UNIVERSIDADE FEDERAL DA BAHIA
FACULDADE DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA
DOUTORADO EM ECONOMIA APLICADA

EVERLÂNDIA DE SOUZA SILVA

ENSAIOS SOBRE AGLOMERAÇÃO, OCUPAÇÕES E SALÁRIOS: UMA ANÁLISE PARA O MERCADO DE TRABALHO FORMAL DO BRASIL

Salvador

2021

EVERLÂNDIA DE SOUZA SILVA

ENSAIOS SOBRE AGLOMERAÇÃO, OCUPAÇÕES E SALÁRIOS: UMA ANÁLISE PARA O MERCADO DE TRABALHO FORMAL DO BRASIL

Tese apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia da Universidade Federal da Bahia como requisito parcial para a obtenção do grau de Doutora em Economia.

Área de atuação: Economia aplicada.

Orientadora: Prof. Dr. Diana Lúcia Gonzaga da Silva.
Coorientadora: Prof. Dr. Roberta de Moraes Rocha.

Salvador

2021

Ficha catalográfica elaborada por Vânia Cristina Magalhães CRB 5- 960

Silva, Everlândia de Souza.
S586 Ensaios sobre aglomeração, ocupações e salários: uma análise para o mercado de trabalho formal do Brasil./ Everlândia de Souza Silva. - 2021.
278 f. il.; tab.; fig.
Tese (doutorado) – Universidade Federal da Bahia. Faculdade de Economia, Salvador, 2021.
Orientadora: Profa. Dra. Diana Lúcia Gonzaga da Silva.
Co-orientadora: Profa. Dra. Roberta de Moraes Rocha.

1.Economia de aglomeração. 2. Salários. 3. Economia urbana. 4. Economia do trabalho. 5.Trabalho formal. I. Silva, Diana Lúcia Gonzaga. II. Rocha, Roberta de Moraes. III.Título. IV. Universidade Federal da Bahia. Faculdade de Economia.

CDD: 331. 981

Universidade Federal da Bahia
Faculdade de Economia
Programa de Pós-Graduação em Economia
Mestrado e Doutorado em Economia

TERMO DE APROVAÇÃO

Everlândia de Souza Silva

"ENSAIOS SOBRE AGLOMERAÇÃO, OCUPAÇÕES E SALÁRIOS: UMA ANÁLISE PARA O MERCADO DE TRABALHO FORMAL DO BRASIL"

Tese de Doutorado aprovada como requisito parcial para obtenção do Grau de Doutora em Economia no Programa de Pós-Graduação em Economia da Faculdade de Economia da Universidade Federal da Bahia, pela seguinte banca examinadora:

Salvador, 05 de novembro de 2021.

Profª Drª Diana Lúcia Gonzaga da Silva
(Orientadora - PPGE/ECO/UFBA)

Profª Drª Roberta de Moraes Rocha
(UFPE-Coorientadora)

Prof. Dr. Gervásio Ferreira dos Santos
(UFBA)

Prof. Dr. Carlos Roberto Azzoni
(USP)

Prof. Dr. Ricardo da Silva Freguglia
(UFJF)

AGRADECIMENTOS

Agradeço a Deus por ter me dado condições para a realização deste trabalho. Aos meus pais, Quitéria Maria e Manoel Casciano, em especial minha mãe, que em todos esses anos foi meu alicerce. Aos meus irmãos, Elitainne e Éder, e à minha prima, Emily, pelo incentivo e apoio prestados. Ao meu noivo, Robson, pela compreensão e dedicação, me dando força nos momentos mais difíceis dessa jornada.

Devo os meus sinceros agradecimentos à minha orientadora Diana Gonzaga por todo ensinamento, ajuda e compreensão. Sem a mesma a concretização deste trabalho seria impossível. À minha coorientadora Roberta Rocha, sempre serei grata pela sua disposição em me ajudar. Agradeço ainda a todo o corpo docente da PPGE/UFBA, em especial aos professores Gervásio Ferreira, Vinícius Mendes e Gisele Tiryaki. Ao Repositório de dados do Instituto de Saúde Coletiva da UFBA por disponibilizar o acesso ao servidor, possibilitando o tratamento dos dados para a realização da pesquisa. À Banca Examinadora por aceitar o convite e pelas suas contribuições para a melhoria deste trabalho.

Aos meus amigos, em especial Adelson Santos, conselheiro de todas as horas. Agradeço a todos os meus colegas de curso pelo companheirismo e amizade, em especial, Fernanda, Firmino, Kécia e Silvana.

Por fim, agradeço a todos os meus amigos e familiares pelo carinho, compreensão e disposição em contribuir, direta ou indiretamente, para a finalização deste trabalho.

RESUMO

Esta tese é dividida em dois ensaios, complementares, e tem por objetivo avaliar a aglomeração ocupacional e os seus efeitos sobre os diferenciais de salários locais. Para o desenvolvimento dos mesmos, utiliza-se os microdados individuais da RAIS para o período de 2009 a 2018, considerando como subunidade regional os Arranjos Populacionais brasileiros. No primeiro ensaio, busca-se estudar os padrões de aglomeração ocupacional no Brasil e os seus determinantes. A literatura mostra que a aglomeração ocupacional pode resultar das externalidades positivas no mercado de trabalho associadas ao compartilhamento de habilidades similares. Neste ensaio, as externalidades associadas aos mecanismos de *sharing* serão mensuradas a partir de um indicador de *labor pooling*. A partir do indicador proposto, o efeito do *labor pooling* poderá ser mensurado quantitativamente e qualitativamente, ao considerar as habilidades dos trabalhadores. Essa é a principal contribuição desse ensaio para a literatura. A estratégia empírica utiliza o estimador de efeitos fixos em um modelo de dados em painel para identificar o efeito do labor *pooling* sobre a aglomeração ocupacional no Brasil. Os resultados sugerem que o *labor pooling* apresenta efeito positivo e estatisticamente significante ao controlar a sua interação com as habilidades, mas a sua magnitude é inferior aos efeitos das habilidades. No que se refere a principal contribuição dessa pesquisa, o efeito do *labor pooling* apresentou comportamento heterogêneo quando classificado pelas habilidades dos trabalhadores, sendo maior para as habilidades de Atenção e menor para as habilidades Clericais. O segundo ensaio busca verificar como a especialização ocupacional, por categorias de habilidades dos trabalhadores (cognitivas, sociais e manuais), pode impactar no diferencial salarial individual, considerando a endogeneidade da escolha local e a heterogeneidade dos trabalhadores e firmas. Adicionalmente, busca-se decompor o prêmio salarial da especialização ocupacional por categorias de habilidades dos trabalhadores. A principal contribuição desse ensaio é realizar a decomposição dos efeitos da especialização de habilidades considerando os seus condicionantes urbanos, como a densidade populacional. A estratégia empírica utiliza dados em painel e propõe controlar as heterogeneidades não observadas individuais e de firmas no modelo de efeitos fixos, além de lidar com a endogeneidade a partir do uso de variáveis instrumentais. Os resultados do prêmio de especialização, com o controle da endogeneidade da escolha local pelo uso de variáveis instrumentais, indicaram que as estimativas para o Brasil estavam subestimadas. O modelo 2SLS (IV Históricas) apresentou um aumento no prêmio da especialização cognitiva e social. Todavia, o prêmio da especialização manual desapareceu. No que se refere ao prêmio salarial da especialização de habilidades com controle da heterogeneidade não observada dos trabalhadores e firmas, os resultados apontaram uma diminuição nas estimativas do prêmio salarial da especialização cognitiva, social e manual. Na análise da decomposição do prêmio de especialização por habilidades, a densidade apresentou um impacto negativo na determinação dos prêmios salariais da especialização cognitiva e social. As evidências para o Brasil sugerem que o efeito competição, inerente aos mercados de trabalho mais densos, pode superar o efeito positivo das interações.

Palavras-chave: economias de aglomeração; especialização ocupacional; l*abor pooling*; habilidades; salários.

ABSTRACT

This dissertation is divided into two complementary essays that aim to evaluate the occupational agglomeration and its effects on the differentials of local wages. To develop these essays, the research employed individual microdata from the Annual Report on Social Information (RAIS – *Relação Anual de Informações Sociais*) in the interval from 2009 to 2018, considering as regional subunit the Brazilian Populational Arrangements. In the first essay, the focus relies on the patterns of occupational agglomeration in Brazil and its determiners. The literature on the topic shows that occupation agglomeration can be a result of positive externalities from the labor market associated with the sharing of similar skills. In this essay, the measurement of externalities associated with mechanisms of sharing relies on an indicator of labor pooling. Based on this indicator, the labor pooling effect can be assessed both quantitatively and qualitatively by considering the workers' skills. This is the essay's main contribution to the literature. The empiric strategy uses the fixed effects model to identify the effect of labor pooling on the occupational agglomeration in Brazil. The results suggest that labor pooling presents a statistically significant positive effect after controlling the interaction with the workers' skills, however it is inferior in magnitude to the effects of the skills. Concerning this research main contribution, the labor pooling effect presented heterogeneous behavior when classified according to the workers' skills, thus being higher for Attention skills and lower for Clerical skills. The second essay aims to verify how the occupational specialization, given the categories of workers' skills (cognitive, social, and physical), can affect the differential for individual wages, considering the endogeneity of local choices and the heterogeneity of workers and firms. Furthermore, the study aims to decompose the wage premium of occupational specialization according to the categories of workers' skills. The main contribution of this essay is conducting the decomposition of the effects of specialization on skills, considering their urban conditions, such as populational density. The empiric strategy uses a fixed effect model and proposes to control the non-observed heterogeneity for individuals and firms. It also deals with the endogeneity bias by using instrumental variables. The results for the specialization wage premium with control of the local choice endogeneity using instrumental variables indicated that estimations for Brazil were underestimated. The IV-2SLS model showed an increase in the premium for cognitive and social specialization whereas the premium for physical specialization disappeared. Regarding the premium for skills specialization with control of non-observed heterogeneity for workers and firms, the results showed a decrease in the estimations for wage premium of cognitive, social, and physical specializations. In the analysis of the decomposition of the skill specialization premium, density had a negative impact on the determination of cognitive and social specialization premium. The evidences for Brazil suggest that the effect of competitivity inherent to more dense labor markets can surpass the positive effect of the interactions.

Keywords: agglomeration economies; occupational specialization; labor pooling; skills; wages.

LISTA DE FIGURAS

Figura 2.1 – Distribuição espacial do número de municípios e da densidade nos Arranjos Populacionais.....41
Figura 2.2 - Tendências gerais da Labor Pooling no Brasil (2009-2018).....67
Figura 2.3 - Tendências gerais do Labor Pooling ponderado pelas Habilidades no Brasil (2009-2018).....69
Figura 2.4 – Tendências gerais da concentração ocupacional entre 2009 e 2018.....73

Figura 3. 1 – Parcela dos trabalhadores formais alocados em Arranjos Populacionais 173
Figura 3. 2 – Composição dos trabalhadores formais por tamanho dos Arranjos Populacionais 174
Figura 3. 3 – Composição dos trabalhadores formais por categorias de habilidades alocados em Arranjos Populacionais.....174
Figura 3. 4 – Distribuição espacial do salário-hora real por habilidades nos Arranjos Populacionais.....177
Figura 3. 5 – Distribuição espacial da especialização ocupacional categorizada por habilidades nos Arranjos Populacionais 182
Figura 3. 6 – Salário-hora real e densidade nos Arranjos Populacionais (2009-2018).....183
Figura 3. 7 – Salário-hora real e especialização ocupacional categorizada por habilidades nos Arranjos Populacionais (2009-2018).....185
Figura 3. 8 – Especialização ocupacional categorizada por habilidades e densidade nos Arranjos Populacionais (2009-2018).....189

LISTA DE TABELAS

Tabela 2. 1 - Crescimento do emprego ocupacional e perfil de habilidades (2009-2018) 63
Tabela 2. 2 - Participação ocupacional no emprego total e perfil de habilidades (2009-2018) 64
Tabela 2. 3 - Distribuição regional das ocupações com maior participação no emprego total (2009-2018) 64
Tabela 2. 4 - Labor Pooling e perfil de habilidades (2009-2018) 66
Tabela 2. 5 - Ocupações menos concentradas para o ano de 2009 75
Tabela 2. 6 - Ocupações mais concentradas para o ano de 2009 76
Tabela 2. 7 - Ocupações menos concentradas para o ano de 2018 78
Tabela 2. 8 - Ocupações mais concentradas para o ano de 2018 79
Tabela 2. 9 - Estatística Descritivas das variáveis 86
Tabela 2. 10 - *Labor Pooling* como determinante da aglomeração ocupacional no Brasil: 2009 a 2018 89
Tabela 2. 11 - Determinantes da aglomeração ocupacional no Brasil: 2009 a 2018 91
Tabela 2. 12 - Determinantes da aglomeração ocupacional no Brasil (2009 a 2018): *Labor Pooling* ponderado pelas habilidades e Efeito Marginal do *Labor Pooling* completo 96

Tabela 3. 1 - Características dos trabalhadores alocados em Arranjos Populacionais 180
Tabela 3. 2 - Mobilidade de trabalhadores e firmas 191
Tabela 3. 3 - Prêmio salarial da especialização ocupacional por habilidades entre os grupos de trabalhadores (Cognitiva, Social e Manual): POLS e 2SLS 193
Tabela 3. 4 - Prêmio salarial da aglomeração por habilidades entre os grupos de trabalhadores (Cognitiva, Social e Manual): POLS e 2SLS 200
Tabela 3. 5 - Prêmio salarial da especialização ocupacional por habilidades entre os grupos de trabalhadores (Cognitiva, Social e Manual): Efeitos Fixos 203
Tabela 3. 6 - Prêmio salarial da aglomeração entre os grupos de trabalhadores (Cognitiva, Social e Manual): Efeitos Fixos 208
Tabela 3. 7 - Prêmio salarial da especialização ocupacional por habilidades entre os grupos de trabalhadores (Cognitiva, Social e Manual): Resumo 210
Tabela 3. 8 - Mobilidade de trabalhadores e firmas nas amostras 211
Tabela 3. 9 - Decomposição do efeito da especialização cognitiva sobre os salários 215
Tabela 3. 10 - Decomposição do efeito da especialização social sobre os salários 220
Tabela 3. 11 - Decomposição do efeito da especialização manual sobre os salários 222

SUMÁRIO

1 INTRODUÇÃO 10

2 AGLOMERAÇÃO OCUPACIONAL E SEUS DETERMINANTES: UMA ANÁLISE DO *LABOR POOLING* E DAS HABILIDADES DOS TRABALHADORES 16
2.1 INTRODUÇÃO 16
2.2 ECONOMIAS DE AGLOMERAÇÃO E DISTRIBUIÇÃO OCUPACIONAL ESPACIAL 21
2.2.1 Externalidades de aglomeração 24
2.2.2 *Labor Pooling*: Modelo Teórico 27
2.2.3 Evidências Empíricas 31
2.2.3.1 Aglomeração industrial 31
2.2.3.2 Aglomeração industrial no Brasil 32
2.2.3.3 Aglomeração ocupacional 34
2.2.3.4 Aglomeração ocupacional no Brasil 36
2.3 ESTRATÉGIA EMPÍRICA 38
2.3.1 Índice de aglomeração ocupacional 38
2.3.2 Modelo econométrico: Determinantes da aglomeração ocupacional 45
2.3.3 Estratégia de identificação 50
2.3.4 Dados 53
2.3.4.1 RAIS 54
2.3.4.2 Categorias de habilidade 56
2.3.4.3 Tratamento dos Dados 59
2.4 *LABOR POOLING* OCUPACIONAL E HABILIDADES DOS TRABALHADORES: O MERCADO DE TRABALHO FORMAL BRASILEIRO (2009-2018) 61
2.5 RESULTADOS 71
2.5.1 Padrões de aglomeração ocupacional no Brasil 71
2.5.2 Aglomeração ocupacional e *labor pooling* no Brasil 85
2.6 CONSIDERAÇÕES FINAIS 99

3 ESPECIALIZAÇÃO OCUPACIONAL CATEGORIZADA POR HABILIDADES: PRÊMIO SALARIAL E DECOMPOSIÇÃO SALARIAL 105
3.1 INTRODUÇÃO 105
3.2 HABILIDADES E OCUPAÇÃO NA DETERMINAÇÃO DOS SALÁRIOS URBANOS 111
3.2.1 O Modelo Ricardiano do Mercado de Trabalho 114
3.2.2 O Modelo de difusão de conhecimentos nas cidades (Transmissão de habilidades) 124
3.2.3 Evidências empíricas 126
3.2.3.1 Evidências empíricas internacionais 126
3.2.3.2 Evidências empíricas nacionais 138
3.3 ESTRATÉGIA EMPÍRICA 143

3.3.1 Especialização ocupacional.....144
3.3.2 Modelagem empírica e prêmio salarial da especialização ocupacional.....148
3.3.2.1 Estratégia de identificação.....149
3.3.2.2 Métodos de Estimação.....155
3.3.3 Decomposição salarial.....160
3.3.4 Dados.....166
3.4 ESPECIALIZAÇÃO OCUPACIONAL CATEGORIZADA POR HABILIDADES: O MERCADO DE TRABALHO FORMAL BRASILEIRO (2009-2018).....172
3.5 RESULTADOS.....190
3.5.1 Prêmio salarial da especialização ocupacional categorizada por habilidades.....190
3.5.1.1 Prêmio salarial da especialização e aglomeração por habilidades: Modelo de variáveis instrumentais.....191
3.5.1.2 Prêmio salarial da especialização e aglomeração por habilidades: Modelo de dados em painel.....202
3.5.2 Decomposição do prêmio salarial da especialização ocupacional categorizada por habilidades.....210
3.6 CONSIDERAÇÕES FINAIS.....223

4 CONCLUSÃO.....230
REFERÊNCIAS.....234
APÊNDICE A – Resultados complementares do Capítulo 2.....243
APÊNDICE B – Resultados complementares do Capítulo 3.....255
ANEXO A – Diferencial salarial no Modelo Ricardiano do Mercado de Trabalho de Acemoglu e Autor (2011).....275

1 INTRODUÇÃO

O termo aglomeração envolve três variáveis fundamentais para a teoria econômica: trabalhadores, empresas e espaço; reflete como os trabalhadores e empresas estabelecem sua localização ótima. Neste sentido, o termo aglomeração é indissociável das cidades. Duranton e Puga (2004) indicam que os retornos crescentes da aglomeração são essenciais para se compreender por que as cidades existem, enfatizam ainda que a desigualdade espacial não pode ser explicada somente considerando a heterogeneidade do espaço.

De acordo com Puga (2010), desde Smith (1776) e Marshall (1890), as vantagens da aglomeração já eram percebidas, todavia, o autor indica que, nos últimos 30 anos, os economistas da Economia Urbana tiveram bastante êxito em explicar as vantagens produtivas de aglomerações urbanas, ao evidenciarem que o agrupamento das atividades produtivas não se dá de modo aleatório ou não reflete as heterogeneidades no espaço e que salários e aluguéis diferem em centros urbanos densos. De modo geral, Duranton e Puga (2004) classificam economias de aglomeração urbana como derivadas de interações entre o mercado de trabalho, fornecedores de bens intermediários e finais e de transbordamentos de conhecimento. Apesar de considerarem que esses eventos estão intrinsecamente integrados, os autores distinguem três tipos de microfundamentos, baseados em *sharing* (compartilhamento de insumos e trabalhadores), *matching* (correspondência entre trabalhadores e firmas) e *learning* (aprendizado e absorção de ideias).

A literatura voltada para a economia urbana distingue dois tipos de fontes de economias de aglomeração (EBERTS; MCMILLEN, 1999; GOLDMAN; KLIER; WALSTRUM, 2019; WHEATON; LEWIS, 2002): economias de localização e economias de urbanização. As economias de localização surgem quando existe um agrupamento e concentração de atividades semelhantes, levando a aglomeração industrial. No que se refere as economias de urbanização, essas emergem do agrupamento ou concentração de atividades não relacionadas, possibilitando maior diversificação, o que pode justificar o surgimento das cidades dada a variedade de produtos e serviços disponíveis. Wheaton e Lewis (2002) refere-se às economias de urbanização como externalidades de Jacobs[1], onde as grandes cidades acabam por oferecer diversos serviços de apoio, e as economias de localização como externalidades de MAR

[1] Devido ao trabalho seminal de Jacobs (1968) intitulado “The Economy of Cities”.

(Marshall, Arrow e Romer), onde empresas e trabalhadores que desenvolvem as mesmas atividades podem transferir conhecimento, intensificando as taxas de crescimento dos salários e da produção.

De modo geral, os estudos de Economia Urbana se debruçam na análise das economias de localização e de urbanização mas, uma vez que o salário é uma medida de bem-estar geral, as economias de localização acabam ganhando mais atenção (GOLDMAN; KLIER; WALSTRUM, 2019; WHEATON; LEWIS, 2002). De acordo com Goldman, Klier e Walstrum (2019), as economias de localização podem ser fundamentais na explicação da concentração de ocupações, dado que grandes firmas podem optar por separar suas atividades, como por exemplo: atividades produtivas, administrativas e atividades de pesquisa e desenvolvimento (P&D). Assim, as economias de localização estariam intensificando as divisões das atividades produtivas no espaço econômico. Essa dinâmica gera uma polarização das ocupações no espaço.

Acemoglu e Autor (2011) buscam explicar essa composição do emprego no espaço não só avaliando a demanda por ocupações, mas a estrutura das indústrias. Os autores enfatizam que os setores industriais estão usando cada vez mais ocupações cognitivas e menos ocupações rotineiras. Assim, avaliam que a mudança no emprego geral em cada ocupação se dá a partir das mudanças na composição interindustrial e intraindustrial. Considerando 10 categorias ocupacionais, para os períodos de 1959 a 1979 e de 1979 a 2007, observaram que a parcela de ocupações profissionais, técnicas e gerenciais aumentou rapidamente em ambos os sexos, no entanto, no último período considerado, esse aumento teve maior explicação na decomposição pela participação do emprego dentro das indústrias do que entre as indústrias. Logo, o emprego em ocupações mais qualificadas deve-se às mudanças internas nas empresas. Essas mudanças, por sua vez, intensificam a concentração de ocupações no espaço, modificando as condições de oferta e demanda de ocupações e consequentemente alterando a estrutura de salários.

Assim, os salários deixam de ser explicados apenas pelas características dos trabalhadores, como escolaridade e experiência, e passam a depender em maior grau das condições de escolha local de firmas (demanda) e trabalhadores (oferta) condicionados às ocupações desempenhadas, isto é, os salários passam a ser determinados a partir do grau de concentração e especialização ocupacionais. Vale ainda enfatizar que a escolha ocupacional e locacional dos trabalhadores é determinada pelas habilidades dos mesmos. Maciente (2013) ressalta que no Brasil os padrões

de concentração parecem estar mais associados às externalidades de conhecimento do que às externalidades de insumo-produto, assim as habilidades dos trabalhadores desempenham um papel central para a determinação da concentração ocupacional e dos salários. O autor, ao criar uma métrica de 21 categorias de habilidades para o Brasil, indica que certas habilidades podem estar negativamente ou positivamente relacionadas a aglomeração, logo sugere que certos tipos de habilidades podem estar escassos no Brasil. Rossi-Hansberg, Sarte e Schwartzman (2019) resumem que a literatura econômica aponta para um aumento da concentração espacial, acompanhado pelo aumento do prêmio salarial urbano e o aumento do prêmio salarial das habilidades nos grandes centros. Desse modo, o estudo da concentração e especialização ocupacional deve estar condicionado às habilidades desempenhadas pelos trabalhadores, a fim de se contrapor com o atual perfil de demanda de habilidades no mercado de trabalho, dada a dinâmica econômica.

É notório que o dinamismo econômico é uma característica central de qualquer economia desenvolvida. Os agentes econômicos devem estar preparados para avaliar novos cenários e estabelecer políticas públicas para se adequar às novas conjunturas econômicas. No mercado de trabalho esta preocupação não deve ser diferente. Schwab (2016) evidencia que a economia mundial passa atualmente pela quarta revolução industrial e muita incerteza ainda há com relação a este cenário, o que provocará mudanças nas estruturas produtivas e, por sua vez, no mercado de trabalho. Para o autor, as novas tecnologias mudarão a natureza do trabalho entre indústrias e ocupações e a incerteza envolvida nesse cenário corresponde ao grau de substituição do trabalho pela automação, o tempo dessa mudança e o seu escopo (SCHWAB, 2016). Em um trabalho realizado após o início da terceira revolução industrial, Rumberger e Levin (1985) já enfatizavam as mudanças que a tecnologia provocou no processo produtivo, indicando que, no mercado de trabalho, seriam requeridos trabalhadores com maiores habilidades e lideranças e o desemprego tecnológico levaria a exclusão de atividades rotineiras e manuais. Nesse sentido, é relevante entender, no contexto da quarta revolução industrial, qual será o perfil das mudanças que devem ocorrer no mercado de trabalho brasileiro.

De acordo com dados do Banco Mundial (2016), o Brasil ocupa a quinta posição de maior país do mundo, o que pode ser explicado pela sua vasta extensão territorial e população, todavia, também é um dos países mais desiguais. Essa desigualdade advém de um passado histórico, mas persiste, mesmo com as evidências de redução da pobreza e desigualdades na última década, em especial a partir de 2013. Não obstante, no Brasil, além das políticas de

transferências governamentais ajudarem na redução das desigualdades regionais, o mercado de trabalho é um fator chave. Para a maioria dos brasileiros, o emprego formal possibilita maior prosperidade. Logo, o desenvolvimento do mercado de trabalho e a distribuição dos seus ganhos são os principais determinantes para a redução da pobreza e desigualdade no Brasil (BANCO MUNDIAL, 2016). Com esse cenário, é fundamental analisar como as economias de localização, atreladas à concentração e especialização ocupacional, podem impactar na dinâmica do mercado de trabalho brasileiro, em especial na determinação dos salários locais.

No Brasil, autores com Maciente, Rauen e Kubota (2019) afirmam que, apesar de no Brasil a quarta revolução industrial ainda ser embrionária, é fundamental analisar os desafios que o futuro do mercado de trabalho enfrentará. Os autores avaliam as mudanças nas estruturas dos requisitos de habilidades dos trabalhadores, no período de 2003 a 2017, a fim de compreender melhor a composição de habilidades do Brasil. Todavia, se faz necessária uma análise mais complexa da estrutura ocupacional do mercado de trabalho brasileiro, buscando entender a aglomeração ocupacional, suas causas e efeitos nos salários, o que pode contribuir para a formulação de políticas públicas frente ao cenário incerto que a quarta revolução industrial apresenta. Nesse sentido, esta tese busca compreender a estrutura ocupacional brasileira, ressaltando o papel das habilidades dos trabalhadores e as especificidades de cada região, a fim de avaliar a aglomeração ocupacional. Este estudo não permitirá avaliar os efeitos das novas tecnologias no mercado de trabalho brasileiro, uma vez que estas ainda não estão bem definidas e seu estágio de implantação no país ainda é inicial, mas pode contribuir para compreender como a estrutura ocupacional brasileira está organizada, permitindo a elaboração de políticas públicas mais efetivas neste cenário.

A compreensão da dinâmica do mercado de trabalho brasileiro requer entender o funcionamento das aglomerações e seus efeitos diversos regionalmente. O objetivo central desta pesquisa foi avaliar os determinantes da aglomeração ocupacional e os efeitos da aglomeração urbana e ocupacional sobre os diferenciais de salários locais. Desse modo, a tese é composta de dois ensaios que buscam analisar a relação entre aglomeração, ocupação, habilidades e salários.

No primeiro ensaio, estudou-se os padrões de aglomeração ocupacional no Brasil considerando o índice Ellison-Glaeser (1997), modificado por Gabe e Abel (2009). Posteriormente, verificou-se os principais determinantes desta aglomeração, onde atenção especial foi dada a uma nova

medida de *labor pooling*[2] ponderada pelas habilidades dos trabalhadores. Para isso, o ensaio utilizou um banco de dados em painel extraído dos microdados individuais da RAIS para o período 2009 a 2018 e como subunidade regional considerou-se os Arranjos Populacionais brasileiros. A fim de encontrar o efeito do *labor pooling* e das habilidades sobre a concentração[3] ocupacional, este trabalho estimou um conjunto de modelos de dados em painel ao nível ocupacional que permite captar a heterogeneidade não observada das ocupações.

No Brasil, apenas a pesquisa de Andrade, Gonçalves e Freguglia (2014) avalia a aglomeração ocupacional no Brasil, todavia, os autores utilizam o índice de Gini, que apresenta algumas limitações, e não utilizam nenhuma medida de *labor pooling* para explicar a determinação da aglomeração ocupacional. Essa pesquisa contribui para a literatura brasileira, pois avalia a aglomeração ocupacional pelo índice Ellison-Glaeser (1997), modificado por Gabe e Abel (2009), e utiliza uma medida de *labor pooling* que considera os choques idiossincráticos do mercado de trabalho. Como contribuição para a literatura internacional, esse ensaio utilizou uma medida de *labor pooling* que associa os choques idiossincráticos, metodologia adotada por Overman e Puga (2012), e as habilidades dos trabalhadores, metodologia utilizada por Gabe e Abel (2009), sendo possível isolar os efeitos de *sharing*, *matching* e *learning* e obter parâmetros mais robustos.

No segundo ensaio, verificou-se como a aglomeração ocupacional, determinada a partir da especialização ocupacional condicionada às habilidades dos trabalhadores[4] (cognitivas, sociais e manuais), tal qual definida por Wheaton e Lewis (2002) e Rossi-Hansberg, Sarte e Schwartzman (2019), pode impactar no diferencial salarial dos trabalhadores brasileiros, considerando a endogeneidade da escolha local e a heterogeneidade dos trabalhadores e firmas. Para tanto, considerando como subunidade regional os Arranjos Populacionais brasileiros, utilizou-se os microdados individuais da RAIS para o período 2009 a 2018. No controle da endogeneidade sobre os salários, o ensaio utilizou duas estratégias de identificação: variáveis instrumentais (variáveis instrumentais históricas e o instrumento Bartik) e estimações em painel

[2] O *labor pooling* é mensurado considerando a metodologia proposta por Overman e Puga (2010), onde se obtém a média por ocupação das diferenças entre a variação percentual do emprego ocupacional no estabelecimento e a variação percentual do emprego na ocupação.
[3] A medida de concentração ocupacional utilizada neste estudo é o índice de aglomeração Ellison-Glaeser (1997), modificado por Gabe e Abel (2009), mensurado a partir da participação do emprego em uma ocupação e Arranjo Populacional em relação ao emprego no Brasil.
[4] Utilizou-se a medida de especialização, sugerida por Rossi-Hansberg, Sarte e Schwartzman (2019), que expressa a participação dos trabalhadores de determinada habilidade e Arranjo Populacional no total de trabalhadores do respectivo Arranjo Populacional.

com o controle de efeitos fixos. Em um segundo momento, utilizando a abordagem em dois estágios proposta no estudo de Combes, Duranton e Gobillon (2008) e aplicando a metodologia proposta por Abowd, Kramarz e Margolis (1999), decompõe-se o prêmio salarial da especialização ocupacional por habilidades, avaliando os efeitos dos indivíduos, das firmas e da densidade populacional, de modo a compreender melhor a dinâmica de determinação dos salários considerando a aglomeração ocupacional.

No Brasil, apenas a pesquisa de Ehrl e Monasterio (2020) considera uma medida de aglomeração de habilidades. Todavia, os autores utilizam uma medida de aglomeração que detecta a quantidade média de trabalhadores com determinadas habilidades por município. Nesse sentido, é importante ressaltar a contribuição desse ensaio para a literatura brasileira, pois estimou-se o prêmio salarial da especialização ocupacional categorizada por habilidades, onde a medida de especialização é uma variável mais completa para avaliar os efeitos de externalidades, pois permite verificar a proporção de trabalhadores em cada categoria de habilidades dentro da unidade regional considerada. Em termos internacionais, a pesquisa contribui ao avaliar a decomposição dos salários locais devido aos efeitos brutos e líquidos (da heterogeneidade não observada de trabalhadores e firmas) da especialização de habilidades. Desse modo, foi possível avaliar a relação entre a densidade e a especialização de habilidades no mercado de trabalho, uma contribuição desta pesquisa.

É importante enfatizar que, as contribuições empíricas dos dois ensaios ajudam a compreender a dinâmica do mercado de trabalho brasileiro e, indiretamente, obter algumas evidências acerca da desigualdade regional brasileira, uma vez que a concentração ou especialização ocupacional categorizada por habilidades é um canal que explica a concentração da atividade produtiva. Monteiro Neto (2014) indica que, na dinâmica brasileira, há um processo de reconcentração territorial da atividade produtiva nas regiões Sul e Sudeste, em detrimento das regiões Norte e Nordeste. Desse modo, o autor mostra que as desigualdades inter-regionais e interestaduais podem ser impactadas pelo processo de concentração da atividade produtiva.

