acesso às oportunidades de empregos diminui a probabilidade do indivíduo estar desempregado, sendo os efeitos do CBD semelhantes àqueles obtidos para os subcentros. Tomando o melhor resultado considerando o instrumento exógeno para subcentros, observou-se que um aumento de 1% no índice de acessibilidade reduziu a probabilidade de estar desempregado em 0,17%. Na análise por gênero, observou-se claramente que o efeito do coeficiente foi mais forte para as mulheres do que para os homens.

Para a informalidade, observou-se que os resultados novamente corroboraram com a literatura do *mismatch* espacial, sugerindo que o melhor acesso às oportunidades de empregos diminui a probabilidade de um indivíduo estar no setor informal. As evidências mostraram que o aumento de 10% no índice de acessibilidade está associado a uma redução de quase 1% na chance de se estar na informalidade. Na análise por gênero observou-se que a relação entre o melhor acesso físico ao CBD para homens não foi significativo, enquanto para as mulheres

se obteve uma relação inversa e significativa. Quando analisado os subcentros como oportunidades de empregos, a relação foi significativa para ambos os gêneros, sendo bem mais forte para o sexo feminino. Além disto, os efeitos marginais obtidos para as mulheres foram maiores que aqueles para amostra total, sugerindo que existe uma segmentação de gênero nos mercados de trabalho, com predominância da informalidade para as mulheres. As evidências encontradas por gênero corroboram com a Teoria da Responsabilidade Familiar.

O estudo também buscou analisar a relação da acessibilidade ao emprego com a chance do indivíduo ser *overeducated*, ou seja, buscando entender a alocação dos trabalhadores no mercado de trabalho da RMSP. Os resultados encontrados mostraram-se significativos e com sinais esperados apenas para os subcentros de empregos, onde se acredita ser melhor referência para uma área como a RMSP. O pior acesso aos subcentros aumenta a chance de ser *overeducated*. Fazendo a análise por gênero, percebeu-se que as evidências obtidas estavam em linha com resultados anteriores, ou seja, acesso é muito mais importante para as mulheres.

As descobertas da pesquisa são bastante informativas para os formuladores de políticas e os planejadores urbanos. Os resultados deste trabalho enfatizam a necessidade de analisar resultados do mercado de trabalho na Economia Urbana. Foram encontradas evidências importantes com disparidades espaciais em termos de acesso a oportunidades de emprego nas áreas de ponderação da RMSP, bem como efeitos de gênero heterogêneos na conectividade nos resultados trabalhistas.

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 e, tendo em vista maiores dificuldades das mulheres, provavelmente maior oferta de creches nas proximidades das residências.

A acessibilidade às oportunidades de emprego é um dos muitos desafios. Famílias mais pobres residentes 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. Entretanto, na literatura ainda não há consenso sobre a forma adequada de modelar essas equações estruturais, bem como as decisões de localização residencial e local de trabalho.

APÊNDICE

Tabela 2.1 – Estimações para a probabilidade de participação no mercado (regressão para homens e mulheres) – 1º Estágio

Variáveis	HOMENS		MULHERES	
	CBD	Subcentros	CBD	Subcentros
	MPL_IV	MPL_IV	MPL_IV	MPL_IV
LnACi_IV	0,9153*** (0,0006)	0,8287*** (0,0008)	0,9094*** (0,0005)	0,8179*** (0,0008)
Idade	0,0004*** (0,0001)	0,0002** (0,0001)	-0,0002 (0,0001)	0,0001 (0,0001)
Idade ²	-0,0000*** (0,0000)	-0,0000** (0,0000)	-0,0000 (0,0000)	-0,0000*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0003 (0,0007)	-0,0006 (0,0008)	-0,0004 (0,0007)	0,0008 (0,0008)
Alta	0,0010 (0,0006)	-0,0033*** (0,0008)	-0,0002 (0,0004)	-0,0032*** (0,0008)
Bem elevada	-0,0026*** (0,0009)	0,0037*** (0,0011)	-0,0047*** (0,0009)	0,0015 (0,0010)
Casado	-0,0093*** (0,0006)	-0,0057*** (0,0007)	-0,0051*** (0,0005)	-0,0039*** (0,0006)
Chefe	0,0039*** (0,0006)	0,0027*** (0,0007)	0,0041*** (0,0006)	0,0016** (0,0007)
Branco	Ref.	Ref.	Ref.	Ref.
Negro	-0,0010 (0,0010)	-0,0049*** (0,0012)	0,0015 (0,0010)	-0,0034*** (0,0012)
Amarelo	-0,0325*** (0,0006)	-0,0247*** (0,0024)	-0,0365*** (0,0019)	-0,0245*** (0,0023)
Pardo	0,0002 (0,0006)	-0,0014** (0,0007)	0,0009* (0,0005)	-0,0016** (0,0006)
Muito Carente	Ref.	Ref.	Ref.	Ref.
Carente	0,0172*** (0,0011)	0,0068*** (0,0013)	0,0190*** (0,0011)	-0,0026** (0,0013)
Média	0,0221*** (0,0013)	0,0322*** (0,0015)	0,0272*** (0,0012)	0,0267*** (0,0015)
Desenvolvida	0,1088*** (0,0015)	0,1742*** (0,0017)	0,1125*** (0,0014)	0,1691*** (0,0017)
Muito Desenvolvida	0,5187*** (0,0021)	0,4242*** (0,0024)	0,5366*** (0,0019)	0,4261*** (0,0023)
Tamanho da família	-0,0017*** (0,0001)	-0,0012*** (0,0001)	-0,0018*** (0,0001)	-0,0008*** (0,0001)
Saneamento	0,0009 (0,0008)	0,0175*** (0,0009)	0,0017** (0,0007)	0,0240*** (0,0009)
Coleta	0,0083*** (0,0011)	0,0036*** (0,0013)	0,0075*** (0,0011)	0,0037*** (0,0014)
<i>Constante</i>	0,8235*** (0,0067)	2,0831*** (0,0113)	0,8928*** (0,0065)	2,2232*** (0,0112)
Estat. F/Wald	6167,54	6167,54	4344,74	4344,74
R ² / Pseudo R ²	0,9248	0,8258	0,9251	0,8168
Observações	401.311	401.311	434.493	434.493

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base nas estimativas.

Tabela A2.2 – Efeitos marginais para a probabilidade de participação no mercado (regressões para homens)

Variáveis	CBD				Subcentros			
	PEA	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	0,0018 (0,0013)	0,0003 (0,0013)	0,0022 (0,0014)	-0,0013 (0,0015)	-0,0036** (0,0016)	-0,0054*** (0,0016)	-0,0044** (0,0020)	-0,0030 (0,0084)
Idade	0,0544*** (0,0002)	0,0429*** (0,0002)	0,0544*** (0,0002)	0,0429*** (0,0002)	0,0545*** (0,0002)	0,0429*** (0,0002)	0,0545*** (0,0002)	0,0429*** (0,0002)
Idade ²	-0,0007*** (0,0000)	-0,0005*** (0,0001)	-0,0007*** (0,0000)	-0,0005*** (0,0001)	-0,0007*** (0,0000)	-0,0005*** (0,0001)	-0,0007*** (0,0000)	-0,0005*** (0,0001)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0460*** (0,0016)	0,0412*** (0,0015)	0,0460*** (0,0016)	0,0412*** (0,0015)	0,0460*** (0,0016)	0,0411*** (0,0015)	0,0460*** (0,0016)	0,0411*** (0,0015)
Alta	0,1361*** (0,0015)	0,1265*** (0,0014)	0,1361*** (0,0015)	0,1265*** (0,0014)	0,1360*** (0,0015)	0,1263*** (0,0014)	0,1361*** (0,0015)	0,1263*** (0,0014)
Bem elevada	0,1585*** (0,0021)	0,1692*** (0,0023)	0,1585*** (0,0021)	0,1692*** (0,0023)	0,1585*** (0,0021)	0,1693*** (0,0023)	0,1585*** (0,0021)	0,1693*** (0,0023)
Casado	0,0297*** (0,0014)	0,0429*** (0,0014)	0,0297*** (0,0014)	0,0429*** (0,0014)	0,0295*** (0,0014)	0,0428*** (0,0014)	0,0295*** (0,0014)	0,0428*** (0,0014)
Chefe	0,0483*** (0,0013)	0,0569*** (0,0014)	0,0483*** (0,0013)	0,0569*** (0,0014)	0,0483*** (0,0013)	0,0568*** (0,0014)	0,0483*** (0,0013)	0,0568*** (0,0014)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0202*** (0,0022)	0,0201*** (0,0022)	0,0202*** (0,0022)	0,0201*** (0,0022)	0,0203*** (0,0022)	0,0202*** (0,0022)	0,0203*** (0,0022)	0,0202*** (0,0022)
Amarelo	-0,0357*** (0,0044)	-0,0368*** (0,0043)	-0,0357*** (0,0044)	-0,0368*** (0,0043)	-0,0357*** (0,0044)	-0,0368*** (0,0043)	-0,0357*** (0,0044)	-0,0368*** (0,0043)
Pardo	0,0192*** (0,0013)	0,0181*** (0,0012)	0,0192*** (0,0013)	0,0181*** (0,0012)	0,0193*** (0,0013)	0,0182*** (0,0012)	0,0193*** (0,0013)	0,0182*** (0,0012)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0169*** (0,0025)	-0,0081*** (0,0016)	-0,0168*** (0,0024)	-0,0081*** (0,0016)	-0,0185*** (0,0015)	-0,0061*** (0,0016)	-0,0187*** (0,0024)	-0,0061*** (0,0016)
Média	-0,0063*** (0,0028)	-0,0158*** (0,0023)	-0,0060*** (0,0028)	-0,0158*** (0,0023)	-0,0099*** (0,0028)	-0,0175*** (0,0023)	-0,0105*** (0,0029)	-0,0175*** (0,0023)
Desenvolvida	-0,0013*** (0,0033)	-0,0154*** (0,0023)	-0,0018*** (0,0034)	-0,0154*** (0,0023)	-0,0042*** (0,0035)	-0,0117*** (0,0022)	-0,0050*** (0,0035)	-0,0117*** (0,0022)
Muito Desenvolvida	-0,0291*** (0,0048)	-0,0506*** (0,0041)	-0,0299*** (0,0049)	-0,0506*** (0,0041)	-0,0208*** (0,0047)	-0,0446*** (0,0039)	-0,0197*** (0,0049)	-0,0446*** (0,0039)
Tamanho da família	-0,0001 (0,0003)	-0,0001 (0,0002)	-0,0000 (0,0003)	-0,0001 (0,0002)	-0,0001 (0,0003)	-0,0001 (0,0012)	-0,0000 (0,0003)	-0,0001 (0,0012)
Saneamento	-0,0038** (0,0017)	-0,0028* (0,0017)	-0,0038** (0,0017)	-0,0028* (0,0017)	-0,0031* (0,0017)	-0,0021 (0,0017)	-0,0030* (0,0017)	-0,0021 (0,0017)
Coleta	0,0297*** (0,0027)	0,0263*** (0,0024)	0,0297*** (0,0025)	0,0263*** (0,0024)	0,0292*** (0,0025)	0,0257*** (0,0024)	0,0291*** (0,0025)	0,0257*** (0,0024)
Constante	-0,2906*** (0,0148)	-	-0,0297*** (0,0159)	-	-0,2276*** (0,0211)	-	-0,2172*** (0,0252)	-
Estat. F/Wald	3993,23	59633,15	3993,26	56038,89	3993,41	59643,54	3993,42	53043,93
R ² / Pseudo R ²	0,1519	0,1494	0,1519	-	0,1519	0,1494	0,1519	-
Observações	401.311	401.311	401.311	401.311	401.311	401.311	401.311	401.311

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base nas médias dos efeitos marginais

Tabela A2.3 – Efeitos marginais para a probabilidade de participação no mercado (regressões para mulheres).

Variáveis	CBD				Subcentros			
	PEA	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	0,0235*** (0,0015)	0,0236*** (0,0016)	0,0259*** (0,0017)	0,0133*** (0,0052)	0,0256*** (0,0019)	0,0248*** (0,0019)	0,0325*** (0,0024)	0,0115*** (0,0024)
Idade	0,0483*** (0,0003)	0,0448*** (0,0003)	0,0483*** (0,0003)	0,0448*** (0,0003)	0,0483*** (0,0003)	0,0449*** (0,0003)	0,0483*** (0,0003)	0,0449*** (0,0003)
Idade ²	-0,0006*** (0,0000)	-0,0006*** (0,0001)	-0,0006*** (0,0000)	-0,0006*** (0,0001)	-0,0006*** (0,0000)	-0,0006*** (0,0001)	-0,0006*** (0,0000)	-0,0006*** (0,0001)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0641*** (0,0019)	0,0580*** (0,0019)	0,0641*** (0,0019)	0,0580*** (0,0019)	0,0641*** (0,0019)	0,0579*** (0,0019)	0,0641*** (0,0019)	0,0579*** (0,0019)
Alta	0,1882*** (0,0018)	0,1732*** (0,0017)	0,1883*** (0,0018)	0,1732*** (0,0018)	0,1881*** (0,0018)	0,1731*** (0,0017)	0,1882*** (0,0018)	0,1731*** (0,0017)
Bem elevada	0,2993*** (0,0024)	0,3120*** (0,0024)	0,2994*** (0,0024)	0,3120*** (0,0024)	0,2987*** (0,0024)	0,3114*** (0,0024)	0,2987*** (0,0024)	0,3114*** (0,0024)
Casado	-0,0894*** (0,0015)	-0,0896*** (0,0015)	-0,0894*** (0,0015)	-0,0896*** (0,0015)	-0,0897*** (0,0015)	-0,0899*** (0,0015)	-0,0896*** (0,0015)	-0,0899*** (0,0015)
Chefe	0,0402*** (0,0016)	0,0408*** (0,0016)	0,0408*** (0,0016)	0,0408*** (0,0016)	0,0408*** (0,0016)	0,0410*** (0,0016)	0,0404*** (0,0016)	0,0410*** (0,0016)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0632*** (0,0028)	0,0636*** (0,0028)	0,0631*** (0,0028)	0,0636*** (0,0028)	0,0635*** (0,0028)	0,0638*** (0,0028)	0,0632*** (0,0028)	0,0638*** (0,0028)
Amarelo	-0,0153*** (0,0052)	-0,0167*** (0,0053)	-0,0154*** (0,0052)	-0,0167*** (0,0053)	-0,0147*** (0,0052)	-0,0160*** (0,0053)	-0,0147*** (0,0052)	-0,0160*** (0,0053)
Pardo	0,0253*** (0,0015)	0,0252*** (0,0015)	0,0258*** (0,0015)	0,0252*** (0,0015)	0,0259*** (0,0015)	0,0252*** (0,0015)	0,0258*** (0,0015)	0,0252*** (0,0015)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0103*** (0,0029)	-0,0058*** (0,0019)	-0,0097*** (0,0029)	-0,0058*** (0,0019)	-0,0086*** (0,0029)	-0,0036*** (0,0018)	-0,0065*** (0,0030)	-0,0036*** (0,0018)
Média	-0,0039*** (0,0034)	-0,0094*** (0,0028)	-0,0022*** (0,0034)	-0,0094*** (0,0028)	-0,0042*** (0,0034)	-0,0080*** (0,0029)	-0,0001*** (0,0035)	-0,0080*** (0,0029)
Desenvolvida	-0,0119*** (0,0039)	-0,0225*** (0,0026)	-0,0146*** (0,0040)	-0,0225*** (0,0026)	-0,0089*** (0,0039)	-0,0174*** (0,0025)	-0,0154*** (0,0041)	-0,0174*** (0,0025)
Muito Desenvolvida	-0,0443*** (0,0056)	-0,0582*** (0,0048)	-0,0487*** (0,0058)	-0,0582*** (0,0048)	-0,0335*** (0,0054)	-0,0449*** (0,0045)	-0,0426*** (0,0057)	-0,0449*** (0,0045)
Tamanho da família	-0,0084 (0,0003)	-0,0085*** (0,0003)	-0,0084*** (0,0003)	-0,0085*** (0,0003)	-0,0084*** (0,0003)	-0,0085*** (0,0003)	-0,0084*** (0,0003)	-0,0085*** (0,0003)
Saneamento	0,0023 (0,0021)	0,0029 (0,0020)	0,0020 (0,0021)	0,0029 (0,0020)	0,0017 (0,0021)	0,0024 (0,0020)	0,0008 (0,0021)	0,0024 (0,0020)
Coleta	0,0014 (0,0031)	0,0025*** (0,0030)	0,0016 (0,0031)	0,0025*** (0,0030)	0,0014 (0,0021)	0,0023 (0,0030)	0,0019 (0,0031)	0,0023 (0,0030)
Constante	-0,4676*** (0,0176)	-	-0,4911*** (0,0189)	-	-0,5416*** (0,0249)	-	-0,6252*** (0,0302)	-
Estat. F/Wald	3821,22	62253,10	3821,46	56232,27	3818,11	62198,34	3818,60	56213,74
R ² / Pseudo R ²	0,1367	0,1091	0,1367	-	0,1366	0,1090	0,1366	-
Observações	434.493	434.493	434.493	434.493	434.493	434.493	434.493	434.493