Esta tese está dividida em três capítulos, além desta introdução. O Capítulo 2 apresenta o primeiro ensaio intitulado “Aglomeração ocupacional e seus determinantes: uma análise do *labor pooling* e das habilidades dos trabalhadores”. O Capítulo 3 apresenta o segundo ensaio intitulado “Especialização ocupacional categorizada por habilidades: prêmio salarial e decomposição salarial”. Por fim, o Capítulo 4 apresenta as conclusões da tese.

2 AGLOMERAÇÃO OCUPACIONAL E SEUS DETERMINANTES: UMA ANÁLISE DO *LABOR POOLING* E DAS HABILIDADES DOS TRABALHADORES

2.1 INTRODUÇÃO

Marshall (1920) enfatiza os efeitos que a aglomeração causa na economia a partir do *labor pooling, intermediate inputs* e *knowledge externalities*. As contribuições do autor no que se refere à teoria da localização foram objeto de estudo de duas vertentes: a organização industrial e os estudos sobre os distritos industriais. No que se refere ao primeiro grupo, Williamson (1975) fez uma definição precisa de aglomeração produtiva, considerando essa como um espaço territorial onde os agentes econômicos buscam interagir e cooperar entre si, a fim de reduzir os custos de transação (VALE; CASTRO, 2010). Fujita, Krugman e Venables (1999) definem aglomeração como um "agrupamento de atividades econômicas, criado e sustentado por uma lógica circular", indicando ainda que essa pode ocorrer em vários níveis, desde locais até mundiais (FUJITA; KRUGMAN; VENABLES, 1999, p. 1, "tradução nossa").

Puga (2010) salienta que a concentração se explica inicialmente pelo caráter irregular do espaço, pois o mesmo não é uniforme e apresenta características geográficas distintas (como recursos naturais). Todavia, o autor dá atenção especial às economias de escala externas para explicar a aglomeração, as quais permitem que os estabelecimentos consigam produzir com mais eficiência e incentivam a formação de centros urbanos maiores e mais densos. Puga (2010) indica que a maneira mais eficiente de medir economias de aglomeração é através da matriz insumo-produto no espaço considerado, sendo esse, então, um fator que explicaria a aglomeração.

É vasta a literatura sobre aglomeração e indústria (ELLISON; GLAESER, 1997; MAUREL; SÉDILLOT, 1999; DURANTON; OVERMAN, 2005; ROSENTHAL; STRANGE, 2001; RESENDE; WYLLIE, 2004; ELLISON; GLAESER; KERR, 2010; OVERMAN; PUGA, 2010; BURKI; KHAN, 2012), mas no que se refere à aglomeração ocupacional encontram-se na literatura mundial apenas os trabalhos de Gabe e Abel (2009) e Goldman, Klier e Walstrum (2019). No Brasil, esta limitação é ainda maior. A pesquisa de Andrade, Gonçalves e Freguglia (2014) avalia a aglomeração ocupacional, porém não faz qualquer referência às habilidades dos trabalhadores que permitem mensurar alguns dos efeitos gerais da aglomeração, como *sharing, matching e learning*.

Gabe e Abel (2009) consideram que os benefícios da aglomeração são mais intensos para ocupação do que para a indústria, uma vez que a aglomeração permite que os indivíduos compartilhem ideias e conhecimentos. Neste sentido, a proximidade entre trabalhadores que desenvolvem a mesma função traz mais benefícios do que a proximidade de trabalhadores na mesma indústria.

A explicação de Puga (2010) sobre a concentração se aplica mais facilmente à aglomeração industrial, pois seu foco está sobre as interdependências tecnológicas de entrada e saída de insumos e produtos. Goldman, Klier e Walstrum (2019) buscam explicar a concentração, enfatizando os motivos que levam a aglomeração ocupacional. De acordo com os autores, existem duas razões para a concentração ocupacional. Primeiro, as grandes empresas podem achar vantajoso separar suas atividades produtivas em estabelecimentos diferentes, comumente podem ser separadas as atividades produtivas, administrativas e de Pesquisa e Desenvolvimento (P&D), o que afetaria a localização ótima dos trabalhadores por ocupação desempenhada. Além disso, os autores indicam que alguns estabelecimentos são naturalmente especializados em determinadas ocupações. Assim, mesmo que a empresa opte por não separar suas atividades, sua natureza intrínseca já traz uma tendência para a aglomeração ocupacional em determinadas regiões.

Goldman, Klier e Walstrum (2019) indicam, ainda, a tendência crescente para o desenvolvimento de atividades inovadoras nos grandes centros. Neste sentido, seria esperado que grupos de ocupação, mesmo que de diferentes indústrias, se concentrem para aproveitar os benefícios dos *spillovers* de conhecimento obtendo maior produtividade. No Brasil, existe um grande dinamismo do emprego ocupacional. De acordo com dados da RAIS, apesar da diminuição de 3,22% no nível geral de emprego formal entre 2009 e 2018, em algumas ocupações houve um crescimento expressivo no nível de emprego, além do surgimento de postos de trabalho em novas ocupações. Ao longo dos anos, novas ocupações, não observadas em 2009, foram contabilizadas em 2018: um total de 181 ocupações.

Considerando as ocupações que, entre 2009 e 2018, apresentaram crescimento mais expressivo (superior a 100%), destaca-se: “Analista de planejamento e orçamento” (1.243,96%), “Pizzaiolo” (1.262,69%), “Petrógrafo” (1.400%), “Monitor de dependente químico” (1.417,14%), “Classificador de grãos” (1.607,52%), “Biomédico” (2.087,06%), “Intérprete de língua de sinais” (2.360%), “Sushiman” (2.754,05%), “Especialista de políticas públicas e

gestão governamental" (4.100%) e "Operador de máquina de preparação de matéria prima para produção de cigarros" (4.750%).

É importante investigar se esse crescimento do emprego ocupacional estaria direcionado espacialmente, isto é: existe um maior crescimento em algumas regiões e/ou mudanças na distribuição regional dessas ocupações? Essa análise permite tecer algumas evidências iniciais da concentração ocupacional no Brasil. No que se refere às 10 ocupações que tiveram um maior crescimento, houve algumas mudanças na participação regional do emprego ocupacional. Considerando a ocupação "Especialista de políticas públicas e gestão governamental", verifica-se que, em 2009, 100% dos seus trabalhadores estavam empregados na macrorregião Sudeste. Em 2018, a maioria desses trabalhadores estavam empregados na região Centro-Oeste (52,38%) e a região Sudeste contava com apenas 23,81% do total desses trabalhadores. Convém ainda mencionar que, em 2018, somente a região Sul não apresentou trabalhadores com esse perfil ocupacional. Destaca-se ainda as ocupações "Classificador de grãos" e "Biomédico", cuja participação no emprego na região Sudeste diminuiu, o que provocou um aumento na participação das regiões Sul, Nordeste e Norte.

Esses fatos não são suficientes para estabelecer que a concentração ocupacional diminuiu. Como um exemplo, em outro extremo, para a ocupação de "Petrógrafo", em 2009, 100% dos seus trabalhadores estavam empregados na macrorregião Sudeste, e, em 2018, esse padrão se manteve. Além disso, para as ocupações "Intérprete de língua de sinais" e "Operador de máquina de preparação de matéria prima para produção de cigarros", cujas maiores participações no nível de emprego em 2009 eram para a região Sudeste, observou-se que, em 2018, essa participação aumentou ainda mais, passando para 42,04% e 26,80%, respectivamente, em detrimento das demais regiões.

Verifica-se que a dinâmica do emprego ocupacional é bastante complexa e esses fatos podem interferir na concentração ocupacional brasileira. Todavia, não é possível apenas com essas informações inferir se a concentração ocupacional no Brasil aumentou ou diminuiu no período considerado. Nesse sentido, um maior rigor metodológico é necessário. Essa pesquisa faz uso do índice de aglomeração ocupacional de Ellison-Glaeser (1997), adaptado por Gabe e Abel (2009), para avaliar o padrão da concentração ocupacional no Brasil. A complexidade da dinâmica do emprego ocupacional deve-se, ainda, ao problema da causalidade reversa: ao mesmo tempo em que a localização das firmas pode determinar a escolha de localização dos

trabalhadores, impactando diretamente na concentração ocupacional, as firmas tentem a se localizar em grandes centros onde o mercado de trabalho apresenta uma oferta de habilidades mais intensa.

Deste modo, esta pesquisa estudou os padrões de aglomeração ocupacional de modo a conhecer as vantagens econômicas que esta tendência pode propiciar na atividade econômica. A análise da aglomeração foi determinada a partir da ocupação que o trabalhador desempenha, não da indústria em que está inserido. Dada a dificuldade em mensurar e decompor os efeitos de *matching*, *sharing* e *learnig* (COMBES; GOBILLON, 2015), conhecida como "equivalência marshalliana", este ensaio avaliou o efeito de *sharing* a partir de variáveis *proxy* exógenas que são capazes de representar o agrupamento ocupacional de trabalhadores.

A variável *labor pooling* definida por Overman e Puga (2010), com base no modelo teórico de Krugman (1991), permite isolar os efeitos os efeitos de produtividade do agrupamento de trabalhadores e contornar o problema da causalidade reversa, uma vez que considera os choques heterogêneos e idiossincráticos do mercado de trabalho. Além disso, são considerados os efeitos de *sharing* (*labor pooling*) através do perfil de habilidades dos trabalhadores, avaliando não só o carácter quantitativo do agrupamento de trabalhadores, mas o seu carácter qualitativo. Portanto, o objetivo deste ensaio foi analisar a contribuição do *labor pooling* e das habilidades para a aglomeração ocupacional.

Para tanto, calculou-se o índice Ellison-Glaeser (1997), todavia, modificado por Gabe e Abel (2009) para a análise ocupacional. Com o uso da base de dados da Relação Anual de Informações Sociais - RAIS foram calculados índices para as ocupações delimitadas pela Classificação Brasileira de Ocupações (CBO-2002), obtidos em nível de Arranjos Populacionais, de acordo com a classificação do Instituto Brasileiro de Geografia e Estatística (IBGE). Com isso, torna-se possível verificar as peculiaridades da aglomeração ocupacional e prognosticar cenários, dada a nova conjuntura da quarta revolução tecnológica que, conforme já mencionado por Maciente, Rauen e Kubota (2019), terá impactos significativos na estrutura ocupacional de emprego no Brasil.

Em um segundo momento, avaliou-se os determinantes da aglomeração ocupacional enfatizando o papel do *labor pooling* e das habilidades dos trabalhadores, através de um painel de dados ao nível ocupacional, o que permite captar a heterogeneidade não observada das

ocupações. Isto é, verificou-se como o *labor pooling* e a distribuição de habilidades podem determinar a concentração ocupacional e qual tipo de habilidade permite maior aglomeração. Nesse sentido, a pesquisa busca responder a perguntas como: o *labor pooling*, que propicia a concentração ocupacional, é mais forte quando se considera as habilidades manuais, sociais ou cognitivas?

Dessa forma, explica-se a aglomeração ocupacional com uma medida alternativa de *labor pooling*, com a influência das habilidades requeridas por ocupação e não apenas da quantidade de trabalhadores em uma ocupação. A medida sugerida por Overman e Puga (2010) para o *labor pooling* é a diferença entre a variação percentual do emprego no estabelecimento e a variação percentual do emprego na indústria. Todavia, a ideia de *labor pooling* expressa por Marshall (1920) é a de que a firma tende a se aglomerar especialmente nos locais com muitos trabalhadores com habilidades necessárias para a sua produção. A medida proposta por Overman e Puga (2010), apesar de ser uma variável exógena que contorna o problema da causalidade reversa, não traz nenhuma indicação sobre o perfil destes trabalhadores aglomerados, logo não se pode estabelecer, por essa medida, uma representação das habilidades dos indivíduos.

Alguns autores, como Rosenthal e Strange (2008) e Ellison e Glaeser (1999), ainda consideram que a concentração de trabalhadores especializados pode ser mensurada a partir da porcentagem de mão de obra com ensino superior. Todavia, essa estratégia não capta o *labor pooling* de forma direta e não serve como *proxy* para habilidades. No que se refere ao *labor pooling*, a variável pode estar endogenamente determinada, não sendo possível isolar os efeitos de *sharing*, *matching* e *learning*. Além disso, sabe-se que habilidade é diferente de escolaridade, conforme explanado por Heckman, Stixrudb e Urzuau (2006) e Maciente (2013).

Desse modo, aplicou-se uma medida mais robusta de *labor pooling* que permita criar uma variável exógena de choques idiossincráticos no mercado de trabalho e captar a diferença entre escolaridade e habilidades, a partir da base de dados de Maciente (2013) que estabelece categorias de habilidades por ocupação, definindo assim um conjunto de conhecimentos especializados dos trabalhadores. Incluiu-se, ainda, na modelagem empírica de determinação da concentração ocupacional desse ensaio, o percentual de trabalhadores com nível superior de escolaridade no subgrupo ocupacional, visando enfatizar a diferença entre habilidade e escolaridade, eliminando possíveis endogeneidades. Além disso, outras variáveis de controle

baseadas na literatura foram consideradas. Maciente (2013) usou as medidas de habilidade para explicar a aglomeração em termos de setores (índice de Ellison-Glaeser), mas espera-se que o impacto das habilidades na aglomeração ocupacional seja mais forte, conforme indicado por Gabe e Abel (2009), que enfatizam que os efeitos de externalidades do conhecimento são provavelmente mais fortes entre ocupações do que entre setores.

Este ensaio está dividido em cinco seções, além desta introdução. Na seção seguinte, é apresentada uma estrutura teórica que permite compreender a aglomeração, ressaltando os fatores que impactam na concentração ocupacional, em especial o modelo de *labor pooling* de Krugman (1991). Além disso, evidências empíricas são apontadas na tentativa de conhecer os estudos existentes na literatura acerca da concentração. Na seção 3 apresenta-se a estratégia empírica desenvolvida para analisar a concentração ocupacional, bem como estabelece a relação com os seus determinantes. Além disso, é feita uma apresentação da estratégia de identificação do modelo empírico e dos dados utilizados. A seção 4 apresenta uma análise do mercado de trabalho formal brasileiro para o período de 2009-2018, considerando o contexto da estrutura ocupacional brasileira e o *labor pooling* ocupacional, categorizados por habilidades. Os resultados e discussões são expostos na seção 5, que é subdividida em Padrões de aglomeração ocupacional no Brasil e Determinantes da aglomeração ocupacional no Brasil. Por fim, tem-se as considerações finais dessa pesquisa.

2.2 ECONOMIAS DE AGLOMERAÇÃO E DISTRIBUIÇÃO OCUPACIONAL ESPACIAL

A Economia Urbana e a Economia do Trabalho formam o aparato teórico necessário ao desenvolvimento deste trabalho. As duas áreas indicadas encontram-se interligadas a partir do conceito de economia de aglomeração e seus efeitos no mercado de trabalho.

A preocupação central da Geografia Econômica é estabelecer critérios de localização ótima de indústrias e indivíduos considerando os custos de transportes indissociáveis em contextos urbanos. A Geografia Econômica busca explicar a localização das atividades econômicas, indicando o sucesso ou o fracasso destas atividades em determinadas regiões (CRUZ *et al.*, 2011). Esse sucesso ou fracasso determina se a aglomeração de pessoas e indústrias vai se concretizar. Segundo Fujita, Krugman e Venables (1999) esta área da teoria econômica foi por muito tempo negligenciada, entretanto, diversos trabalhos surgiram na tentativa de aprimorar os conhecimentos sobre a aglomeração espacial (KRUGMAN, 1979; KRUGMAN, 1991;

KRUGMAN; VENABLES, 1995; DIXIT; STIGLIZ, 1997; FUJITA; THISSE, 2002). Um fato que merece ainda mais atenção no estudo sobre aglomeração é a desigualdade inerente a este processo econômico (CRUZ *et al.*, 2011). Diante disto, uma atenção especial deve ser dada ao Brasil, uma vez que este é um país historicamente desigual em termos produtivos e de rendimentos, principalmente no que se refere ao mercado de trabalho. Nesse sentido, diferenciais salariais podem perpetuar ou intensificar essa conjuntura desigual.

A Economia Urbana teve grande avanço a partir de Krugman (1991), que se destacou ao inserir o espaço em na teoria econômica. Em geral, o espaço não era considerado como uma variável econômica relevante, dado a supremacia da Escola Neoclássica que tinha como hipótese principal a concorrência perfeita. A introdução do espaço na economia, por sua vez, aponta a necessidade de incorporar estruturas de concorrência monopolística, estabelecendo a ideia de um oligopólio espacial, permitindo incrementos crescentes de escala. Além disso, os condicionantes matemáticos de então não permitiam análises mais apuradas do espaço. Krugman e Venables (1995) avançaram na análise tentando verificar como a globalização, impactando diretamente nos custos de transporte, mudaria a conjuntura da aglomeração industrial, ao modificar as condições de operação no comércio entre nações. Dixit e Stigliz (1997) incorporaram o espaço num modelo estruturado de equilíbrio geral, possibilitando avanços no campo da economia espacial e urbana, avaliando se os custos de transporte em localizações múltiplas e com concorrência monopolística seriam mais adequados à realidade econômica.

Todos estes autores partiram do conceito basilar de “Economias Externas” elaborado por Marshall (1920). Marshall (1920) diferencia os conceitos de economias internas e de economias externas, sendo este último termo criado por ele. Enquanto a divisão do trabalho, os ajustes no processo produtivo e o gerenciamento de atividades permitiam ganhos de escala nas empresas, o autor observou que condicionantes sistêmicos fora do âmbito da empresa possibilitavam impactos diretos e indiretos na lucratividade das empresas, o que ele denominou de economias externas, as quais dependem do desenvolvimento geral das indústrias em determinada localidade. Estes condicionantes externos poderiam ser representados pela melhor comunicação entre as empresas e entres os trabalhadores, além das facilidades de crédito. Desse modo, as indústrias tenderiam a se localizarem estrategicamente no espaço próximas às fontes de matérias primas, do mercado consumidor e do mercado de trabalho potencial.

Segundo Marshall (1920) a concentração geográfica seria benéfica por três razões: (i) As ligações de insumo produto ficariam mais fáceis, ou seja, a compra e venda de bens e matérias primas seria facilitada; (ii) A concentração das firmas levaria a uma concentração de trabalhadores, haveria um agrupamento de trabalhadores heterogêneos que poderiam atender às empresas lá instaladas facilitando ajustamentos no mercado de trabalho; (iii) A aglomeração de indústrias e trabalhadores facilitaria a disseminação de informações. A partir de então, o tripé basilar de Marshall (1920), *intermediate inputs*, *labor pooling* e *knowledge externalities*, passou a fundamentar os estudos urbanos, especialmente aqueles relacionados às aglomerações.

Os estudos teóricos sobre os efeitos das economias de aglomeração, todavia, se intensificaram com as contribuições teóricas de Perroux (1955) e Hirschman (1977). Tratando centro aglomerativo e polo de crescimento como sinônimos, os autores identificam três formas em que a proximidade das empresas, caracterizadas em indústrias motrizes e secundárias, poderiam provocar o crescimento econômico: efeitos técnicos, efeitos de rendas e efeitos geográficos ou psicológicos. Os efeitos técnicos, sendo o principal deles, estariam atrelados às interdependências tecnológicas existentes entre as empresas, que permitiriam a compra e venda de insumo com mais facilidade e a preços mais competitivos. Os efeitos de renda referem-se aos ganhos refletidos no mercado de trabalho que poderiam intensificar a demanda das próprias indústrias ali estabelecidas. Os efeitos psicológicos ou geográficos criariam uma atmosfera favorável, aumentando o processo de imitação das empresas, tanto dos métodos produtivos como tecnológicos, o que incentivaria a inserção de mais empresas na localidade (PERROUX, 1955). Hirschman (1977) ainda indica que, a aglomeração de indústria provoca repercussões econômicas diretas entre as regiões, apresentando tanto efeitos positivos como negativos. Todavia, o autor acredita que os efeitos positivos superariam os efeitos negativos. Dessa forma, o autor é favorável à política de investimentos públicos nos polos de crescimento ou centros aglomerativos, uma vez que, considera a forte presença dos efeitos de transbordamentos para outros centros ou regiões, expandindo a indústria e levando a um maior crescimento econômico.

No que se refere à relação entre aglomeração e investimentos públicos, Combes, Duranton e Gobillon (2011) indicam que identificar a aglomeração e seus efeitos permite iniciativas públicas como a construção de clusters, intensificando o polo aglomerativo com a entrada de novas grandes indústrias, impactando novamente na aglomeração e criando um ciclo virtuoso de crescimento. Neste sentido, compreender melhor a dinâmica aglomerativa, objetivo deste trabalho, é de suma importância no direcionamento de investimentos públicos.

Entretanto, a aglomeração não traz apenas benefícios. Combes, Duranton e Gobillon (2011) enfatizam a dicotomia entre economias de aglomeração e custos urbanos. Os autores indicam que dada a existência de economias de aglomeração e custos urbanos, estes custos dominam em grandes cidades e as economias aglomerativas nas pequenas, isto é, quanto maior a cidade, maiores serão os seus custos e, dada a livre mobilidade de indivíduos entre as cidades e a conjuntura econômica destas, esta dinâmica determinará o tamanho das cidades.

Com uma abordagem similar, Henderson (1980) estabelece que as economias externas podem ainda determinar o tamanho das cidades. O autor tinha como preocupação básica estudar o tamanho ótimo das cidades, considerando economias e deseconomias externas, deseconomias essas associadas aos custos de transporte. Henderson (1980) considera a existência de assimetrias entre economias e deseconomias externas, onde os custos de transportes impactariam no tamanho das cidades. As cidades deveriam conter indústrias onde os efeitos mútuos de transbordamento fossem mais intensos, pois, assim, seria vantajoso, em termos econômicos, a proximidade de indústrias afins, gerando uma especialização. Além disso, o autor enfatiza que os efeitos das economias externas podem ser pequenos. Diante desta conjuntura, as deseconomias externas poderiam limitar o crescimento das cidades.

Mesmo tendo isso em vista, a densidade confere um nível de produtividade acima da média tanto para os trabalhadores como para as empresas nas grandes cidades, possibilitando, ainda, emergir inovações organizacionais e tecnológicas (PUGA, 2010).

2.2.1 Externalidades de aglomeração

Baseado no tripé basilar de Marshall (1920), Duranton e Puga (2004) e Puga (2010) explicam os efeitos e mecanismos propulsores da aglomeração a partir de uma nova roupagem, utilizando os conceitos de *sharing*, *matching* e *learning*: (i) A aglomeração permite o compartilhamento de infraestrutura, fornecedores de insumos e trabalhadores qualificados. O *sharing*, assim, possibilita retornos crescentes de escala; (ii) Uma melhor correspondência no mercado de trabalho pode ser verificada em centros aglomerativos, onde um equilíbrio mais eficiente entre empresas e trabalhadores pode ser obtido. Esse mecanismo pode gerar um equilíbrio mais eficiente no mercado de insumos, ou seja, o *matching* melhora as condições gerais de eficiência na economia dadas as maiores possibilidades de tentativa e erro entre os fatores; (iii) A aglomeração possibilita, ainda, melhorar o aprendizado a partir das externalidades positivas de

conhecimento entre trabalhadores e empresas, sendo o *learning* um importante impulsionador de desenvolvimento tecnológico.

De acordo com Puga (2010), o *sharing* pode ser analisado com mais cautela tanto no que se refere aos ganhos de especialização individual como o *labor pooling.* Os ganhos individuais são aqueles que possibilitam aumentar a produtividade do trabalhador, uma vez que este está inserido em uma cidade maior, pelo contato direto com um maior número de trabalhadores, tal como previsto por Smith (1976). Entretanto, a especialização, apontada por Puga (2010), acontece de forma mais específica, não apenas restrita aos trabalhadores em tarefas diferentes (especialização ou ganhos de escala geral), mas a partir do contato direto com trabalhadores exercendo tarefas próximas. Logo, a especialização, em mercados de trabalho maiores, ganha uma ótica voltada para as ocupações e os profissionais que realizam tarefas similares.

No que se refere ao *labor pooling*, a evidência favorável a aglomeração foi apontada desde Marshall (1920), ao enfatizar que indústrias em mercados densos possuem uma maior e constante oferta de trabalhadores habilidosos. Krugman (1991) faz uma análise de como o *labor pooling* pode impactar na lucratividade da empresa, a partir de choques idiossincráticos. Esses choques são representados por qualquer perturbação econômica que modifique as condições no mercado de trabalho e que venham a alterar os salários locais. As empresas devem responder a esses choques e sua postura no centro aglomerativo pode impactar diretamente nos seus lucros, podendo alterar a matriz de localização ótima das firmas, aproximando ou dispersando as indústrias. Quanto mais isolada a empresa e quanto menores suas interdependências tecnológicas no centro aglomerativo menores os impactos, em nível de produção e emprego, devidos aos choques idiossincráticos que podem ser tanto positivos quanto negativos. Logo, considerando o contexto dinâmico da economia e as diferenças de mudanças e composições de empregos entre as empresas, torna-se vantajoso para as mesmas uma localização estratégica em regiões que tenham um agrupamento de trabalhadores com as habilidades das quais necessitam (PUGA, 2010).

O *matching* é outro efeito importante da aglomeração na interação firmas e trabalhadores, e como o *sharing* e *learning* encontra-se intrinsecamente relacionado com as habilidades dos trabalhadores. Essas habilidades são demandadas pelas firmas e, considerando a heterogeneidade das firmas e dos trabalhadores, em mercados mais densos é mais fácil de obter

uma melhor correspondência entre firmas e trabalhadores, reduzindo os custos de incompatibilidades, como especializações e treinamentos (PUGA, 2010).

Na relação entre aglomeração e trabalhadores, destaque-se o *learning.* As cidades densamente povoadas transmitem conhecimento entre os trabalhadores de forma natural e intencional, a partir de cursos de especialização de conhecimentos para determinadas ocupações. Dessa forma, as cidades facilitam e criam novos conhecimentos, através de externalidades positivas de conhecimento, promovendo o desenvolvimento das habilidades cognitivas, sociais e manuais dos trabalhadores (PUGA, 2010).

De modo geral, Fujita, Krugman e Venables (1999), Combes e Duranton (2015) e Puga (2010) atentam para o fato de que os efeitos de *sharing*, *matching* e *learning* conjuntamente são difíceis de modelar de maneira formal. Puga (2010) indica que, estudos sobre o *learning* são pouco abordados na literatura teórica e empírica. Nesse sentido, o autor destaca que já existem modelos tratáveis de aglomeração para o *sharing* e o *matching*, mas ainda é insuficiente o conhecimento sobre o *learning* nas cidades (PUGA, 2010).

Dentro desta perspectiva, esta pesquisa é uma tentativa de aprofundar empiricamente essa lacuna, avaliando o efeito de *sharing* (*labor pooling*) através de variáveis que captem as habilidades dos trabalhadores. A ênfase principal foi dada ao *labor pooling* e optou-se por adotar uma ótica ocupacional e não industrial, pois as externalidades de conhecimento e de habilidades se estabelecem mais fortemente quando se considera a aglomeração ocupacional, conforme enfatizado por Gabe e Abel (2009), e a possibilidade de *sharing*, *learning* e *matching* é maior nos grandes centros urbanos. Conforme verificado, esses três microfundamentos tem um elo em comum: as habilidades dos trabalhadores. Dentro deste cenário, a análise da aglomeração ocupacional atrelada às habilidades dos trabalhadores é uma importante contribuição para a literatura econômica brasileira.

Empiricamente, a medida de *labor pooling* de Overman e Puga (2010) que, originalmente, capta a média da diferença entre taxa de crescimento dos empregos nas empresas e a taxa de crescimento dos empregos no setor, explicaria a maior concentração industrial. A fim de se compreender melhor o agrupamento de habilidades no mercado de trabalho como fonte de aglomeração e aumento da produtividade, os autores se basearam nos preceitos teóricos de Krugman (1991), cujo modelo será apresentado abaixo.

2.2.2 *Labor Pooling*: Modelo Teórico

Krugman (1991) buscou explicar os microfundamentos econômicos do agrupamento dos trabalhadores (*labor pooling*) nos centros urbanos como fonte de economias de aglomeração que geram aumento de produtividade nas firmas, possibilitando que essas possam aproveitar as vantagens econômicas da concentração. Intuitivamente, percebe-se que os estabelecimentos que apresentam mudanças frequentes nos seus níveis de emprego, em relação ao demais, encontram vantagem na localização próxima ao centro de empregos, onde há uma maior variabilidade de trabalhadores no que se refere às suas habilidades. Desse modo, Krugman (1991) estabelece que, considerando determinados setores, haverá maior concentração espacial quando os estabelecimentos apresentam maior volatilidade idiossincrática.

O modelo empírico de Krugman (1991) considera um conjunto de setores ($s = 1, \dots, S$) e cada setor é composto por um número discreto de estabelecimentos indexados por *i* ($i = 1, \dots, I$) que tem como fator de produção principal os trabalhadores, com habilidades específicas. Os estabelecimentos ao determinarem sua localização estão sujeitos à choques de produtividade ε_i. Estatisticamente, esses choques não apresentam correlação entre os estabelecimentos e estão distribuídos de forma idêntica com média zero e variância σ_s, ou seja, assume-se $\varepsilon_i \sim (0, \sigma^2)$. Dado que a sobrevivência dos estabelecimentos e, indiretamente, sua localização, depende dos lucros econômicos auferidos pelas mesmas, o modelo de Krugman (1991) tenta prever como o *labor pooling* no espaço impacta nos lucros dos estabelecimentos. Nessas configurações, a função de produção de uma empresa *i* é definida como:

$$y_i = (\beta + \varepsilon_i) l_i - \frac{1}{2} \psi l_i^2 \tag{2.1}$$

Em que, l_i representa o total de trabalhadores empregados no estabelecimento, β representa a produtividade marginal do trabalho, ε_i capta os choques de produtividade da empresa e ψ representa os retornos da produção. A função de custo da empresa é representada pelo montante de salários pagos aos trabalhadores ($w l_i$). Os salários individuais de cada estabelecimento podem ser definidos como:

$$w = \beta + \varepsilon_i - \psi l_i \tag{2.2}$$

Considerando o equilíbrio no mercado de trabalho, após os choques de produtividade, a demanda por trabalho do estabelecimento é derivada de (2.2):

$$l_i = \frac{\beta + \varepsilon_i - w}{\psi} \tag{2.3}$$

Desse modo, a força de trabalho de uma determinada localidade, ou as horas de trabalho efetivas, é obtida a partir do somatório dos trabalhadores de cada estabelecimento. Logo:

$$L_i = \sum_{i=1}^{n} l_i = \sum_{i=1}^{n} \frac{\beta + \varepsilon_i - w}{\psi} = \frac{1}{\psi}(n\beta + \sum_{i=1}^{n} \varepsilon_i + nw) \tag{2.4}$$

Pode-se, então, obter o salário médio da localidade, rearranjando os termos da equação (2.4), isolando w e dividindo pelo número de estabelecimentos dessa localidade (n):

$$w = \beta - \psi \frac{L}{n} + \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i \tag{2.5}$$

Aplicando a esperança matemática na expressão (2.5), considerando que $\varepsilon_i \sim (0, \sigma^2)$, conforme já mencionado, o salário médio esperado é definido pela seguinte expressão:

$$E(w) = \beta - \psi \frac{L}{n} \tag{2.6}$$

Verifica-se que, o salário médio esperado depende negativamente do número de trabalhadores e positivamente do número de estabelecimentos, isto é, quanto maior a relação entre trabalhadores e estabelecimentos maior é a pressão da oferta no mercado de trabalho, diminuindo os salários. Além disso, apresenta retornos de produção (ψ) decrescentes, uma vez que o aumento do número dos trabalhadores impacta negativamente na produtividade marginal decrescente dos mesmos. Observa-se ainda que, a variância do salário e a covariância do salário com os choques de produtividade dependem do número de estabelecimentos, respectivamente, tem-se:

$$var(w) = \frac{\sigma^2}{n} \tag{2.7}$$

$$cov(w, \varepsilon_i) = \frac{\sigma^2}{n} \tag{2.8}$$

Por fim, a função de lucro pode ser definida como:

$$\pi_i = y_i - wl_i \tag{2.9}$$

Considerando as equações (2.1) e (2.2), tem-se o lucro como uma função convexa dos choques idiossincráticos:

$$\pi_i = (\beta + \varepsilon_i)l_i - \frac{1}{2}\psi l_i^2 - (\beta + \varepsilon_i - \psi l_i)\, l_i$$

$$\pi_i = \frac{\beta - w - \varepsilon_i}{2\psi} \tag{2.10}$$

O lucro esperado é representado por:

$$E(\pi_i) = \frac{[\beta - E(w)]^2 + var(\varepsilon_i - w)}{2\psi} \tag{2.11}$$

Ao substituir (2.6) na equação de lucro esperado, obtém-se:

$$E(\pi_i) = \frac{\psi}{2}\left(\frac{L}{n}\right)^2 + \frac{var(\varepsilon_i) + var(w) - 2cov(w, \varepsilon_i)}{2\psi} \tag{2.12}$$

Dado o primeiro termo da equação acima, na ausência de choques de produtividade dos estabelecimentos, verifica-se que os lucros dos estabelecimentos dependem da quantidade de trabalhadores e estabelecimentos, ou da relação entre essas variáveis que capta a interação entre quantidade ofertada e demandada de trabalhadores. Assim, quanto maior a relação (L/N) maior o lucro, via diminuição nos salários, devido a uma maior oferta de trabalho. O segundo termo capta o efeito do *labor pooling*. Observa-se que os lucros dependem positivamente da variância dos choques de produtividade idiossincráticos dos estabelecimentos e da variância do salário, dado a convexidade dos lucros aos choques idiossincráticos. Destaca-se ainda que os lucros são negativamente correlacionados com a covariância do salário com os choques de produtividade.