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base na média dos efeitos marginais.

Tabela A2.4 – Estimações para a probabilidade de desempregados (regressão para homens e mulheres) – 1º Estágio

Variáveis	HOMENS		MULHERES	
	CBD	Subcentros	CBD	Subcentros
	MPL_IV	MPL_IV	MPL_IV	MPL_IV
LnACi_IV	0,9175*** (0,0006)	0,8271*** (0,0009)	0,9127*** (0,0007)	0,8119*** (0,0011)
Idade	-0,0005 (0,0001)	0,0000 (0,0001)	-0,0005*** (0,0001)	-0,0001 (0,0001)
Idade ²	0,0000 (0,0000)	-0,0001 (0,0000)	0,0000*** (0,0000)	0,0000 (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0009 (0,0008)	-0,0003 (0,0010)	-0,0004 (0,0010)	0,0005 (0,0011)
Alta	0,0016** (0,0007)	-0,0031*** (0,0008)	0,0002 (0,0009)	-0,0046*** (0,0010)
Bem elevada	-0,0009 (0,0010)	0,0032*** (0,0012)	-0,0057*** (0,0011)	-0,0019 (0,0012)
Casado	-0,0090*** (0,0006)	-0,0061*** (0,0008)	-0,0062*** (0,0007)	-0,0047*** (0,0008)
Chefe	0,0043*** (0,0006)	0,0031*** (0,0008)	0,0045*** (0,0007)	0,0021*** (0,0009)
Branco	Ref.	Ref.	Ref.	Ref.
Negro	-0,0017 (0,0011)	-0,0059*** (0,0013)	0,0011 (0,0013)	-0,0034** (0,0015)
Amarelo	-0,0323*** (0,0022)	-0,0269*** (0,0026)	-0,0366*** (0,0025)	-0,0219*** (0,0029)
Pardo	0,0007 (0,0006)	-0,0012* (0,0007)	0,0009 (0,0007)	-0,0013 (0,0008)
Muito Carente	Ref.	Ref.	Ref.	Ref.
Carente	0,0158*** (0,0012)	0,0042*** (0,0015)	0,0169*** (0,0014)	-0,0033** (0,0017)
Média	-0,0202*** (0,0014)	0,0292*** (0,0017)	0,0253*** (0,0016)	0,0279*** (0,0020)
Desenvolvida	0,1076*** (0,0017)	0,1734*** (0,0019)	0,1153*** (0,0019)	0,1766*** (0,0022)
Muito Desenvolvida	0,5204*** (0,0023)	0,4252*** (0,0026)	0,5361*** (0,0025)	0,4316*** (0,0028)
Tamanho da família	-0,0017*** (0,0001)	-0,0010*** (0,0001)	-0,0020*** (0,0010)	-0,0008*** (0,0002)
Saneamento	0,0012 (0,0009)	0,0202*** (0,0010)	0,0018* (0,0010)	0,0266*** (0,0012)
Coleta	0,0105*** (0,0013)	0,0095*** (0,0015)	0,0099*** (0,0015)	0,0062*** (0,0017)
<i>Constante</i>	0,8065*** (0,0075)	2,1011*** (0,0125)	0,8655*** (0,0083)	2,3010*** (0,0141)
Estat. F/Wald	6167,54	6167,53	2833,88	2833,88
R ² / Pseudo R ²	0,9245	0,8262	0,9231	0,8168
Observações	333.347	333.347	283.407	283.407

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base nas estimativas.

Tabela A2.5 – Efeitos marginais para a probabilidade de desempregados (regressões para homens).

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
Desemprego	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	-0,0056*** (0,0009)	-0,0054*** (0,0010)	-0,0056*** (0,0010)	-0,0056*** (0,0010)	-0,0046*** (0,0012)	-0,0043*** (0,0012)	-0,0058*** (0,0014)	-0,0021*** (0,0014)
Idade	-0,0119*** (0,0001)	-0,0080*** (0,0001)	-0,0119*** (0,0001)	-0,0080*** (0,0001)	-0,0119*** (0,0001)	-0,0080*** (0,0001)	-0,0119*** (0,0001)	-0,0080*** (0,0001)
Idade ²	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0047*** (0,0012)	0,0011 (0,0011)	0,0047*** (0,0012)	0,0011 (0,0011)	0,0047*** (0,0012)	-0,0011 (0,0011)	0,0047*** (0,0012)	-0,0011 (0,0011)
Alta	-0,0196*** (0,0011)	-0,0167*** (0,0010)	-0,0196*** (0,0011)	-0,0167*** (0,0010)	-0,0195*** (0,0011)	-0,0166*** (0,0010)	-0,0195*** (0,0011)	-0,0166*** (0,0010)
Bem elevada	-0,0283*** (0,0015)	-0,0391*** (0,0017)	-0,0283*** (0,0015)	-0,0391*** (0,0017)	-0,0281*** (0,0012)	-0,0389*** (0,0017)	-0,0281*** (0,0011)	-0,0389*** (0,0017)
Casado	-0,0215*** (0,0009)	-0,272*** (0,0010)	-0,0215*** (0,0009)	-0,272*** (0,0010)	-0,0213*** (0,0009)	-0,0271*** (0,0010)	-0,0214*** (0,0009)	-0,0271*** (0,0010)
Chefe	-0,0187*** (0,0009)	-0,0222*** (0,0010)	-0,0187*** (0,0009)	-0,0222*** (0,0010)	-0,0187*** (0,0009)	-0,0222*** (0,0010)	-0,0187*** (0,0009)	-0,0222*** (0,0010)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0049*** (0,0016)	0,0052*** (0,0015)	0,0049*** (0,0016)	0,0052*** (0,0015)	0,0048*** (0,0016)	0,0052*** (0,0015)	0,0049*** (0,0016)	0,0052*** (0,0015)
Amarelo	-0,0002 (0,0032)	-0,0004 (0,0036)	-0,0002 (0,0032)	-0,0004 (0,0036)	-0,0004 (0,0032)	-0,0005 (0,0036)	-0,0004 (0,0032)	-0,0005 (0,0036)
Pardo	0,0014 (0,0009)	0,0018** (0,0009)	0,0014 (0,0009)	0,0018** (0,0009)	0,0013 (0,0009)	0,0018** (0,0009)	0,0014 (0,0009)	0,0018** (0,0009)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0117*** (0,0018)	-0,0032*** (0,0011)	-0,0117*** (0,0018)	-0,0032*** (0,0011)	-0,0116*** (0,0018)	-0,0041*** (0,0148)	-0,0120*** (0,0015)	-0,0041*** (0,0148)
Média	-0,0086*** (0,0021)	-0,0103*** (0,0173)	-0,0087*** (0,0021)	-0,0103*** (0,0173)	-0,0076*** (0,0021)	-0,0102*** (0,0017)	-0,0083*** (0,0022)	-0,0102*** (0,0017)
Desenvolvida	-0,0069*** (0,0024)	-0,0080*** (0,0208)	-0,0069*** (0,0025)	-0,0080*** (0,0208)	-0,0046** (0,0024)	-0,0102*** (0,0016)	-0,0058** (0,0025)	-0,0102*** (0,0016)
Muito Desenvolvida	-0,0098* (0,0035)	-0,0062* (0,0034)	-0,0099*** (0,0036)	-0,0062* (0,0034)	-0,0052* (0,0034)	-0,0108*** (0,0032)	-0,0068** (0,0035)	-0,0108*** (0,0032)
Tamanho da família	0,0049*** (0,0002)	0,0030*** (0,0002)	0,0049*** (0,0002)	0,0030*** (0,0002)	0,0049*** (0,0002)	0,0030*** (0,0002)	0,0049*** (0,0002)	0,0030*** (0,0002)
Saneamento	-0,0009 (0,0013)	-0,0009 (0,0012)	-0,0009*** (0,0013)	-0,0009 (0,0012)	-0,0009 (0,0013)	-0,0010 (0,0012)	-0,0008*** (0,0013)	-0,0010 (0,0012)
Coleta	-0,0060*** (0,0019)	-0,0053*** (0,0018)	-0,0061*** (0,0019)	-0,0053*** (0,0018)	-0,0059*** (0,0019)	-0,0051*** (0,0018)	-0,0060*** (0,0019)	-0,0051*** (0,0018)
Constante	0,3869*** (0,0107)	-	0,3871*** (0,0115)	-	0,3860*** (0,0154)	-	0,4010*** (0,0184)	-0,0080*** (0,0001)
Estat. F/Wald	926,50	14311,00	926,22	13297,58	925,37	14294,93	925,43	13289,09
R ² / Pseudo R ²	0,0476	0,0873	0,0476	-	0,0476	0,0869	0,0476	-
Observações	333.347	333.347	333.347	333.347	333.347	333.347	333.347	333.347

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base na média dos efeitos marginais.

Tabela A2.6 – Efeitos marginais para a probabilidade de desempregados (regressão para mulheres).

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
Desemprego	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	-0,0159*** (0,0013)	-0,0178*** (0,0013)	-0,0179*** (0,0014)	-0,0073*** (0,0086)	-0,0170*** (0,0016)	-0,0169*** (0,0016)	-0,0236*** (0,0020)	-0,0042*** (0,0116)
Idade	-0,0141*** (0,0002)	-0,0100*** (0,0002)	-0,0141*** (0,0002)	-0,0100*** (0,0002)	-0,0141*** (0,0002)	-0,0100*** (0,0002)	-0,0141*** (0,0002)	-0,0100*** (0,0002)
Idade ²	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0131*** (0,0017)	0,0056*** (0,0016)	0,0131*** (0,0017)	0,0056*** (0,0016)	0,0132*** (0,0017)	0,0057*** (0,0016)	0,0131*** (0,0017)	0,0057*** (0,0016)
Alta	-0,0257*** (0,0015)	-0,0226*** (0,0015)	-0,0258*** (0,0015)	-0,0226*** (0,0015)	-0,0256*** (0,0015)	-0,0223*** (0,0015)	-0,0257*** (0,0015)	-0,0223*** (0,0015)
Bem elevada	-0,0597*** (0,0019)	-0,0812*** (0,0022)	-0,0598*** (0,0019)	-0,0812*** (0,0022)	-0,0592*** (0,0019)	-0,0807*** (0,0022)	-0,0593*** (0,0019)	-0,0807*** (0,0022)
Casado	-0,0054*** (0,0012)	-0,0063*** (0,0013)	-0,0055*** (0,0012)	-0,0063*** (0,0013)	-0,0052*** (0,0012)	-0,0061*** (0,0013)	-0,0053*** (0,0012)	-0,0061*** (0,0013)
Chefe	0,0020 (0,0013)	0,0003 (0,0014)	0,0021 (0,0013)	0,0003 (0,0014)	0,0019 (0,0013)	0,0002 (0,0014)	0,0020 (0,0013)	0,0002 (0,0014)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0098*** (0,0023)	0,0548*** (0,0129)	0,0099*** (0,0023)	0,0548*** (0,0129)	0,0097*** (0,0023)	0,0093*** (0,0022)	0,0099*** (0,0023)	0,0093*** (0,0022)
Amarelo	-0,0059 (0,0043)	-0,0138*** (0,0052)	-0,0059 (0,0043)	-0,0138*** (0,0052)	-0,0064 (0,0043)	-0,0142*** (0,0052)	-0,0063 (0,0043)	-0,0142*** (0,0052)
Pardo	0,0143*** (0,0013)	0,0130*** (0,0012)	0,0144*** (0,0013)	0,0130*** (0,0012)	0,0143*** (0,0013)	0,0129*** (0,0012)	0,0144*** (0,0013)	0,0129*** (0,0012)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0084*** (0,0025)	-0,0148*** (0,0016)	-0,0089*** (0,0025)	-0,0148*** (0,0016)	-0,0094*** (0,0025)	-0,0171*** (0,0015)	-0,0114*** (0,0026)	-0,0171*** (0,0015)
Média	-0,0075*** (0,0029)	-0,0082*** (0,0023)	-0,0061*** (0,0029)	-0,0082*** (0,0023)	-0,0081*** (0,0029)	-0,0086*** (0,0023)	-0,0039*** (0,0030)	-0,0086*** (0,0023)
Desenvolvida	-0,0062* (0,0033)	-0,0195*** (0,0023)	-0,0039* (0,0034)	-0,0195*** (0,0023)	-0,0088*** (0,0033)	-0,0246*** (0,0022)	-0,0025*** (0,0035)	-0,0246*** (0,0022)
Muito Desenvolvida	-0,0088** (0,0046)	-0,0079*** (0,0045)	-0,0125*** (0,0047)	-0,0079*** (0,0045)	-0,0007 (0,0044)	-0,0202*** (0,0042)	-0,0094** (0,0047)	-0,0202*** (0,0042)
Tamanho da família	0,0063*** (0,0003)	-0,0050*** (0,0003)	0,0063*** (0,0003)	-0,0050*** (0,0003)	0,0063*** (0,0003)	0,0049*** (0,0003)	0,0063*** (0,0003)	0,0049*** (0,0003)
Saneamento	-0,0123*** (0,0018)	-0,0089*** (0,0016)	-0,0121*** (0,0018)	-0,0089*** (0,0016)	-0,0120*** (0,0018)	-0,0088*** (0,0016)	-0,0111*** (0,0018)	-0,0088*** (0,0016)
Coleta	-0,0125*** (0,0026)	-0,0108*** (0,0024)	-0,0127*** (0,0026)	-0,0108*** (0,0024)	-0,0124*** (0,0026)	-0,0105*** (0,0024)	-0,0129*** (0,0026)	-0,0105*** (0,0024)
Constante	0,6103*** (0,0144)	-	0,6301*** (0,0026)	-	0,6554*** (0,0209)	-	0,7360*** (0,0255)	-
Estat. F/Wald	915,98	15258,06	916,42	13852,05	913,44	15198,89	914,92	13837,23
R ² / Pseudo R ²	0,0550	0,0782	0,0550	-	0,0548	0,0779	0,0548	-
Observações	283.407	283.407	283.407	283.407	283.407	283.407	283.407	283.407

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base na média dos efeitos marginais.