Overman e Puga (2010) enfatizam que o termo de covariância entre os salários e os choques $(cov(w, \varepsilon_i))$ traz a intuição chave do modelo de Krugman (1991) na determinação microeconômica do papel do *labor pooling* como fonte de externalidade positiva da aglomeração: se o salário é maior quando um estabelecimento deseja aumentar a produção e o emprego em resposta a um choque idiossincrático positivo, ou se o salário é menor quando um

estabelecimento deseja diminuir a produção e o emprego em resposta a um choque idiossincrático negativo, os lucros tornam-se menos convexos. Logo, o lucro esperado diminui. Isso acontece se a covariância entre salário e choques tiver uma relação positiva como o lucro esperado. Na equação (2.12), como os lucros são negativamente correlacionados com a covariância do salário com os choques de produtividade, essa situação não acontece, e a empresa consegue ajustar o nível de produção sem pressões sobre os lucros, pois os choques de produtividade são minimizados em vez de refletidos nos salários locais. Portanto, os estabelecimentos acabam sendo favorecidos pelo compartilhamento do mercado de trabalho frente à choques de produtividade idiossincráticos.

A fim de avaliar a heterogeneidade dos choques idiossincráticos e simplificar a equação (2.12), considerando $\varepsilon_i \sim (0, \sigma^2)$ e substituindo (2.7) e (2.8), verifica-se que quanto maiores os choques idiossincráticos maior o impacto do *labor pooling* sobre o lucro médio dos estabelecimentos em um mesmo setor e localização:

$$E(\pi) = \frac{\psi}{2}\left(\frac{L}{n}\right)^2 + \left(1 - \frac{1}{n}\right)\frac{\sigma^2}{2\psi} \tag{2.13}$$

Nesta situação, verifica-se que as externalidades positivas da concentração são intensificadas quanto maior for a variância dos choques idiossincráticos, estimulando a concentração industrial. Isto é, a probabilidade da concentração setorial é maior quanto mais heterogêneos são os choques de produtividade.

Considerando os microfundamentos econômicos apresentados no modelo de Krugman (1991), Overman e Puga (2010) estabeleceram uma medida empírica que permita isolar e captar o efeito do *labor pooling*. A variável exogenamente determinada pelos autores é definida como: a diferença entre a variação percentual no emprego do estabelecimento e a variação percentual no emprego do setor (em valor absoluto). Desse modo, tem-se uma medida direta dos choques idiossincráticos das alterações nos níveis de emprego de qualquer estabelecimento. Percebe-se que essa medida de *labor pooling* permite captar a intuição chave do modelo de Krugman (1991): tem-se um valor alto quando os estabelecimentos estão expandindo seu nível de emprego para se ajustar a choques positivos, enquanto o restante da economia está contraindo o emprego, impedindo pressões sobre os salários, ou vice-versa. Esta estratégia empírica é adotada neste ensaio utilizando os microdados longitudinais da RAIS, de modo a captar os

choques idiossincráticos das alterações nos níveis de emprego dos estabelecimentos em nível ocupacional, dado o objetivo desta pesquisa. Isto é, obtém-se o *labor pooling* ocupacional.

2.2.3 Evidências Empíricas

Esta pesquisa busca avaliar o *sharing* (*labor pooling*), a partir da aglomeração ocupacional. Do ponto de vista empírico, vários trabalhos na literatura buscam explicar a aglomeração industrial. Porém, é escassa, na literatura econômica, a utilização de um arcabouço teórico e empírico que busque avaliar e explicar a aglomeração ocupacional, em especial considerando as habilidades dos indivíduos, pelos efeitos de *sharing* (*labor pooling*), *learning* e *matching.*

2.2.3.1 Aglomeração industrial

Ellison e Glaeser (1997) enfatizaram a importância de aprofundar empiricamente a análise da aglomeração industrial e fizeram uma contribuição empírica fundamental, ao criar um índice capaz de medir com precisão esse nível de aglomeração. O índice foi elaborado considerando o processo de otimização das decisões de lucratividade e localização das empresas. Para tanto, consideraram na elaboração da modelagem, as vantagens naturais e os *spillovers* específicos dos setores, que os levam a se aglomerarem. Além disso, consideraram os efeitos de transbordamentos tecnológicos, que permitem maiores ganhos dado a proximidade das empresas. Esse modelo permitiu elaborar um índice que possui várias propriedades desejáveis, permitindo a comparação entre indústrias e regiões diferentes. Além de propor um índice de aglomeração, os autores elaboraram um índice de coaglomeração industrial.

Os autores supracitados observaram, ao utilizar dados de 459 indústrias manufatureiras, classificadas pelo *Census Bureau* de 50 Estados dos Estados Unidos, além do Distrito de Columbia, que todas as indústrias apresentavam uma concentração excessiva. As indústrias que mostraram maior concentração foram: Produtos de Peles (0,63) e Vinhos (0,49). No que se refere a coaglomeração industrial, verificou-se considerável coaglomeração das indústrias de tabaco, têxtil e madeira.

Considerando o índice de aglomeração industrial de Ellison e Glaeser (1997), Overman e Puga (2010) buscaram explicar como o *labor pooling* pode determinar a aglomeração industrial. Para justificar como o agrupamento no mercado de trabalho pode vir a determinar a aglomeração

industrial, os autores utilizam os embasamentos teóricos propostos por Krugman (1991). Outras variáveis, além do *labor pooling*, foram utilizadas pelos autores para explicar a aglomeração industrial. Os dados foram extraídos do *Annual Respondent Database (ARD)* que deriva do *Annual Census of Production* do Reino Unido, correspondendo ao período de 1994 a 2003. Como resultado principal, observou-se que o *labor pooling* apresenta um coeficiente positivo e significativo (0,1167), corroborando a hipótese dos autores sobre a aglomeração industrial.

2.2.3.2 Aglomeração industrial no Brasil

Na literatura brasileira, Maciente (2013) buscou avaliar a aglomeração e coaglomeração industrial e seus determinantes considerando dois tipos de externalidades: as externalidades de insumo-produto e as externalidades de conhecimento, causadas pelo *labor pooling*. É importante mencionar que a medida de *labor pooling* utilizada é diferente da de Overman e Puga (2010). O autor ainda enfatiza que as vantagens naturais locais são indispensáveis na análise da aglomeração industrial de acordo com a Nova Geografia econômica, logo, insere no seu modelo variáveis que captem essas vantagens naturais. Para tanto, Maciente (2013) também utilizou o índice de Ellison e Glaeser (1997) e, além desse, avaliou a aglomeração pelo índice de Gini e pelos índices de Moran univariados e bivariados.

No que se refere às externalidades de insumo-produto, o autor considerou o modelo de Venables (1996) da formulação de matriz insumo produto, com encadeamentos para frente e para trás nos setores por ele considerados. Maciente (2013) utilizou como base de dados as matrizes de insumo produto brasileiras disponibilizadas pelo IBGE, considerando 55 indústrias e 110 *commodities*, para os anos de 2000 a 2009. Quanto às externalidades de conhecimento devido ao *labor pooling*, o autor criou medidas de habilidades para os trabalhadores brasileiros, sendo essa uma grande contribuição do autor não só para a análise da aglomeração industrial no Brasil, mas para os estudos na Economia Urbana e Economia do Trabalho, em geral. Para criar medidas de habilidade, o autor utilizou a base de dados da *Occupational Information Network* - ONET[5] e a base de dados da RAIS, que apresentam variáveis relacionadas às ocupações desempenhadas pelos trabalhadores, e fez uso de uma metodologia de Análise Fatorial para compatibilizar a base de dados dos EUA com a do Brasil. Após esse processo de compatibilização, o autor

[5] A ONET é um sistema nacional de classificação ocupacional desenvolvido pelo Departamento de Trabalho, Emprego e Treinamento, vinculado ao Governo dos Estados Unidos (U.S. Department of Labor/Employment and Training Administration - USDOL/ETA).

conseguiu classificar 21 medidas de habilidades para os trabalhadores inseridos no mercado de trabalho formal brasileiro.

Maciente (2013) ainda avança na discussão ao mostrar, a partir de medidas de correlação, que as categorias de habilidades não estão relacionadas com a escolaridade dos trabalhadores, medida alternativa utilizada na literatura para definir nível de habilidade. As correlações obtidas pelo autor corroboram a sua hipótese inicial de que a escolaridade dos trabalhadores não é uma boa variável para representar os níveis de habilidades. Habilidade e escolaridade são condicionantes diferentes e impactam de maneira diversa na produtividade do trabalho dos indivíduos. Esse resultado serviu de respaldo para o uso da métrica de habilidades definidas pelo autor.

Empiricamente, Maciente (2013) comparou a estimação de três equações de determinação de aglomeração e coaglomeração industrial: a primeira contendo apenas variáveis de externalidades de insumo-produto, a segunda contendo apenas externalidades de habilidades de mão de obra e uma terceira incluindo todas as variáveis conjuntamente, contendo variáveis que representem ameninades locais. De modo geral, o autor percebeu que, comparando as duas primeiras equações com a última, as externalidades de conhecimento apresentaram maior magnitude e significância na determinação da aglomeração e coaglomeração industrial. As variáveis naturais são significantes, porém, apresentam pequeno impacto na determinação do modelo. Diante dessas evidências, o autor enfatiza que, considerando a análise municipal feita e o nível de desagregação industrial, as habilidades são mais importantes na economia brasileira para a determinação da aglomeração industrial.

A partir da principal conclusão da pesquisa de Maciente (2013), reforça-se a importância de estudar e avaliar os determinantes da aglomeração pela ótica da ocupação, considerando os efeitos do *learning*, *sharing (labor pooling) e matching* apresentados na literatura de aglomeração urbana. Como contribuição à análise do autor supracitado, pretende-se avaliar o poder das externalidades de mão de obra para a análise da aglomeração ocupacional.

Almeida e Rocha (2018) aplica a medida de *labor pooling* de Overman e Puga (2010) a fim de verificar os determinantes da aglomeração da indústria de transformação brasileira, considerando um painel de dados em nível setorial. Para tanto, o autor usa os dados da RAIS, para o período compreendido entre 2002 e 2014, e mensura a aglomeração através do Índice de

Ellison e Glaeser (calculado para as microrregiões brasileiras). O autor encontrou um índice médio, para os setores e anos, de 0,0244, sugerindo que a escolha local das firmas não se dá de forma aleatória. Além disso, verificou uma tendência de queda da concentração da indústria de transformação no Brasil no período considerado. Em 2014, Almeida e Rocha (2018) observou que os setores de alta tecnologia estão entre as 10 primeiras posições no *ranking*, destacando-se: o setor de fabricação de coque, refino de petróleo, elaboração de combustíveis nucleares e produção de álcool.

No que se refere ao *labor pooling*, o autor verificou que esse se relaciona positivamente com a aglomeração da indústria da transformação no Brasil, isto é, este setor se beneficia da aglomeração dos trabalhadores no espaço. A especificação em painel de dados inclui variáveis de controle como: Gastos de pessoal, Gastos da indústria com compras de energia elétrica e consumo de combustíveis, Gastos da indústria com aluguéis e arrendamentos e Dispêndios da indústria com impostos e taxas. Nesse contexto, observou-se que o efeito médio do *labor pooling* foi de 0,022, bem menor do que o encontrado por Overman e Puga (2010), mais ainda positivo e significante.

2.2.3.3 Aglomeração ocupacional

No que se refere à aglomeração ocupacional a literatura econômica é bastante escassa (GABE; ABEL, 2009; ANDRADE; GONÇALVES; FREGUGLIA, 2014; GOLDMAN; KLIER; WALSTRUM, 2019). Gabe e Abel (2009) indicam que a sua pesquisa era a primeira tentativa de estudar a aglomeração de ocupação.

Gabe e Abel (2009) são enfáticos ao considerar que a localização geográfica das ocupações é fundamental para a análise das aglomerações, pois avaliam com mais propriedade o que as pessoas fazem nos seus empregos e as tarefas desempenhadas. A proximidade dos indivíduos que desempenham a mesma ocupação gera maiores *spillovers* de conhecimentos do que empregos com requisitos genéricos, como na análise da aglomeração industrial. Logo, os autores esperam que as ocupações com características semelhantes se aglomerem. Destacam, ainda, que a aglomeração ocupacional pode impactar mais nos efeitos marshallianos de *labor pooling* e *knowledge externalities* e apresentam o exemplo a seguir para esclarecer o argumento:

> Um programador de computador, por exemplo, supostamente se beneficia mais da proximidade com outras pessoas envolvidas em atividades cotidianas similares (por exemplo, interagindo com computadores, usando tecnologia) do que trabalhando junto a outras pessoas na mesma indústria (por exemplo, recepcionista de uma empresa de software, especialista em recursos humanos ou diretor executivo). (GABE; ABEL, 2009, p. 2, "tradução nossa").

A fim de avaliar a aglomeração ocupacional os autores fazem uso do índice Gini, já indicando suas limitações, e de uma versão modificada do índice Ellison-Glaeser (1997) para captar as peculiaridades das ocupações, uma vez que essas não apresentam a mesma estrutura organizacional e metodológica dos setores de atividade econômica. Utilizando dados do Censo dos EUA de 2000, consideram 324 áreas metropolitanas, calculam o índice para cada uma das 468 ocupações presentes na ONET. Os autores observaram que as ocupações que apresentam maior concentração são: Ocupações de Agricultura, Pesca e Silvicultura; e Ocupações da Vida, Física e Ciências Sociais.

No que se refere à determinação da aglomeração ocupacional, os autores utilizaram como variáveis explicativas para o índice de Ellison-Glaeser modificado: Conhecimento especializado; Equipamento especializado; Atualização do conhecimento; Interação com o público (ou interação social), Tamanho médio dos estabelecimentos e Participação das pessoas que trabalham em ocupações na agricultura e mineração quando comparadas a indústria.

A variável "Conhecimento especializado" representa uma *proxy* para o *labor pooling*, sendo esta a principal variável de interesse. A construção desta variável tenta estabelecer como o perfil de conhecimento de cada ocupação difere da média de conhecimento dos EUA. Para calculá-la, foram utilizados os dados e informações presentes na ONET, relacionadas as habilidades e domínios requeridos em cada ocupação. Gabe e Abel (2009) detectaram a importância do *labor pooling* para a determinação da aglomeração ocupacional: os coeficientes estimados indicam que uma duplicação do Conhecimento especializado aumenta em 40% o índice de aglomeração ocupacional nos EUA.

Goldman, Klier e Walstrum (2019) avaliam a aglomeração ocupacional através de um instrumental metodológico diferente, considerando uma vertente intraindústria. Todavia, os autores indicam que os resultados empíricos obtidos permitem ter maior embasamento para explicar as fontes de aglomeração na indústria, dada a análise intrasetorial. Dessa forma, considerando indústria e ocupação, é possível detectar como dentro das indústrias as ocupações

se aglomeram. Para tanto, os autores focam em três grupos de ocupação específicos (ocupação de produção, ocupação de administração e ocupação de P&D) e avaliam a aglomeração ocupacional para todos os setores, mas dão atenção especial às indústrias manufatureiras.

Utilizando os dados do *US Census Bureau's American Community Survey (ACS)*, Goldman, Klier e Walstrum (2019) aplicam duas estratégias metodológicas. Primeiro, calculam os coeficientes locacionais para cada grupo de ocupação, considerando os mercados de trabalhos locais e a indústria como um todo, e, em um segundo momento, calculam o índice de Duranton e Overman (2005) para cada grupo ocupacional, utilizando a medida de aglomeração da indústria como contrafactual. O índice foi calculado ao nível de trabalhador e é obtido a partir de uma função de densidade de probabilidade, não paramétrica, considerando as distâncias espaciais entre os trabalhadores. Os resultados obtidos pelos autores também contribuem para o estudo da inovação, uma vez que, como resultado principal, detectou-se que, das três categorias de ocupação consideradas na pesquisa, aquelas atreladas à P&D foram as que apresentaram maior aglomeração, tanto pelo uso do coeficiente locacional como pelo índice de Duranton e Overman (2005), sendo seguidos pelos trabalhadores administrativos e trabalhadores de produção. Os autores justificam esse resultado indicando que o mesmo corrobora a teoria marshalliana que prevê maiores *spillovers* de conhecimento. Assim, como as ocupações de P&D exigem habilidades cognitivas e de criatividade intensas, a aglomeração favoreceria o desempenho dessas atividades.

2.2.3.4 Aglomeração ocupacional no Brasil

No Brasil, o trabalho de Andrade, Gonçalves e Freguglia (2014) representa a primeira tentativa de avaliar a aglomeração ocupacional no Brasil. Baseados no trabalho de Gabe e Abel (2009), os autores buscam explicar os determinantes da aglomeração ocupacional com uma modelagem empírica bem próxima da pesquisa de Gabe e Abel (2009). Entretanto, não utilizam nenhuma medida de *labor polling*, conhecimento especializado ou habilidade. A variável dependente principal é o nível tecnológico da ocupação (alta tecnologia, média tecnologia e baixa tecnologia). A medida utilizada para detectar e medir a aglomeração ocupacional foi o índice de Gini que, conforme enfatizado por Ellison e Glaeser (1997) e Gabe e Abel (2009), apresenta as seguintes limitações: altos níveis de aglomeração podem ser captados mesmo em situações com um padrão de aglomeração aleatório e disperso se os setores apresentarem poucas

indústrias, mas de grandes dimensões. O índice de Ellison e Glaeser (1997) supera essa limitação, uma vez que inclui informações sobre a distribuição e tamanho das empresas.

Andrade, Gonçalves e Freguglia (2014) utilizam a base de dados da RAIS para o período de 2003 a 2008, avaliando a aglomeração ocupacional a partir de um painel de dados. Mesmo sendo a variável de intensidade tecnológica fixa no tempo, para contornar esse problema os autores fizeram uma interação da *dummies* de intensidade tecnológica e *dummies* de tempo. Verificou-se que algumas ocupações relacionadas às atividades agrícolas, florestais e pesqueiras evidenciam níveis de concentração elevados e que a alta tecnologia tem um efeito positivo e significante sobre a concentração das ocupações. Observou-se ainda, que o impacto da alta tecnologia apresenta efeitos decrescentes ao longo dos anos considerados.

Neste momento, convém indicar algumas contribuições desta pesquisa para a literatura empírica até aqui apresentada. Os efeitos de *labor pooling*, indicados por Marshall, requerem uma proximidade de trabalhadores com as habilidades requeridas pelas empresas. Todavia, a medida proposta por Overman e Puga (2012) não traz nenhuma referência sobre as características deste agrupamento em termos das habilidades requeridas desses trabalhadores. Logo, essa medida de *labor pooling*, apesar de exógena e contornar o problema da causalidade reversa, não permite inferir nada sobre a habilidade dos trabalhadores e quais tipos de habilidades são essas. Ademais, dado que cada estabelecimento requer trabalhadores com perfis de escolaridade e habilidades cognitivas, sociais e manuais diferentes, informações sobre habilidades são de suma importância para uma melhor compreensão do *labor pooling* ocupacional.

Conforme já indicado, Gabe e Abel (2009) utilizaram como *proxy* para *labor pooling* a variável Conhecimento Especializado. Neste sentido, não incorporam no modelo a dinâmica de contratação das empresas em mercados densamente aglomerados, ou seja, apenas consideram se as habilidades (conhecimento especializado) impactam na determinação da aglomeração ocupacional, não incluindo os choques idiossincráticos do mercado de trabalho que podem modificar a estrutura setorial e ocupacional das empresas. Neste sentido, esta pesquisa avança em relação à literatura já mencionada, pois além de avaliar a concentração ocupacional no Brasil pelo índice de Ellison-Glaeser modificado por Gabe e Abel (2009), utiliza uma medida de *labor pooling* que considerada tanto os choques idiossincráticos como as habilidades dos trabalhadores, sendo possível isolar os efeitos de *sharing*, *matching* e *learning*, obtendo

parâmetros mais robustos. Essa proposta empírica será explicada com mais detalhes na próxima seção.

2.3 ESTRATÉGIA EMPÍRICA

Esta seção busca apresentar a abordagem empírica utilizada a fim de medir a aglomeração ocupacional no Brasil e avaliar seus determinantes, tendo como foco principal os efeitos marshallianos de *sharing* (*labor pooling*), *matching* e *learning*. A medida utilizada para medir a aglomeração ocupacional é o índice de Ellison-Glaeser (1997) modificado e adaptado por Gabe e Abel (2009). No que se refere à determinação da aglomeração ocupacional, a estratégia empírica propõe utilizar um painel de dados no nível da ocupação.

2.3.1 Índice de aglomeração ocupacional

O índice Ellison-Glaeser (1997) modificado por Gabe e Abel (2009) foi calculado para o Brasil a partir dos Arranjos Populacionais Brasileiros (IBGE, 2016). As pesquisas empíricas listadas neste ensaio, na seção anterior, normalmente analisam os índices considerando como categoria geográfica as macrorregiões brasileiras. Esta pesquisa contribui para a literatura brasileira avaliando os índices em uma unidade regional mais desagregada, os Arranjos Populacionais, os quais definem os mercados de trabalhos locais do país. No que se refere à categoria ocupacional, foi utilizada a Classificação Brasileira de Ocupações (CBO-2002) considerando o menor nível de desagregação possível, as ocupações. Desse modo, é possível obter uma análise mais minuciosa e profunda da aglomeração ocupacional.

Como o objetivo da pesquisa é avaliar a aglomeração ocupacional deve-se considerar a concentração urbana no território nacional. Considerando a crescente urbanização, torna-se fundamental avaliar como essa concentração urbana se dá e seus efeitos econômicos, em especial, no mercado de trabalho. Neste sentido, a presente pesquisa, a fim de obter uma medida mais fidedigna da aglomeração ocupacional considerará o conceito de Arranjos Populacionais definido pelo IBGE (2016).

Os Arranjos Populacionais constituem um conjunto integrado de municípios que apresentam características semelhantes, definidas pela proximidade espacial e organizacional, sinalizando uma unidade de urbanização que atende aos critérios de contiguidade e integração econômica.

Para chegar a esta unidade territorial econômica, Arranjos Populacionais, o IBGE (2016) avalia movimentos pendulares de trabalho e estudo capturados a partir de um índice de intensidade relativa, que considera: Densidade demográfica no município, Porcentagem da população economicamente ativa nos setores de atividades secundário e terciário; Porcentagem de pessoas residentes que trabalham fora do município em relação à população economicamente ativa; Pessoal ocupado em relação à população economicamente ativa, para as indústrias de transformação e extração mineral; Fluxos telefônicos; Linhas de transporte coletivo; Fluxos de mercadorias e capitais, entre outros (IBGE, 2016).

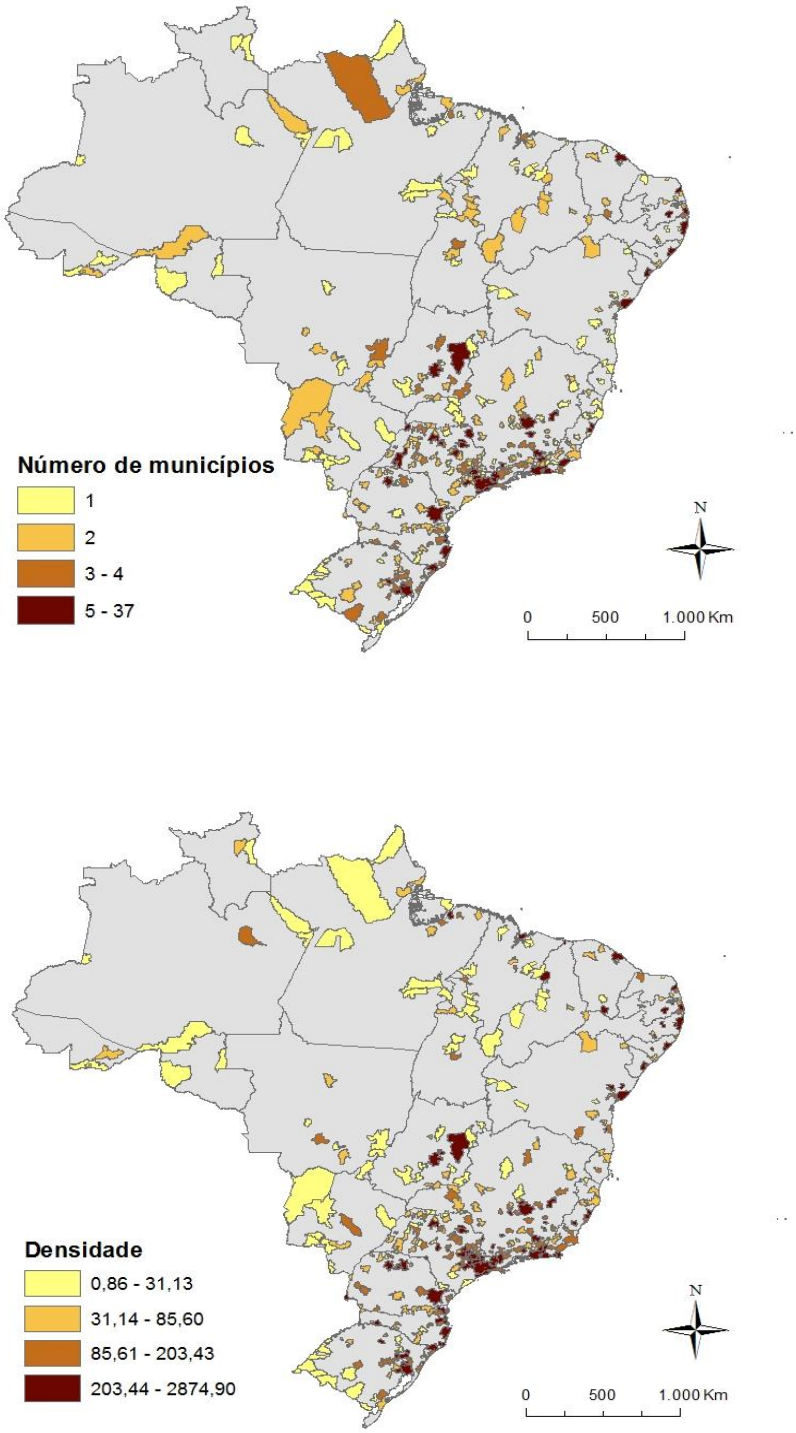
Com estes critérios o IBGE (2016) definiu 294 Arranjos Populacionais no país compostos por 953 municípios. Estes arranjos populacionais estão distribuídos entre as macrorregiões brasileiras da seguinte forma: Região Sudeste (116), Regiões Sul (87), Nordeste (51), Centro-Oeste (24) e Norte (16). Os Arranjos Populacionais no Brasil ainda são classificados em grandes, médios e pequenos. Para esta classificação o IBGE (2016) estabelece que arranjos com uma população menor que 100 mil são considerados pequenos, entre 100 e 750 mil habitantes configura-se a categoria de arranjos médios e uma população maior que 750 mil define os grandes arranjos. Dessa forma, verificou-se que no Brasil existem 189 arranjos pequenos, 81 médios e 24 grandes. Assim, de um total de 5.565 municípios, somente 953 pertencem a Arranjos populacionais[6], o que representa 55,7% da população residente para o ano de 2010.

Além disso, o IBGE (2016) ainda define o que seriam as Médias Concentrações e Grandes Concentrações. As médias concentrações urbanas são os municípios isolados (que não formam arranjos) e os Arranjos Populacionais que possuem de 100.000 a 750.000 habitantes, já as grandes concentrações urbanas são definidas como municípios isolados e Arranjos Populacionais com população acima de 750.000 habitantes. Do total de concentrações médias, 80 são municípios isolados e do total de concentrações grandes, 2 municípios são isolados. Assim, considerando todos os Arranjos Populacionais e os municípios isolados das médias e grandes concentrações, totalizou-se 376 unidades geográficas a serem analisadas nessa pesquisa, que são chamados genericamente de Arranjos Populacionais, a partir dessa seção.

[6] Todos os municípios que pertencem a Arranjos Populacionais, separados por Estados, encontram-se presentes na Tabela 1.1 do estudo "Arranjos Populacionais e Concentrações Urbanas do Brasil" do IBGE (2016).

A Figura 2.1 abaixo mostra a distribuição espacial dos Arranjos Populacionais brasileiros, considerando o número de municípios e a densidade populacional dos mesmos. Os maiores Arranjos Populacionais (que possuem maior número de municípios) e com maior densidade populacional são litorâneos, em sua maioria, e se concentram nas macrorregiões Sul e Sudeste.

Figura 2. 1 – Distribuição espacial do número de municípios e da densidade nos Arranjos Populacionais



Fonte: Elaboração própria (2021), com base nos dados do IBGE e IBGE (2016).

Convém ainda mencionar algumas características e propriedades do índice de Ellison-Glaeser (1997), bem como indicar as modificações propostas por Gabe e Abel (2009) para avaliar a aglomeração ocupacional. Ellison-Glaeser (1997) avaliam a concentração do emprego

industrial mensurando a participação do emprego em um setor específico em uma determinada região em relação ao emprego como um todo. Todavia, é notório que o tamanho das empresas nesses setores pode enviesar o resultado desta participação e não refletir apenas a concentração industrial: um número reduzido de grandes indústrias poderia superestimar a concentração. Como solução Ellison-Glaeser (1997) ponderam esta participação pela distribuição de tamanho das plantas nas indústrias, a partir do índice de Herfindahl:

$$H = \sum_{j=1}^{N} z_j^2 \tag{2.14}$$

Em que, N representa o número total de firmas (j) de cada setor e z_j é a parcela de emprego de cada firma (j). Dessa forma, é possível atribuir pesos de acordo com o tamanho de cada firma nos setores e o índice de Ellison-Glaeser (1997) pode ser definido como:

$$IEG_s = \frac{[G-(1-\sum_{i=1}^{n} {t_i}^2)H]}{(1-\sum_{i=1}^{n} t_i^2)(1-H)}$$
$$G = \sum_{i=1}^{n}(s_i - t_i)^2 \tag{2.15}$$

Em que,
t_ié participação do emprego total na região (i)
s_i é participação do emprego setorial na região (i)

Dessa forma, Ellison-Glaeser (1997) indicam que este índice é uma estimativa da presença de economias de aglomeração em um setor, numa escala em que 0 indica falta de aglomeração industrial e valores positivos em sua magnitude indicam presença de aglomeração. Assim, este índice possui algumas propriedades desejáveis: facilidade no cálculo, dada a disponibilidade dos dados; a escala do índice permite fazer comparações entre setores, sejam eles não aglomerados ou aglomerados (e em diferentes dimensões). Isso é possível pois assume-se que os dados são obtidos a partir de um modelo de escolha aleatória da localização, sem levar em consideração vantagens naturais ou *spillovers* específicos da indústria; E, ainda, o índice permite comparações entre setores cuja distribuição do tamanho das firmas difere. Todavia, convém citar uma limitação do índice: apresenta sensibilidade à escolha da escala geográfica de análise.