Tabela 2.7 – Estimações para a probabilidade de informalidade (regressão para homens e mulheres) – 1º Estágio

Variáveis	HOMENS		MULHERES	
	CBD	Subcentros	CBD	Subcentros
	MPL_IV	MPL_IV	MPL_IV	MPL_IV
LnACi_IV	0,9187*** (0,0007)	0,8265*** (0,0010)	0,9134*** (0,0090)	0,8098*** (0,0011)
Idade	-0,0001 (0,0001)	-0,0001 (0,0001)	-0,0005*** (0,0001)	-0,0001 (0,0001)
Idade ²	0,0000 (0,0000)	0,0001*** (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0011 (0,0009)	-0,0003 (0,0010)	-0,0004 (0,0011)	0,0008 (0,0012)
Alta	0,0016** (0,0008)	-0,0029*** (0,0009)	0,0003 (0,0009)	-0,0046*** (0,0011)
Bem elevada	-0,0014 (0,0010)	0,0030** (0,0012)	-0,0058*** (0,0011)	-0,0023* (0,0013)
Casado	-0,0101*** (0,0007)	-0,0064*** (0,0008)	-0,0070*** (0,0007)	-0,0053*** (0,0008)
Chefe	0,0041*** (0,0007)	0,0028*** (0,0008)	0,0043*** (0,0008)	0,0019** (0,0009)
Branco	Ref.	Ref.	Ref.	Ref.
Negro	-0,0015 (0,0012)	-0,0062*** (0,0014)	0,0011 (0,0014)	-0,0040*** (0,0016)
Amarelo	-0,0328*** (0,0023)	-0,0262*** (0,0027)	-0,0369*** (0,0026)	-0,0219*** (0,0030)
Pardo	0,0011* (0,0007)	-0,0008 (0,0008)	0,0011 (0,0008)	-0,0014 (0,0009)
Muito Carente	Ref.	Ref.	Ref.	Ref.
Carente	0,0158*** (0,0013)	0,0043*** (0,0015)	0,0173*** (0,0016)	0,0026** (0,0009)
Média	0,0202*** (0,0015)	0,0301*** (0,0018)	0,0260*** (0,0018)	0,0300*** (0,0021)
Desenvolvida	0,1076*** (0,0017)	0,1751*** (0,0020)	0,1152*** (0,0021)	0,1789*** (0,0024)
Muito Desenvolvida	0,5195*** (0,0017)	0,4259*** (0,0027)	0,5359*** (0,0026)	0,4341*** (0,0030)
Tamanho da família	-0,0018*** (0,0001)	-0,0013*** (0,0002)	-0,0019*** (0,0002)	-0,0007*** (0,0002)
Saneamento	0,0009 (0,0009)	0,0199*** (0,0011)	0,0017 (0,0011)	0,0270*** (0,0013)
Coleta	0,0103*** (0,0014)	0,0091*** (0,0016)	0,0093*** (0,0016)	0,0048*** (0,0150)
<i>Constante</i>	0,7955*** (0,0078)	2,1111*** (0,0130)	0,8598*** (0,0090)	2,3292*** (0,0150)
Estat. F/Wald	616,754	616,753	252,532	252,532
R ² / Pseudo R ²	0,9241	0,8262	0,9217	0,8168
Observações	310.912	310.912	252.551	252.551

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base nas estimativas.

Tabela A2.8 – Efeitos marginais para a probabilidade de trabalhadores informais (regressões para homens).

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
Informalidade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	-0,0014 (0,0013)	-0,0035 (0,0067)	-0,0016 (0,0014)	-0,0046 (0,0073)	-0,0044*** (0,0016)	-0,0188** (0,0084)	-0,0083*** (0,0020)	-0,0378*** (0,0101)
Idade	-0,0127*** (0,0002)	-0,0512*** (0,0011)	-0,0127*** (0,0002)	-0,0512*** (0,0011)	-0,0127*** (0,0002)	-0,0512*** (0,0011)	-0,0127*** (0,0002)	-0,0511*** (0,0011)
Idade ²	0,0001*** (0,0000)	0,0005*** (0,0000)	0,0001*** (0,0000)	0,0005*** (0,0000)	0,0001*** (0,0000)	0,0005*** (0,0000)	0,0001*** (0,0000)	0,0005*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	-0,0287*** (0,0017)	-0,1396*** (0,0081)	-0,0287*** (0,0017)	-0,1296*** (0,0257)	-0,0287*** (0,0017)	-0,1297*** (0,0081)	-0,0287*** (0,0017)	-0,1297*** (0,0257)
Alta	-0,0782*** (0,0015)	-0,3647*** (0,0076)	-0,0783*** (0,0015)	-0,3647*** (0,0076)	-0,0783*** (0,0015)	-0,3649*** (0,0076)	-0,0784*** (0,0015)	-0,3654*** (0,0076)
Bem elevada	-0,0962*** (0,0020)	-0,5158*** (0,0113)	-0,0962*** (0,0020)	-0,5158*** (0,0113)	-0,0961*** (0,0020)	-0,5156*** (0,0113)	-0,0961*** (0,0020)	-0,5155*** (0,0113)
Casado	-0,0215*** (0,0009)	-0,1704*** (0,0069)	-0,0308*** (0,0013)	-0,1705*** (0,0069)	-0,0308*** (0,0013)	-0,1706*** (0,0069)	-0,0309*** (0,0013)	-0,1711*** (0,0069)
Chefe	-0,0308*** (0,0013)	-0,0742*** (0,0068)	-0,0128*** (0,0013)	-0,0743*** (0,0068)	-0,0129*** (0,0013)	-0,0744*** (0,0068)	-0,0129*** (0,0013)	-0,0746*** (0,0068)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0081*** (0,0023)	0,0383*** (0,0112)	0,0081*** (0,0023)	0,0383*** (0,0112)	0,0082*** (0,0023)	0,0385*** (0,0112)	0,0082*** (0,0023)	0,0387*** (0,0112)
Amarelo	-0,0003 (0,0044)	-0,0076 (0,0237)	-0,0003 (0,0044)	-0,0076 (0,0237)	-0,0004 (0,0044)	-0,0078 (0,0237)	-0,0005 (0,0044)	-0,0082 (0,0237)
Pardo	0,0023* (0,0013)	0,0111* (0,0066)	0,0023* (0,0013)	0,0112* (0,0066)	0,0023* (0,0013)	0,0113* (0,0066)	0,0023* (0,0013)	0,0115* (0,0066)
Muito	-0,0109*** (0,0025)	-0,0506*** (0,0120)	-0,0109*** (0,0025)	-0,0503*** (0,0120)	-0,0099*** (0,0025)	-0,0457*** (0,0121)	-0,0087*** (0,0025)	-0,0398*** (0,0122)
Carente	-0,0012 (0,0029)	-0,0079 (0,0141)	-0,0010 (0,0029)	-0,0071 (0,0142)	0,0005 (0,0029)	0,0016 (0,0143)	0,0030 (0,0030)	0,0139 (0,0147)
Média	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Desenvolvida	0,0050 (0,0034)	0,0198 (0,0167)	0,0052 (0,0034)	0,0211 (0,0170)	0,0074** (0,0034)	0,0336** (0,0167)	0,0212*** (0,0036)	0,0520*** (0,0176)
Muito	0,0015 (0,0048)	0,0035 (0,0251)	0,0019 (0,0049)	0,0055 (0,0257)	0,0045 (0,0047)	0,0216 (0,0242)	0,0098** (0,0049)	0,0467* (0,0253)
Tamanho da família	0,0031*** (0,0003)	0,0120*** (0,0016)	0,0031*** (0,0003)	0,0120*** (0,0016)	0,0031*** (0,0003)	0,0120*** (0,0016)	0,0031*** (0,0003)	0,0119*** (0,0016)
Saneamento	-0,0144*** (0,0018)	-0,0635*** (0,0087)	-0,0143*** (0,0014)	-0,0634*** (0,0087)	-0,0139*** (0,0018)	-0,0614*** (0,0087)	-0,0134*** (0,0018)	-0,0588*** (0,0087)
Coleta	-0,0116*** (0,0027)	-0,0468*** (0,0127)	-0,0116*** (0,0027)	-0,0469*** (0,0127)	-0,0118*** (0,0027)	-0,0480*** (0,0127)	-0,0121*** (0,0027)	-0,0494*** (0,0127)
Constante	0,4828*** (0,0149)	0,2965*** (0,0739)	0,4844*** (0,0160)	0,3075*** (0,0793)	0,5216*** (0,0214)	0,4904*** (0,1063)	0,5697*** (0,0257)	0,7227*** (0,1271)
Estad. F/Wald	596,16	10012,90	596,17	9646,69	596,49	10017,63	597,02	9660,16
R ² / Pseudo R ²	0,0334	0,0422	0,0334	-	0,0334	0,0422	0,0334	-
Observações	310.912	310.912	310.912	310.912	310.912	310.912	310.912	310.912

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base na média dos efeitos marginais.

Tabela A2.9 – Efeitos marginais para a probabilidade de trabalhadores informais (regressões para mulheres).

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
Informalidade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	-0,0102*** (0,0017)	-0,0387*** (0,0069)	-0,0104*** (0,0018)	-0,0397*** (0,0074)	-0,0148*** (0,0002)	-0,0571*** (0,0086)	-0,0224*** (0,0027)	-0,0858*** (0,0105)
Idade	-0,0055*** (0,0003)	-0,0195*** (0,0011)	-0,0055*** (0,0003)	-0,0195*** (0,0011)	-0,0055*** (0,0003)	-0,0195*** (0,0011)	-0,0055*** (0,0003)	-0,0195*** (0,0011)
Idade ²	0,0001*** (0,0000)	0,0002*** (0,0000)	0,0006*** (0,0000)	0,0002*** (0,0000)	0,0006*** (0,0000)	0,0002*** (0,0000)	0,0001*** (0,0000)	0,0002*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	-0,0715*** (0,0024)	-0,2156*** (0,0086)	-0,0715*** (0,0024)	-0,2156*** (0,0086)	-0,0715*** (0,0024)	-0,2154*** (0,0086)	-0,0715*** (0,0024)	-0,2155*** (0,0086)
Alta	-0,1588*** (0,0021)	-0,5335*** (0,0078)	-0,1588*** (0,0021)	-0,5336*** (0,0078)	-0,1588*** (0,0015)	-0,5636*** (0,0078)	-0,1590*** (0,0021)	-0,5342*** (0,0078)
Bem elevada	-0,2151*** (0,0025)	-0,8674*** (0,0106)	-0,2151*** (0,0025)	-0,8674*** (0,0106)	-0,2148*** (0,0025)	-0,8664*** (0,0106)	-0,2149*** (0,0025)	-0,8667*** (0,0106)
Casado	-0,0181*** (0,0016)	-0,0761*** (0,0065)	-0,0181*** (0,0016)	-0,0762*** (0,0065)	-0,0181*** (0,0016)	-0,0760*** (0,0065)	-0,0182*** (0,0016)	-0,0765*** (0,0065)
Chefe	0,0264*** (0,0018)	0,0974*** (0,0069)	0,0264*** (0,0018)	0,0975*** (0,0069)	0,0264*** (0,0018)	0,0973*** (0,0069)	0,0264*** (0,0018)	0,0976*** (0,0069)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0234*** (0,0031)	0,0837*** (0,0115)	0,0234*** (0,0031)	0,0837*** (0,0115)	0,0234*** (0,0031)	0,0838*** (0,0115)	0,0235*** (0,0031)	0,0843*** (0,0115)
Amarelo	-0,0056 (0,0056)	-0,0267 (0,0239)	-0,0056 (0,0056)	-0,0267 (0,0239)	-0,0059 (0,0056)	-0,0278 (0,0239)	-0,0058 (0,0056)	-0,0278 (0,0239)
Pardo	0,0198*** (0,0017)	0,0716*** (0,0066)	0,0198*** (0,0017)	0,0717*** (0,0066)	0,0198*** (0,0017)	0,0718*** (0,0066)	0,0199*** (0,0017)	0,0721*** (0,0066)
Muito	-0,0022 (0,0017)	-0,0050 (0,0127)	-0,0022 (0,0034)	-0,0048 (0,0127)	-0,0004 (0,0035)	0,0021 (0,0129)	0,0019 (0,0035)	0,0111 (0,0130)
Carente	-0,0108*** (0,0039)	-0,0403*** (0,0147)	-0,0106*** (0,0039)	-0,0396*** (0,0149)	-0,0086*** (0,0040)	-0,0317*** (0,0149)	-0,0037 (0,0041)	-0,0132 (0,0154)
Média	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Desenvolvida	-0,0086*** (0,0045)	-0,0387*** (0,0172)	-0,0084*** (0,0046)	-0,0376*** (0,0176)	-0,0066 (0,0045)	-0,0303*** (0,0173)	0,0007 (0,0048)	-0,0024*** (0,0183)
Muito	0,0019 (0,0061)	0,0092 (0,0247)	0,0022 (0,0006)	0,0110 (0,0253)	0,0018 (0,0059)	0,0100 (0,0238)	0,0120*** (0,0063)	0,0478*** (0,0251)
Tamanho da família	0,0048*** (0,0004)	0,0170*** (0,0016)	0,0048*** (0,0004)	0,0170*** (0,0016)	0,0048*** (0,0004)	0,0170*** (0,0016)	0,0048*** (0,0004)	0,0170*** (0,0016)
Saneamento	-0,0256*** (0,0024)	-0,0845*** (0,0090)	-0,0256*** (0,0024)	-0,0844*** (0,0090)	-0,0248*** (0,0024)	-0,0816*** (0,0090)	-0,0238*** (0,0024)	-0,0779*** (0,0090)
Coleta	-0,0142*** (0,0035)	-0,0458*** (0,0132)	-0,0142*** (0,0035)	-0,0459*** (0,0132)	-0,0144*** (0,0035)	-0,0469*** (0,0132)	-0,0150*** (0,0035)	-0,0490*** (0,0132)
Constante	0,5410*** (0,0193)	0,3351*** (0,0762)	0,5428*** (0,0209)	0,3450*** (0,0818)	0,6175*** (0,0281)	0,6384*** (0,1093)	0,7115*** (0,0344)	0,9897*** (0,1326)
Estat. F/Wald	823,34	14364,35	823,08	13467,59	823,92	14376,81	825,13	13506,34
R ² / Pseudo R ²	0,0554	0,0581	0,0554	-	0,0555	0,0581	0,0554	-
Observações	252.551	252.551	252.551	252.551	252.551	252.551	252.551	252.551

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base na média dos efeitos marginais.

Tabela 2.10 – Estimações para a probabilidade de se estar na condição de overeducated (regressão para homens e mulheres) – 1º Estágio

Variáveis	HOMENS		MULHERES	
	CBD	Subcentros	CBD	Subcentros
	MPL_IV	MPL_IV	MPL_IV	MPL_IV
LnACi_IV	0,9187*** (0,0007)	0,8161*** (0,0010)	0,8769*** (0,0091)	0,8050*** (0,0011)
Idade	-0,0001 (0,0001)	-0,0001 (0,0001)	-0,0005*** (0,0001)	-0,0002 (0,0001)
Idade ²	0,0000 (0,0000)	0,0000 (0,0000)	0,0000*** (0,0000)	0,0000 (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.
Intermediária	0,0011* (0,0007)	-0,0033*** (0,0008)	0,0005 (0,0008)	-0,0051*** (0,0009)
Alta	-0,0019** (0,0010)	0,0027** (0,0012)	-0,0056*** (0,0010)	-0,0028** (0,0012)
Casado	-0,0101*** (0,0007)	-0,0068*** (0,0008)	-0,0070*** (0,0007)	-0,0052*** (0,0009)
Chefe	0,0041*** (0,0007)	0,0027*** (0,0008)	0,0043*** (0,0008)	0,0018** (0,0009)
Branco	Ref.	Ref.	Ref.	Ref.
Negro	-0,0015 (0,0012)	-0,0061*** (0,0014)	0,0011 (0,0014)	-0,0040** (0,0016)
Amarelo	-0,0328*** (0,0023)	-0,0262*** (0,0027)	-0,0369*** (0,0026)	-0,0219*** (0,0030)
Pardo	0,0011* (0,0007)	-0,0007 (0,0008)	0,0011 (0,0008)	-0,0014 (0,0009)
Muito Carente	Ref.	Ref.	Ref.	Ref.
Carente	0,0044*** (0,0008)	0,0278*** (0,0010)	0,0086*** (0,0009)	0,0328*** (0,0011)
Média	-0,0158*** (0,0013)	0,0034** (0,0016)	-0,0173*** (0,0016)	-0,0003 (0,0019)
Desenvolvida	0,0918*** (0,0011)	0,1749*** (0,0013)	0,0979*** (0,0013)	0,1831*** (0,0014)
Muito Desenvolvida	0,5036*** (0,0020)	0,4272*** (0,0022)	0,5186*** (0,0021)	0,4390*** (0,0023)
Tamanho da família	-0,0018*** (0,0001)	-0,0012*** (0,0002)	-0,0019*** (0,0002)	-0,0007*** (0,0002)
Saneamento	0,0009 (0,0009)	0,0257*** (0,0011)	0,0017 (0,0011)	0,0277*** (0,0013)
Coleta	0,0103*** (0,0016)	0,0059*** (0,0016)	0,0093*** (0,0016)	0,0046*** (0,0019)
Constante	0,8122*** (0,0079)	2,2449*** (0,0138)	0,8769*** (0,0091)	2,3884*** (0,0155)
Estat. F/Wald	99999,00	80451,31	99999,00	64937,94
R ² / Pseudo R ²	0,9241	0,8148	0,9217	0,8138
Observações	310.912	310.912	252.551	252.552

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base nas estimativas.

Tabela A2.11 – Efeitos marginais para a probabilidade de se estar na condição de overeducated (regressões para homens).

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
Overeducated	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	0,0049*** (0,0017)	0,0058*** (0,0016)	0,0023 (0,0018)	0,0187*** (0,0018)	-0,0044** (0,0021)	-0,0027 (0,0021)	-0,0067** (0,0026)	-0,0158* (0,0089)
Idade	-0,0039*** (0,0003)	-0,0037*** (0,0003)	-0,0039*** (0,0003)	-0,0037*** (0,0003)	-0,0038*** (0,0003)	-0,0037*** (0,0003)	-0,0038*** (0,0003)	-0,0037*** (0,0003)
Idade ²	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)	0,0001*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,3656*** (0,0017)	0,3317*** (0,0014)	0,3656*** (0,0017)	0,3317*** (0,0014)	0,3654*** (0,0017)	0,3315*** (0,0014)	0,3654*** (0,0017)	0,3315*** (0,0014)
Alta	0,2709*** (0,0024)	0,2643*** (0,0022)	0,2709*** (0,0024)	0,2643*** (0,0022)	0,2708*** (0,0024)	0,2642*** (0,0022)	0,2708*** (0,0024)	0,2642*** (0,0022)
Casado	-0,0054*** (0,0017)	-0,0038** (0,0017)	-0,0054*** (0,0017)	-0,0038** (0,0017)	-0,0057*** (0,0017)	-0,0041*** (0,0017)	-0,0057*** (0,0017)	-0,0041*** (0,0017)
Chefe	-0,0027 (0,0017)	-0,0015 (0,0017)	-0,0027 (0,0017)	-0,0015 (0,0017)	-0,0028* (0,0017)	-0,0016 (0,0017)	-0,0028* (0,0017)	-0,0016 (0,0017)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0054* (0,0029)	0,0023 (0,0029)	0,0054* (0,0029)	0,0023 (0,0029)	0,0055* (0,0029)	0,0024 (0,0029)	0,0055* (0,0029)	0,0024 (0,0029)
Amarelo	-0,0106* (0,0056)	-0,0066 (0,0054)	-0,0106* (0,0056)	-0,0066 (0,0054)	-0,0106* (0,0056)	-0,0065 (0,0054)	-0,0106* (0,0056)	-0,0065 (0,0054)
Pardo	0,0128*** (0,0017)	0,0100*** (0,0017)	0,0128*** (0,0017)	0,0100*** (0,0017)	0,0129*** (0,0017)	0,0100*** (0,0017)	0,0129*** (0,0017)	0,0100*** (0,0017)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0166*** (0,0021)	-0,0138*** (0,0020)	-0,0166*** (0,0021)	-0,0138*** (0,0020)	-0,0129*** (0,0020)	-0,0102*** (0,0020)	-0,0129*** (0,0020)	-0,0102*** (0,0020)
Média	-0,0022 (0,0032)	-0,0029 (0,0033)	-0,0022 (0,0032)	-0,0029 (0,0033)	-0,0049 (0,0032)	-0,0053 (0,0033)	-0,0049 (0,0032)	-0,0053 (0,0033)
Desenvolvida	-0,0689*** (0,0029)	-0,0589*** (0,0028)	-0,0689*** (0,0029)	-0,0589*** (0,0028)	-0,0615*** (0,0028)	-0,0517*** (0,0027)	-0,0615*** (0,0028)	-0,0517*** (0,0027)
Muito Desenvolvida	-0,1591*** (0,0051)	-0,1396*** (0,0050)	-0,1591*** (0,0051)	-0,1396*** (0,0050)	-0,1465*** (0,0048)	-0,1272*** (0,0047)	-0,1465*** (0,0048)	-0,1272*** (0,0047)
Tamanho da família	0,0010** (0,0004)	0,0007* (0,0004)	0,0010** (0,0004)	0,0007* (0,0004)	0,0009** (0,0044)	0,0007* (0,0004)	0,0009** (0,0044)	0,0007* (0,0004)
Saneamento	0,0125*** (0,0023)	0,0148*** (0,0023)	0,0125*** (0,0023)	0,0148*** (0,0023)	0,0137*** (0,0023)	0,0160*** (0,0024)	0,0137*** (0,0023)	0,0160*** (0,0024)
Coleta	0,0081** (0,0034)	0,0103*** (0,0035)	0,0081** (0,0034)	0,0103*** (0,0035)	0,0074** (0,0034)	0,0097*** (0,0035)	0,0074** (0,0034)	0,0097*** (0,0035)
Constante	0,1627*** (0,0192)	-	0,1627*** (0,0192)	-	0,2701*** (0,0278)	-	0,2701*** (0,0278)	-
Observações	310.912	310.912	310.912	310.912	310.912	310.912	310.912	310.912
Estat. F/Wald	3145,34	47148,84	53466,69	43812,38	3145,05	47138,17	53470,93	43814,64
R ² / Pseudo R ²	0,1468	0,1267	0,1467	-	0,1467	0,1267	0,1467	-

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base na média dos efeitos marginais.

Tabela A2.12 – Efeitos marginais para a probabilidade de se estar na condição de overeducated (regressões para mulheres).

Variáveis	CBD				Subcentros			
	MPL	Probit	MPL_IV	Probit_IV	MPL	Probit	MPL_IV	Probit_IV
Overeducated	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnACi	0,0013 (0,0019)	0,0018 (0,0019)	-0,0012 (0,0021)	-0,0025 (0,0068)	-0,0053** (0,0024)	-0,0045** (0,0024)	-0,0105*** (0,0030)	-0,0048 (0,0030)
Idade	-0,0074*** (0,0003)	-0,0066*** (0,0001)	-0,0074*** (0,0003)	-0,0066*** (0,0001)	-0,0074*** (0,0003)	-0,0066*** (0,0003)	-0,0074*** (0,0003)	-0,0066*** (0,0003)
Idade ²	0,0000*** (0,0000)	0,0000*** (0,0000)	0,0000*** (0,0000)	0,0000*** (0,0000)	0,0000*** (0,0000)	0,0001*** (0,0000)	0,0000*** (0,0000)	0,0001*** (0,0000)
Baixa	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Intermediária	0,3895*** (0,0020)	0,3616*** (0,0016)	0,3895*** (0,0020)	0,3616*** (0,0016)	0,3884*** (0,0020)	0,3614*** (0,0016)	0,3884*** (0,0020)	0,3614*** (0,0016)
Alta	0,2488*** (0,0025)	0,2566*** (0,0024)	0,2488*** (0,0025)	0,2566*** (0,0024)	0,2487*** (0,0025)	0,2565*** (0,0024)	0,2487*** (0,0025)	0,2565*** (0,0024)
Casado	-0,0082*** (0,0018)	-0,0061*** (0,0018)	-0,0082*** (0,0018)	-0,0061*** (0,0018)	-0,0083*** (0,0018)	-0,0063*** (0,0018)	-0,0083*** (0,0018)	-0,0063*** (0,0018)
Chefe	-0,0018 (0,0020)	-0,0007 (0,0020)	-0,0018 (0,0020)	-0,0007 (0,0020)	-0,0017 (0,0020)	-0,0006 (0,0020)	-0,0017 (0,0020)	-0,0006 (0,0020)
Branco	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Negro	0,0086** (0,0034)	0,0046 (0,0035)	0,0086** (0,0034)	0,0046 (0,0035)	0,0087*** (0,0034)	0,0047 (0,0035)	0,0087*** (0,0034)	0,0047 (0,0035)
Amarelo	-0,0014 (0,0062)	0,0014 (0,0061)	-0,0014 (0,0062)	0,0014 (0,0061)	-0,0013 (0,0062)	0,0014 (0,0061)	-0,0013 (0,0062)	0,0014 (0,0061)
Pardo	0,0236*** (0,0019)	0,0199*** (0,0019)	0,0236*** (0,0019)	0,0199*** (0,0019)	0,0236*** (0,0019)	0,0199*** (0,0019)	0,0236*** (0,0019)	0,0199*** (0,0019)
Muito Carente	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Carente	-0,0329*** (0,0023)	-0,0299*** (0,0023)	-0,0329*** (0,0023)	-0,0299*** (0,0023)	-0,0305*** (0,0023)	-0,0276*** (0,0023)	-0,0305*** (0,0023)	-0,0276*** (0,0023)
Média	-0,0099*** (0,0038)	-0,0097** (0,0039)	-0,0099*** (0,0038)	-0,0097** (0,0039)	-0,0120*** (0,0039)	-0,0116*** (0,0040)	-0,0120*** (0,0039)	-0,0116*** (0,0040)
Desenvolvida	-0,0777*** (0,0031)	-0,0700*** (0,0031)	-0,0777*** (0,0031)	-0,0700*** (0,0031)	-0,0730*** (0,0031)	-0,0654*** (0,0030)	-0,0730*** (0,0031)	-0,0654*** (0,0030)
Muito Desenvolvida	-0,1374*** (0,0054)	-0,1258*** (0,0054)	-0,1374*** (0,0054)	-0,1258*** (0,0054)	-0,1297*** (0,0051)	-0,1183*** (0,0051)	-0,1297*** (0,0051)	-0,1183*** (0,0051)
Tamanho da família	0,0028*** (0,0005)	0,0028*** (0,0005)	0,0028*** (0,0005)	0,0028*** (0,0005)	0,0028*** (0,0005)	0,0028*** (0,0005)	0,0028*** (0,0005)	0,0028*** (0,0005)
Saneamento	0,0043 (0,0027)	0,0059** (0,0027)	0,0043 (0,0027)	0,0059** (0,0027)	0,0051* (0,0027)	0,0067*** (0,0028)	0,0051* (0,0027)	0,0067*** (0,0028)
Coleta	0,0032 (0,0039)	0,0055 (0,0041)	0,0032 (0,0039)	0,0055 (0,0041)	0,0027 (0,0039)	0,0051 (0,0041)	0,0027 (0,0039)	0,0051 (0,0041)
Constante	0,2966*** (0,0218)	-	0,2966*** (0,0218)	-	0,3782*** (0,0319)	-	0,3782*** (0,0319)	-
Observações	252.551	252.551	252.551	252.551	252.551	252.551	252.551	252.551
Estat. F/Wald	2842,27	42971,19	48321,58	38928,59	2842,57	42973,73	48333,37	38939,33
R ² / Pseudo R ²	0,1606	0,1364	0,1606	-	0,1606	0,1364	0,1606	-

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%.