O desafio de Gabe e Abel (2009) foi aplicar o índice de Ellison-Glaeser (1997) para a análise da aglomeração ocupacional sem perder suas propriedades. Para tanto os autores adaptaram o índice de Herfindahl para captar a distribuição de trabalhadores em uma ocupação nos principais setores industriais:

$$I = \sum_{j=1}^{m} {}^{y_j}/_{e_k} \tag{2.16}$$

Em que,

y_j é participação do setor (*j*) no emprego ocupacional (*k*)

e_k é emprego ocupacional total (*k*) nos Arranjos Populacionais

m é o número das principais categorias setoriais

Assim, se um setor tem um maior número de trabalhadores em uma ocupação específica, a robustez do índice não será prejudicada ou o mesmo não será superestimado. Com essas considerações, o índice de Ellison-Glaeser (1997) adaptado por Gabe e Abel (2009), considerando o período de 2009 a 2018, apresenta a seguinte representação para cada categoria ocupacional (k):

$$INDEX_{kt} = \frac{[G_t - (1 - \sum_{i=1}^{n} t_{it}^2) I_t]}{[(1 - \sum_{i=1}^{n} t_{it}^2)(1 - I_t)} \tag{2.17}$$

$$G_t = \sum_{i=1}^{n} (s_{it} - t_{it})^2$$

$$I_t = \sum_{j=1}^{m} {}^{y_{jt}}/_{e_{kt}}$$

Em que,

t_{it}é participação do emprego total no Arranjo Populacional (i) no período (t)

s_{it} é participação do emprego ocupacional (k) no Arranjo Populacional (i) no período (t)

y_{jt} é participação do setor (j) no emprego ocupacional (k) no período (t)

e_{kt} é emprego ocupacional (k) total nos Arranjos Populacionais (i) no período (t)

m é o número das principais categorias setoriais

n é o número das sub-regiões (Arranjos Populacionais)

Vale ressaltar que foi realizada uma adaptação do indicador para que seus componentes variassem ao longo do tempo, considerando a organização dos dados em painel. Assim, foi obtido um índice para cada ocupação e ano, a partir do cálculo desagregado de 376 Arranjos Populacionais. Neste momento cabe esclarecer melhor as classificações de ocupação consideradas para o cálculo do índice acima.

No que se refere à delimitação das ocupações, conforme já mencionado, foi utilizada a CBO-2002. A CBO realizada em 2002 tem por finalidade nomear e classificar as ocupações além de descrever as suas características. Sua estrutura organiza as ocupações em dez grandes Grupos: (0) Membros das forças armadas, policiais e bombeiros militares; (1) Membros superiores do poder público, dirigentes de organizações de interesse público e de empresas, gerentes; (2) Profissionais das ciências e das artes; (3) Técnicos de nível médio; (4) Trabalhadores de serviços administrativos; (5) Trabalhadores dos serviços, vendedores do comércio em lojas e mercados; (6) Trabalhadores agropecuários, florestais e da pesca; (7) Trabalhadores da produção de bens e serviços industriais (I); (8) Trabalhadores da produção de bens e serviços industriais (II); (9) Trabalhadores em serviços de reparação e manutenção. Esses "Grupos" originam 47 "Subgrupos principais" que, por sua vez, promovem a classificação de 192 "Subgrupos". Os "Subgrupos" originam 596 "Famílias" e, a partir destas, agrupam-se 2.422 ocupações e 7.258 títulos sinônimos (BRASIL, 2017).

É importante salientar que, desde a elaboração da CBO em sua última versão realizada em 2002, essa vem sofrendo alterações, geralmente promovendo a inclusão de novas famílias, ocupações e sinônimos, a fim de acompanhar a dinâmica do mercado de trabalho. Assim, considerando os últimos históricos de alteração da CBO disponível (2012, 2013 e 2016), tem-se atualmente um total de 2.649 ocupações[7]. Conforme já salientado, a base de dados da RAIS contém informações referentes a ocupação dos trabalhadores obedecendo a essa classificação da CBO, isto é, a classificação é dada por 6 dígitos que correspondem aos títulos das ocupações.

Dessa forma, considerando a equação (2.17), compreende-se que o índice bruto, determinado pelo componente $G_t = \sum_{i=1}^{n}(s_{it} - t_{it})^2$, tem maior destaque na determinação do índice total.

[7] Para consultar as informações sobre estrutura e listagem da CBO, acessar o endereço eletrônico http://www.mtecbo.gov.br/cbosite/pages/downloads.jsf. A última listagem foi realizada em 17 de março de 2020, já estando incluída nessa as alterações realizadas ao longo dos anos. Para verificar o histórico de alterações, acessar o endereço eletrônico http://www.mtecbo.gov.br/cbosite/pages/DownloadHistorico.jsf

O componente G_t, por sua vez, é ponderado pelas diferenças no emprego total nos Arranjos Populacionais, $(1 - \sum_{i=1}^{n} {t_{it}}^2)$, e pela distribuição do emprego ocupacional entre os setores de atividade econômica, $I_t = \sum_{j=1}^{m} {y_{jt}}/{e_k}$. Essa ponderação permite que o índice não seja subestimado ou superestimado dada a dinâmica do mercado de trabalho em determinadas subunidades geográficas ou pela participação das ocupações em determinados setores da atividade econômica, respectivamente. Dessa forma, as ocupações que estão concentradas em um ou poucos setores da atividade econômica terão valores mais baixos para o índice, pois verifica-se que, neste caso, o alto nível de concentração não se daria somente por uma questão espacial, mas devido a organização desta ocupação nos setores econômicos, o contrário ocorre para ocupações que estão dispersas entre os diversos setores da atividade econômica.

De modo geral, a tendência comportamental de G_t reflete a magnitude do índice, quanto maior o valor de G_t, maior a aglomeração ocupacional. Percebe-se que um valor igual a zero da diferença $s_{it} - t_{it}$, indica que a participação do emprego em determinada ocupação e em determinado Arranjo Populacional (s_{it}) é igual a participação do emprego total no mesmo Arranjo Populacional (t_{it}), ou seja, a distribuição do emprego ocupacional é a mesma não havendo evidências para estabelecer uma concentração espacial daquela ocupação neste espaço geográfico. Para valores positivos, percebe-se que a participação do emprego nesta ocupação e nesse Arranjo Populacional (s_{it}) supera a participação total do emprego neste Arranjo Populacional (t_{it}), indicando que essa ocupação tende a ser concentrada nesse espaço geográfico. Após somar essa diferença para todos os Arranjos Populacionais, tem-se uma medida geral de como essa ocupação se concentra no Brasil, dada a análise minuciosamente feita a nível de Arranjos Populacionais. Dessa forma, valores próximos de zero indicam que os trabalhadores de determinada ocupação estão dispersos nos Arranjos Populacionais e valores positivos indicam que os trabalhadores de determinada ocupação estão concentrados em poucos Arranjos Populacionais.

2.3.2 Modelo econométrico: Determinantes da aglomeração ocupacional

Nesta seção será apresentado o modelo econométrico utilizado para avaliar os determinantes da aglomeração ocupacional no mercado de trabalho formal brasileiro, considerando como foco da análise os efeitos marshallianos de *sharing* (*labor pooling*), *matching* e *learning*.

Conforme indicado, a principal contribuição desta pesquisa é explicar a determinação da aglomeração ocupacional, a partir de uma medida mais robusta de *labor pooling* adotada por Overman e Puga (2010). Essa medida capta os choques idiossincráticos no mercado de trabalho que podem modificar a estrutura setorial e ocupacional das empresas. Por sua vez, foi utilizada, ainda, a medida de *labor pooling* adotada em Gabe e Abel (2009), que utilizaram a variável conhecimento especializado, ou seja, um indicador da participação ocupacional do fator de habilidade.

Utilizando os microdados da RAIS-ME[8], para o período de 2009 a 2018, o efeito do *labor pooling* sobre a aglomeração ocupacional foi estimado considerando como variável dependente o índice de Ellison-Glaeser (1997) adaptado por Gabe e Abel (2009), através de um painel de dados em nível das categorias ocupacionais, o que permite captar a heterogeneidade não observada das mesmas[9], a partir do seguinte modelo econométrico:

$$INDEX_{kt} = \alpha + \delta L_{kt} + \gamma_h H_{kt}^h + \mu_h\left(H_{kt}^h * L_{kt}\right) + \beta_1 ES_{kt} + \beta_2 TE_{kt} + \beta_3 EXP_{kt} + \beta_4 PA_{kt} + \beta_5 PEM_{kt} + \emptyset_k + T_t + \varepsilon_{kt} \quad (2.18)$$

Em que, $INDEX_{kt}$ representa o índice de aglomeração ocupacional (k) no período (t), L_{kt} representa a medida de *labor pooling* por ocupação (k) no período (t) como uma adaptação da medida estabelecida por Overman e Puga (2010), H_{kt}^h representa os indicadores da participação ocupacional do fator de habilidade das 21 categorias (h) por ocupação (k) definidas por Maciente (2013), ES_{kt} representa a participação de trabalhadores com ensino superior por ocupação (k) no período (t), TE_{kt} representa o tamanho médio dos estabelecimentos por ocupação (k) no período (t), EXP_{kt} representa a experiência média dos trabalhadores por ocupação (k) e no período (t), PA_{kt} representa participação de pessoas na ocupação (k) que trabalham no setor da agropecuária em relação à indústria no período (t), PEM_{kt} representa a participação de pessoas na ocupação (k) que trabalham no setor de extração mineral em relação à indústria no período (t), $\emptyset_k$ representa o efeito fixo do grupo ocupacional, T_t representa as *dummies* anuais e ε_{kt} representa o termo de erro do modelo.

[8] Os microdados identificados da Relação Anual de Informações Sociais (RAIS) estão, desde 2019, hospedados na Secretaria do Trabalho do Ministério da Economia (ME).
[9] O uso de dados em painel representa uma contribuição em relação à pesquisa empírica dos autores Gabe e Abel (2009), pois os mesmos não estimaram o modelo a partir de um painel de dados.

A variável *labor pooling* (L_{kt}) foi obtida tal como definido por Overman e Puga (2010), todavia, utilizando uma abordagem ocupacional, isto é, a média por ocupação da seguinte expressão:

$$L_{kt} = \left|\Delta\%emp_{kE,t} - \Delta\%emp_{k,t}\right| \tag{2.19}$$

Em que, $\Delta\%emp_{kE,t}$ é a variação percentual do emprego ocupacional no estabelecimento e $\Delta\%emp_{k,t}$ é a variação percentual do emprego ocupacional. Assim, tem-se a média por ocupação das diferenças entre a variação percentual do emprego ocupacional no estabelecimento e a variação percentual na ocupação. Dada sua localização, conforme estabelecido por Overman e Puga (2010), os estabelecimentos sofrem choques de produtividade de acordo com o mercado de trabalho local, e, a partir desse cenário, tomam decisões referentes a contratação de mão de obra.

A variável usada como indicador da participação ocupacional do fator de habilidade (H_{kt}^{h}), é baseada nos dados de Maciente (2013), os quais estabelecem 21 categorias de habilidades[10], então, γ_h é um vetor que apresenta 21 coeficientes diferentes. Maciente (2013) determinou 21 fatores para cada ocupação. Neste sentido, o autor determina o uso ou importância de cada fator de habilidade para o desempenho da função do trabalhador. A fim de avaliar a proporção e o comportamento tendencial desses fatores de habilidades por ocupação ao longo dos anos, de modo a estimar um painel de dados, esta pesquisa criou um indicador da participação ocupacional do fator de habilidade (PFH_{ikt}), estabelecido pela seguinte expressão:

$$PFH_{ikt} = \frac{emp_{kt} \times F_{ik}}{|\sum_k emp_{kt} \times F_{ik}|} = \frac{emp_{kt} \times F_{ik}}{|(emp_{1t} \times F_{i1}) + \cdots + (emp_{Kt} \times F_{iK})|} \tag{2.20}$$

Em que, emp_{kt} é o emprego na ocupação $k = 1, \ldots, K$, no ano $t = 2009, \ldots, 2018$ e F_{ik} é o fator de habilidade $i = 1, \ldots, I$ da ocupação k. Assim, tem-se a proporção no uso do fator de habilidade i na ocupação k em relação ao uso do fator da habilidade i no total de empregos no Brasil para cada ano e ocupação. Logo, tem-se a participação ocupacional do fator de habilidade por ano. Por exemplo, considerando o vetor de habilidades cognitivas, pode-se verificar a intensidade do uso dessa habilidade para a ocupação de "Físico" no Brasil em relação a intensidade do uso da mesma habilidade para o emprego total no Brasil. Dessa forma, é possível

[10] Ver Tabela A.1 do Apêndice A.

verificar o uso de fator de habilidade para cada ocupação, ao longo do tempo, sempre ponderando esse uso de habilidade em relação ao total usado no país (ou seja, quando considera-se seu uso em todas as ocupações).

Verifica-se ainda que a variável *labor pooling*, como sugerido por Overman e Puga (2010), equação (2.19), é ponderada pelo vetor de habilidades, uma a medida de *labor pooling* similar à adotada em Gabe e Abel (2009), permitindo que o *labor pooling* ponderado capte choques idiossincráticos de acordo com as habilidades dos trabalhadores. Nesse aspecto reside a contribuição empírica dessa pesquisa que dá maior robustez as medidas sugeridas pelos autores, mantendo a exogeneidade da variável, controlando o problema da causalidade reversa e permitindo obter uma medida não apenas quantitativa de *labor pooling*, mas uma medida qualitativa, cujo papel das habilidades dos trabalhadores é fundamental. Assim, é possível verificar, por exemplo, se a medida de *labor pooling* ponderada pelo indicador da participação ocupacional do fator de habilidade cognitiva dos trabalhadores, tem maior impacto na aglomeração ocupacional do que as demais.

Inseriu-se as variáveis separadamente, pois quando se deseja captar o efeito de interação de duas quaisquer variáveis se faz fundamental incluir as variáveis isoladas. Essas variáveis isoladas influenciam o índice, porém existe uma dependência ou complementariedade entre as mesmas. Se a complementariedade não existe, o coeficiente estimado será estatisticamente insignificante. Dessa forma, a análise do efeito marginal de cada variável será igual ao seu coeficiente somado com o coeficiente da variável de interação.

Verifica-se que as variáveis propostas nesta pesquisa são utilizadas para analisar como o *sharing* (*labor pooling*) pode determinar a concentração ocupacional. Conforme já mencionado, esses efeitos da aglomeração são indissociáveis do conceito de habilidades e interagem entre si a partir desse conceito, neste aspecto reside a dificuldade em isolar os efeitos. Desse modo, o *labor pooling* representa uma medida de *sharing*, conforme metodologia proposta por Krugman (1991). Adicionalmente, pode-se estabelecer que o vetor do indicador da participação ocupacional do fator de habilidade é *proxy* para o *matching*, dado que essa variável mede quantitativamente os requisitos de habilidades para o desempenho de cada ocupação no Brasil, podendo captar, então, a interação entre oferta e demanda por ocupações e a correspondência entre firmas e trabalhadores no que se refere ao quesito habilidade. A interação entre as duas variáveis, $\left(H_{kt}^{h} * L_{kt}\right)$, pode captar, então, a interdependência entre os

efeitos de *sharing*, *matching* e *learning*. Essa estratégia empírica permite um modelo ainda mais robusto uma vez que, ao inserir esses controles, o efeito do *labor pooling* estará mais consistente, eliminando possíveis endogeneidades que ainda possam existir.

Silva (2019) apresenta três abordagens para a mensuração do *matching*. A primeira tem como suporte principal a realização de entrevistas com os trabalhadores que indicam como suas qualificações correspondem as suas ocupações; a segunda, por sua vez, prioriza o levantamento de informações dos requisitos educacionais em cada ocupação, através de institutos estatísticos e analistas de mercado de trabalho. A terceira abordagem é mais minuciosa e verifica, em cada ocupação, a composição dos trabalhadores por escolaridade, de modo a determinar os requisitos primordiais e as habilidades necessárias para o desempenho de cada ocupação. Verifica-se que essa última abordagem reflete o método de mensuração utilizado nessa pesquisa, de acordo com a metodologia proposta por Maciente (2013) para obtenção dos fatores de habilidades.

Dado que os efeitos de *sharing*, *matching* e *learning* são intrinsecamente relacionados, espera-se que os coeficientes do *labor pooling*, do vetor de habilidades e das variáveis de interação sejam significativos, demonstrando dependência e complementariedade estatísticas, e comprovando empiricamente a relação indissociável entre os efeitos. Conforme verificado na equação (2), além das variáveis de interesse, *labor pooling*, habilidades, e interação entre *labor pooling* e habilidades, variáveis de controle foram consideradas. Cabe, agora, justificar o uso das variáveis de controle selecionadas.

A variável de "participação de trabalhadores com ensino superior por ocupação", é considerada por alguns autores, como Rosenthal e Strange (2001) e Ellison e Glaeser (1999), como uma medida onde é possível capturar as habilidades dos trabalhadores. Porém, autores como Heckman, Stixrudb e Urzuau (2006) e Maciente (2013) asseguram que habilidade e escolaridade são medidas distintas e utilizar escolaridade como *proxy* para habilidades pode incorrer em vieses de estimação. Como a utilização da base de dados de Maciente (2013) possibilita incluir uma medida mais adequada e efetiva para um conjunto diverso de habilidades, optou-se por incluir a "participação de trabalhadores com ensino superior por ocupação" a fim de diferenciar os efeitos de habilidade e escolaridade e eliminar possíveis vieses de endogeneidade entre as variáveis. Espera-se que a variável "participação de trabalhadores com ensino superior por ocupação" exerça um efeito positivo na determinação da aglomeração ocupacional.

A variável “tamanho médio dos estabelecimentos por ocupação”, foi utilizada no modelo empírico de Gabe e Abel (2009) visando captar os efeitos de economias de escala. Logo, se o grupo ocupacional pertencer a grandes empresas, as possibilidades de concentração industrial e, neste caso, de concentração ocupacional aumentam. Assim, espera-se um efeito positivo desta variável na determinação da aglomeração ocupacional.

A variável “experiência média dos trabalhadores por ocupação” estabelece o tempo médio no emprego que a ocupação proporciona aos trabalhadores. É sabido que, quanto maior o tempo no emprego menor a rotatividade no mercado de trabalho, o que permite maiores chances de investimento em treinamento geral e específico, que aumentam a produtividade do trabalhador e, na média, a produtividade do grupo ocupacional. Esta maior experiência possibilita adquirir habilidades para desempenhar com maior eficiência as atividades exigidas na ocupação, logo, os *spillovers* de conhecimento são mais intensos. Dessa forma, espera-se um efeito positivo desta variável na determinação da aglomeração ocupacional.

As variáveis “participação de pessoas na ocupação que trabalham no setor da agropecuária em relação à indústria” e “participação de pessoas na ocupação que trabalham no setor de extração mineral em relação à indústria” permitem integrar a análise setorial na determinação da aglomeração ocupacional, estas também foram incluídas na modelagem empírica de Gabe e Abel (2009). Os autores justificam o uso das mesmas, pois permitem captar “a importância de vantagens naturais e do uso de matérias-primas no processo de aglomeração” (GABE; ABEL, 2009, p. 12, “tradução nossa”). Como permitem identificar possíveis vantagens naturais, espera-se um efeito positivo desta variável na determinação da aglomeração ocupacional. Por fim, as *dummies* temporais devem controlar possíveis choques macroeconômicos ou cíclicos, na economia brasileira, que possam ter ocorrido ao longo do período analisado.

2.3.3 Estratégia de identificação

Conforme já mencionado, a análise da concentração ocupacional, ao considerar as externalidades positivas da aglomeração, pode ser determinada pelos efeitos de *sharing*, *matching* e *learning*, sendo difícil, empiricamente, isolar esses efeitos dada a “equivalência marshalliana”. Nesse sentido, é comum que na determinação da concentração ocupacional, considerando os impactos do *labor pooling*, possa existir o problema da causalidade reversa, isto é: ao mesmo tempo que a localização das firmas pode determinar a escolha de localização

dos trabalhadores, impactando diretamente na concentração ocupacional, as firmas tentem a se localizar em grandes centros onde o mercado de trabalho apresenta uma oferta de habilidades mais intensa.

Desse modo, a estratégia de identificação deste trabalho consiste no uso da variável exógena de *labor pooling*, com base na definição de Overman e Puga (2010), que capta choques idiossincráticos no mercado de trabalho, para controlar o problema da causalidade reversa e captar (ou isolar) a diferença entre escolaridade e habilidades. Além disso, utiliza-se um modelo longitudinal que possibilita controlar o viés de heterogeneidade não observada das ocupações.

O método de estimação utilizado aplica dados em painel de Efeitos Fixos. A principal literatura, Overman e Puga (2010), usa o método de MQO. Apesar dos autores possuírem dados longitudinais, os mesmos indicam que o modelo se ajustou melhor após obtenção de uma média anual das variáveis utilizadas. Os trabalhos de Almeida e Rocha (2018) e Andrade, Gonçalves e Freguglia (2014) aplicados para o Brasil usam o método de painel de dados de Efeitos Fixos à nível de setores e ocupação, respectivamente.

O uso de dados em painel possui vantagens econométricas, pois possibilita o controle das características não observadas, permitindo a obtenção de parâmetros mais robustos e eficientes. Na estratégia empírica desta pesquisa, a estimação considerou a heterogeneidade não observada das ocupações. Conforme indicado por Gabe e Abel (2009), características não observadas, como o nível tecnológico em cada ocupação e a natureza de interação com o público, podem determinar a concentração ocupacional. Além disso, preferências individuais e características regionais, como instituições e políticas governamentais, podem impactar na decisão de escolha ocupacional e escolha local dos trabalhadores, o que pode ter reflexos na concentração ocupacional. Todavia, na maioria das vezes, essas são características e variáveis que não são observadas ou mensuradas (ANDRADE; GONÇALVES; FREGUGLIA, 2014). Além disso, é imprescindível controlar a heterogeneidade não observada das ocupações, uma vez que essas podem explicar uma parte substancial dos padrões de concentração ocupacional a partir da sua correlação com as habilidades e a educação.

No que se refere a variável *labor pooling*, Overman e Puga (2010) estabeleceram como medida empírica a diferença entre a variação percentual no emprego do estabelecimento e a variação percentual no emprego do setor. Essa estratégia seria suficiente para captar os choques

idiossincráticos no mercado de trabalho. Conforme já mencionado na seção de referencial teórico desta pesquisa, o modelo de *labor pooling* de Krugman (1991), ao considerar que a economia pode passar por choques de produtividade afetando os estabelecimentos e, portanto, a lucratividade dos mesmos, assume que esses choques seriam idiossincráticos e os lucros dos estabelecimentos seriam convexos em relação a esses choques.

Os fundamentos microeconômicos do autor apresentam o lucro esperado como uma função, que depende positivamente da proporção de trabalhadores por estabelecimento e do agrupamento no mercado de trabalho, isto é, o *labor pooling*, que representa os choques de produtividade dos estabelecimentos. Ou seja, um maior nível de mão de obra disponível para cada estabelecimento provoca uma diminuição nos salários e um aumento na lucratividade. Esses choques idiossincráticos podem ser positivos ou negativos e os mesmos afetarão os lucros a partir de mudanças nos salários. Esse efeito é maior quando os estabelecimentos estão em locais com grande número de trabalhadores com as habilidades requeridas pelas empresas. Logo, espera-se que, quanto maior a volatilidade idiossincrática maior será a aglomeração, pois a empresa consegue fazer ajustes na sua produção sem pressões sobre os salários.

Não obstante, o modelo teórico de Krugman (1991) e a estratégia empírica de Overman e Puga (2010) considerarem um contexto de aglomeração setorial, nessa pesquisa aplica-se esses fundamentos para uma análise ocupacional. Neste sentido, a configuração econômica do modelo de Krugman (1991), já mencionada, considera uma série de setores e cada um com um número discreto de estabelecimentos, nos quais há um conjunto de trabalhadores com habilidades específicas para cada setor. Neste último ponto reside a ligação com a variável *labor pooling* ocupacional: cada ocupação faz parte da estrutura produtiva dos estabelecimentos, onde cada estabelecimento apresenta um contínuo de trabalhadores com ocupações e habilidades específicas para o mercado de trabalho ocupacional local.

A medida de *labor pooling* considerada está representada na equação (2.19) desta seção. É importante salientar que, ao considerar uma média do *labor pooling* por ocupação, as propriedades de exogenidade da variável foram mantidas, uma vez que o modelo de Krugman (1991) capta o comportamento dos lucros dos estabelecimentos dados os choques idiossincráticos no mercado de trabalho, e esses choques podem ser avaliados por uma perspectiva setorial ou ocupacional. Pois, a formulação da variável requer uma mudança no nível de emprego do estabelecimento dadas as habilidades requeridas.

Nesta pesquisa, capta-se os choques idiossincráticos do mercado de trabalho no que se refere as ocupações, isto é, como o agrupamento de trabalhadores na mesma ocupação pode vir a impactar nos lucros dos estabelecimentos. As ocupações podem captar choques idiossincráticos no mercado de trabalho de forma mais específica no que se refere as habilidades de fato requeridas pelos estabelecimentos, uma vez que setor de atividade é uma agregação de ocupações.

O foco principal de Krugman (1991) é captar como os choques impactam nos estabelecimentos, estabelecendo como conclusão de que quanto maior a heterogeneidade do choque maiores serão o *labor pooling* e a concentração, independente desses choques serem analisados por uma ótica setorial ou ocupacional, dado que a variável mais importante é o perfil de habilidades requeridas pelas empresas. Como a dinâmica ocupacional é mais intensa, observa-se choques idiossincráticos mais heterogêneos. Assim, uma medida de *labor pooling* ocupacional é uma variável exógena e se adequa ao modelo originalmente proposto, uma vez que esse modelo prevê que o efeito *do labor pooling* pode ser maior quando os estabelecimentos estão localizados em centros urbanos com muitos trabalhadores com as habilidades requeridas pelas empresas.

Adicionalmente, a modelagem empírica de determinação da concentração ocupacional, representada pela equação (2.18), ao controlar as habilidades dos trabalhadores, permite controlar a causalidade reversa do modelo empírico, isolando o efeito das habilidades sobre a concentração ocupacional. Além disso, é possível verificar se há complementariedade entre os efeitos de *sharing*, *matching* e *learning* nesse modelo quando se faz a interação entre as variáveis *labor pooling* e habilidades. Dessa forma, a estratégia empírica desse ensaio permite obter um efeito mais robusto do *labor pooling* sobre a concentração ocupacional.

2.3.4 Dados

Esta seção tem por objetivo apresentar e descrever as bases de dados utilizadas para identificar o efeito do *labor pooling* sobre a aglomeração ocupacional, bem como analisar os seus principais determinantes. As duas bases de dados a serem utilizadas encontram-se relacionadas, o Relatório Anual de Informações Sociais (RAIS) e a base de dados de Maciente (2013), já que esta última utilizou informações disponíveis na Classificação Brasileira de Ocupações (CBO-2002) para a sua construção, também presente na RAIS. Dessa forma, ressalta-se a importância

e a riqueza de informações disponíveis nas duas bases de dados, para análises associadas ao mercado de trabalho formal do Brasil.

2.3.4.1 RAIS

O Relatório Anual de Informações Sociais (RAIS) do Ministério da Economia (ME) constitui uma das principais fontes de dados do mercado de trabalho formal brasileiro. A RAIS é obtida a partir de um cadastro de registros administrativos, onde estabelecimentos/empresas formais, tanto públicos e privados, devem fornecer informações referentes a cada um dos seus empregados. Essa base de dados abrange todo o território nacional e por isso pode fornecer informações com um maior nível de desagregação, até mesmo em nível municipal, o que contribui ainda mais para a sua relevância. Dessa forma, a RAIS tem como objetivo suprir as necessidades de controle da atividade trabalhista formalizada no país, prover os dados necessários para a preparação de estatísticas do trabalho e fornecer informações às entidades governamentais (BRASIL, 2019).

Além de fornecer informações precisas sobre variáveis socioeconômicas e trabalhistas, a RAIS possibilita o acompanhamento (geográfico, setorial e ocupacional) dos trabalhadores e empresas ao longo dos anos, através de um código de identificação específico para o trabalhador e o estabelecimento, o PIS (ou CPF) e o CNPJ, respectivamente. Utiliza-se, nesta pesquisa, a base de microdados individuais da RAIS, compreendendo um período de 2009 a 2018 para o Brasil.

As variáveis descritas no Quadro 2.1, utilizadas no modelo econométrico, foram obtidas a partir da base de dados da RAIS, exceto as categorias de habilidades que foram obtidas a partir da base de dados de Maciente (2013).

Quadro 2. 1 - Descrição das variáveis

Variável	Descrição	Fonte
Índice de Aglomeração Ocupacional ($INDEX_{kt}$)	Índice de aglomeração da ocupação (k) no período t de acordo com a metodologia de Ellison-Glaeser (1997) adaptado por Gabe e Abel (2009)	RAIS
Labor Pooling (L_{kt})	Média da diferença entre taxa de crescimento dos empregos por ocupação nas empresas e a taxa de crescimento dos empregos totais por ocupação (Medida de variação idiossincrática adaptada a partir da metodologia de Overman e Puga (2010))	RAIS
Indicador de Habilidades ocupacionais (H_{kt}^{h})	Indicador da participação ocupacional do fator de habilidade (21 fatores) obtidos a partir da análise fatorial com dados da O*NET e da CBO de acordo com a metodologia de Maciente (2013)	Maciente (2013) e RAIS
Participação de trabalhadores com ensino superior (ES_{kt})	Número de trabalhadores com ensino superior em cada ocupação dividido pelo total de trabalhadores em cada ocupação	RAIS
Tamanho médio dos estabelecimentos (TE_{kt})	Para cada ocupação obtém-se uma média do tamanho dos estabelecimentos	RAIS
Experiência média dos trabalhadores (EXP_{kt})	Para cada ocupação obtém-se uma média do tempo no emprego (em meses) dos trabalhadores	RAIS
Participação de trabalhadores no setor da agropecuária em relação à indústria (PA_{kt})	Para cada ocupação obtém-se a razão entre o número de trabalhadores pertencentes ao setor da agricultura e o número de trabalhadores do setor industrial	RAIS
Participação de trabalhadores no setor da indústria da extração mineral em relação à indústria (PEM_{kt})	Para cada ocupação obtém-se a razão entre o número de trabalhadores pertencentes ao setor da indústria da extração mineral e o número de trabalhadores do setor industrial	RAIS

Fonte: Elaboração própria (2021)

Para a construção da variável "Índice de aglomeração ocupacional" a pesquisa fez uso da quantidade de trabalhadores em cada ocupação, de acordo com a CBO-2002, desagregando-os por Arranjos Populacionais, conforme já mencionado. Quanto ao indicador da participação ocupacional do fator de habilidade, utilizou-se os fatores de habilidade que compõem cada ocupação, de acordo com Maciente (2013), e o emprego na ocupação em cada ano da pesquisa, a partir dos dados da RAIS, conforme já mencionado.

A variável “Tamanho médio do estabelecimento” por ocupação foi obtida a partir da classificação da RAIS sobre o tamanho das firmas. Na RAIS, a classificação dos estabelecimentos baseada no tamanho utiliza o critério de número de trabalhadores empregados. São considerados micro estabelecimentos aquelas empresas que empregavam entre 1 a 19 trabalhadores, assumindo valor 1; pequenos estabelecimentos entre 20 e 99 trabalhadores, assumindo valor 2; médios estabelecimentos entre 100 e 499 trabalhadores, assumindo valor 3; e grandes estabelecimentos aquelas empresas que empregam mais de 500 trabalhadores, assumindo valor 4.

No que se refere à “Experiência média dos trabalhadores” por ocupação, a variável utilizada, presente na RAIS, foi o tempo no emprego, que registra o tempo do emprego, em meses, no vínculo atual do trabalhador. Logo, a variável tempo no emprego serve como uma *proxy* para experiência. As variáveis “Participação de trabalhadores no setor da agropecuária em relação à indústria” e “Participação de trabalhadores no setor de extração mineral em relação à indústria” por ocupação foram obtidas a partir da CNAE 1.0, observando o setor de atividade dos trabalhadores e calculando essa participação percentual.

2.3.4.2 Categorias de habilidade

A base de dados de Maciente (2013) representa um avanço para as análises empíricas brasileiras, não só na área da Economia do Trabalho e Economia Urbana, pois foi a primeira base de dados elaborada para o Brasil que permite estabelecer categorias de habilidades para os trabalhadores. Até então, no Brasil, só era possível inferir habilidades a partir de variáveis como percentual de trabalhadores com ensino superior, por exemplo. Maciente (2013) rompe com essa limitação e ainda mostra empiricamente, através de uma matriz de correlação entre as categorias de habilidades e níveis de escolaridade, que habilidades e escolaridade não apresentam correlação perfeita. As habilidades cognitivas apresentam forte relação com níveis elevados de escolaridade, enquanto as habilidades motoras e sociais não apresentam correlações elevadas. Isso corrobora a ideia de que habilidade é um termo amplo, com várias dimensões e categorizações, e a escolaridade não pode explicá-la na sua totalidade. Logo, escolaridade e habilidades são variáveis diferentes.

Para a construção desta métrica de habilidades, Maciente (2013) faz uso da *The Occupational Information Network* (ONET) e tenta integrá-la à Classificação Brasileira de Ocupações (CBO-

2002). O primeiro passo feito é equiparar as ocupações brasileiras às ocupações dos EUA e, depois, a partir do conteúdo de informações (ou domínios) que apresentam indicações sobre as habilidades utilizadas em cada ocupação presentes na ONET, aplica-se uma Análise Fatorial destas informações, resultando assim em categorias de habilidades. Por fim, se obtém os escores a serem utilizados nas análises de regressão. Para compreender o procedimento se faz necessário esclarecer como as duas classificações estão organizadas.

A CBO-2002 apresenta uma estrutura de ramificações: Grandes grupos; Subgrupo principal; Subgrupos; Famílias; e Ocupações. A metodologia utilizada por Maciente (2013) dá atenção especial às classificações de 4 e 5 dígitos, ou seja, Família e Ocupações, respectivamente. As Famílias têm 613 classificações, que constituem um conjunto de ocupações similares, e cada família permite, ainda, estabelecer ocupações sinônimas. Desagregando cada Família se obtém 2.557 ocupações. Conforme já mencionado, essas últimas ocupações também possuem sinônimos de 5 dígitos, os quais, por sua vez, são similares aos "títulos de ocupação" presentes na ONET, a partir dos quais a compatibilização de ocupações é feita. Essa compatibilização se faz necessária visto que a CBO não tem nenhuma informação adicional sobre as habilidades exigidas nas ocupações desempenhadas[11].

Convém, ainda, mencionar que a compatibilização não foi realizada para as 2.557 ocupações da CBO, dado que são excluídas do banco de dados as informações sobre o primeiro Grande Grupo: "Membros das forças armadas, policiais e bombeiros militares". Com isso tem-se apenas 2.449 ocupações. Além disso, entre 2002 a 2013 (período em que a base foi desenvolvida), algumas ocupações tiveram seus códigos alterados. Maciente (2013) optou por realizar algumas correspondências entre estas ocupações, que foram modificadas, levando a manter os códigos novos e antigos compatibilizados ao mesmo conjunto de códigos ONET, totalizando 2.702 ocupações na CBO.

Os dados da ONET possuem a finalidade de definir um conjunto de ocupações. É um banco de dados disponibilizado pelo Departamento do Trabalho dos EUA, em 1998, e apresenta 1.090 ocupações (excetuando as militares). Esta base, além de classificar as ocupações, coleta

[11] Para a classificação das ocupações brasileiras, utiliza-se uma metodologia de desenvolvimento curricular, o DACUM (Desenvolvimento de um Currículo), que é definido por um grupo de discussão de especialistas supervisionados. Todavia, estas discussões não avançam no sentido de determinar as habilidades para cada ocupação, devido aos elevados custos envolvidos (MACIENTE, 2013).

informações referentes a essas a partir de seis grupos principais de dados, denominados de domínios, são estes:

1. Características do Trabalhador (Habilidades, Interesses ocupacionais, Valores de Trabalho e Estilos de trabalho);
2. Requisitos do Trabalhador (Habilidades Básicas, Habilidades interfuncionais, Conhecimento e Educação);
3. Requisitos de Experiência (Experiência de Trabalho Relacionada e Treinamento no local ou na fábrica);
4. Requisitos Ocupacionais (Atividades de Trabalho Generalizadas e Contexto de trabalho);
5. Características da força de trabalho;
6. Informações específicas da ocupação.

Com esse rico grupo de informações é possível criar um vetor de habilidades, da seguinte maneira: considerando o domínio de Características do Trabalhador (habilidades), para cada ocupação são listadas 35 habilidades gerais (exemplo: Compreensão de leitura, Matemática, Pensamento crítico, Percepção Social, Seleção de Equipamentos, Operação e Controle, Gestão de recursos humanos, entre outros); um conjunto de analistas define, então, uma escala de valores de precisão que especificam o grau de necessidade da habilidade na ocupação, sendo apresentados, ainda, os erros-padrão de cada estimativa de precisão. Dessa forma, verifica-se a relevância dessa habilidade para a categoria ocupacional (O*NET RESOURSE CENTER, 2018).

Percebe-se que, para o desempenho de uma ocupação, vários tipos de habilidades podem ser requisitados, em maior ou menor intensidades. Neste contexto, aplica-se a Análise Fatorial para determinar os graus de habilidades mais requisitados, esta métrica de habilidade é condicionada às ocupações. Consideradas as características das duas classificações de ocupação, cabe indicar como Maciente (2013) estabeleceu a compatibilização das ocupações. Para cada sinônimo de ocupações foi utilizada, como ponte entre a ONET e a CBO, a *International Classification of Occupations* (ISCO 88). Verificou-se, em alguns casos, que para um código da CBO mais de um código da ONET estaria relacionado e, neste caso, as compatibilizações recebiam um peso maior.

Feitas as compatibilizações e dado o intenso conjunto de informações de habilidade, apresentando grande correlação, optou-se por utilizar um procedimento de Análise Fatorial (MAROCO, 2007; MACIENTE, 2013). Esta técnica foi aplicada às informações das variáveis da ONET, que apresentam diferentes dimensões de domínios, com aproximadamente 264 variáveis iniciais que podem representar habilidades, o que permitiu a Maciente (2013) obter 21 fatores de habilidades, a saber: cognitivas; de manutenção e operação; de assistência; gerenciais; de design e engenharia; de transporte; artísticas; de precisão e automação; de trabalho supervisionadas; de ensino e ciências sociais; força física; de telecomunicação; de independência; em ciências naturais; de atenção; experiência no trabalho; de gestão de conflitos; de trabalho em grupo; de vendas; de monitoramento e conformidade e clericais. Essas habilidades estão descritas na Tabela A.1 do Apêndice A. Os fatores escolhidos correspondem a 90% da variabilidade das 264 variáveis, apresentando um bom ajuste. Os escores obtidos foram, então, aplicados às ocupações brasileiras, dada a correspondência já realizada com as ocupações da ONET.

Todavia, é fundamental apontar as limitações desta métrica de habilidades criada por Maciente (2013). Primeiro, não existe uma correspondência perfeita entre as ocupações dos EUA e as do Brasil. Conforme salientado por Maciente (2013), as semelhanças entre as classificações de ocupações podem não ser perfeitas se forem consideradas as tarefas que dependem de equipamentos ou tecnologias diversas, dado que a estrutura produtiva e organizacional das empresas e os níveis de escolaridade dos trabalhadores são diferentes nos dois países. A segunda limitação está atrelada a suposição de que as características e requisitos do trabalho sejam estáveis no tempo, o que consiste em uma hipótese fraca. Nesse sentido, cabe ressaltar que, foi utilizada a versão 17 da ONET atualizada em 2011. Todavia, grandes mudanças no mercado de trabalho, tanto no que se refere à criação ou destruição de ocupações, quanto à composição dos requisitos do trabalhador para realizar determinadas funções, são bastante lentas. Desta forma, Maciente (2013) indica que estas limitações não invalidam a riqueza de informações que esta métrica de habilidades dispõe.

2.3.4.3 Tratamento dos Dados

A proposta empírica desta pesquisa foi calcular o índice de concentração ocupacional a partir da metodologia proposta por Gabe e Abel (2009) para cada ano compreendido no período de 2009 e 2018, e após esta etapa inicial estabelecer econometricamente os determinantes dessa

concentração através de um painel em nível de ocupações, onde as principais variáveis a serem considerada é o *labor polling*, habilidades e o *labor polling* ponderado por habilidades. Conforme, já mencionado foram utilizadas as bases de dados da RAIS e a base de dados de Maciente (2013) que traz os fatores de habilidades por ocupação para o Brasil. No que se refere a RAIS, apesar das metodologias desta pesquisa considerarem os dados de forma agregada por ocupação, para a obtenção dos índices e variáveis utilizadas foi necessário manusear os microdados da RAIS em nível de indivíduos (PIS) e estabelecimentos (CNPJ). Esse manuseio foi feito anualmente e após isso gerou-se o painel de dados necessário para a segunda etapa da pesquisa.

Considerando que a base de dados da RAIS pode apresentar inconsistências e erros de preenchimento, a limpeza da base foi realizada anualmente. Inicialmente, considerando apenas os trabalhadores com vínculos ativos, foram mantidos na base somente os trabalhadores do setor privado, dado que a dinâmica de determinação local e salarial dos trabalhadores do setor público obedece à critérios peculiares. Além disso, como é comum em análises do mercado de trabalho, considerou-se somente a população em idade ativa, isto é, trabalhadores entre 18 e 65 anos. Adicionalmente, para uma análise mais consistente do mercado de trabalho, no que se refere as condições em que este é estabelecido, exclui-se os trabalhadores com carga horária semanal menor que 20 horas, com tempo no emprego igual a zero meses e aqueles que recebem menos do que o salário mínimo vigente no respectivo ano.

Considerando possíveis erros de preenchimento, exclui-se da base os vínculos que apresentavam PIS inexistente ou igual a zero, remuneração inexistente ou igual a zero, municípios e ocupações sem registro, dado que estas variáveis são fundamentais para obtenção dos resultados dessa pesquisa. A fim de evitar possíveis erros de preenchimentos e ter uma análise mais consistente do mercado de trabalho, foram excluídos da base trabalhadores que possuíam no mesmo ano vínculos (PIS) multiplicados, isto é, que possuíam mais de um vínculo, e inconsistências. Inicialmente, foram excluídos aqueles trabalhadores que possuíam mais de 5 vínculos, pois certamente tratava-se de erro de preenchimentos, posteriormente exclui-se aqueles trabalhadores com gêneros diferentes entre os vínculos (PIS) no mesmo ano. Além disso, foram mantidos os vínculos com mais horas contratadas e com data de admissão mais antiga. Após essas exclusões, objetivando manter apenas um vínculo por trabalhador (PIS), exclui-se aleatoriamente os PIS repetidos na base. Além disso, exclui-se aqueles vínculos cujo CNPJ dos estabelecimentos não possuíam registro (eram inexistentes ou igual a zero).

Por fim, exclui-se aqueles vínculos que estão localizados em municípios que não participam de nenhum Arranjo Populacional ou que não sejam caracterizados como Média Concentração e Grande Concentração, uma vez que, conforme já mencionado, a análise da concentração ocupacional nesta pesquisa considera como subunidades regionais os Arranjos Populacionais, dada a sua importância na dinâmica econômica do país.

Após essas correções, o painel inicial em nível de trabalhadores e estabelecimentos apresentou 260.660.553 observações, referentes a 58.635.427 trabalhadores formais e 4.257.279 estabelecimentos[12]. Convém ainda mencionar que, esta pesquisa possui ampla representatividade para analisar o mercado de trabalho formal nos Arranjos Populacionais brasileiros, uma vez que não gerou uma amostra aleatória, já que os macrodados posteriormente gerados foram obtidos da base de dados completa da RAIS, apenas foram excluídas as inconsistências, os vínculos adicionais ou em tempo parcial e aqueles que não faziam parte de Arranjos Populacionais, conforme já mencionado acima.

Após essa limpeza na base de dados anual, foi possível gerar, através desses microdados anuais, os macrodados considerados nessa pesquisa, que foram obtidos a partir de manipulações algébricas (somatórios, variações percentuais, médias, etc.) desses microdados por ocupação, Arranjo Populacional e também por estabelecimentos. Após essa etapa, as bases anuais foram agregadas, possibilitando formar um painel de dados não balanceado com 25.037 observações, referente a 2.590 ocupações para o período 2009 a 2018.

2.4 *LABOR POOLING* OCUPACIONAL E HABILIDADES DOS TRABALHADORES: O MERCADO DE TRABALHO FORMAL BRASILEIRO (2009-2018)

O mercado de trabalho formal brasileiro apresentou, entre o período de 2009 e 2018, uma certa estabilidade, o que já era esperado, dado que as maiores oscilações no emprego no Brasil estão concentradas no mercado informal. O número de trabalhadores formais considerados nesta

[12] Em nível anual, após as correções feitas obteve-se o seguinte número de observações: em 2009, 23.486.109; em 2010, 25.586.993; em 2011, 26.910.983; em 2012, 27.925.934; em 2013, 28.683.625; em 2014, 28.908.122; em 2015, 27.574.673; em 2016, 24.570.906; em 2017, 25.170.161; e em 2018, 21.843.047.

pesquisa[13] oscilou entre 29.277.891 (2018) e 36.944.478 (2014). Entre 2009 e 2018, houve uma variação de -3,22% no total de trabalhadores formais (RAIS, ME).

Ao longo dos anos, algumas ocupações não observadas em 2009 foram contabilizadas em 2018: um total de 181 ocupações. Esse cenário mostra grande dinamismo na estrutura ocupacional do mercado de trabalho formal brasileiro. Dessas ocupações, 10 foram incluídas no grupo (1) Membros superiores do poder público, dirigentes de organizações de interesse público e de empresas, gerentes (exemplo, Tecnólogo em gestão hospitalar), 123 no grupo (2) Profissionais das ciências e das artes (exemplo, Engenheiro de controle e automação), 16 no grupo (3) Técnicos de nível médio (exemplo, Doula), 2 no grupo (4) Trabalhadores de serviços administrativos (exemplo, Despachante de trânsito); 9 no grupo (5) Trabalhadores dos serviços, vendedores do comércio em lojas e mercados (exemplo, Agente indígena de saúde); 8 no grupo (6) Trabalhadores agropecuários, florestais e da pesca (exemplo, Criador de jacarés) e 13 do grupo (7) Trabalhadores da produção de bens e serviços industriais (I) (exemplo, Bloqueiro (trabalhador portuário)).

A Tabela 2.1 mostra as 10 ocupações que apresentaram maior crescimento, desde 2009. Todas as ocupações apresentadas tiveram um crescimento superior a 100%, com grande destaque para "Operador de máquina de preparação de matéria prima para produção de cigarros", cuja habilidade mais requisitada para o desempenho da ocupação é a Habilidade de Gestão de conflitos. De modo geral, percebe-se que, dado esse grande crescimento das ocupações, o perfil de habilidades no mercado de trabalho brasileiro tem apresentado alterações, com grande solicitação das habilidades de Gestão de conflitos, Gerenciais, Vendas, Ensino e ciências sociais, Ciências naturais e Força física.

[13] Considera-se apenas os trabalhadores do setor privado com idades entre 18 e 65 anos, carga horária semanal maior do que 20 horas semanais e remuneração positiva acima do salário mínimo vigente em cada ano.

Tabela 2. 1 - Crescimento do emprego ocupacional e perfil de habilidades (2009-2018)

Ocupação	Crescimento do emprego (%)	Habilidade
Operador de máquina de preparação de matéria prima para produção de cigarros	4.750	Gestão de conflitos
Especialista de políticas públicas e gestão governamental – EPPGG	4.100	Gerenciais
Sushiman	2.754,05	Gerenciais
Intérprete de língua de sinais	2.360	Vendas
Biomédico	2.087,06	Ensino e ciências sociais
Classificador de grãos	1.607,52	Ciências naturais
Monitor de dependente químico	1.417,14	Ensino e ciências sociais
Petrógrafo	1.400	Ciências naturais
Pizzaiolo	1.262,69	Força física
Analista de planejamento e orçamento – APO	1.243,96	Gerenciais

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

Todavia, não é possível estabelecer o perfil da estrutura ocupacional e intensidade de habilidades requeridas através desse crescimento. Para traçar um perfil da estrutura ocupacional no Brasil, deve-se verificar as ocupações que tem maior representatividade, isto é, as ocupações mais expressivas no período de análise (percentual no emprego total), além de verificar o perfil de habilidades e a distribuição espacial por macrorregiões dessas ocupações. Essas informações estão expostas nas Tabelas 2.2 e 2.3.

Verifica-se que a ocupação “Vendedor de comércio varejista” tem a maior participação no mercado de trabalho brasileiro, com grande requisição de habilidades de vendas. Destaca-se ainda, “Auxiliar de escritório, em geral”, “Assistente administrativo”, “Faxineiro”, “Alimentador de linha de produção”, “Operador de caixa”, “Servente de obras”, “Motorista de caminhão”, “Vigilante” e “Porteiro de edifícios”. Desse modo, percebe-se que a estrutura ocupacional brasileira é pouco complexa em habilidades cognitivas, destacando-se habilidades menos complexas como: Vendas, Independência, Precisão e automação, Força física, Manutenção e operação, Transporte e Gestão de conflitos (TABELA 2.2).

Tabela 2. 2 - Participação ocupacional no emprego total e perfil de habilidades (2009-2018)

Ocupação	Participação no emprego (%)	Habilidades
Vendedor de comércio varejista	5,63	Vendas
Auxiliar de escritório, em geral	4,47	Independência
Assistente administrativo	3,47	Precisão e automação
Faxineiro	2,86	Força física
Alimentador de linha de produção	2,32	Manutenção e operação
Operador de caixa	2,25	Vendas
Servente de obras	2,11	Força física
Motorista de caminhão (rotas regionais e internacionais)	2,03	Transporte
Vigilante	1,90	Gestão de conflitos
Porteiro de edifícios	1,71	Força física

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

Tabela 2. 3 - Distribuição regional das ocupações com maior participação no emprego total (2009-2018)

Ocupação	Norte (%)	Nordeste (%)	Sudeste (%)	Sul (%)	Centro-Oeste (%)
Vendedor de comércio varejista	4,67	18,06	53,7	16,44	7,14
Auxiliar de escritório, em geral	3,74	14,11	58,4	16,3	7,45
Assistente administrativo	3,59	14,93	58,9	15,34	7,25
Faxineiro	4,26	11,63	62,01	13,83	8,26
Alimentador de linha de produção	3,07	9,43	56,88	25,37	5,24
Operador de caixa	5,06	16,39	55,01	15,37	8,17
Servente de obras	6,44	28,15	46,33	11,32	7,76
Motorista de caminhão (rotas regionais e internacionais)	3,58	12,38	53,94	21,55	8,54
Vigilante	7,27	21,68	48,45	13,22	9,39
Porteiro de edifícios	3,29	19,25	61,46	9,76	6,24

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

No que se refere a distribuição regional das ocupações mais expressivas do Brasil, para o período de análise, verifica-se que essa obedece a distribuição geral do mercado de trabalho no Brasil, onde a maioria dos trabalhadores tem seus postos de trabalho na macrorregião Sudeste. De acordo com dados da RAIS, para o período compreendido entre 2009 e 2018, 56,63% dos trabalhadores brasileiros encontram-se na região Sudeste, sendo seguidas pela região Sul (16,05%), Nordeste (15,99%), Centro-Oeste (7,19%) e Norte (4,15%). Categorizando os trabalhadores por ocupações, observa-se que as ocupações “Auxiliar de escritório, em geral”,

"Assistente administrativo", "Faxineiro", "Alimentador de linha de produção" e "Motorista de caminhão" seguem a mesma tendência. Nas demais ocupações, a maioria dos trabalhadores tem seus postos de trabalho nas regiões Sudeste e Nordeste, respectivamente (TABELA 2.3).

É importante salientar que, a dinâmica do mercado de trabalho brasileiro, no período considerado, provocou uma queda no nível de emprego geral ao mesmo tempo em que estimulou o surgimento de novas ocupações. Esse contexto permite que choques idiossincráticos no mercado de trabalho aconteçam com maior intensidade. Esses choques captam mudanças no mercado de trabalho em termos de contratação e desligamento de trabalhadores por perfis ocupacionais. A variável *labor pooling* ocupacional capta esses choques, isto é, permite verificar como os estabelecimentos contratam ou demitem trabalhadores com relação à média de contratação ou desligamento geral dessa ocupação. A Tabela 2.4 mostra as 10 ocupações com maior *labor pooling* médio para o período de 2009 a 2018. Convém mencionar que um maior *labor pooling* indica maior dinamismo das ocupações, tanto no que se refere a contratações como a desligamentos. Logo, valores maiores da variável *labor pooling* indicam que os estabelecimentos estão expandindo o emprego em determinada ocupação enquanto o emprego total dessa ocupação está se reduzindo e/ou vice-versa. Dessa forma, não necessariamente as ocupações que mais cresceram terão um maior *labor pooling*. Essa evidência pode ser notada ao comparar as informações contidas nas Tabelas 2.1 e 2.4.

Verifica-se que a ocupação "Agente indígena de saúde" é a que apresenta o maior *labor pooling* médio para o período considerado, e a maior requisição de habilidade para o desempenho dessa ocupação é Ciências naturais. Seguindo o mesmo perfil de habilidade, tem-se a ocupação "Agente indígena de saneamento", ocupando o terceiro maior *labor pooling* médio, e "Técnico de enfermagem da estratégia de saúde da família", na oitava posição. Destaca-se, ainda, a ocupação de "Maquiador de caracterização", com o segundo maior *labor pooling* médio, cuja habilidade intensiva é Artística. Observa-se, ainda, que as habilidades de Assistência são primordiais para as ocupações "Socorrista (exceto médicos e enfermeiros)" e "Empregado doméstico nos serviços gerais", as quais estão na sexta e décima posição entre as ocupações com maior *labor pooling*. As demais ocupações com maior labor *pooling* apresentam perfis de habilidades distintos, a saber: Monitoramento e conformidade, Vendas, Transporte e Força Física. Assim, de modo geral, observa-se que as ocupações com maior *labor pooling* não obedecem a um padrão claro no que se refere às habilidades requeridas mais intensivas (TABELA 2.4).

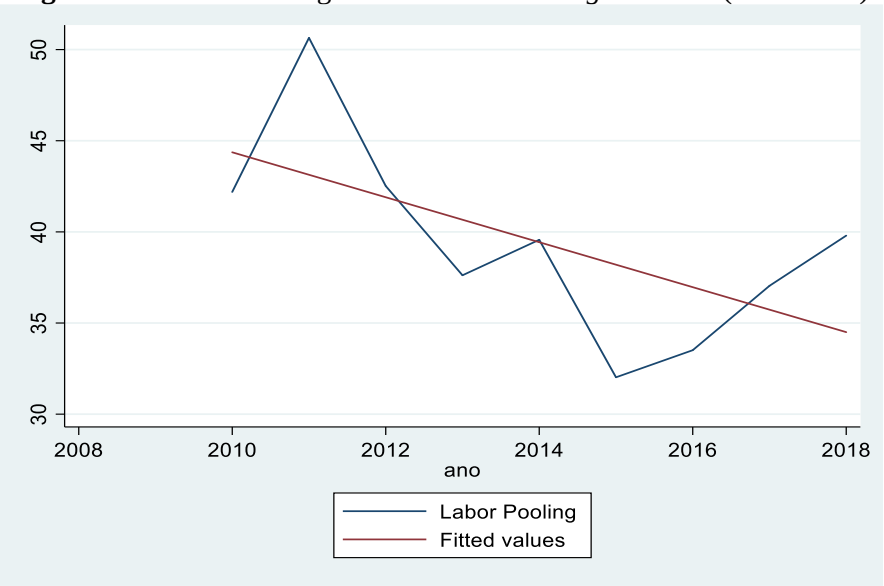
Tabela 2. 4 - *Labor Pooling* e perfil de habilidades (2009-2018)

Ocupação	***Labor Pooling***	Habilidade
Empregado doméstico nos serviços gerais	258,11	Assistência
Pescador artesanal de lagostas	269,33	Força Física
Técnico de enfermagem da estratégia de saúde da família	391,72	Ciências naturais
Trabalhador da exploração de cipós produtores de substâncias aromáticas, medicinais e tóxicas	400	Transporte
Socorrista (exceto médicos e enfermeiros)	469,97	Assistência
Esteticista	492,1	Vendas
Agente de inteligência	570,43	Monitoramento e conformidade
Agente indígena de saneamento	1.013,91	Ciências naturais
Maquiador de caracterização	1.029,54	Artísticas
Agente indígena de saúde	1.034,71	Ciências naturais

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

O comportamento tendencial do *labor pooling* no Brasil, no período considerado, pode ser ilustrado pelas Figuras 2.2 e 2.3. Os resultados indicam que houve uma diminuição média do *labor pooling* no Brasil entre 2010 e 2018 (FIGURA 2.2). Todavia, verifica-se que os choques idiossincráticos ocupacionais oscilaram bastante no período considerado, refletindo grande dinamismo do mercado de trabalho ocupacional no que se refere à contratações e desligamentos. É importante mencionar que, um maior *labor pooling* facilita os ajustes no mercado de trabalho ocupacional sem pressões sobre os lucros das empresas, permitindo maior aproveitamento das externalidades positivas da aglomeração. De modo geral, constata-se que, apesar da grande tendência de queda até 2015, o *labor pooling* ocupacional quase retorna aos patamares iniciais em 2018.

Figura 2. 2 - Tendências gerais da *Labor Pooling* no Brasil (2009-2018)



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

O *labor pooling* apresenta um aumento, em 2011, fruto da recuperação econômica após a crise de 2009, mudando de tendência a partir de 2012. Somente, em 2016, observa-se uma tendência crescente. Cabe considerar que, no Brasil, no ano de 2014, tem-se início uma crise econômica, onde o produto brasileiro caiu entre 2014 e 2016 (IBGE, 2021). Considerando a variação anual do PIB no Brasil, para o período de 2009 a 2018, de acordo com Instituto Brasileiro de Geografia e Estatística, Sistema de Contas Nacionais (IBGE/SCN Anual)[14], pode-se inferir que, o *labor pooling* no Brasil está captando com mais intensidade as demissões nos períodos de crise, isto é, os choques negativos no mercado de trabalho, pois a partir de 2014 a tendência comportamental do *labor poling* é a mesma da variação anual do PIB brasileiro (IBGE, 2021). Não obstante, de 2010 a 2013, o *labor pooling* poderia estar captando com mais intensidade as contratações. Isso indica grande uma sensibilidade do *labor pooling* à dinâmica do mercado de bens e serviços em períodos de recessões econômicas.

Ao analisar a Figura 2.3, constata-se que os choques idiossincráticos ocupacionais são bastantes heterogêneos quando se considera a sua interação com o indicador de intensidade de uso do fator de habilidade nas ocupações. Assim, a Figura 2.3 indica como a evolução da intensidade do uso de cada habilidade no emprego ocupacional impacta nos choques idiossincráticos ocupacionais, ao se considerar o uso específico de determinada habilidade. Observa-se que o

[14] https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9300-contas-nacionais-trimestrais.html?=&t=series-historicas&utm_source=landing&utm_medium=explica&utm_campaign=pib#evolucao-taxa

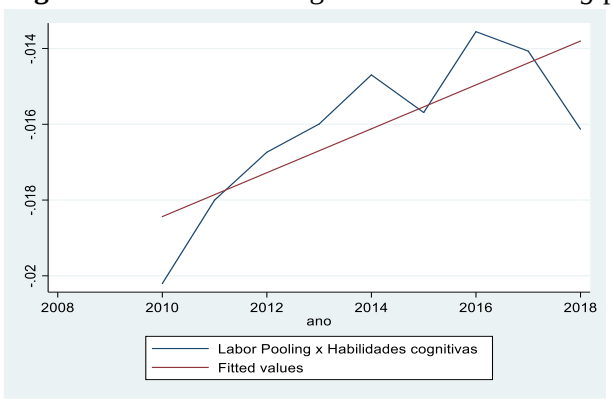
intervalo de variação do *labor pooling* interagido com o indicador de participação de habilidades é menor quando comparado com o intervalo de variação do *labor pooling* isolado (FIGURA 2.2). Essa tendência se dá porque, nessas circunstâncias, a interação capta movimentos ou choques mais específicos no mercado de trabalho.

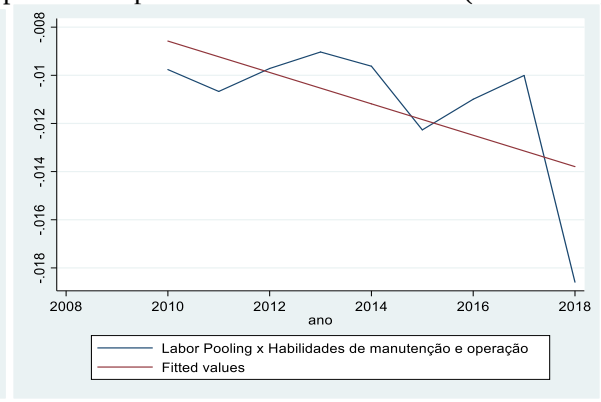
Os resultados indicam que, entre 2010 a 2018, houve aumento no *labor pooling* associado às habilidades cognitivas, gerenciais, artísticas, de precisão e automação, de trabalho supervisionado, de ensino em ciências sociais, de telecomunicação, de independência, de ciências naturais, de atenção, em gestão de conflito, de trabalho em grupo, de vendas e clericais. Esse comportamento indica que choques idiossincráticos no Brasil são mais intensos para empregos intensivos nessas habilidades em detrimentos das demais. Nesse sentido, as externalidades positivas podem ser mais bem aproveitadas para mercados de trabalho que fazem uso mais intenso dessas habilidades, dado que quanto maior o *sharing* maiores serão os benefícios da aglomeração ocupacional.

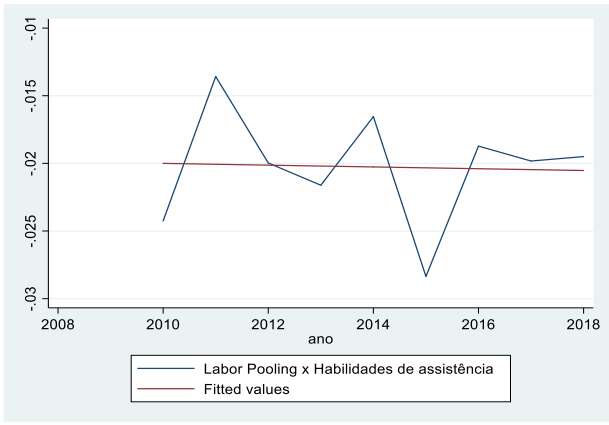
No entanto, esses resultados precisam ser interpretados com cuidado, pois além do *labor pooling* captar contratações e/ou desligamentos no mercado de trabalho, a evolução tendencial dos indicadores de participação de habilidade no Brasil não indica que essas habilidades são as maiores demandas, mas são aquelas em que os choques no mercado de trabalho (positivos e/ou negativos) afetaram de forma mais intensa.

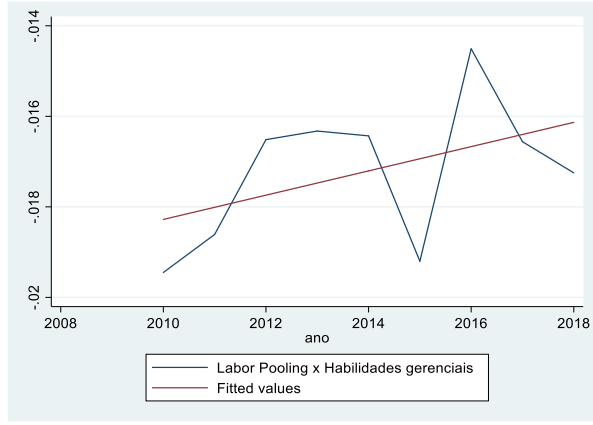
É importante mencionar que, em 2014, os choques idiossincráticos no mercado de trabalho ocupacional apresentaram uma diminuição para a maioria das categorias de habilidades (FIGURA 2.3), exceto para as habilidades de telecomunicação, de independência, de experiência, de gestão de conflito, de trabalho em grupo e de vendas. Maciente (2012) indica que os fatores ocupacionais de habilidades mais utilizados no Brasil são Vendas, Gestão de conflitos, Força física, Assistência e saúde, Ensino e ciências sociais, Telecomunicações e Transporte. Dessa forma, verifica-se que os choques idiossincráticos foram menos intensos nas categorias de habilidade mais utilizadas no Brasil no período de crise. Isto é, os efeitos da crise foram minimizados para essas categorias de habilidades, dado que essas estão entre os fatores mais utilizados na estrutura produtiva brasileira.