Fonte: Elaboração própria com base na média dos efeitos marginais.

Tabela A2.13 – Resultado do 1º estágio para a probabilidade de participação, desempregados, informalidade e condição de *overeducated* (regressões com jovens adultos que moram com os pais).

	PEA			Desemprego	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,8922*** (0,0012)	0,8120*** (0,0021)	LnACi_IV	0,8933*** (0,0014)	0,8127*** (0,0023)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	79.037	79.037	Observações	61.480	61.480
R ² / Pseudo R ²	0,9275	0,7998	R ² / Pseudo R ²	0,9281	0,7957
Estat. F	56164,86	17540,26	Estat. F	44057,84	13296,76
Durbin (score, <i>Chi</i>)	0,37	1,29	Durbin (score, <i>Chi</i>)	1,21	1,30
Wu-Hausman (F)	0,37	1,29	Wu-Hausman (F)	1,21	1,30
	Informalidade			Overeducated	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,8939*** (0,0015)	0,8121*** (0,0026)	LnACi_IV	0,8939*** (0,0015)	0,8120*** (0,0026)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	52.610	52.610	Observações	52.610	52.610
R ² / Pseudo R ²	0,9270	0,7944	R ² / Pseudo R ²	0,9270	0,7944
Estat. F	37091,99	11286,50	Estat. F	39274,12	11949,85
Durbin (score, <i>Chi</i>)	2,14	0,82	Durbin (score, <i>Chi</i>)	10,70***	20,60***
Wu-Hausman (F)	2,14	0,82	Wu-Hausman (F)	10,70***	20,60***

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs.: Controles adicionais: Idade, Idade², nível de escolaridade, *dummy* para masculino, raça, tipos de vizinhança, *dummy* para infraestrutura, termo constante.

Tabela A2.14 – Resultado do 1º estágio para a probabilidade de participação, desempregados, informalidade e condição de *overeducated* (regressões com 50% mais pobres da amostra).

	PEA			Desemprego	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,9112*** (0,0004)	0,8271*** (0,0006)	LnACi_IV	0,9140*** (0,0005)	0,8260*** (0,0008)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	621.342	621.342	Observações	61.480	61.480
R ² / Pseudo R ²	0,9279	0,8066	R ² / Pseudo R ²	0,9275	0,8038
Estat. F	420891,75	136371,28	Estat. F	290689,34	93068,25
Durbin (score, <i>Chi</i>)	7,93***	0,68	Durbin (score, <i>Chi</i>)	4,13**	13,04***
Wu-Hausman (F)	7,93***	0,68	Wu-Hausman (F)	4,13**	13,04***
	Informalidade			Overeducated	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,9149*** (0,0005)	0,8259*** (0,0008)	LnACi_IV	0,9149*** (0,0005)	0,8259*** (0,0008)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	382.892	382.892	Observações	382.892	382.892
R ² / Pseudo R ²	0,9270	0,8050	R ² / Pseudo R ²	0,9270	0,8050
Estat. F	256072,46	83165,86	Estat. F	270299,40	87786,36
Durbin (score, <i>Chi</i>)	0,04	24,67***	Durbin (score, <i>Chi</i>)	10,56***	2,79*
Wu-Hausman (F)	0,04	24,67***	Wu-Hausman (F)	10,56***	2,78*

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs.: Controles adicionais: Idade, Idade², nível de escolaridade, *dummy* para masculino, raça, tipos de vizinhança, *dummy* para infraestrutura, termo constante.

Tabela A2.15 – Resultado do 1º estágio para a probabilidade de participação, desempregados, informalidade e condição de *overeducated* (regressões com indivíduos que moram em casa própria).

	PEA			Desemprego	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,8955*** (0,0004)	0,8146*** (0,0007)	LnACi_IV	0,8982*** (0,0005)	0,8114*** (0,0008)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	602.773	602.773	Observações	438.308	438.308
R ² / Pseudo R ²	0,9277	0,8123	R ² / Pseudo R ²	0,9171	0,8135
Estat. F	407010,39	137257,64	Estat. F	293213,19	100649,43
Durbin (score, <i>Chi</i>)	3,05*	3,41**	Durbin (score, <i>Chi</i>)	3,73**	21,01***
Wu-Hausman (F)	3,05*	3,41**	Wu-Hausman (F)	3,73**	21,01***

	Informalidade			Overeducated	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,8985*** (0,0006)	0,8102*** (0,0009)	LnACi_IV	0,8985*** (0,0006)	0,8102*** (0,0009)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	401.342	401.342	Observações	401.342	401.342
R ² / Pseudo R ²	0,9265	0,8141	R ² / Pseudo R ²	0,9265	0,8141
Estat. F	266158,56	92504,45	Estat. F	280935,92	97643,22
Durbin (score, <i>Chi</i>)	0,15	16,94***	Durbin (score, <i>Chi</i>)	11,11***	5,08**
Wu-Hausman (F)	0,15	16,94***	Wu-Hausman (F)	11,11***	5,07**

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs.: Controles adicionais: Idade, Idade², nível de escolaridade, *dummy* para masculino, raça, tipos de vizinhança, *dummy* para infraestrutura, termo constante.

Tabela A2.16 – Resultado do 1º estágio para a probabilidade de participação, desempregados, informalidade e condição de *overeducated* (regressões considerando -0,5 no grau de decaimento).

	PEA			Desemprego	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,2225*** (0,0016)	1,8148*** (0,0013)	LnACi_IV	0,2167*** (0,0018)	1,8084*** (0,0015)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	835.804	835.804	Observações	616.754	616.754
R ² / Pseudo R ²	0,0400	0,8363	R ² / Pseudo R ²	0,0383	0,8363
Estat. F	1831,57	224651,82	Estat. F	1293,33	165795,18
Durbin (score, <i>Chi</i>)	131,6***	38,17***	Durbin (score, <i>Chi</i>)	128,14***	59,74***
Wu-Hausman (F)	131,7***	38,17***	Wu-Hausman (F)	128,16***	59,75***

	Informalidade			Overeducated	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,2158*** (0,0019)	1,8067*** (0,0016)	LnACi_IV	0,2158*** (0,0019)	1,8067*** (0,0016)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	563.463	563.463	Observações	563.463	563.463
R ² / Pseudo R ²	0,0381	0,8365	R ² / Pseudo R ²	0,0381	0,8365
Estat. F	1174,40	151678,09	Estat. F	1239,57	160104,76
Durbin (score, <i>Chi</i>)	15,85***	24,27***	Durbin (score, <i>Chi</i>)	0,57	4,11**
Wu-Hausman (F)	15,85***	24,27***	Wu-Hausman (F)	0,57	4,11**

Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs.: Controles adicionais: Idade, Idade², nível de escolaridade, *dummy* para masculino, raça, tipos de vizinhança, *dummy* para infraestrutura, termo constante.

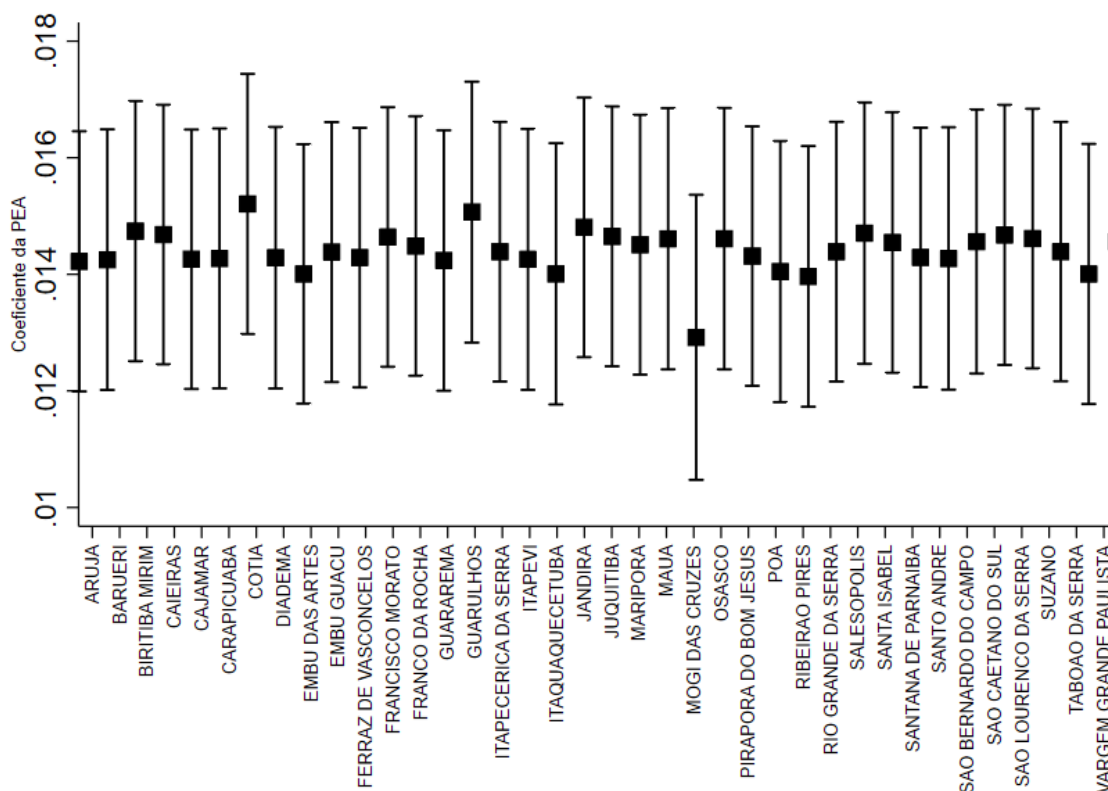
Tabela A2.17 – Resultado do 1º estágio para a probabilidade de participação, desempregados, informalidade e condição de *overeducated* (regressões considerando -1,5 no grau de decaimento).

	PEA		LnACi_IV	Desemprego	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,2225*** (0,0016)	1,3000*** (0,0025)	LnACi_IV	0,2167*** (0,0018)	1,2933*** (0,0029)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	835.804	835.804	Observações	616.754	616.754
R ² / Pseudo R ²	0,0400	0,4664	R ² / Pseudo R ²	0,0383	0,4673
Estat. F	1831,57	38441,36	Estat. F	1293,33	28477,18
Durbin (score, <i>Chi</i>)	131,6***	12,81***	Durbin (score, <i>Chi</i>)	128,14***	30,34***
Wu-Hausman (F)	131,7***	12,81***	Wu-Hausman (F)	128,16***	30,34***

	Informalidade		LnACi_IV	Overeducated	
	CBD	Subcentros		CBD	Subcentros
	MPL_IV	MPL_IV		MPL_IV	MPL_IV
LnACi_IV	0,2158*** (0,0019)	1,2907*** (0,0031)	LnACi_IV	0,2158*** (0,0019)	1,2903*** (0,0031)
Controles	Sim	Sim	Controles	Sim	Sim
Observações	563.463	563.463	Observações	563.463	563.463
R ² / Pseudo R ²	0,0381	0,4679	R ² / Pseudo R ²	0,0381	0,4679
Estat. F	1174,40	26080,63	Estat. F	1239,57	27529,54
Durbin (score, <i>Chi</i>)	15,85***	22,45***	Durbin (score, <i>Chi</i>)	0,57	19,26***
Wu-Hausman (F)	15,85***	22,45***	Wu-Hausman (F)	0,57	19,26***

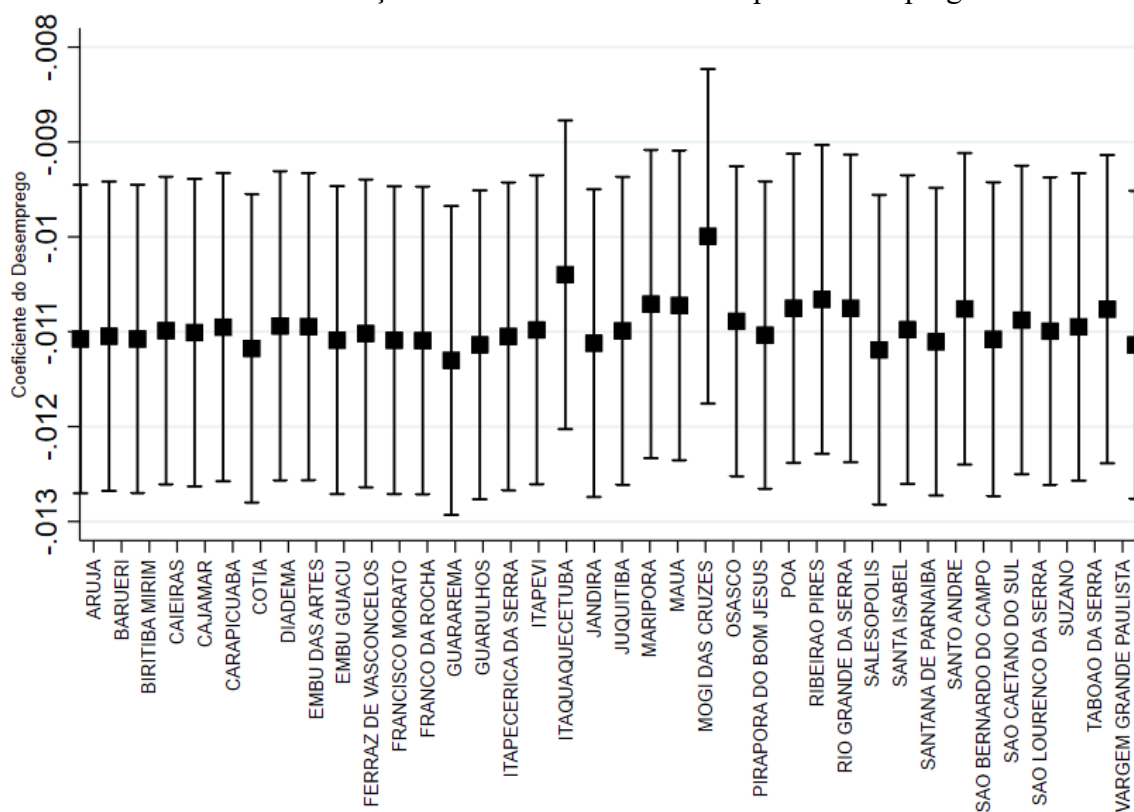
Notas: *** Significativo ao nível de 1%; ** 5% e * 10%. Obs.: Controles adicionais: Idade, Idade², nível de escolaridade, *dummy* para masculino, raça, tipos de vizinhança, *dummy* para infraestrutura, termo constante.