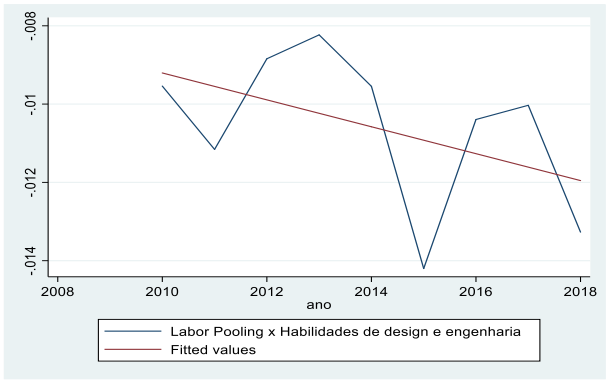
Figura 2. 3 - Tendências gerais do *Labor Pooling* ponderado pelas Habilidades no Brasil (2009-2018)

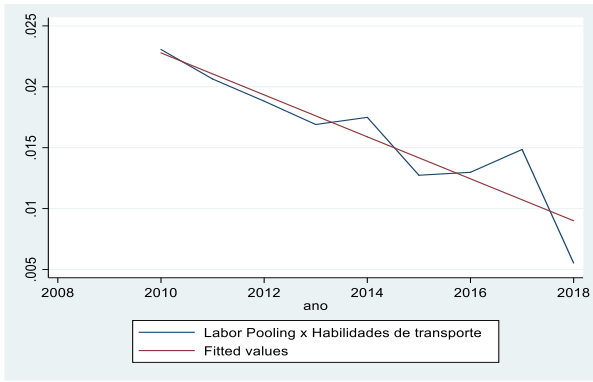


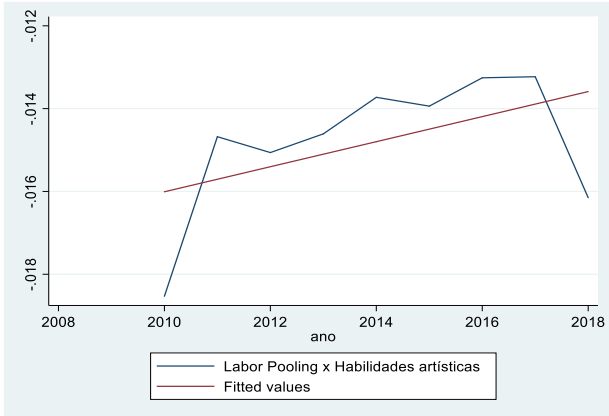


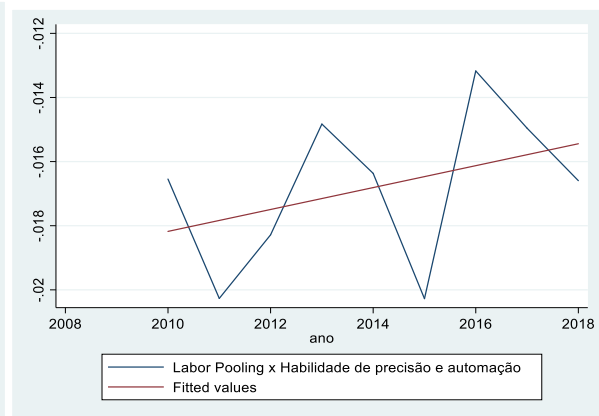






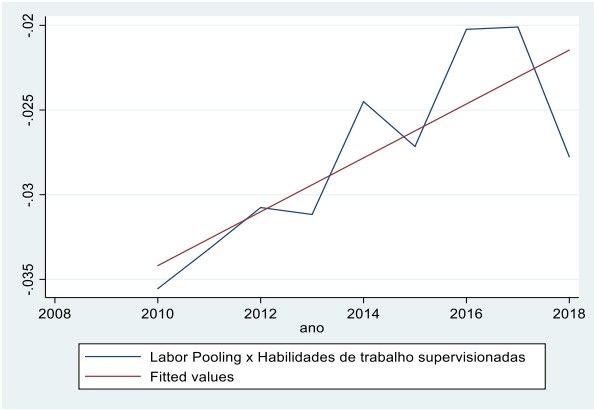


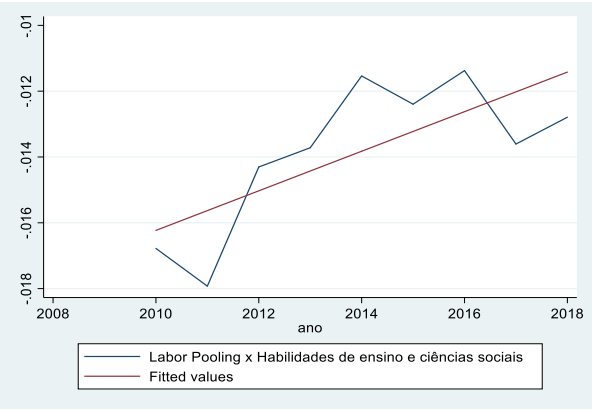


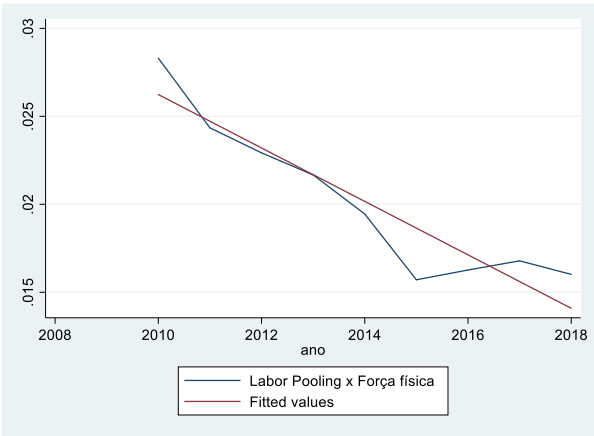


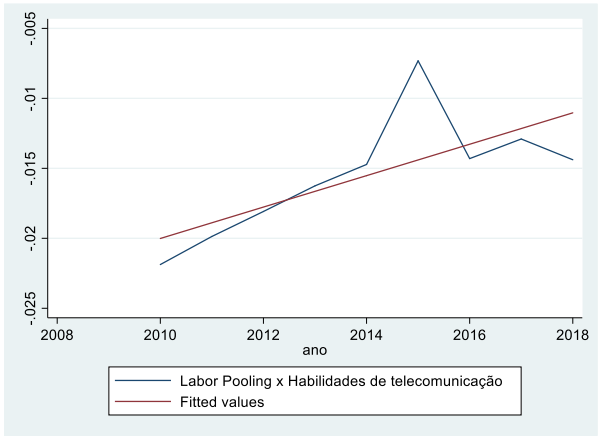
(continua)

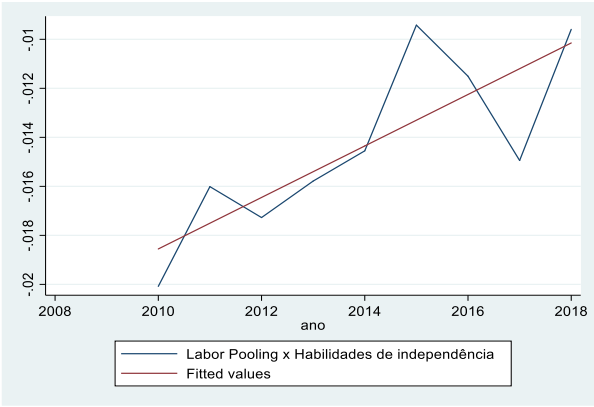
(continuação)

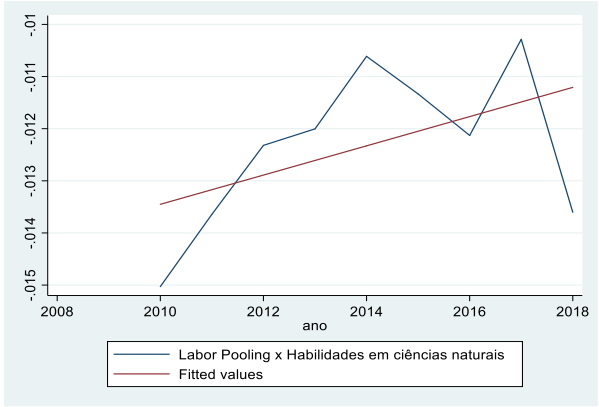


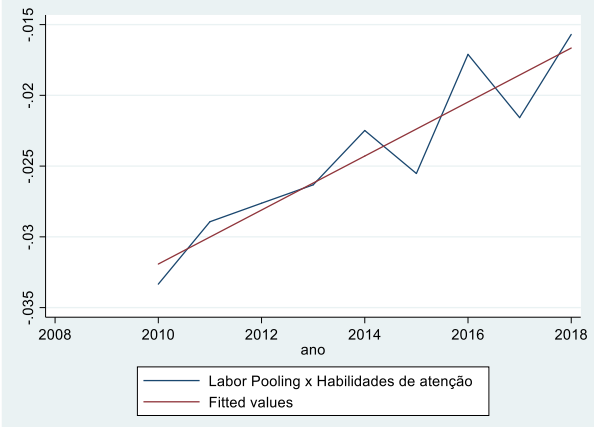


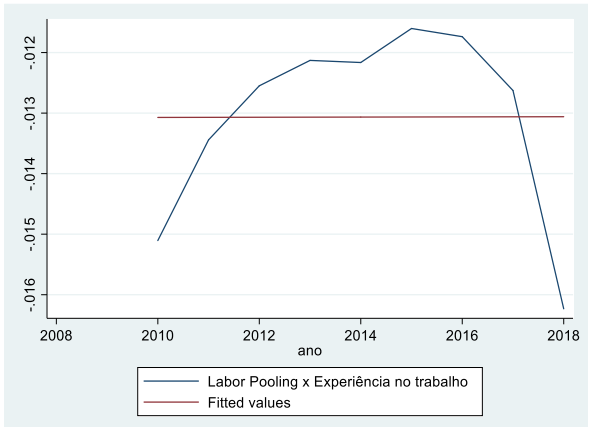






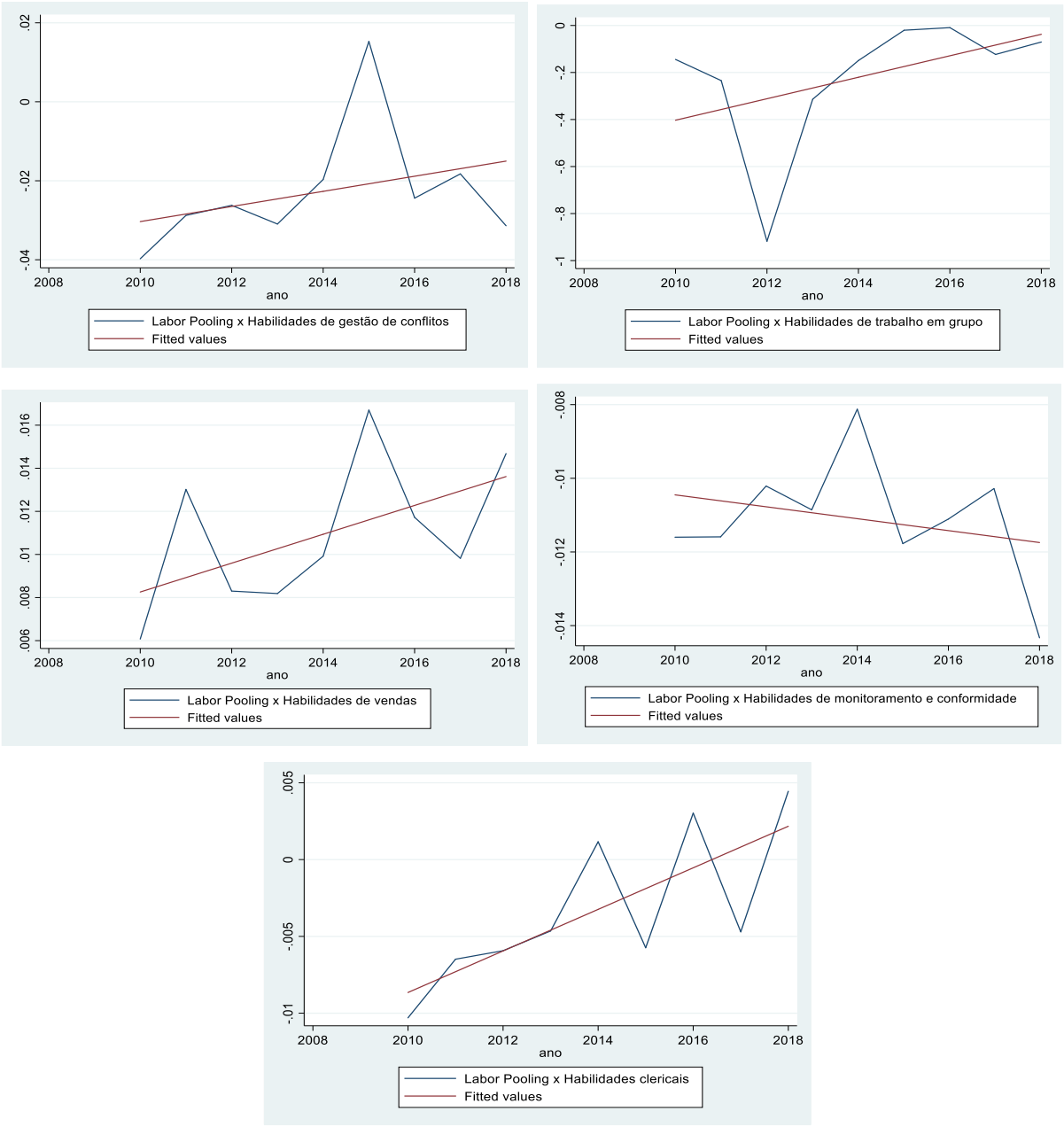






(continuação)

(conclusão)



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

2.5 RESULTADOS

2.5.1 Padrões de aglomeração ocupacional no Brasil

Conforme mencionado na metodologia o índice de aglomeração ocupacional é indicado para verificar o comportamento espacial dos trabalhadores categorizados por ocupações. É importante salientar que o valor deste índice depende da subunidade regional considerada e da categorização ocupacional. Esta pesquisa utilizou como subunidade geográfica os Arranjos

Populacionais brasileiros delimitados pelo IBGE (2016), um total de 376 subunidades geográficas e, como categoria ocupacional, as 2.557 ocupações categorizadas pela CBO (2002). Este índice foi calculado anualmente de 2009 a 2018.

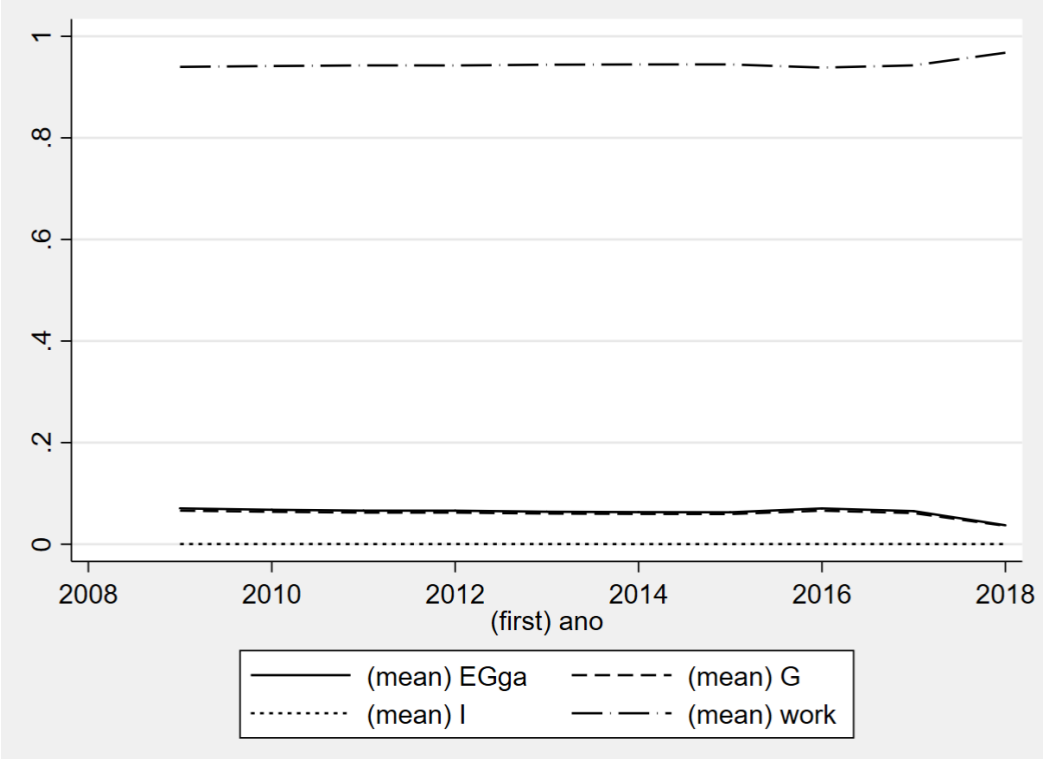
A Figura 2.4 ilustra o comportamento tendencial do índice de aglomeração ocupacional para o Brasil no período considerado. Como se trata de uma média anual, a Figura 2.4 não serve para fazer inferências acerca da concentração por ocupação, entretanto, tem a finalidade de verificar o comportamento tendencial da concentração ocupacional no Brasil, verificando a dinâmica das partes componentes do índice ao longo do tempo.

No painel A da Figura 2.4 é possível ver o comportamento tendencial do índice e de todos os seus componentes. De modo geral observa-se um comportamento constante do componente I_t, indicando que a distribuição das ocupações entre os setores de atividade econômica se manteve o mesmo ao longo dos anos. No que se refere ao componente *"work"*, que reflete o emprego total nos Arranjos Populacionais, pela expressão $(1 - \sum_{i=1}^{n} {t_{it}}^2)$, verifica-se que este se manteve praticamente constante até 2017, apresentando um leve aumento em 2018, isso reflete uma pequena diminuição no emprego total dos Arranjos Populacionais no Brasil. No que se refere a concentração bruta G_t, observa-se uma leve diminuição em 2018. O índice de Ellison-Glaeser (1997) adaptado por Gabe e Abel (2009) acompanha essa tendência, pois conforme já mencionado, o índice final tem maior determinação pelo índice de concentração bruta ($G_t = \sum_{i=1}^{n}(s_{it} - t_{it})^2$).

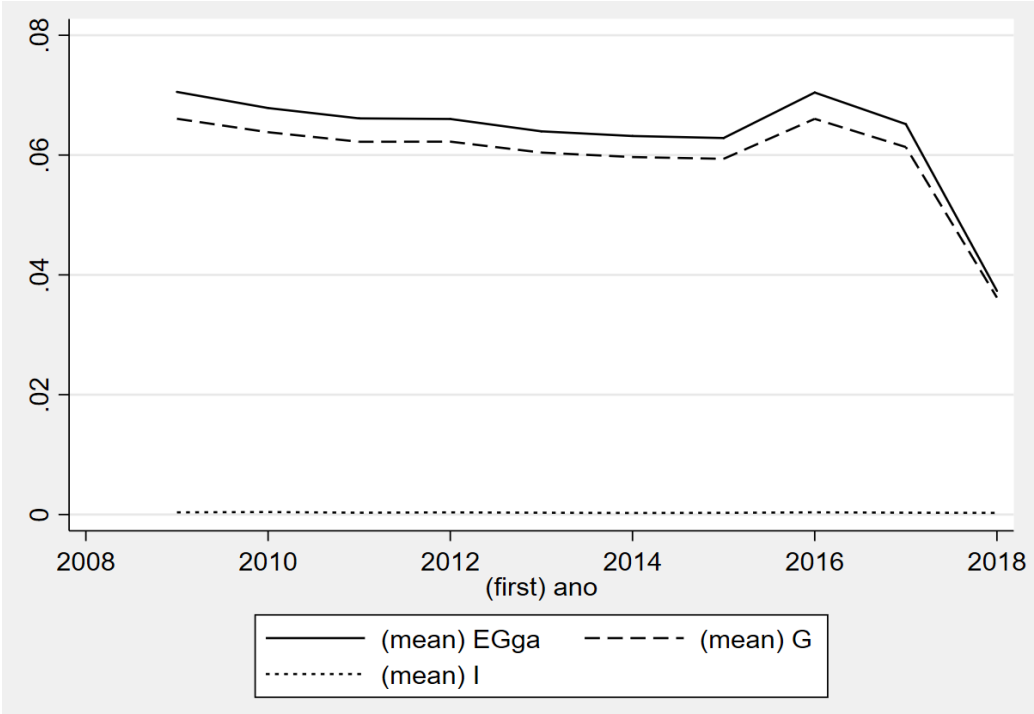
No painel B foi excluído do gráfico o componente *"work"*, cuja escala é maior, a fim de permitir uma visualização mais clara da tendência comportamental entre o índice bruto e o índice final. Nessa escala é evidente que o índice Ellison-Glaeser (1997), adaptado por Gabe e Abel (2009), está acima do índice de concentração bruta G, ou seja, a concentração bruta geral estava sendo subestimada, e o componente *"work"* teve maior peso na ponderação final do índice, controlando a concentração ocupacional pelo tamanho do mercado de trabalho das subunidades geográficas consideradas. A redução no índice de concentração bruta e, consequentemente, no índice de Ellison-Glaeser (1997) adaptado por Gabe e Abel (2009), para o ano de 2018, indica que a participação do emprego ocupacional dos Arranjos Populacionais brasileiros está convergindo para a participação no emprego total dos mesmos Arranjos.

Figura 2. 4 – Tendências gerais da concentração ocupacional entre 2009 e 2018

Painel A



Painel B



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se os trabalhadores em Arranjos Populacionais e do setor privado.

Convém mencionar que, na abordagem de Maciente (2013), realizada a fim de captar o índice de concentração setorial no Brasil, no período de 1994 a 2010, o índice bruto esteve acima do índice final, e esse comportamento foi gerado a partir do índice de Herfindahl, cuja tendência

era descendente. Desse modo, o autor observou que, considerando a concentração setorial na análise do comportamento anual do índice e de seus componentes, o componente G_t tem maior peso na determinação do índice total. Os demais componentes, conforme estabelecido na elaboração do indicador originalmente desenvolvido por Ellison-Glaeser (1997), têm por finalidade ponderar o índice bruto de modo que este não seja superestimado. Esse comportamento também foi observado para a análise de concentração ocupacional para o Brasil realizada nessa pesquisa.

Para fornecer uma análise mais detalhada do índice de concentração ocupacional de Ellison-Glaeser (1997), adaptado por Gabe e Abel (2009), o comportamento do indicador será avaliado para as 25 ocupações menos concentradas e as 25 ocupações mais concentradas, para os anos inicial e final do período de análise considerado (2009 e 2018).

A Tabela 2.5 mostra as ocupações menos concentradas para o ano de 2009. De modo geral, percebe-se que, o índice de concentração ocupacional é praticamente zero para as 25 ocupações consideradas. Isso indica total dispersão destes trabalhadores nos Arranjos Populacionais Brasileiros. A ocupação menos concentrada no Brasil para o ano de 2009 é "Trabalhador na cultura de mamona", seguido por "Trabalhador da exploração de pinhão" e "Trabalhador da exploração de Ouricuri", ocupações tipicamente agrícolas e que requerem baixo nível de escolaridade. Estas características são comuns para a maioria das ocupações menos concentradas, conforme pode ser visualizado na Tabela 2.5. Todavia, é importante salientar que entre as 25 ocupações menos concentradas, para este ano, encontra-se também ocupações que não estão relacionadas à atividades agrícolas, como: "Físico (térmica)", "Cirurgião dentista - reabilitador oral", "Cirurgião dentista – odontogeriatra", "Físico (matéria condensada)", "Médico legista" e "Professor de zootecnia do ensino superior". Essas ocupações requerem ensino superior e alto grau de complexidade.

Tabela 2. 5 - Ocupações menos concentradas para o ano de 2009

Código da Ocupação	Ocupação	Índice de Concentração (EGga)
622725	Trabalhador na cultura de mamona	0,000000222
632415	Trabalhador da exploração de pinhão	0,00000128
632355	Trabalhador da exploração de Ouricuri	0,00000133
213175	Físico (térmica)	0,00000174
223264	Cirurgião dentista - reabilitador oral	0,00000199
612740	Produtor da cultura de soja	0,00000615
613410	Criador de animais produtores de veneno	0,00000643
612620	Produtor de fumo	0,00000735
623320	Trabalhador da cunicultura	0,0000112
622705	Trabalhador na cultura de amendoim	0,0000281
316335	Desincrustador (poços de petróleo)	0,0000318
213435	Petrógrafo	0,0000318
223228	Cirurgião dentista – odontogeriatra	0,0000328
213140	Físico (matéria condensada)	0,0000357
632340	Trabalhador da exploração de malva (pãina)	0,0000395
612625	Produtor de guaraná	0,0000415
223137	Médico legista	0,0000477
612810	Produtor de plantas aromáticas e medicinais	0,0000479
742135	Confeccionador de órgão	0,0000692
234460	Professor de zootecnia do ensino superior	0,0000761
631320	Criador de ostras	0,0000924
631315	Criador de mexilhões	0,0000933
622720	Trabalhador na cultura de dendê	0,0001072
632510	Trabalhador da exploração de cipós produtores de substâncias aromáticas, medicinais e tóxicas	0,000111
632420	Trabalhador da exploração de pupunha	0,0001201

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se os trabalhadores em Arranjos Populacionais e do setor privado.

No que se refere as ocupações mais concentradas para o ano de 2009, 25 delas estão listadas na Tabela 2.6. A ocupação mais concentrada é “Vendedor de comércio varejista”, cujo índice chega a “5,668944”, outras ocupações como “Trabalhador da cultura de cana-de-açúcar” e “Auxiliar de escritório, em geral” apresentam índice maior do que 1. Percebe-se que as ocupações mais concentradas não apresentam similaridade das características dos trabalhadores, tem-se trabalhadores que desenvolvem atividades de serviços, agrícolas e administrativas, respectivamente, e observa-se que essas requerem diferentes níveis de escolaridade. As ocupações mais concentradas indicam que os trabalhadores estão aglomerados em poucos Arranjos Populacionais. As 25 ocupações mais concentradas apresentadas na Tabela 2.6 são bem diversificadas, entretanto convém citar as ocupações “Costureiro, a máquina na confecção em série” e “Costureiro na confecção em série” que são típicas do setor de produção têxtil. Nos EUA, conforme relatado por Gabe e Abel (2009), entre as ocupações mais

concentradas estão aquelas com aspectos da produção têxtil. Neste aspecto, cita-se também Ellison e Glaeser (1997) e Duranton e Overman (2005) que na análise da aglomeração setorial verificaram que as indústrias têxteis se encontravam entre as mais aglomeradas. Assim, verifica-se que o Brasil acompanha essa tendência.

Tabela 2. 6 - Ocupações mais concentradas para o ano de 2009

Código da Ocupação	Ocupação	Índice de Concentração (EGga)
992225	Auxiliar geral de conservação de vias permanentes (exceto trilhos)	0,1378519
513205	Cozinheiro geral	0,1446235
415205	Carteiro	0,1456678
354125	Assistente de vendas	0,1463227
514320	Faxineiro	0,1498832
848505	Abatedor	0,1640167
641015	Tratorista agrícola	0,170661
763215	Costureiro, a máquina na confecção em série	0,170981
521135	Frentista	0,2044441
521125	Repositor de mercadorias	0,2055182
622505	Trabalhador no cultivo de árvores frutíferas	0,2453942
421125	Operador de caixa	0,248203
717020	Servente de obras	0,2628658
763210	Costureiro na confecção em série	0,3020355
514225	Trabalhador de serviços de limpeza e conservação de áreas públicas	0,3415658
421105	Atendente comercial (agência postal)	0,3671418
848520	Magarefe	0,408939
784105	Embalador, a mão	0,4411148
412205	Contínuo	0,5807176
764005	Trabalhador polivalente da confecção de calçados	0,6546536
782510	Motorista de caminhão (rotas regionais e internacionais)	0,6590222
784205	Alimentador de linha de produção	0,9452128
411005	Auxiliar de escritório, em geral	1,006771
622110	Trabalhador da cultura de cana-de-açúcar	1,752206
521110	Vendedor de comércio varejista	5,668944

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se os trabalhadores em Arranjos Populacionais e do setor privado.

Considerando o ano de 2018, verifica-se, na Tabela 2.7, que o índice de concentração para as 25 ocupações menos concentradas se manteve com valor de praticamente zero, mas com muitas alterações entre as ocupações quando comparadas com o ano de 2009, o que já era esperado dado o dinamismo que pode ocorrer no mercado de trabalho após nove anos. A ocupação menos concentrada em 2018 é “Catador de caranguejos e siris”, sendo seguido por “Preparador de fumo na fabricação de charutos” e “Trabalhador da exploração de piaçava”, ocupações que em 2009 não estavam entre as 25 ocupações menos concentradas. Apenas as ocupações

“Trabalhador da cunicultura”, “Trabalhador na cultura de amendoim” e “Trabalhador da exploração de malva (pãina)” mantiveram-se entre as ocupações menos concentradas em 2018. Como em 2009, entre as ocupações menos aglomeradas a maioria é tipicamente agrícola, todavia, ocupações como “Procurador do estado”, “Subprocurador-geral da república”, “Tecnólogo em rochas ornamentais”, “Cirurgião dentista - protesiólogo bucomaxilofacial”, “Professor de literatura comparada”, “Promotor de justiça” e “Médico hiperbarista” são atreladas à indústria e setores diversos, as quais exigem maior nível de escolaridade e complexidade das funções desempenhadas.

Quanto às 25 ocupações mais concentradas de 2018, permanece a ocupação “Vendedor de comércio varejista” como a mais concentrada, tal como em 2009, todavia, apresentando um índice bem menor (2,943979), o que corrobora a Figura 2.4, indicando uma tendência decrescente para a concentração ocupacional anual no Brasil. Os demais índices de concentração tiveram valor abaixo de 1. Convém mencionar que entre as ocupações mais concentradas poucas mudanças ocorreram, apenas nove novas ocupações foram substituídas no intervalo dos nove anos considerados na análise (“Trabalhador na produção de mudas e sementes”, “Auxiliar de corte (preparação da confecção de roupas)”, “Marinheiro de convés (marítimo e fluviário)”, “Armazenista”, “Trabalhador de extração florestal, em geral”, “Motorista operacional de guincho”, “Trabalhador volante da agricultura”, “Assistente administrativo” e “Preparador de calçados”), isto é, 16 ocupações permaneceram entre as mais concentradas no intervalo de tempo considerado, mas com mudanças nas ordens e na magnitude dos índices, como pode ser visualizado a partir de comparações entre as Tabelas 2.6 e 2.8. Como poucas alterações ocorreram, permaneceram entre as ocupações mais concentradas aquelas do setor têxtil, característica que também ocorre nos EUA, como mencionado acima.

Tabela 2. 7 - Ocupações menos concentradas para o ano de 2018

Código da ocupação	Ocupação	Índice de Concentração (EGga)
631005	Catador de caranguejos e siris	0,00000000359
842205	Preparador de fumo na fabricação de charutos	0,000000505
632365	Trabalhador da exploração de piaçava	0,000000571
516705	Astrólogo	0,000000784
241220	Procurador do estado	0,000000791
842225	Celofanista na fabricação de charutos	0,000000815
613210	Criador de ovinos	0,000000955
623320	Trabalhador da cunicultura	0,00000102
632340	Trabalhador da exploração de malva (pãina)	0,00000301
242245	Subprocurador-geral da república	0,0000031
214750	Tecnólogo em rochas ornamentais	0,00000509
223252	Cirurgião dentista - protesiólogo bucomaxilofacial	0,00000518
234640	Professor de literatura comparada	0,00000566
377225	Árbitro de futebol de salão	0,00000762
376230	Equilibrista	0,00000766
613215	Criador de suínos	0,00000792
516710	Numerólogo	0,00000897
242235	Promotor de justiça	0,0000107
376220	Contorcionista	0,0000119
225345	Médico hiperbarista	0,0000154
612515	Produtor de espécies frutíferas trepadeiras	0,0000194
622615	Trabalhador da cultura de erva-mate	0,0000354
612610	Produtor de cacau	0,0000451
622705	Trabalhador na cultura de amendoim	0,0000467
613310	Cunicultor	0,0000522

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se os trabalhadores em Arranjos Populacionais e do setor privado.

Tabela 2. 8 - Ocupações mais concentradas para o ano de 2018

Código da ocupação	Ocupação	Índice de Concentração (EGga)
622015	Trabalhador na produção de mudas e sementes	0,0799439
763105	Auxiliar de corte (preparação da confecção de roupas)	0,0801092
782705	Marinheiro de convés (marítimo e fluviário)	0,0854129
414110	Armazenista	0,0857158
632125	Trabalhador de extração florestal, em geral	0,0898061
782515	Motorista operacional de guincho	0,0977142
763215	Costureiro, a máquina na confecção em série	0,1010757
514225	Trabalhador de serviços de limpeza e conservação de áreas públicas	0,1020669
622020	Trabalhador volante da agricultura	0,1150567
717020	Servente de obras	0,1408324
521135	Frentista	0,1631397
641015	Tratorista agrícola	0,1794829
411010	Assistente administrativo	0,2046803
848520	Magarefe	0,2430896
521125	Repositor de mercadorias	0,2462368
764115	Preparador de calçados	0,2881428
514320	Faxineiro	0,2967279
622505	Trabalhador no cultivo de árvores frutíferas	0,3220155
764005	Trabalhador polivalente da confecção de calçados	0,3599749
421125	Operador de caixa	0,438738
622110	Trabalhador da cultura de cana-de-açúcar	0,4939147
411005	Auxiliar de escritório	0,6375124
782510	Motorista de caminhão (rotas regionais e internacionais)	0,8257594
784205	Alimentador de linha de produção	0,8962209
521110	Vendedor de comércio varejista	2,943979

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Notas: Considera-se os trabalhadores em Arranjos Populacionais e do setor privado.

Convém, ainda, mencionar e explicar algumas evidências gerais da concentração ocupacional no Brasil, considerando as Tabelas apresentadas acima. Ao analisar as 25 ocupações menos e mais aglomeradas e seu comportamento tendencial entre os anos, verifica-se uma maior variabilidade entre as ocupações menos aglomeradas. Percebe-se que para os dois anos o índice de concentração das 25 ocupações menos aglomeradas apresentou um valor praticamente zero. Ao analisar o índice de concentração bruto $G_t = \sum_{i=1}^{n}(s_{it} - t_{it})^2$, a explicação inicial para esse valor zero indica que a participação do emprego na ocupação considerada em determinado Arranjo Populacional (s_{it}) é igual a participação do emprego total no mesmo Arranjo Populacional (t_{it}), e o somatório desse índice bruto dessa ocupação em todos os Arranjos Populacionais considerados também seria zero. Essa é a explicação mais didática para analisar os valores mínimos para o índice final. Todavia, uma análise minuciosa dos dados obtidos para o Brasil indica que estas 25 ocupações menos concentradas (que apresentaram um valor de

zero) têm algumas características que completam a definição acima. Esse valor de zero obtido se dá devido à dinâmica no mercado de trabalho, onde poucos trabalhadores nessas ocupações estão contratados formalmente. No ano de 2009, considerando o painel de dados utilizado nessa pesquisa, o total de trabalhadores das 25 ocupações menos aglomeradas estava entre 1 e 115 trabalhadores, com valor máximo para a ocupação "Trabalhador na cultura de dendê". Já no ano de 2018, esse intervalo esteve entre 1 e 127, com valor máximo para a ocupação "Trabalhador da cultura de erva-mate".

Como pode-se perceber esses valores são muito baixos quando se considera a dimensão do mercado de trabalho do Brasil como um todo. Vale ressaltar que, quando existe apenas 1 trabalhador na ocupação este estará alocado em um Arranjo Populacional, logo a existência de 1 trabalhador remete a uma situação de dispersão, não de aglomeração. No caso da ocupação "Trabalhador da cultura de erva-mate", os 127 trabalhadores estão distribuídos em 6 Arranjos Populacionais de forma equilibrada, logo também não se pode falar em aglomeração, mas dispersão. De modo geral, percebe-se que a dispersão também pode ser explicada pelo pouco dinamismo da ocupação no mercado de trabalho do país, especificamente, nos casos supracitados, além de ser obtido um valor próximo de zero para $(s_{it} - t_{it})$, o somatório seria definido com base em 1 ou poucos Arranjos Populacionais. Logo, esse somatório seria algo bem próximo de zero, explicando a dispersão pelo baixo dinamismo em contratações nessas ocupações no mercado de trabalho formal. Então, esse índice além de permitir mostrar as ocupações mais e menos concentradas, no que se refere às menos concentradas, o mesmo pode oferecer um indicador das ocupações menos comuns.

Outra evidência geral é que a maioria das ocupações menos concentradas são ocupações agrícolas, isso pode ser explicado pela característica do Brasil como um país tipicamente agrícola. De acordo com dados da *Food and Agriculture Organization of the United Nations* (FAO), a Produção de cereais no Brasil cresce continuamente, desde 1961, atingindo em 2017 seu ponto máximo com aproximadamente 117 mil hectares. Em 2018, essa produção diminuiu levemente, passando a representar 103 mil hectares. Entre 2009 e 2018 a produção de cereais no Brasil aumentou cerca de 45,33%, dados que podem ser utilizados para justificar o *slogan* nacional de "celeiro do mundo" (FAO, 2018). Nesse sentido é comum que atividades agrícolas estejam presentes em todo território nacional de forma dispersa, e essa dispersão vai se intensificando cada vez mais, conforme observado por Vieira Filho (2016) que indica que nas

últimas quatro décadas essa expansão se deu através do bioma do cerrado e aproximação da região amazônica.

Entretanto, considerando a tendência de que ocupações tipicamente agrícolas são menos concentradas, convém mencionar uma exceção, os "Trabalhador da cultura de cana-de-açúcar" que, para os dois períodos de tempo analisados, esteve entre as 5 ocupações mais concentradas. De acordo com os dados da RAIS, considerando o painel de dados utilizado nessa pesquisa, a ocupação "Trabalhador da cultura de cana-de-açúcar", representava 0,26% e 0,14% do total de trabalhadores do Brasil, para os anos de 2009 e 2018, respectivamente. Em 2009, 28,99% dos trabalhadores desta ocupação estavam concentrados na região Nordeste e 26,75% na região Sudeste. Em 2018, a concentração destes trabalhadores se manteve nessas duas macrorregiões, mas com alterações significativas, 72,30% na região Nordeste e 18,93% na região Sudeste.

É notório que a produção de culturas agrícolas é determinada pela especificidade de solo ou clima e, desde os primórdios da produção de cana-de-açúcar no Brasil, o Nordeste apresentou grande potencial de vantagem natural para sua produção. Todavia, o desenvolvimento tecnológico e as políticas públicas desenvolvidas até então, em especial o Plano Estratégico de longo prazo (2006-2015) do Ministério da Agricultura, Pecuária e Abastecimento (MAPA), principalmente no que se refere a culturas não alimentares, voltadas para ampliar a participação da agroenergia e da matriz energética brasileira, fez com que a sua produção se dissipasse, principalmente para o Sudeste (FAO, 2013).

Dado esse atual contexto, o Sudeste, apesar de não ter o maior número de trabalhadores na ocupação "Trabalhador da cultura de cana-de-açúcar" no mercado de trabalho formal, conforme verificado a partir dos dados da RAIS considerados nessa pesquisa, possui maior produtividade fazendo com que o Estado de São Paulo ganhe destaque na produção nacional, com mais da metade da produção nacional, conforme destacado por Vieira Filho (2016). No que se refere a produtividade da produção da cana-de-açúcar, de acordo com dados da Companhia Nacional de Abastecimento (CONAB), no período da safra 2017/2018, o Nordeste apresentou a menor produtividade (48.849 kg/ha), enquanto o Sudeste se destaca com maior produção (417.470,7 toneladas) e maior produtividade (76.622 kg/ha), já o Nordeste em termos de produção ficou em terceiro lugar (com uma produção de 41.140,5 toneladas), superando apenas a produção do Sul e Norte (CONAB, 2018). Estes dados estão de acordo com a observação feita por Vieira Filho (2016) de que a produção do Nordeste foi perdendo espaço para a produção mecanizada

do Centro-Oeste. Desse modo, todo esse cenário para a cultura de cana-de-açúcar explicaria a maior concentração da atividade e da ocupação "Trabalhador da cultura de cana-de-açúcar" no Brasil.

Pode-se ainda verificar como evidência geral que a ocupação "Vendedor de comércio varejista" foi a mais concentrada nos dois períodos considerados, isto é, em um intervalo de 10 anos os trabalhadores dessa ocupação concentram-se em poucos Arranjos Populacionais brasileiros. Por ser uma ocupação comum que não requer altos níveis de escolaridade, esperava-se que essa ocupação fosse dispersa pelo território nacional, todavia, uma análise mais minuciosa do mercado de trabalho brasileiro explica essa tendência. De acordo com dados do IBGE (2020) é notório que o setor de serviços é o setor de maior representatividade econômica e o que mais cresce no Brasil, com taxa de crescimento entre 2009 e 2018 de 112,87%. Em termos de representatividade econômica, em 2009 o setor de serviços representava 69,34% do Valor adicionado a preços básicos e, em 2018, esse percentual chegou a 75,45%. Considerando a metodologia básica do IBGE, que divide o setor de Serviços em subsetores (Comércio, Transporte, armazenagem e correio, Informação e comunicação, Atividades financeiras, de seguros e serviços relacionados, Atividades imobiliárias, Outras atividades de serviços e Administração, saúde e educação públicas e seguridade social), convém mencionar que, o setor de Comércio apresenta a maior representatividade em relação aos demais tanto para 2009 (18,47%) como para 2018 (18,38%), também apresentando um crescimento expressivo de 111,87% para o período considerado.

Considerando esse cenário, sabe-se que a ocupação "Vendedor de comércio varejista" está totalmente atrelada ao setor de Comércio, setor esse que apresenta um crescimento significativo dentro da economia brasileira, todavia, é fundamental esclarecer que essa representatividade e crescimento são concentrados. Logo, seria comum esperar uma maior concentração desta ocupação. De acordo com a Pesquisa Anual do Comércio (PAC)[15], em 2009, de um total de 1.570.350 empresas comerciais, 48,70% dessas estavam localizadas na região Sudeste, sendo seguida pelas regiões Sul (22,83%), Nordeste (18,21%), Centro-Oeste (8,38%) e Norte

[15] É importante considerar que a PAC é realizada pelo IBGE desde 1988 e a mesma guarda relação com a base de dados da RAIS desde a sua formação inicial. Essa base teve como intuito inicial representar todo o setor formal da atividade comercial, considerando microempresas, pequenas, médias e grandes empresas. Todavia, dados os problemas orçamentários que surgiram, a partir de 1992 a PAC passou a ter uma amostra somente com as médias e grandes empresas presentes no Censo Econômico de 1985, sendo atualizada anualmente pelas novas empresas da RAIS que possuem 50 ou mais trabalhadores. Para mais informações, ver o relatório metodológico da PAC, disponibilizado pelo IBGE (2000).

(1,86%). Quanto ao número de trabalhadores no setor de Comércio, o Brasil apresenta um total de 8.800.834, apresentando a seguinte distribuição entre as regiões: Sudeste - 52,57%, Sul - 20,03%, Nordeste - 15,97%, Centro-Oeste - 8,45% e Norte - 2,97%. No ano de 2018, mesmo com o crescimento do setor de Comércio, a distribuição de empresas e trabalhadores entre as regiões se manteve praticamente estável. Com 1.652.660 empresas no setor, 49,44% estavam localizadas na região Sudeste, sendo seguida pelas regiões Sul (21,61%), Nordeste (18,58%), Centro-Oeste (8,19%) e Norte (2,16%). No que se refere ao pessoal ocupado em 2018, dos 10.212.427 de trabalhares considerados, 51,61% encontram-se no Sudeste, sendo seguida pelas regiões Sul (19,55%), Nordeste (17,02%), Centro-Oeste (8,70%) e Norte (3,12%).

A análise indica que o setor do Comércio atua com mais força na região Sudeste, nesse sentido, é importante verificar se os trabalhadores categorizados como "Vendedor de comércio varejista" estão mais concentrados nessa região. Uma análise mais minuciosa para a localização dos trabalhadores dessa ocupação pode ser feita a partir dos dados da RAIS. No Brasil, considerando o painel de dados utilizado nessa pesquisa para os anos de 2009 e 2018, existiam 1.410.348 e 1.118.821 trabalhadores classificados como "Vendedor de comércio varejista", respectivamente, representando 6,0% e 5,12% do total de trabalhadores do Brasil. Considerando o ano de 2009, do total de trabalhadores classificados como "Vendedor de comércio varejista" a maioria estava localizada na região Sudeste (55,82%), e em 2018, essa tendência se mantem, com um percentual de 41,65%. Logo, apesar de se caracterizar como uma ocupação comum, dada a própria dinâmica econômica e a conjuntura do mercado de trabalho, nesta pesquisa, avaliada a partir de Arranjos Populacionais, a ocupação "Vendedor de comércio varejista" constitui-se na ocupação mais concentrada no Brasil.

Além disso, convém lembrar que, a análise da concentração ocupacional nesta pesquisa se restringe ao mercado de trabalho formal, e esse fator pode contribuir para explicar os resultados acima. De modo geral, ao avaliar todas as considerações, percebe-se que a concentração ocupacional no Brasil não é facilmente explicada por variáveis como escolaridade dos trabalhadores ou setor de atividade econômica, isto é, a determinação de ocupações mais concentradas ou menos concentradas não obedece a um padrão comum. Como preconizado por Marshall (1920), Duranton e Puga (2004) e Puga (2010) a aglomeração permitiria o compartilhamento, tanto de conhecimento, dado as externalidades positivas de conhecimento (*learning*), quanto de infraestrutura, fornecedores de insumos e trabalhadores qualificados (*sharing),* afetando tanto as empresas como os trabalhadores, gerando *spillover* tecnológicos e

melhor *matching* no mercado de trabalho, permitindo um aumento da produtividade dos trabalhadores.

Considerando este arcabouço teórico, poderia ser esperado que as ocupações com maior nível de escolaridade e complexidade cognitiva fossem mais aglomeradas, fato que, conforme já mencionado, não foi verificado no Brasil. Gabe e Abel (2009) verificaram que a ocupação mais concentrada dos EUA foi a de "Economistas", o que pode indicar que esta ocupação está se beneficiando do compartilhamento de conhecimento teórico e especializado concentrado espacialmente. Entretanto, o estudo encontrou, também, entre as ocupações mais concentradas, a de "Atores", "Motoristas de táxi e motoristas", "Instaladores, operadores e editores de máquinas de corte têxteis", ou seja, ocupações que não têm exigências de altos níveis de escolaridade ou que estejam relacionadas por setores de atividade, tal qual observado nesta pesquisa. Entretanto, o compartilhamento de conhecimento pode ser verificado tanto nas atividades menos complexas quanto nas mais complexas. Conforme já mencionado, a participação de trabalhadores com ensino superior (utilizada nas pesquisas de autores como Rosenthal e Strange (2001) e Ellison e Glaeser (1999)) não constitui uma variável adequada e robusta para medir as habilidades dos trabalhadores (HECKMAN; STIXRUDB; URZUAU, 2006; MACIENTE, 2013)

Dessa forma, o entendimento dos padrões de concentração ocupacional requer avaliar o *labor pooling* e as habilidades dos trabalhadores em suas diferentes esferas, considerando a categorização de Maciente (2013). Além disso, foi possível verificar que essa concentração ocupacional, apesar de não estar claramente determinada pela escolaridade ou setor de atividade, guarda forte relação com a dinâmica do mercado de trabalho, onde as ocupações menos concentradas geralmente são aquelas onde poucos trabalhadores estão contratados formalmente. Esse cenário pode ser melhor avaliado considerando os choques idiossincráticos das ocupações no mercado de trabalho, com base em uma medida de *labor pooling*, proposta por Overman e Puga (2012).

Logo, é fundamental avaliar como o *labor pooling* e as habilidades dos trabalhadores, variáveis de interesse desse ensaio, determinam a concentração ocupacional, ou de forma mais precisa, como a interação dessas duas variáveis pode explicar essa concentração, sendo essa a principal contribuição desta pesquisa e que será objeto de estudo da próxima seção.

2.5.2 Aglomeração ocupacional e *labor pooling* no Brasil

As estatísticas descritivas das variáveis de interesse e de controle utilizadas no modelo de determinação da concentração ocupacional encontram-se na Tabela 2.9, a seguir. Trata-se de um painel de dados não balanceado com 25.037 observações, referente a 2.590 ocupações para o período 2009 a 2018.

Considerando todas as ocupações no período estudado, a média do índice de concentração ocupacional foi de 0,0633. Um valor médio inferior ao encontrado por Gabe e Abel (2009) para o mercado de trabalho dos EUA, a média encontrada pelos autores foi de 0,1872. Esse resultado já era esperado uma vez que os EUA possuem um maior dinamismo no mercado de trabalho e, conforme já relatado, a aglomeração é um importante fator de aumento de produtividade. Dessa forma, um maior desempenho econômico vem acompanhado de uma melhora do mercado de trabalho, dada a relação intrínseca entre o mercado de bens e serviços e o mercado de trabalho. Assim, a aglomeração leva a uma maior produtividade, que por sua vez intensifica a aglomeração, levando a um ciclo vicioso. Além disso, é possível ainda verificar a variação do indicador, observando os valores mínimos e máximos, o que já foi analisado na seção anterior.

Tabela 2. 9 - Estatística Descritivas das variáveis

Variável	Média	Desvio	Mínimo	Máximo
EGga	0,0633	0,0966	1,18e-10	5,669
Labor Pooling	39,40	101,3	0	8.983
Habilidades cognitivas	-0,000406	0,00382	-0,108	0,0285
Habilidades de manutenção e operação	-0,000406	0,00870	-0,271	0,132
Habilidades de assistência	-0,000406	0,0116	-0,188	0,301
Habilidades gerenciais	-0,000406	0,00549	-0,115	0,0944
Habilidades de design e engenharia	-0,000406	0,00821	-0,207	0,135
Habilidades de transporte	0,000406	0,0206	-0,299	0,998
Habilidades artísticas	-0,000406	0,00444	-0,145	0,0401
Habilidade de precisão e automação	-0,000406	0,0252	-1,126	0,729
Habilidades de trabalho supervisionadas	-0,000406	0,0204	-0,604	0,441
Habilidades de ensino e ciências sociais	-0,000406	0,00671	-0,154	0,0904
Força física	0,000406	0,00648	-0,0648	0,197
Habilidades de telecomunicação	-0,000406	0,0111	-0,295	0,236
Habilidades de independência	-0,000406	0,0147	-0,369	0,312
Habilidades em ciências naturais	-0,000406	0,00628	-0,166	0,146
Habilidades de atenção	-0,000406	0,0165	-0,657	0,422
Experiência no trabalho	-0,000406	0,00433	-0,0859	0,0499
Habilidades de gestão de conflitos	-0,000406	0,0294	-0,632	1,037
Habilidades de trabalho em grupo	8,12e-05	0,937	-54,55	82,92
Habilidades de vendas	0,000406	0,0197	-0,224	1,008
Habilidades de monitoramento e conformidade	-0,000406	0,00620	-0,132	0,101
Habilidades clericais	0,000406	0,0271	-0.649	0.990
Participação de trabalhadores com ensino superior	0,295	0,385	0	1
Participação de trabalhadores no setor da agropecuária	0,0299	0,124	0	1
Participação de trabalhadores no setor de extração mineral	0,0175	0,0837	0	1
Média do tamanho do estabelecimento	2,341	0,659	1	4
Experiência média	54,70	27,43	0.300	435.4

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se os trabalhadores em Arranjos Populacionais e do setor privado.

Verifica-se que o índice de concentração apresentou um baixo desvio padrão, apresentando pouca variabilidade nos dados, o que não aconteceu para a variável *labor pooling* que apresentou o maior desvio padrão entre todas as variáveis (101.3), com uma média de 39,40 e o intervalo de 0 a 8.983. Desse modo, convém dar mais destaque a essa variável no que se refere a sua interpretação. Conforme já indicado essa variável capta os choques idiossincráticos do mercado de trabalho em termos de contratação e desligamento dos trabalhadores: verifica-se o comportamento médio dos estabelecimentos por ocupação, isto é, como esses estabelecimentos contratam ou demitem trabalhadores com relação à média de contratação ou desligamento geral dessa ocupação. Assim, quanto maior o valor da variável, maior o dinamismo para essa ocupação. Um valor igual a zero indica nenhum dinamismo para esta ocupação entre os anos considerados, ou seja, não existiram mudanças no total de emprego para essa ocupação e não houve variação contratual entre os estabelecimentos, logo, a composição do emprego desta ocupação nos estabelecimentos se manteve.

Pode-se citar como exemplo a ocupação “Diretor de crédito rural” que em 2018 apresentou um valor de 0, todavia, o comportamento ao longo dos anos varia, dado que essa mesma ocupação apresentou um maior dinamismo em 2015 com um *labor pooling* de 200. No que se refere ao valor máximo da variável *labor pooling*, houve maior dinamismo para a ocupação “Maquiador de caracterização” no ano de 2012 (8.983), todavia, em 2014 esse dinamismo foi bem menor (16,47). A partir destes exemplos percebe-se a importância em utilizar um painel de dados para variáveis como *labor pooling*, pois o dinamismo de contratação e desligamentos entre as ocupações no mercado de trabalho brasileiro é muito volátil. Um painel de dados permite capturar variações idiossincráticas considerando o comportamento tendencial e o dinamismo anual.

Para o vetor de habilidades, todos os 21 fatores apresentaram média praticamente zero, valores mínimos negativos e valores máximos positivos. Todavia, é comum a variabilidade destes valores mínimos e máximos entre os fatores, o que indica um comportamento peculiar para cada fator de habilidade em cada ocupação. Conforme já mencionado, cada ocupação utiliza em maior ou menor grau diversas habilidades, categorizadas nesta pesquisa a partir dos 21 fatores de habilidades de Maciente (2013).

No que se refere as variáveis de controle, a variável participação de trabalhadores com ensino superior por ocupação apresentou uma média de 0,295, isto é, considerando todas as ocupações 29,5% dos trabalhadores têm ensino superior no Brasil para o período considerado. A Participação média de trabalhadores no setor da agropecuária em relação ao setor da Indústria por ocupação é 2,99%, enquanto a Participação média de trabalhadores no setor da extração mineral em relação ao setor da Indústria por ocupação é 1,75%. Como estas variáveis referem-se a participações, seus valores mínimos são 0 e os máximos são 1.

Quanto à média (geral) do tamanho do estabelecimento por ocupação tem-se um valor de 2,341 o que indica que a maioria das ocupações estão inseridas em pequenos estabelecimentos. A experiência média (geral) por ocupação é 54,7 meses.

Convém mencionar que essa estatística descritiva reflete toda a base de dados, considerando todas as 2.590 ocupações para cada ano considerado. Uma análise mais minuciosa exigiria avaliar essa estatística por ocupação e ano, o que é impossível em termos de espaço para uma

pesquisa como essa, todavia, a análise dos valores máximos e mínimos, apresentada na Tabela 10, permite ter uma visão mais ampla dos dados.

Conforme verificado, as variáveis utilizadas nesse modelo apresentam unidades de medidas diferentes e dispersões bastante heterogêneas, a fim de minimizar possíveis problemas de estimação, optou-se por normalizar as variáveis, a partir da distribuição Normal Padrão com média 0 e desvio padrão 1. Dessa forma, obtém-se dados paramétricos, estando os valores mais próximos da média. A literatura utiliza geralmente como principal forma de normalização o logaritmo natural das variáveis, todavia, no presente modelo as variáveis possuem valores não positivos, dessa forma, optou-se por utilizar a distribuição normal padrão tal como proposto e utilizado por Gabe e Abel (2009) e Ellison, Glaeser e Kerr (2010). As estatísticas descritivas básicas das variáveis explicativas utilizadas no modelo de determinação da concentração ocupacional, após a normalização, encontram-se na Tabela A.2 do Apêndice A. Os resultados dessa pesquisa consideram as variáveis normalizadas, todavia, todas as estimações em nível estão no Apêndice A.

Considerando o modelo econométrico apresentado na Equação (2.18), essa pesquisa apresenta três resultados que merecem destaque. O primeiro deles é o efeito do *labor pooling* na determinação da concentração ocupacional. Um valor positivo indica que, quanto maiores os choques idiossincráticos no mercado de trabalho maior será o índice de concentração, isto é, quanto mais dinâmico o mercado de trabalho para determinada ocupação maior será a concentração. O segundo resultado de interesse é o efeito das habilidades para determinação da concentração ocupacional. Um coeficiente positivo indica que determinada habilidade contribui para aumentar a concentração ocupacional e um negativo demonstra que a habilidade não favorece a concentração. O último resultado de interesse é o efeito marginal do *labor pooling*. O efeito completo do *labor pooling*, isto é, o efeito marginal do *labor pooling*, é obtido a partir da soma do parâmetro estimado do *labor pooling* com o parâmetro estimado da interação entre o *labor pooling* e os fatores de habilidade. Conforme já mencionado esses efeitos marginais se aproximam mais da definição de Marshall (1920), Duranton e Puga (2004) e Puga (2010) sobre os efeitos da aglomeração, podendo mensurar e isolar os efeitos de *sharing* (*labor pooling*) *e learnig* e *matching* (através do perfil de habilidades dos trabalhadores).

Todavia, antes de avaliar esse efeito final, algumas especificações do modelo da equação (2.18) foram avaliadas, estas estão presentes na Tabela 2.10 e 2.11 que apresentam tanto os resultados

considerando um *Pooled* de dados como os resultados em painel. Na Tabela 2.10 a especificação (I) considera apenas a variável *labor pooling* como determinante da concentração ocupacional. Na Tabela 2.11, a especificação (II) considera apenas as habilidades como determinante da concentração ocupacional e a (III) considera a variável *labor pooling*, o vetor de habilidades e as variáveis de controle.

De modo geral, observa-se que o controle da heterogeneidade não observada das ocupações faz com que a magnitude dos coeficientes estimados diminua, mantendo os efeitos significativos. Logo, os resultados sugerem que parte do efeito dos determinantes da concentração se deve à heterogeneidade não observada das ocupações (como nível tecnológico em cada ocupação, natureza de interação com o público, preferências individuais, institucionalidades e políticas governamentais) ou outros fatores não controlados. Assim, verifica-se que o estimador de Efeito Fixo, se adequa melhor ao modelo econométrico proposto nesta pesquisa, o que pode ser verificado pelo teste de Hausman que rejeitou a hipótese de Efeitos Aleatórios, exceto para a especificação (I), apresentada na Tabela 2.10, cuja hipótese nula não foi rejeitada. Todavia, convém mencionar que a especificação (I) não apresentou significância para os coeficientes de *labor pooling*, conforme exposto na Tabela abaixo.

Tabela 2. 10 - *Labor Pooling* como determinante da aglomeração ocupacional no Brasil: 2009 a 2018

Variável Dependente: EGga		
	***Pooled* OLS**	**Efeito Aleatório (EA)**
	(I)	**(I)**
Labor Pooling (LP)	-0,0055	-0,0000094
	(0,006536)	(0,001588)
Constante	0,06171***	0,04865**
	(0,02004)	(0,01905)
Dummies de ano	SIM	SIM
N. observações	21.804	21.804
R^2 global	0,0101	0,0100
R^2 within		0,1910
R^2 between		0,0000
Prob>F	0,0000	0,0000
Hausman		$\chi^2(9)= 1,95$
Prob > χ^2		0,9923

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: *** significativo a 1%; ** significativo a 5%; * significativo a 10%. Erro padrão entre parênteses.
Obs.: Todas as variáveis foram normalizadas com média 0 e desvio padrão de 1.

No que se refere a especificação (II) presente na Tabela 2.11, todas os fatores de habilidade foram estatisticamente significantes ao nível de 1% no modelo de Efeito Fixo, com exceção da Habilidade de Ensino e ciências sociais. As habilidades que têm efeito positivo na determinação da concentração ocupacional são as Habilidades Gerenciais, de Transporte, Artísticas, de Força

física, de Independência, de Ciências naturais, de Atenção, de Gestão de conflitos, de Vendas e Clericais. Conforme verificado na Tabela 2.11, a habilidade que tem maior efeito na determinação da concentração ocupacional é a Habilidade de Vendas (0,4122). Esse resultado corrobora a análise anterior das ocupações mais concentradas para os anos de 2009 e 2018 que foram “Vendedores do comércio varejista”, isto é, conforme previsto, mais do que o setor de atividade econômica ou o perfil de escolaridade da ocupação, as habilidades têm maior poder de determinação na concentração ocupacional.

As habilidades que apresentaram efeitos negativos na concentração ocupacional foram: Cognitivas, de Manutenção e operação, de Assistência, de Design e engenharia, de Precisão e automação, de Trabalho supervisionadas, de Telecomunicação, de Experiência no trabalho e de Monitoramento e conformidade. Dentre as habilidades que exercem um efeito negativo na concentração ocupacional, a que apresentou a maior magnitude no coeficiente estimado foi Habilidades de Experiência no trabalho (-0,5433), isso pode ser explicado pelo Teoria do Capital Humano que indica que quanto maior a experiência no trabalho menor a rotatividade. Em termos de Economia Urbana é possível indicar que as ocupações intensivas em habilidades de experiência no trabalho são menos dinâmicas no que se refere à mobilidade espacial, dada a baixa rotatividade no emprego.

Tabela 2. 11 - Determinantes da aglomeração ocupacional no Brasil: 2009 a 2018

Variável Dependente: EGga				
Variável	***Pooled* OLS**		**Efeito Fixo (EF)**	
	(II)	**(III)**	**(II)**	**(III)**
Labor Pooling (LP)		-0,00446*		0,0003
		(0,00265)		(0,00116)
Habilidades				
Cognitivas	-0,2661***	-0,261***	-0,2415***	-0,3375***
	(0,01043)	(0,0106)	(0,0289)	(0,02927)
Manutenção e operação	0,2357***	0,239***	-0,0839***	-0,1062***
	(0,00908)	(0,00934)	(0,0191)	(0,01808)
Assistência	-0,0584***	-0,0494***	-0,1055***	-0,1109***
	(0,00488)	(0,00475)	(0,0072)	(0,00678)
Gerenciais	-0,0367***	-0,0347***	0,3067***	0,2673***
	(0,00584)	(0,00572)	(0,0223)	(0,02233)
Design e engenharia	-0,1201***	-0,117***	-0,2159***	-0,2191***
	(0,00918)	(0, 00891)	(0,016)	(0,01667)
Transporte	0,0247***	0,0307***	0,1029***	0,1082***
	(0,00582)	(0, 00575)	(0,0129)	(0,01204)
Artísticas	-0,00984	-0,00964	0,1453***	0,2408***
	(0,0093)	(0,00913)	(0,02103)	(0,02115)
Precisão e automação	-0,1221***	-0,115***	-0,4606***	-0,3986***
	(0,00578)	(0,00588)	(0,00725)	(0,00705)
Trabalho supervisionadas	0,1657***	0,153***	-0,0992***	0,0257
	(0,00546)	(0,00536)	(0,01647)	(0,01616)
Ensino e ciências sociais	-0,1412***	-0,141***	0,0036	0,0261
	(0,00483)	(0,00484)	(0,01577)	(0,01634)
Força física	0,0823***	0,0837***	0,4322***	0,4304***
	(0,00828)	(0,00818)	(0,01855)	(0,01788)
Telecomunicação	-0,0976***	-0,0941***	-0,0489***	-0,0428***
	(0,004)	(0,00389)	(0,00669)	(0,0064)
Independência	0,2016***	0,208***	0,1807***	0,2539***
	(0,0049)	(0,00483)	(0,0133)	(0,0124)
Ciências naturais	0,0714***	0,0716***	0,1382***	0,1425***
	(0,00599)	(0,00584)	(0,01223)	(0,01207)
Atenção	0,0327***	0,0364***	0,3116***	0,2897***
	(0,00672)	(0,00657)	(0,013)	(0,01232)
Experiência no trabalho	-0,1349***	-0,137***	-0,5433***	-0,5019***
	(0,0068)	(0,00672)	(0,0157)	(0,01478)
Gestão de conflitos	-0,1236***	-0,113***	0,0684***	0,0381***
	(0,00605)	(0,00583)	(0,00574)	(0,00529)
Trabalho em grupo	-0,0144***	-0,00267	-0,0085***	-0,0044***
	(0,00327)	(0,00311)	(0,00156)	(0,00141)
Vendas	0,7032***	0,682***	0,4122***	0,3514***
	(0,0099)	(0,00984)	(0,022)	(0,02084)
Monitoramento e conformidade	0,1082***	0,107***	-0,1469***	-0,0651***
	(0,00471)	(0,00473)	(0,0131)	(0,01327)
Clericais	0,0443***	0,0428***	0,0649***	0,0199*
	(0,00593)	(0,00605)	(0,01025)	(0,01061)
Variáveis de Controle				
Participação de trabalhadores		-0,00312		-0,0114
com ensino superior		(0,00282)		(0,00839)
Participação de trabalhadores no		-0,0144***		-0,0057
setor da agropecuária		(0,00261)		(0,00383)
Participação de trabalhadores no		-0,0084***		0,0009
setor de extração mineral		(0,00258)		(0,00367)
Média do tamanho do		0,0294***		0,0007
estabelecimento		(0,00288)		(0,00282)

(continua)

(conclusão)

Variável	*Pooled* OLS		Efeito Fixo (EF)	
	(III)	**(II)**	**(III)**	**(II)**
Experiência média		-0,0000123		-0,0006
		(0,00278)		(0,00234)
Constante	0,0727***	0,0592***	0,0736***	0,0574***
	(0,0082)	(0,00786)	(0,00358)	(0,00319)
Dummies de ano	SIM	SIM	SIM	SIM
N. observações	21.804	21.804	21.804	21.804
R^2 global	0,8369	0,8481	0,6575	0,7014
R^2 within			0,5554	0,5871
R^2 between			0,6653	0,7077
Prob>F	0,0000	0,0000	0,0000	0,0000
Hausman			$\chi^2(30)=4650$	$\chi^2(35)= 2940,34$
Prob > χ^2			0,0000	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: *** significativo a 1%; ** significativo a 5%; * significativo a 10%. Erro padrão entre parênteses.
Obs.: Todas as variáveis foram normalizadas com média 0 e desvio padrão de 1.