Gráfico A2.1 – Validação Cruzada “leave-one-out” para PEA - CBD



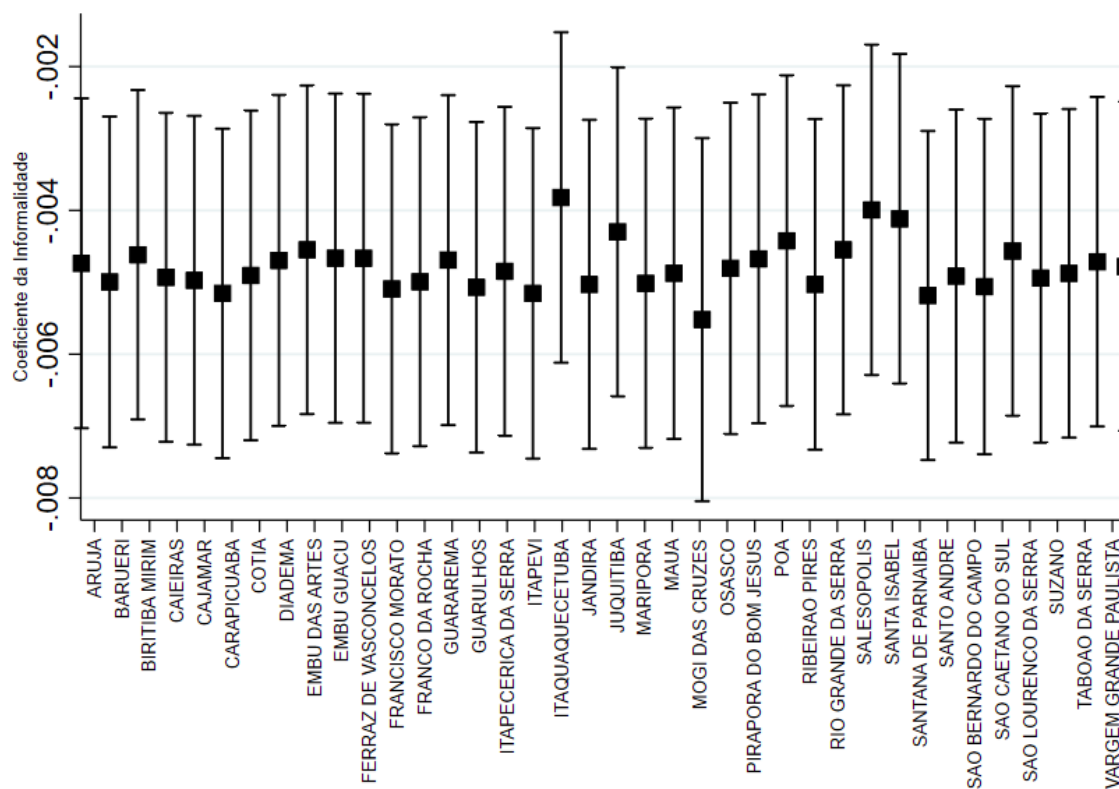
Fonte: Elaboração própria com base nas estimativas

Gráfico A2.2 – Validação Cruzada “leave-one-out” para Desemprego - CBD

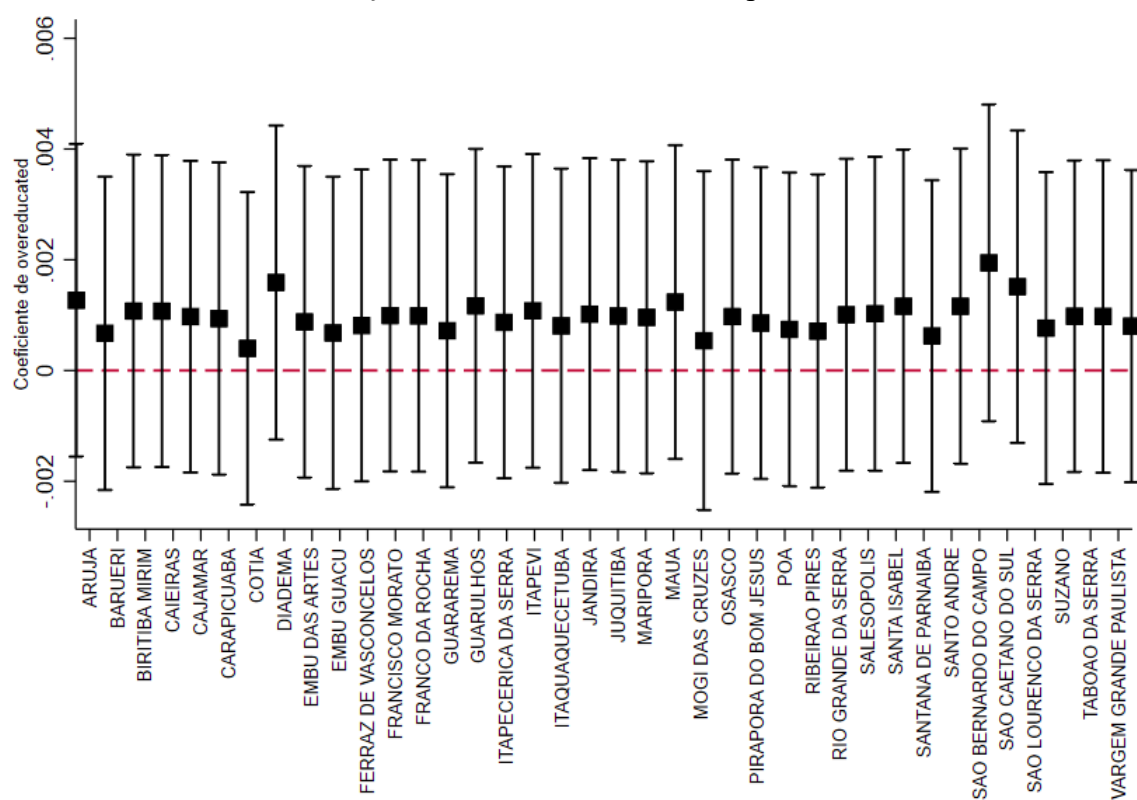


Fonte: Elaboração própria com base nas estimativas

Gráfico A2.3 – Validação Cruzada “leave-one-out” para Informalidade - CBD



Fonte: Elaboração própria com base nas estimativas

Gráfico A2.4 – Validação Cruzada “leave-one-out” para *overeducated* - CBD

Fonte: Elaboração própria com base nas estimativas

Referências

- ABANE, A., 1993. Mode choice for the journey to work among formal sector employees in Accra, Ghana. *J. Transp. Geogr.* 1 (4), 219–229.
- ALONSO, W. Location and land use. Toward a general theory of land rent. 1964.
- ANDERSSON, Fredrik, et al. Job displacement and the duration of joblessness: The role of spatial mismatch. *Review of Economics and Statistics*, 2018, 100.2: 203-218.
- ANGRIST, J. AND EVANS, W. (1998). “Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, 88(3): 450-477.
- ANGRIST, Joshua D.; PISCHKE, Jorn-Steffen. **Mostly Harmless Econometrics: an Empiricist's Companion**. Princeton: Princeton University Press, 2008.
- ASLUND, O; ÖSTH, J. AND ZENOU, Y. (2006) How important is access to jobs? Old question— improved answer. Working Paper No. 2006-1, Institute for Food Studies and Agro-industrial Development (IFAU), Hørsholm.
- AVNER, P.; LALL, S. (2016) Matchmaking in Nairobi: The Role of Land Use. Policy Research Working Paper No. 7904. World Bank, Washington, DC.
- BARBOSA, M. R. M; SILVEIRA NETO, R. M. "Adensamento urbano como condicionante da mobilidade: o caso da região metropolitana do Recife." *Revista Brasileira de Estudos Regionais e Urbanos* 11.2 (2017): 233-250.
- BARUFI, A. M. B; HADDAD, E.A. Spatial mismatch, wages and unemployment in metropolitan areas in Brazil. *Region*, Vol. 4, No. 3, pp. 175-200, 2017.
- BAUE, T. K. Educational mismatch and wages: a panel analysis. *Economics of Education Review*. Cambridge, v.21, n.3, p. 221-229, jun. 2002.
- BAUM, Charles L. The effects of vehicle ownership on employment. *Journal of Urban Economics*, 2009, 66.3: 151-163.
- BAUM-SNOW, Nathaniel. Did highways cause suburbanization? *The Quarterly Journal of Economics*, v.122, n.2, p.775-805, mai.2007.
- BAUM-SNOW, N; BRANDT, L; HENDERSON, J. V; TURNER, M. A; ZHANG, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics*, 99(3), 435-448.
- BELMIRO, C. H. P; RODRIGUES, F. A. C; SILVEIRA NETO, R. M. De Centro Exportador a Polo de Serviços - a Cidade do Recife ainda é Monocêntrica? Uma Investigação Empírica sobre sua Atual Configuração. VI Encontro Pernambucano de Economia. Pensando o Brasil: Perspectivas Pós-crise. Recife. 2017.
- BLUMENBERG, E., ONG, P., 2001. Cars, buses, and jobs-welfare participants and employment access in Los Angeles, Sustainability and environmental concerns in transportation 2001: Planning and administration; energy and environment, pp. 22-31.

BOISJOLY, G., MORENO-MONROY, A. I., & EL-GENEIDY, A. (2017). Informality and accessibility to jobs by public transit: Evidence from the São Paulo Metropolitan Region. *Journal of Transport Geography*, 64, 89-96.

BOLLINGER, C. R. IHLANFELDT, K. R. 2003: The intraurban spatial distribution of employment: Which government interventions make a difference? *Journal of Urban Economics*, 53(3), 396–412.

BORGES, Breno Albuquerque Brandão. *Patrimônio ferroviário e autenticidade: avaliação do Conjunto Ferroviário de Caruaru*. MS thesis. Universidade Federal de Pernambuco, 2014.

BRANCO, A. M. Os custos sociais do transporte urbano brasileiro. *Revista dos Transportes Públicos*, ANTP, São Paulo, n. 84, 1999.

BRUECKNER, J. K.; THISSE, J.; ZENOU, Y. Why Is Central Paris Rich and Downtown Detroit Poor? An Amenity-Based Theory. *European Economic Review*. v. 43, p. 91-107, 1999.

BÜCHEL, Felix; VAN HAM, Maarten. Overeducation, regional labor markets, and spatial flexibility. *Journal of Urban Economics*, 2003, 53.3: 482-493

BUZAR, S., OGDEN, P.E., HALL, R. (2005) “Household matters: the quiet demography of urban transformation”, 29(4): 413-436.

CALV´O-ARMENGOL A, JACKSON M. 2004. The effects of social networks on employment and inequality. *American Economic Review* 94: 426–454.

CAMPOS, Felipe Andrade Souza. A busca por trabalho sob a hipótese de descasamento espacial na cidade de Salvador. 90 p. (Dissertação de Mestrado) — Universidade Federal da Bahia, 2015.

CARVALHO, D. S. A. D. (2017). Transporte e densidade populacional na área urbana de Salvador. *Dissertação de Mestrado*, Programa de Pós-Graduação em Economia da Faculdade de Economia da Universidade Federal da Bahia.

CASCETTA, Ennio; CARTENÌ, Armando; MONTANINO, Marcello. A new measure of accessibility based on perceived opportunities. **Procedia - Social and Behavior al Sciences**, v.87, p.117-132, 2013.

CERVERO, R; WU, K.L. (1997). Polycentrism, commuting, and residential location in the San Francisco Bay Area. *Environment and Planning A* 29(5): 865–886.

CHANDRA, A; THOMPSON, E. Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. **Regional Science and Urban Economics**, v. 30, n. 4, p. 457-490, 2000.

CHAPPLE, K. (2001). Time to work: Job search strategies and commute time for women on welfare in San Francisco. *Journal of Urban Affairs* 23(2): 155–173.

CHARRON, M., 2007. From excess commuting to commuting possibilities: more extension to the concept of excess commuting. *Environ. Plan. A* 39 (5), 1238.

CHEN, Y., JAUPART, P., MORENO-MONROY, A., & PICARELLI, N. (2017). Unequal commutes. Job accessibility e employment in Accra.

CHETTY, RAJ; NATHANIEL, H. “The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects,” *Quarterly Journal of Economics*, 133 (2018), 1107–1162

CHETTY, R; N. HENDREN; P. KLINE; E. SAEZ. “Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States,” *Quarterly Journal of Economics*, 129 (2014), 1553–1623.

CLOGG, C. C.; SHOCKEY, J. W. Mismatch between occupation and schooling: a prevalence measure, recent trends, and demographic analysis. *Demography*. Chicago, v. 21, n. 2, p. 233-257, mai. 1984.

COMBES, P. P., DURANTON, G., GOBILLON, L., & ROUX, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration economics* (pp. 15-66). University of Chicago Press.

CÔRTEZ, M. V. R. **Localização residencial e demanda por recreação**: um modelo para cidades litorâneas. 2008. Dissertação (Mestrado em Engenharia) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2008.

COSTAL, Giovanni; PICKUP, Laurie; DI MARTINO, Vittorio. Commuting—a further stress factor for working people: evidence from the European Community. **International archives of occupational and environmental health**, v. 60, n. 5, p. 377-385, 1988.

CRANE, R. (2007) “Is There a Quiet Revolution in Women’s Travel? Revisiting the Gender Gap in Commuting”, *Journal of the American Planning Association*, Summer 2007, 73: 298-316.

CROCE, Giuseppe; GHIGNONI, Emanuela. Educational mismatch and spatial flexibility in Italian local labour markets. *Education Economics*, 2015, 23.1: 25-46.

CUSTÓDIO, Vanderli. Dos surtos urbanísticos do final do século XIX ao uso das várzeas pelo Plano de Avenidas. *Geosul*, 2004, 19.38: 77-98.

DA COSTA OLIVEIRA, E. N. A Economia do Algodão no Nordeste Brasileiro. VI Encontro Estadual de História da ANPUH/SE. 2018.

DEVILLANOVA, Carlo. Over-education and spatial flexibility: New evidence from Italian survey data. *Papers in Regional Science*, 2013, 92.3: 445-464.

DIAS FILHO, J.M.; CORRAR, L.J. Regressão logística. In: CORRAR, L.J; PAULO, E; DIAS FILHO, J.M. (Orgs.). *Análise multivariada para os cursos de Administração, Ciências Contábeis e Economia*. São Paulo: Atlas, 2012, p. 280-323.

DIAZ, M. D. M.; MACHADO, L. Overeducation e Undereducation no Brasil: Incidência e Retornos. *Estudos Econômicos*. São Paulo, v.38, n.3, p. 431-460, jul.-set. 2008.

DI PAOLO, Antonio; MATAS, Anna; RAYMOND, Josep. Job accessibility, employment and job-education mismatch in the metropolitan area of Barcelona. *Employment and Job-Education Mismatch in the Metropolitan Area of Barcelona* (May 26, 2014), 2014.

DUARTE, José Lins. Recife no tempo da maxambomba (1867-1889): *o primeiro trem urbano do Brasil*. 2005. Master's Thesis. Universidade Federal de Pernambuco.

DUJARDIN, C., SELOD, H. AND THOMAS, I. (2008) Residential segregation and unemployment: the case of Brussels, *Urban Studies*, 45, pp. 89–113.

DUNCAN, G. J.; HOFFMAN, S. D. The Incidence and Wage Effects of Overeducation. *Economics of Education Review*. Cambridge, v. 1, n.1, p. 75-86. 1981.

DURANTON, G; TURNER, M. A. Urban growth and transportation. *The Review of Economic Studies*, 79(4): 1407–1440, 2012.

DURANTON, G. TURNER, M. A. 2011. The fundamental Law of Road congestion: evidence from us cities. *American Economic Review* 101, 2616-2652.

EDLUND, L., MACHADO, L., SVIATSCHI, M.M. (2015) “Bright Minds, Big Rents: Gentrification and the Rising Returns to Skills”, NBER Working Papers Series 21729, <http://www.nber.org/papers/w21729>.

EISENBERG, P. L. (1977) *Modernização sem mudanças: a indústria açucareira de Pernambuco, 1840-1910*; Rio de Janeiro: Paz e Terra; Campinas: Universidade Estadual de Campinas.

ELLWOOD, David T. The spatial mismatch hypothesis: Are there teenage jobs missing in the ghetto?. In: *The black youth employment crisis*. University of Chicago Press, 1986. p. 147-190.

ESTEVEZ, L. A. Incompatibilidade escolaridade-ocupação e salários: evidências de uma empresa industrial brasileira. **Revista brasileira de economia**, Rio de Janeiro, v. 63, n. 2, p. 77-90, abr./jun. 2009.

FABER, Benjamin. Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System. **Review of Economic Studies**, v. 81, n. 3, p. 1046-1070, 2014.

FEITOSA, F. F. et al. (2007) “Global and local spatial indices of urban segregation”, *International Journal of Geographical Information Science*, 21(3): 299-323.

FREEMAN, R. B. *The overeducated American*. New York: Academic Press, 1976. 212p.

FREITAS, M. P; FREIRE, M. E. L; FARIA, M. S. (2019). Os caminhos do açúcar em Pernambuco: reflexões sobre a relação espacial e operacional da ferrovia com a usina de açúcar. *Instituto do Patrimônio Histórico e Artístico Nacional–Pernambuco*. Disponível em: <<https://docplayer.com.br/47107728-Os-caminhos-do-acucar-em-pernambuco-reflexoes-sobrea-relacao-espacial-e-operacional-da-ferrovia-com-a-usina-de-acucar.html>>. Acesso em, 23.

FUJITA, M., & OGAWA, H. (1982). Multiple equilibria and structural transition of nonmonocentric urban configurations. *Regional Science and Urban Economics*, 12(2), 161-196.

GARCIA-LOPEZ, Miquel-Àngel. Urban spatial structure, suburbanization and transportation in Barcelona. *Journal of Urban Economics*, v.72, n.2-3, p.176-190, set./nov. 2012.

GARCIA-LÓPEZ, M. A; HOLL, A; VILADECANS-MARSAL, E. Suburbanization and highways: When the Romans, the Bourbons and the first cars still shape Spanish cities. *Journal of Urban Economics*, 85:52–67, 2015.

GEURS, Karst T.; VAN WEE, Bert. Accessibility evaluation of land-use and transport strategies: review and research directions. **Journal of Transport Geography**, v.12, n. 2, p. 127-140, jun.2004.

GIULIANO, G., SMALL, K.A., 1993. Is the journey to work explained by urban structure? *Urban Stud.* 30 (9), 1485–1500.

GIULIANO, G. Information technology, work patterns and intra-metropolitan location: A case study. *Urban Studies*, 1998, 35.7: 1077-1095.

GLAESER, E. L. (1996) Spatial effects upon employment outcomes: the case of New Jersey teenagers. Discussion, *New England Economic Review*, May/June, pp. 58–64.

GLAESER, E. L.; KAHN, M. E.; RAPPAPORT, J. Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, v. 63, n. 1, p. 1-24, 2008.

GOBILLON, L. AND SELOD, H. (2006). Ségrégation résidentielle, accessibilité aux emplois et chômage: le cas de l’Île de France. Working Paper No. 0605, Laboratoire d’Economie Appliquée, INRA, Paris.

GOBILLON, L; SELOD, H. (2014). Spatial Mismatch, Poverty, and Vulnerable Populations, Post-Print halshs00978439, HAL.

GOBILLON L, MAGNAC T, SELOD H (2011). The effect of location on finding a job in the Paris region. *Journal of Applied Econometrics*, 26(7):1079-1112.

GROOT, W; MAASSEN VAN DEN BRINK, H. Overeducation in the labor market: a meta-analysis. **Economics of education review**, Columbia, v. 19, n. 2, p. 149-158, 2000.

GUBITS, D. B. (2004). Commuting, Work Hours, and the Metropolitan Labor Supply Gradient, Mimeo.

HADDAD, E. A; 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.

HAIR, J. F., et al. Análise multivariada de dados. Trad. Adonai S. Sant’Anna e Anselmo C. Neto. 5 ed. Porto Alegre: Bookman, 2005.

HAIR, J. F. J.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. Análise multivariada de dados. 6ª edição. Porto Alegre: Bookman, 2009.

HANSEN, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners*, 25(2), 73-76.

HANSON, S; PRATT, G, (1995), *Gender, Work, and Space*. London: Psychology Press.

HANSSON, Erik et al. Relationship between commuting and health outcomes in a cross-sectional population survey in southern Sweden. **BMC public health**, v. 11, n. 1, p. 834, 2011.

HARTOG, J. Over-education and earnings: where are we, where should we go? *Economics of Education Review*, Cambridge, Mass., v.19, n. 2, p. 131-147, Abr. 2000.

HENSEN, Maud M.; DE VRIES, M. Robert; CÖRVERS, Frank. The role of geographic mobility in reducing education-job mismatches in the Netherlands. *Papers in Regional Science*, 2009, 88.3: 667-682.

HOLZER H, IHLANFELDT K. 1998. Customer discrimination and employment outcomes for minority workers. *Quarterly Journal of Economics* 113: 835–867

HOLZER, Harry J.; IHLANFELDT, Keith R.; SJOQUIST, David L. Work, search, and travel among white and black youth. *Journal of Urban Economics*, 1994, 35.3: 320-345.

HOLZER, H. J., QUIGLEY, J. M. AND RAPHAEL, S. (2003) Public transit and the spatial distribution of minority employment: evidence from a natural experiment, *Journal of Policy Analysis and Management*, 22, pp. 415–423.

HORNER, M.W., MURRAY, A.T., 2002. Excess commuting and the modifiable areal unit problem. *Urban Stud.* 39 (1), 131–139.

HSU, Wen-Tai; ZHANG, Hongliang. The fundamental law of highway congestion revisited: Evidence from national expressways in Japan. **Journal of Urban Economics**, v. 81, p. 65-76, 2014.

HUBER, Peter. Do commuters suffer from job–education mismatch?. *Applied Economics Letters*, 2012, 19.4: 349-352.

IAMMARINO, Simona; MARINELLI, Elisabetta. Education–Job (mis) match and interregional migration: Italian university graduates' transition to work. *Regional Studies*, 2015, 49.5: 866-882.

IBGE. **Base de informações do Censo Demográfico 2010**: Resultados do Universo por área de ponderação. Rio de Janeiro: IBGE, 2011.

IBGE- Instituto Brasileiro de Geografia e Estatística. **Síntese dos Indicadores Sociais: Uma Análise das Condições de Vida da População Brasileira 2013**. Rio de Janeiro, 2013. Disponível em: <http://www.ibge.gov.br/> Acesso em: 08/07/2018.

IHLANFELDT, K. (2006) “A Primer on Spatial Mismatch within Urban Labor Markets”, in *A Companion to Urban Economics*, Edited by Arnott R. J. and McMillen, 404-17.

- IHLANFELDT, Keith R. Intra-urban job accessibility and Hispanic youth employment rates. *Journal of Urban Economics*, 1993, 33.2: 254-271.
- IHLANFELDT, Keith R.; SJOQUIST, David L. Job accessibility and racial differences in youth employment rates. *The American economic review*, 1990, 80.1: 267-276.
- IMMERGLUCK, Daniel. Job proximity and the urban employment problem: do suitable nearby jobs improve neighbourhood employment rates?. *Urban Studies*, 1998, 35.1: 7-23.
- INGRAM, D. R. The concept of accessibility: a search for an operational form. *Regional Studies*, v.5, n.2, p.101-107, fev.1971
- IPEA. INSTITUTO DE PESQUISA ECONÔMICA APLICADA. Atlas da vulnerabilidade social nas regiões metropolitanas brasileiras. Brasília: Ipea, 2017. Disponível em: <<http://ivs.ipea.gov.br/index.php/pt/>>. Acesso em: abr, 2018.
- IPHAN. Inventário do Patrimônio Ferroviário em Pernambuco, v. 3. Recife: Iphan – PE, 2009.
- JAUHIAINEN, Signe. Overeducation in the Finnish regional labour markets. *Papers in Regional Science*, 2011, 90.3: 573-588.
- JOHNSTON-ANUMONWO, Ibipo. The influence of household type on gender differences in work trip distance. *The Professional Geographer*, 1992, 44.2: 161-169.
- JOHNSTON, R. J., JONES, K., BURGESS, S. et al. (2004) Scale, factor analyses, and neighborhood effects, *Geographical Analysis*, 36(4), pp. 350–368.
- JUNG, G; YANG, T-C (2016) “Household Structure and Suburbia Residence in U.S. Metropolitan Areas: Evidence from the American Household Survey”, *Social Sciences*, 5, 74: 1-17.
- KAIN, J. (1968). Housing segregation, negro employment and metropolitan decentralization. *Quarterly Journal of Economics* 82: 175–197.
- KARLSTRÖM, Anders; ISACSSON, Gunnar. (2009). Is sick absence related to commuting travel time?: Swedish Evidence Based on the Generalized Propensity Score Estimator. 2010.
- KASARDA, John D. The implications of contemporary redistribution trends for national urban policy. *Social Science Quarterly*, 1980, 61.3/4: 373-400.
- KAWABATA, M. (2003) “Job access and employment among low-skilled autoless workers in US metropolitan areas”, *Environment and Planning A* 35, 1651-68.
- KIM, J.H, PAGLIARA, F., PRESTON, J. (2005) “The Intention to Move and Residential Location Choice Behavior”, *Urban Studies*, 42(9): 1621-1636.
- LEE B. SONG; MACDONALD J. F. (2003). Determinants of commuting time and distance for Seoul Residents: The impact of family status on the commuting of women. *Urban Studies*, Vol. 40, No. 7, 1283-1302,

LUCAS, R. E; ROSSI-HANSBERG, E. (2002). On the internal structure of cities. *Econometrica*, 70 (4), 1445-1476.

MACHADO, A. F.; OLIVEIRA, A. M. H. C.; CARVALHO, N. F. Tipologia de qualificação da força de trabalho: uma proposta com base na noção de incompatibilidade entre ocupação e escolaridade. **Nova economia**, Belo Horizonte, v. 14, n. 2, p. 11-33, 2004.

MACHADO, D.C; PERO, V; MIHESSEN, V. Mobilidade Urbana e Mercado de Trabalho na Região Metropolitana do Rio de Janeiro. *Revista da ABET*. 2016 Mar 1;14(2).

MADDEN, J. F. Why women work closer to home. *Urban studies*, 1981, 18.2: 181-194.

MARTINCUS, C. V., CARBALLO, J., & CUSOLITO. Routes, exports and employment in developing countries: Following the trace of the inca roads. Mimeo, Inter-American Development Bank, 2013.

MATAS, A., RAYMOND, J. L. AND ROIG, J. L. (2010) “Job Accessibility and Female Employment Probability: The Cases of Barcelona and Madrid”, *Urban Studies* 47(4), 769-87.

MAYER, T; TREVIEN, C. The impact of urban public transportation evidence from the Paris region. **Journal of Urban Economics**, v. 102, p. 1-21, 2017.

MCGUINESS, S. Overeducation in the labor market. *Journal of Economic Surveys*, v. 20, n. 3, p. 387-418. 2006.

MCMILLEN, D. P. (2001). Nonparametric employment subcenter identification. *Journal of Urban economics*, 50(3), 448-473.

MELO, M. L. Metropolização e subdesenvolvimento: o caso do Recife. Recife: UFPE, 1975.

MELO, P. C., & GRAHAM, D. J. (October 2009). Agglomeration economies and labour productivity: Evidence from longitudinal worker data for GB's travel-to-work areas. SERC Discussion Paper, No. 31.

MELO, Josemir Camilo de. Modernização e mudanças: o trem inglês nos canaviais do Nordeste (1852-1902). Tese de Doutorado do Programa de Pós-Graduação em História da Universidade Federal de Pernambuco (UFPE). Recife: 2000.

MEYER, R. M. P; GROSTEIN, M. D; BIDERMAN, C. (2004). São Paulo metrópole. São Paulo: Edusp.

MICHAELS, G. The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, 90(4):683–701, 2008.

MIHESSEN, Vitor; MACHADO, Danielle Carusi; PERO, Valéria. Mobilidade Urbana e Mercado de Trabalho na Região Metropolitana do Rio de Janeiro. In: **Anais do XLII Encontro Nacional de Economia [Proceedings of the 42nd Brazilian Economics Meeting]**. ANPEC-Associação Nacional dos Centros de Pósgraduação em Economia [Brazilian Association of Graduate Programs in Economics], 2016.

MILLS, E. S. An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, p. 197-210, 1967.

MITRA, A., 2005. Women in the urban informal sector: perpetuation of meagre earnings. *Dev. Chang.* 36 (2), 291–316.

MONSUETO, S. E; GOVEIA, J.M.A., CARRIJO, B.C.P.S. (2014) “Tempo de deslocamento do domicílio até o trabalho e salários em Goiás, Brasil”, VI Congreso de la Asociación Latinoamericana de Población, Lima, Peru.

MORAES, A. C. Congestionamento urbano: custos sociais. *Revista ANTP*, São Paulo, n. 135, p. 41-48, 2013.

MORENO-MONROY, A., POSADA, H., 2014. On the effect of transport subsidies on informality rates. CEPESP discussion paper, Fundação Getulio Vargas.

MORENO-MONROY, A., RAMOS, F., 2015. The impact of public transport expansions on informality: The case of the São Paulo Metropolitan Region.

MORENOFF, J. D. SAMPSON, R. (1997) Violent crime and the spatial dynamics of neighborhood transition: Chicago, 1970-1990. *Social Forces*, 76, 31-64.

MOON, M. (2015) “From gentrification to youngthification? The increasing importance of young age in delineating high-density living”, *Urban Studies*, 53(14): 1-18.

MOTTE, B., AGUILERA, A., BONIN, O., NASSI, C.D., 2016. Commuting patterns in the metropolitan region of Rio de Janeiro. What differences between formal and informal jobs? *Journal of Transport Geography* 51, 59-69.

MTPA – MINISTÉRIO DOS TRANSPORTES, PORTOS E AVIAÇÃO CIVIL. Transportes no Brasil – Síntese Histórica. Disponível em: < <http://www.transportes.gov.br/conteudo/136-transportes-no-brasil-sintese-historica.html>>. Acesso em: abr. 2018

MUTH, R. F. Cities and Housing; The Spatial Pattern of Urban Residential Land Use. Chicago, IL: University of Chicago Press, 1969. 355 p.

NADALIN, Vanessa Gapriotti; IGLIORI, Danilo Camargo. Evolução urbana e espraiamento na Região Metropolitana de São Paulo. 2010.

NELEN, Annemarie. 2012. Does commuting affect health?. Working Paper, University of Wuppertal.

NOVACO, Raymond W.; COLLIER, Cheryl. Commuting stress, ridesharing, and gender: Analyses from the 1993 state of the commute study in Southern California. 1994.

OLIVEIRA, Rodrigo et al. Commute duration and health: Empirical evidence from Brazil. **Transportation Research Part A: Policy and Practice**, v. 80, p. 62-75, 2015.

OLIVEIRA, T. G; SILVEIRA NETO, R. M. Segregação residencial na cidade do Recife: um estudo da sua configuração. *Revista Brasileira de Estudos Regionais e Urbanos* 9, no. 1 (2016): 71-92.

ONG, P. M; MILLER, D. (2005). “Spatial and Transportation Mismatch in Los Angeles”, *Journal of Planning and Research* 25, 43-56.

O'REGAN, K. M. QUIGLEY, J. M. (1996) Spatial effects upon employment outcomes: the case of New Jersey teenagers, *New England Economic Review*, May/June, pp. 41–58.

O'REGAN, K. M. QUIGLEY, J. M. (1998) Where youth live: economic effects of urban space on employment prospects, *Urban Studies*, 35(7), pp. 1187–1205.

PATACCHINI, E; ZENOU, Y. (2005) “Spatial mismatch, transport mode and search decisions in England”, *Journal of Urban Economics* 58, 62-90.

PATACCHINI E, ZENOU Y. 2006. Search activities, cost of living, and local labor markets. *Regional Science and Urban Economics* 36: 227–248.

PERO, Valéria; STEFANELLI, Victor. A questão da mobilidade urbana nas metrópoles brasileiras. **Revista de Economia Contemporânea**, v. 19, n. 3, p. 366-402, 2015.

PERO, Valéria; MIHESSEN, Vitor. Mobilidade urbana e pobreza no Rio de Janeiro. **Revista Econômica**, v. 15, n. 2, 2013.

PIERRARD, O. (2008). Commuters, residents and job competition. *Regional Science and Urban Economics* 38(6), 565–577.

PONTUAL, V. (2001) Uma Cidade e dois Prefeitos, *Narrativas do Recife das 1930 a 1950*, Recife: Universidade Federal de Pernambuco.

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*, Programa de Pós-Graduação em Economia da Faculdade de Economia da Universidade Federal da Bahia.

QUIGLEY, J. (1976) “Housing Demand in the Short Run: an Analysis of Polytomous Choice”, in NBER (ed.), *Explorations in Economic Research*, Volume 3, number1.

QUINTANAR, S. A. K. Job accessibility and probability of being employed by educational level and informality in Metropolitan Area of Mexico City. Working Paper. Universitat Autònoma de Barcelona. [http://pagines.uab.cat/appliedeconomics/sites/pagines.uab.cat/appliedeconomics/files/Koi ke,% 20S. _paper. pdf](http://pagines.uab.cat/appliedeconomics/sites/pagines.uab.cat/appliedeconomics/files/Koi%20S._paper.pdf), 2013.

QUINN, M.A.; RUBB, S. Mexico's labor market: the importance of education-occupation matching on wages and productivity in developing countries. *Economics of Education Review*, v. 25, n. 2, p. 147-156, 2006.

RAMOS, Frederico Roman. Identificando subcentros de emprego na região metropolitana de São Paulo: potenciais locais e consequências para a estrutura urbana. *Anais*, 2016, 1-21.

RAMOS, Raul; SANROMÁ, Esteban. Overeducation and Local Labour Markets in Spain. *Tijdschrift voor economische en sociale geografie*, 2013, 104.3: 278-291.

RECIFE. Prefeitura da cidade do Recife. Pesquisa de Mobilidade Urbana da Região de Metropolitana de Recife. Recife: ICPS, 2017. Disponível em: <http://planodemobilidade.recife.pe.gov.br/node/9>. Acesso em: nov, 2017.

REDDING, S. J; TURNER, M. A. Transportation costs and the spatial organization of economic activity. NBER Working Paper 20235, National Bureau of Economic Research, 2014.

ROGERS, Cynthia L. Job search and unemployment duration: Implications for the spatial mismatch hypothesis. *Journal of urban Economics*, 1997, 42.1: 109-132.

ROSSI-HANSBERG, E. (2004). Optimal urban land use and zoning. *Review of Economic Dynamics*, 7, 69-106.

RUBB, S. Overeducation in the labor market: a comment and re-analysis of a meta-analysis. **Economics of education review**, Columbia, v. 22, p. 621-629, 2003.

_____. Overeducation, undereducation, and the theory of career mobility: a comment and a note on underemployment. **Applied economics letters**, Coventry, v. 12, n. 2, p. 115-118, 2005.

SANTOS. C. C. (2012). O Valor do Tempo na Avaliação de Projetos de Transportes. *Dissertação de Mestrado*, Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Rio Grande do Sul.

SANTOS, A. Overeducation no Mercado de trabalho brasileiro. *Revista Brasileira de Economia de Empresas*, v.2, n.2, p. 1-22. 2002

SCHWARTZMAN, S. Equity, quality and relevance in higher education in Brazil. **Anais da Academia Brasileira de Ciências**, v. 76, n. 1, p. 173-188, 2004.

SCOTT, M; URRY, J. (1994). *Economies of Signs and Space*. London: Sage.

SEABRA, D. M. S.; SILVEIRA NETO, RAUL M; MENEZES, T. A. Amenidades Urbanas e Valor das Residências: uma Análise Empírica para a Cidade do Recife. *Economia Aplicada* (Impresso), v. 20, p. 143-169, 2016.

SHEN, Qing. Location characteristics of inner-city neighborhoods and employment accessibility of low-wage workers. *Environment and planning B: Planning and Design*, 1998, 25.3: 345-365.

SILVA. R. A. Planos Ferroviários de 1886. Disponível em: <http://vfco.brazilia.jor.br/Planos-Ferrovios/1886-Plano-Rodrigo-Silva.shtml>. Acesso em mai, 2020.

SILVA, Rubens Lopes Pereira da; SILVEIRA NETO, Raul da Mota; ROCHA, Roberta. Localization patterns within urban areas: evidence from Brazil. *Area Development and Policy*, 2019, 4.2: 157-176.

SILVEIRA NETO, R. M; DUARTE, G. (2016) “Estrutura Familiar e Escolha de Localização de Residência nas Cidades, uma Análise Empírica para o Caso da Região Metropolitana de São Paulo”, *Desenvolvimento em Questão*, 33: 41-64.

SILVEIRA NETO, R., DUARTE, G., Paéz, A. (2015) “Gender and commuting time in São Paulo metropolitan region”, *Urban Studies*, 52(2): 298-313.

SILVEIRA NETO, Raul; MOURA, Klebson. Commuting time, public exposure and victimization: Evidence from Brazilian metropolitan regions. **Papers in Regional Science**, v. 98, n. 2, p. 1159-1175, 2019.

SIMPSON, W; VAN DER VEEN, A. (1992). The economics of commuting and the urban labor market. *Journal of Economic Surveys* 6(1): 45–62.

SLOVIC, Anne Dorothée, et al. The long road to achieving equity: Job accessibility restrictions and overlapping inequalities in the city of São Paulo. *Journal of Transport Geography*, 2019, 78: 181-193.

SMALL, K.A; SONG, S. (1992). Wasteful commuting: A resolution. *Journal of Political Economy* 100(4): 888–898.

SMITH T, ZENOU Y. 2003. Spatial mismatch, search effort and urban spatial structure. *Journal of Urban Economics* 54: 129–156.

STOCK, J. H.; WATSON, M. W. *Econometria*. São Paulo: Pearson AddisonWesley, 2004.

STOCK, James H.; WRIGHT, Jonathan H.; YOGO, Motohiro. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 2002, 20.4: 518-529.

STONE, Arthur A.; SCHNEIDER, Stefan. 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**, v. 42, p. 117-124, 2016.

STORPER, M; WALKER, R. (1983). The theory of labor and the theory of location. *International Journal of Urban and Regional Research* 7(1): 1–43.

STUTZER, A. AND B. S. FREY (2008). Stress that doesn't pay: The commuting paradox. *Scandinavian Journal of Economics* 110(2), 339–366.

SUÁREZ, M; DELGADO, J. (2007). Estructura y eficiencia urbanas: Accesibilidad a empleos, localización residencial e ingreso en la ZMCM 1990–2000. *Economía, Sociedad y Territorio* VI(23): 693–724.

SUÁREZ, M; DELGADO, J. (2010). Patrones de movilidad residencial en la Ciudad de México como evidencia de co-localización de población y empleos. *EURE* 36(107): 67–91.

SUÁREZ, M; MURATA, M; CAMPOS, J. D. "Why do the poor travel less? Urban structure, commuting and economic informality in Mexico City." *Urban studies* 53.12 (2016): 2548-2566.

SUMMERHILL, W. R. (2018). *Trilhos do Desenvolvimento, as ferrovias no crescimento da economia brasileira 1854-1913*, São Paulo: Alfaatar.

TRAVASSOS, L. R. F. C; GROSTEIN, M. D. (2003). A ocupação de fundos de vale urbanos em São Paulo e sua dimensão sócio-ambiental. In: *Novas escalas e estratégias territoriais na gestão ambiental*. X Encontro Nacional da Anpur.

TURNER, T; NIEMEIER, D. Travel to work and household responsibility: new evidence. *Transportation*, 1997, 24.4: 397-419.

VAN OMMEREN, J. N.; GUTIÉRREZ, E. Labour supply and commuting. *Journal of Urban Economics*, v. 68, p. 82-89, 2010.

VAN OMMEREN, J. E GUTIÉRREZ-I-PUIGARNAU, E. (2011). Are Workers with a Long Commute Less Productive? An Empirical Analysis of Absenteeism. *Regional Science and Urban Economics*, v. 41, pp. 1-8.

VERDUGO, R.; VERDUGO, N. The impact of surplus schooling on earnings: some additional findings. *Journal of Human Resources*, Madison, Wis., v. 24, n. 4, 629643, 1989.

VIANNA, Cláudia Hamacek; OLIVEIRA, A. M. H. C. Sobre-escolarização nas ocupações brasileiras: uma análise dos efeitos de idade, período e coorte. *ENCONTRO NACIONAL DE ECONOMIA*, 2010, 38.

VIEIRA, R., & HADDAD, E. A. (2015). An accessibility index for the metropolitan region of São Paulo. In K. Kourtit, P. Nijkamp, & R. R. Stough (Eds.), *The rise of the City: Spatial dynamics in the urban century* (pp. 242e258). Cheltenham, UK: Edward Elgar Publishing.

WALES, T. J. (1978). 'Labour supply and commuting time: an empirical study', *Journal of Econometrics*, vol. 8(2) (October), pp. 215–26.

WANG, C; CHEN, N. (2015). A GIS-based spatial statistical approach to modeling job accessibility by transportation mode: Case study of Columbus, Ohio. *Journal of Transport Geography* 45, 1-11.

WARDMAN, M. (1998). The Value of Travel Time: A Review of British Evidence. *Journal of Transport Economics and Policy*, v 32, n. 3, PP. 285–316.

WEBSITE. The Tramways of Recife Pernambuco states Brazil. Disponível em: <<http://www.tramz.com/br/re/re.html>>. Acesso em: nov, 2017.

WEINBERG, B. A., REAGAN, P. A. AND YANKOW, J. J. (2004) Do neighborhoods affect hours worked? Evidence from longitudinal data, *Journal of Labor Economics*, 22, pp. 891–924.

WENER, Richard E. et al. Running for the 7: 45: The effects of public transit improvements on commuter stress. **Transportation**, v. 30, n. 2, p. 203-220, 2003.

WHEELER, C. H. (2003). Evidence on agglomeration economies, diseconomies, and growth. *Journal of Applied Econometrics* 18: 79–104.

WHITE, M. J. (1988). 'Location choice and commuting behaviour in cities with decentralized employment', *Journal of Urban Economics*, vol. 24(2) (September), pp. 129–52.

WILSON J. 1996. *When Work Disappears: The World of the New Urban Poor*. Alfred A. Knopf: New York.

ZAX, Jeffrey S.; KAIN, John F. Moving to the suburbs: do relocating companies leave their black employees behind?. *Journal of Labor Economics*, 1996, 14.3: 472-504.

ZENOU, Y. (2002). How do Firms Redline Workers? *Journal of Urban Economics*, v. 52, pp. 391-408.

ZHENG, X. P. (1998) Measuring optimal population distribution by agglomeration economies and diseconomies: A case study of Tokyo. *Urban Studies* 35: 95–112.