Comparando a especificação (II) com os resultados da pesquisa de Maciente (2013) para avaliar os determinantes da concentração industrial, verifica-se que o vetor de habilidades se ajusta melhor à análise da aglomeração ocupacional, isto é, habilidades impactam tanto na aglomeração industrial como na aglomeração ocupacional, todavia, possuem maior poder de determinação na última análise, conforme já previsto por Gabe e Abel (2009). Ao analisar os resultados do modelo de Maciente (2013), que inclui, entre outras variáveis, os fatores de habilidade, e considerando a estimação por MQO, especificação que mais se aproxima dessa pesquisa, verifica-se que somente 12 fatores de habilidade foram significantes para explicar a concentração industrial, enquanto que na especificação (II) dessa pesquisa considerando o estimador de *Pooled* MQO, 20 fatores de habilidades foram significantes. Além disso, o ajustamento do modelo captado pelo R^2 global foi de 0,4819 no modelo de Maciente (2013), bem inferior ao obtido nessa pesquisa.

Na especificação (III) de Efeitos Fixos, a direção dos efeitos das habilidades na determinação da concentração ocupacional permanece a mesma, exceto para o efeito das habilidades de Trabalho supervisionadas que passa a ser insignificante, junto com o efeito da habilidade de Ensino e ciências sociais. Além disso, os resultados parecem indicar que a habilidade de Venda (0,3514) é a que possui maior influência positiva na determinação da concentração ocupacional e a habilidade de Experiência no trabalho é a que possui maior influência negativa na determinação da concentração ocupacional. Outro resultado que merece destaque é a ausência de significância estatística do *labor pooling* no modelo de Efeito Fixo, enquanto no *pooled* de dados ele apresentou efeito negativo e significante. Uma explicação para esse resultado é que a inclusão da heterogeneidade não observada das ocupações permitiu lidar com o viés de variável

omitida, o qual se mostrou relevante ao tornar o efeito do *labor pooling* estatisticamente não significante.

No que se refere as variáveis de controle, verifica-se que no modelo *Pooled* OLS os coeficientes das variáveis Participação de trabalhadores no setor da agropecuária e Participação de trabalhadores no setor de extração mineral foram negativos e significantes ao nível de 1%. O efeito da Média do tamanho do estabelecimento foi positivo ao nível de 1%, enquanto as variáveis de experiência e ensino superior não apresentaram significância. A variável Média do tamanho do estabelecimento, que busca verificar a premissa levantada por Marshall de que economias de escala são importantes para a produtividade do estabelecimento, podendo impactar na concentração ocupacional, apresentou um efeito positivo, conforme esperado.

Comparando com a pesquisa de Gabe e Abel (2009), realizada para os EUA com dados do Censo de 2000, verificou-se que apenas a Participação de trabalhadores no setor da agropecuária e a Participação de trabalhadores no setor de extração mineral foram significativas, todavia, com um efeito positivo. Esses resultados evidenciam que os países possuem estruturas produtivas diferentes e essa estrutura impacta na concentração ocupacional. No Brasil, a Participação de trabalhadores no setor da agropecuária e a Participação de trabalhadores no setor de extração mineral exercem um efeito negativo na determinação da concentração ocupacional. Esse resultado pode ser explicado pelo fato de atividades agrícolas e de extração mineral serem desenvolvidas em todo o território nacional, dado que o Brasil é um país tipicamente agrícola quando comparado com os EUA. Nesse último, as atividades tipicamente agrícolas são menos predominantes, logo, justifica-se a concentração ocupacional em determinadas regiões metropolitanas (unidade regional considerada na pesquisa de Gabe e Abel (2009)). No entanto, convém mencionar que essa pesquisa traz como uma contribuição empírica à pesquisa de Gabe e Abel (2009) a análise da determinação da concentração ocupacional considerando a heterogeneidade não observada das ocupações. Assim, ao observar o modelo de Efeito Fixo, verifica-se que todas as variáveis de controle se tornam insignificantes.

Este resultado é condizente com a pesquisa desenvolvida por Andrade, Gonçalves e Freguglia (2014) para o Brasil. Os autores, utilizando um painel de dados para o período de 2003 a 2008, buscaram verificar como se dá a concentração ocupacional nas regiões metropolitanas brasileiras a partir da análise do coeficiente de Gini locacional e qual o papel da intensidade tecnológica da ocupação na determinação do coeficiente de Gini. Na referida pesquisa foram

utilizadas como variáveis de controle a Distribuição da indústria, o Número de empregados, o Tamanho médio das firmas e a Porcentagem dos empregados em agricultura, pesca e indústrias extrativas. Os coeficientes obtidos para o Tamanho médio das firmas e a Porcentagem dos empregados em agricultura, pesca e indústrias extrativas foram positivos e significantes, considerando o modelo Mínimos Quadrados Agrupados, todavia, ao considerar o modelo de Efeitos Fixos esses coeficientes se tornam insignificantes, corroborando os resultados obtidos nessa pesquisa.

No que se refere, a especificação completa proposta na Equação (2.18), que inclui as variáveis de interação, isto é, a variável *labor pooling* ponderada pelas habilidades, novamente verifica-se que o resultado do teste de Hausman permite rejeitar a hipótese nula de que não existe diferença sistemática nos coeficientes estimados, de modo que o modelo de Efeitos Fixo constitui o melhor estimador. Além disso, constatou-se insignificância das variáveis de controle na presença da heterogeneidade não observada. Os resultados estão expostos na Tabela A.3 do Apêndice A[16].

A principal evidência desse modelo é que o efeito do *labor pooling*, apesar da baixa magnitude, tornou-se positivo e significante. Esse resultado indica que avaliar a concentração ocupacional requer isolar os efeitos de *sharing* (*labor pooling*) *e learning* e *matching* (através do perfil de habilidades dos trabalhadores). Desse modo, percebe-se que, no modelo completo as variáveis *labor pooling* e habilidades de fato estavam correlacionadas, logo, a modelagem empírica completa (EQUAÇÃO (2.18)) permite obter parâmetros mais robustos do efeito do *labor pooling* ocupacional. O efeito médio do *labor pooling* sobre a concentração ocupacional é de 0,0057 (TABELA A.3 do APÊNDICE A). Apesar de utilizadas unidades geográficas diferentes, pode ser interessante comparar esse coeficiente com o da pesquisa de Almeida e Rocha (2018), realizada para o Brasil, que fez uma análise da concentração setorial da indústria da transformação no período de 2002 a 2014. Almeida e Rocha (2018) observou que o efeito médio do *labor pooling* foi de 0,022. Nesse sentido, percebe-se que o efeito do *labor pooling* é menor para a determinação da concentração ocupacional no Brasil, mas ainda positivo e significante.

Percebe-se, ainda, que apesar da significância e valor positivo, quando comparado como o vetor de habilidades, o *labor pooling* tem menor poder de determinação, uma vez que o valor do

[16] O resultado completo do modelo estimado, bem como o resultado para todas as modelagens apresentadas nas Tabelas 7 e 8 com as variáveis em nível, isto é, sem normalização, encontram-se no Apêndice A.

coeficiente foi baixo. Além disso, avaliando a literatura sobre aglomeração industrial (OVERMAN; PUGA, 2010; ROSENTHAL; STRANGE, 2001; BURKI; KHAN, 2013; ALKAY; HEWINGS, 2012; ALMEIDA; ROCHA, 2018) percebe-se que o coeficiente da variável *labor pooling* é positivo e, em geral, parece ter maior magnitude e poder de determinação na concentração de indústrias. Isso é uma indicação de que, quando se considera a aglomeração ocupacional, o *labor pooling* é importante, mas as habilidades apresentam maior poder explicativo. Assim, corrobora-se a argumentação de Gabe e Abel (2009) de que as externalidades de conhecimento e de habilidades são mais fortes quando se considera a aglomeração ocupacional, isto é, a proximidade dos indivíduos da mesma ocupação gera maiores *spillovers* de conhecimentos do que empregos com requisitos genéricos, como na análise da aglomeração industrial.

Analisando o estimador de Efeitos Fixos, o padrão de efeitos positivos e negativos do vetor de habilidade se manteve, quando comparado com a especificação (III) de Efeitos Fixos da Tabela 2.11, exceto para a habilidade de Gestão de Conflitos que apresentava efeito positivo e torna-se insignificante e para a habilidade de Ensino e ciências naturais que era insignificante e torna-se positiva na especificação completa, como pode ser visualizada na Tabela 2.12, que apresenta os coeficientes de forma resumida para o modelo em painel de Efeitos Fixos[17].

Desse modo, verifica-se que as habilidades que têm efeito positivo na determinação da concentração ocupacional são: Gerenciais (0,2783), de Transporte (0,1065), Artísticas (0,3474), de Ensino e ciências sociais (0,0468), Força física (0,4305), de Independência (0,1723), de Ciências naturais (0,1216), de Atenção (0,2247), de Vendas (0,5609) e Clericais (0,2547). Já as habilidades que apresentam efeito negativo na concentração ocupacional são: Cognitivas (-0,3549), de Manutenção e operação (-0,1412), de Assistência (-0,1471), de Design e engenharia (-0,2447), de Precisão e automação (-0,414), de Telecomunicação (-0,0265), de Experiência no trabalho (-0,428), de Trabalho em grupo (-0,0214) e de Monitoramento e conformidade (-0,0442). Dentre as habilidades com efeito positivo a maior continua sendo a de Vendas e dentre aquelas com efeito negativo a maior continua sendo a de Experiência no trabalho. Os principais fatores que podem explicar esses resultados já foram listados anteriormente (TABELA 2.12).

[17] A Tabela A.4 do Apêndice A apresenta os coeficientes de forma resumida para o modelo em *Pooled* OLS.

Apesar de objetivar avaliar os determinantes da aglomeração industrial, convém mencionar que, a pesquisa de Maciente (2013) constatou que as habilidades Gerenciais, de Transporte, Força física e as de Vendas também estão relacionadas positivamente com a aglomeração industrial no Brasil.

Tabela 2. 12 - Determinantes da aglomeração ocupacional no Brasil (2009 a 2018): *Labor Pooling* ponderado pelas habilidades e Efeito Marginal do *Labor Pooling* completo

Variável Dependente: EGga			
Efeito Fixo			
Habilidades	$\boldsymbol{\gamma_h}$ **(Habilidades)**	$\boldsymbol{\mu_h}$ **(LP x habilidades)**	$\boldsymbol{\delta + \mu_h}$ $(\boldsymbol{\delta = 0,0057})$
Cognitivas	-0,3549***	-0,0247	-
Manutenção e operação	-0,1412***	0,0377*	0,0434
Assistência	-0,1471***	0,0539***	0,0596
Gerenciais	0,2783***	-0,0194	-
Design e engenharia	-0,2447***	0,0718***	0,0775
Transporte	0,1065***	-0,021*	-0,0153
Artísticas	0,3474***	-0,1279***	-0,1222
Precisão e automação	-0,414***	0,0107***	0,0164
Trabalho supervisionadas	0,0209	-0,0654***	-0,0597
Ensino e ciências sociais	0,0468**	-0,0196**	-0,0139
Força física	0,4305***	-0,1258***	-0,1201
Telecomunicação	-0,0265***	-0,024**	-0,0183
Independência	0,1723***	-0,0127	-
Ciências naturais	0,1216***	0,002	-
Atenção	0,2247***	0,0731**	0,0788
Experiência no trabalho	-0,428***	-0,0149	-
Gestão de conflitos	-0,0054	0,0354***	0,0411
Trabalho em grupo	-0,0214***	0,0163**	0,022
Vendas	0,5609***	-0,1416***	-0,1359
Monitoramento e conformidade	-0,0442***	-0,0272***	-0,0215
Clericais	0,2547***	-0,1541***	-0,1484

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: *** significativo a 1%; ** significativo a 5%; * significativo a 10%.
Obs.: Todas as variáveis foram normalizadas com média 0 e desvio padrão de 1.

Convém salientar que o *labor pooling* impacta positivamente na concentração ocupacional, e capta choques idiossincráticos nas ocupações, isto é, representa o dinamismo do mercado de trabalho ocupacional. Entretanto, uma análise mais precisa da dinâmica do mercado de trabalho requer avaliar como as habilidades requeridas para o desenvolvimento das ocupações podem impactar nesse dinamismo. Logo, o dinamismo do mercado de trabalho ocupacional, pode ser intensificado ou limitado, dado o conjunto de habilidades que está sendo mais requisitado em cada período, exercendo efeitos mais robustos e fidedignos na concentração ocupacional.

Observando a Tabela 2.12, quando se considera o modelo de Efeitos Fixos, houve efeitos positivos para a interação entre *labor pooling* e habilidade para sete dos 21 fatores de

habilidade, a saber: de Manutenção e operação (0,0377), de Assistência (0,0539), de Design e engenharia (0,0718), de Precisão e automação (0,0107), de Atenção (0,0731), de Gestão de conflitos (0,0354) e de Trabalho em grupo (0,0163). Por sua vez, nove fatores apresentam interação negativa com o *labor pooling*: Habilidades de Transporte (-0,021), Artísticas (-0,1279), de Trabalho supervisionadas (-0,0654), de Ensino e ciências sociais (-0,0196), Força física (-0,1258), de Telecomunicação (-0,024), de Vendas (-0,1416), de Monitoramento e conformidade (-0,0272) e Clericais (-0,1541).

Verifica-se, ainda, que o efeito da maioria das variáveis de interação (14 habilidades, a saber: de Manutenção e operação, de Assistência, de Design e engenharia, de Transporte, Artísticas, de Precisão e automação, de Trabalho supervisionadas, de Ensino e ciências sociais, de Força física, de Independência, de Gestão de conflitos, de Trabalho em grupo, de Vendas e Clericais) apresentou efeito contrário ao efeito individual dessa habilidade sobre a concentração ocupacional. Esse é mais um indicador da importância em se avaliar a interação do *labor pooling* com as habilidades, pois o efeito dessa interação na dinâmica do mercado de trabalho representa algo mais realístico, dado que a análise da oferta e demanda de trabalho deve considerar as peculiaridades dos trabalhadores e empresas no que se refere as habilidades requeridas. Empresas e trabalhadores atuam no mercado de trabalho considerando o perfil desses trabalhadores e dessas empresas. A aglomeração possibilita uma melhor correspondência no que se refere ao perfil de habilidades, isto é, o *matching* pode vir a impactar na concentração ocupacional, dado que esta depende das interações de empresas e trabalhadores. Assim, essa variável de interação poderia representar uma variável mais apropriada da dinâmica do mercado de trabalho quando se faz necessário avaliar as habilidades requeridas pelos trabalhadores.

Após essas análises, resta avaliar o efeito marginal do *labor pooling*, considerando o estimador de Efeitos Fixos. Conforme já mencionado, esse efeito é obtido a partir da soma do parâmetro estimado do *labor pooling* ($\delta = 0{,}0057$) com o parâmetro estimado da interação entre o *labor pooling* e os fatores de habilidade (μ_h), apresentado na última coluna da Tabela 2.12.

É importante apontar que, os coeficientes das interações insignificantes são indícios de que as habilidades não têm um efeito maior ou menor do que a média (*labor pooling*) na determinação da concentração ocupacional, isto é, os efeitos dessas habilidades são iguais ao efeito médio do *labor pooling*. Observa-se isso para as seguintes habilidades: Cognitivas, Gerenciais, de

Independência, de Ciências naturais e de Experiência no trabalho. Neste sentido, os coeficientes negativos para a interação indicam um retorno menor que a média (ou uma menor inclinação), isto é, essas habilidades contribuem para uma diminuição no efeito de *labor pooling*. Enquanto, os coeficientes positivos para a interação indicam que essas habilidades contribuem para um aumento no efeito de *labor pooling*, intensificando o efeito positivo do *labor pooling* sobre a concentração ocupacional.

Desse modo, observa-se que o efeito marginal do *labor pooling* superou o efeito médio (0,0057) na concentração ocupacional, quando se considera as seguintes habilidades: de Manutenção e operação (0,0434), de Assistência (0,0596), de Design e engenharia (0,0775), de Precisão e automação (0,0164), de Atenção (0,0788), de Gestão de conflitos (0,0411) e de Trabalho em grupo (0,022). Essas habilidades contribuem para intensificar o efeito médio positivo do *labor pooling* na concentração ocupacional. Enquanto as Habilidades de Transporte (-0,0153), Artísticas (-0,1222), de Trabalho supervisionadas (-0,0597), de Ensino e ciências sociais (-0,0139), Força física (-0,1201), de Telecomunicação (-0,0183), de Vendas (-0,1359), de Monitoramento e conformidade (-0,0215) e Clericais (-0,1484) reduzem o efeito médio do *labor pooling* na concentração ocupacional. Essa redução é tão intensa que o efeito positivo do *labor pooling* passa a ser negativo.

Logo, o efeito marginal do *labor pooling*, quando se considera as habilidades que interagem negativamente, provoca uma diminuição na concentração ocupacional, indicando que a dinâmica do mercado de trabalho ou os choques idiossincráticos podem impactar negativamente na concentração ocupacional. Dessa forma, verifica-se que é fundamental considerar os efeitos das interações para encontrar a direção final dos efeitos de *labor pooling* por ocupação. Dentro desta perspectiva, observa-se que nas ocupações que são intensivas nas habilidades supracitadas, que interagem negativamente com o *labor pooling*, os choques idiossincráticos no mercado de trabalho no período considerado contribuem para reduzir a concentração ocupacional.

Esses efeitos negativos podem estar atrelados às deseconomias de escala. A aglomeração é um fenômeno urbano que envolve empresas e trabalhadores. As deseconomias de escala causadas pela aglomeração, que surgem a partir dos efeitos nocivos da concorrência, afetam tanto trabalhadores como empresas. No que se refere às habilidades, essas podem ser adquiridas a partir do estudo formal, treinamento ou aperfeiçoamentos, e das peculiaridades intrínsecas aos

indivíduos. De modo geral, sabe-se que os custos de adquirir habilidades podem ser distribuídos entre trabalhadores e empresas. As habilidades Cognitivas, por exemplo, são importantes para o desenvolvimento de qualquer empresa. Nesse aspecto, podem ser necessários maiores níveis de esforço individual e investimento em capital humano (escolaridade ou treinamento) no processo de aquisição dessas habilidades.

A competição por trabalhadores qualificados em determinadas habilidades e treinamentos específicos pode ser associada a um fator de deseconomia de aglomeração. Assim, as empresas podem permitir a dispersão desses trabalhadores de modo que os mesmos não beneficiem outras empresas, limitando a concorrência do mercado. Por sua vez, os trabalhadores com determinadas habilidades podem ser beneficiados na decisão de não se aglomerarem, pois uma das deseconomias de escala da aglomeração de trabalhadores é o aumento da oferta de trabalho especializado e sua consequente diminuição dos salários. Dessa forma, o perfil de habilidades que exigem maior investimento em capital humano, ao limitarem a concentração ocupacional, pode favorecer tanto empresas como trabalhadores, minimizando os efeitos das deseconomias de escala para ambos.

2.6 CONSIDERAÇÕES FINAIS

O objetivo principal desse ensaio foi avaliar o efeito do *labor pooling* sobre a aglomeração ocupacional no mercado de trabalho formal do Brasil. A medida utilizada para medir a aglomeração ocupacional foi o índice de Ellison-Glaeser (1997) modificado e adaptado por Gabe e Abel (2009). A estratégia empírica consistiu na estimação de um painel de dados para o período de 2009 a 2018, no nível da ocupação desempenhada pelo trabalhador, tendo como foco principal o efeito *marshalliano* de *sharing* (*labor pooling*) sobre a aglomeração ocupacional.

Esses três microfundamentos têm um elo em comum: as habilidades dos trabalhadores. Dentro deste cenário, a análise da aglomeração ocupacional atrelada ao *labor pooling* e às habilidades dos trabalhadores é uma importante contribuição para a literatura econômica brasileira. Dada a dificuldade em mensurar e decompor os efeitos de *matching, sharing* e *learnig* (COMBES; GOBILLON, 2015), este ensaio buscou avaliar basicamente esses efeitos a partir de variáveis *proxies* que sejam capazes de representar o *labor pooling*, considerando os choques idiossincráticos exógenos, bem como um conjunto de habilidades dos trabalhadores.

O ensaio utilizou os microdados individuais da RAIS, tomando como subunidade geográfica de análise os 376 Arranjos Populacionais brasileiros e municípios isolados que se classificam como médias e grandes concentrações urbanas, tal como delimitados pelo IBGE (2016). O índice de Ellison-Glaeser (1997), modificado e adaptado por Gabe e Abel (2009), foi calculado para o período compreendido entre 2009 a 2018, a partir das 2.557 ocupações categorizadas pela CBO (2002). Comparando os anos de 2009 e 2018, verificou-se que das 25 ocupações menos e mais aglomeradas, houve maior variabilidade entre as ocupações menos aglomeradas. Entre as 25 ocupações menos aglomeradas, nos dois anos ressaltados, a maioria das ocupações apresentou índice igual a zero, isso se deve a dinâmica no mercado de trabalho, uma vez que há poucos trabalhadores contratados formalmente no Brasil para essas ocupações.

Os resultados evidenciaram, ainda, que a maioria das ocupações menos concentradas são ocupações agrícolas, o que pode ser explicado pela característica do Brasil como um país tipicamente agrícola. Todavia, algumas exceções foram fundamentais para se compreender a dinâmica do mercado de trabalho no Brasil, como a ocupação "Trabalhador da cultura de cana-de-açúcar" que esteve entre as ocupações mais concentradas. Essa concentração, além de ser explicada pelas especificidades naturais, fazendo com que a atividade se concentre especialmente nos Arranjos Populacionais pertencentes a região Nordeste, tem forte relação com o plano de desenvolvimento estratégico do Ministério da Agricultura, Pecuária e Abastecimento (MAPA), que intensificou a produção de certas culturas alimentares para a produção agroenergética, intensificando a concentração dessa ocupação na região Sudeste.

No que se refere as ocupações mais concentradas convém destacar aquelas relacionadas ao setor têxtil, tal como verificado por Gabe e Abel (2009) para as regiões metropolitanas dos EUA, e corroborando os resultados para a concentração industrial em geral (ELLISON; GLAESER, 1997; DURANTON; OVERMAN, 2005) e no Brasil (ALMEIDA, ROCHA, 2018). Vale ressaltar, ainda, a ocupação "Vendedor de comércio varejista", a qual se manteve entre a ocupação mais concentrada do país tanto no ano de 2009 como no ano de 2018. Esse resultado pode ser explicado pela estrutura produtiva do Brasil, que tem o setor de serviços e, por sua vez, o setor de comércio, como destaque na composição e crescimento do PIB. Em termos regionais, ao considerar o número de estabelecimentos e o pessoal ocupado, o setor de comércio tem maior destaque na Região Sudeste, de modo que essa concentração ocupacional ocorre principalmente nos Arranjos Populacionais da referida região.

A análise inicial da concentração ocupacional no Brasil, a partir do índice Ellison-Glaeser (1997) modificado e adaptado por Gabe e Abel (2009), sugere que não existe um padrão comum para a concentração no Brasil, no que se refere a escolaridade ou ao setor de atividade econômica dos trabalhadores. Esse comportamento justifica a análise da determinação da concentração ocupacional considerando variáveis mais complexas e que captem melhor a dinâmica do mercado de trabalho bem como as habilidades dos trabalhadores. Neste aspecto, uma contribuição importante dessa pesquisa é ampliar o conjunto de variáveis relevantes para a análise da concentração ocupacional, baseadas na literatura. Neste sentido, este estudo incluiu, como variáveis relevantes, a medida de *labor pooling*, definida por Overman e Puga (2010), e medidas de habilidades dos trabalhadores, a partir da base de dados de Maciente (2013), bem como a interação entre essas variáveis.

Verificou-se que o estimador de Efeitos Fixos se adequa a modelagem empírica desenvolvida nessa pesquisa e que o *labor pooling* só apresenta efeito estatisticamente significante quando se considera a sua interação com os fatores de habilidade ocupacionais na modelagem econométrica. Os resultados indicam a importância em avaliar e isolar os efeitos de *sharing* (*labor pooling*), *learning* e *matching* (através do perfil de habilidades dos trabalhadores). Dado que as variáveis de interação foram significantes, tem-se um indício de complementariedade dos efeitos supracitados. Apesar da significância e efeito positivo do coeficiente do *labor pooling* isolado, verifica-se que sua magnitude é muito baixa quando comparado como o vetor de habilidades. Desse modo, os resultados dessa pesquisa confirmam a principal hipótese levantada por Gabe e Abel (2009) acerca do poder das externalidades de conhecimento e habilidades na determinação da aglomeração ocupacional. A importância das habilidades para avaliar a concentração no Brasil foi enfatizada por Maciente (2013) no seu estudo da concentração industrial e a partir dos resultados dessa pesquisa pode-se verificar que essa importância é ainda maior no contexto ocupacional.

Ao avaliar os efeitos das habilidades é notório a importância das habilidades de Vendas para o aumento da concentração ocupacional. Esta habilidade apresentou o maior efeito positivo, em todas as especificações estimadas, o que se mostrou compatível com o resultado de que a ocupação “Vendedores do comércio varejista” foi a mais concentrada. Desse modo, a análise do índice de Ellison-Glaeser (1997), modificado e adaptado por Gabe e Abel (2009), foi corroborada pela estratégica empírica adotada no presente estudo.

Conforme já mencionado, para uma análise mais precisa e robusta dos efeitos do *labor pooling* e das habilidades, se fez necessário obter seus efeitos marginais, que dependem da significância e magnitude das variáveis de interação entre o *labor pooling* e as habilidades. Essa interação é a variável mais completa do modelo, pois possibilita avaliar os choques idiossincráticos na ocupação, considerando as habilidades dos trabalhadores. Considerando o estimador de Efeitos Fixos, sete conjuntos de habilidades apresentaram coeficientes positivos para a interação, a saber: Habilidades de Manutenção e operação, de Assistência, de Design e engenharia, de Precisão e automação, de Atenção, de Gestão de conflitos e de Trabalho em grupo. Por sua vez, nove habilidades apresentaram efeitos negativos da interação: Habilidades de Transporte, Artísticas, de Trabalho supervisionadas, de Ensino e ciências sociais, Força física, de Telecomunicação, de Vendas, de Monitoramento e conformidade e Clericais.

Verificou-se que 14 variáveis de interação entre *labor pooling* e habilidades (de Manutenção e operação, de Assistência, de Design e engenharia, de Transporte, Artísticas, de Precisão e automação, de Trabalho supervisionadas, de Ensino e ciências sociais, de Força física, de Independência, de Gestão de conflitos, de Trabalho em grupo, de Vendas e Clericais) apresentaram efeitos marginais contrários aos efeitos individuais das habilidades correspondentes, o que ratifica a importância em considerar essas variáveis de interação para a obtenção de efeitos marginais robustos.

De modo geral, verifica-se que a inclusão das variáveis de interação provocou mudanças expressivas na magnitude dos efeitos marginais obtidos para o *labor pooling*. No modelo completo, o efeito médio do *labor pooling* passou a ser significativo e positivo, conforme previsto por Krugman (1991) e Overma e Puga (2010). No que se refere a principal contribuição dessa pesquisa, o efeito marginal do *labor pooling* apresentou comportamento heterogêneo. É importante destacar que o efeito marginal do *labor pooling* passa a ser negativo considerando a interação com as seguintes habilidades: de Transporte, Artísticas, de Trabalho supervisionadas, de Ensino e ciências sociais, Força física, de Telecomunicação, de Vendas, de Monitoramento e conformidade e Clericais. Esses efeitos negativos indicam que a heterogeneidade dos choques nas ocupações não atua aumentando a concentração quando se considera essas habilidades, isto é, o agrupamento da mão-de-obra especializada nessas habilidades não impacta positivamente na concentração ocupacional. Nessa situação, os choques no mercado de trabalho podem gerar deseconomias de escala.

Esses resultados indicam a importância em se utilizar as variáveis de interação na modelagem da concentração ocupacional, uma vez que os efeitos de *sharing*, *learning* e *matching* são intrinsecamente relacionados e, conforme já estabelecido, avaliar seus impactos isoladamente pode não ser suficiente para identificar os canais pelos quais podem explicar a concentração ocupacional (FUJITA; KRUGMAN; VENABLES, 1999; COMBES; DURANTON, 2015; PUGA, 2010). Essa estratégia empírica, além de inédita na literatura nacional, permite obter parâmetros mais robustos.

Dado esse contexto, tem-se um resultado importante frente ao cenário da Quarta Revolução Industrial, onde algumas considerações devem ser elencadas. Conforme enfatizado por Maciente, Rauen e Kubota (2019), apesar da natureza incerta das mudanças provocadas no mercado de trabalho brasileiro, a partir da Quarta Revolução Industrial, alguns fatos são mais consensuais: as ocupações que envolvem trabalho rotineiro e força física perderão importância em detrimento daquelas mais complexas que exijam habilidades cognitivas, de engenharia, interpessoais, gerenciais e habilidades ligadas às ciências. Observou-se que, no Brasil, as habilidades Cognitivas, de Experiência no trabalho, de Manutenção e operação, de Assistência, de Design e engenharia, de Precisão e automação, de Telecomunicação, de Trabalho em grupo e de Monitoramento e conformidade exercem um impacto negativo na concentração ocupacional, ou seja, de modo geral, as ocupações mais concentradas são aquelas que requerem menor complexidade de habilidades. Logo, os Arranjos Populacionais que têm maior concentração ocupacional e, portanto, menor conteúdo de habilidades complexas, podem sofrer impactos negativos frente a nova revolução tecnológica.

Nesse contexto, é fundamental a elaboração de políticas públicas educacionais e políticas públicas de treinamento e profissionalização mais intensas nos Arranjos Populacionais de maior concentração ocupacional, cujas ocupações que mais se destacaram em 2018 são “Vendedor de comércio varejista”, “Alimentador de linha de produção”, “Auxiliar de escritório”, “Trabalhador da cultura de cana-de-açúcar”, “Operador de caixa”, “Trabalhador polivalente da confecção de calçados”, “Assistente administrativo”, “Costureiro, a máquina na confecção em série”, entre outras. Políticas públicas educacionais com foco em habilidades específicas e políticas de treinamento e profissionalização, podem permitir o desenvolvimento de habilidades cognitivas, a fim de minimizar os impactos negativos associados à perda de competitividade diante da chegada da nova revolução tecnológica no Brasil.

A literatura teórica presente nessa pesquisa indica os efeitos favoráveis da aglomeração e, por sua vez, da concentração ocupacional. Entretanto, como já apontado, esses efeitos se dão, principalmente, via trabalhadores, isto é, uma maior densidade de trabalhadores por área permite maiores níveis de *sharing*, *learning* e *matching*. Se a categoria de trabalhadores que gera maior concentração passa a ser a menos demandada no mercado de trabalho, isso impacta na produtividade e nos salários e, por sua vez, na distribuição de renda, e as empresas de determinada localidade não poderão avançar no seu processo de desenvolvimento, pois tanto os efeitos de *sharing* como os de *matching* seriam amortizados. Assim, em um momento futuro, o contexto da concentração ocupacional do Brasil, apresentado nessa pesquisa, poderia provocar um saldo líquido negativo para o crescimento econômico brasileiro, frente às mudanças previstas, tornando-se fundamental o estudo de políticas regionais a fim de melhorar a qualidade da oferta de trabalho e a estrutura produtiva do país.

Por fim, extensões futuras desse trabalho envolvem avaliar a coaglomeração ocupacional que permite verificar a tendência de trabalhadores com perfis de habilidades semelhantes se aglomerarem. A existência da coaglomeração ocupacional permitiria intensificar ainda mais os efeitos de *sharing*, *matching* e, em especial, o efeito de *learning*. Além disso, outra extensão interessante é avaliar o efeito marginal do *labor pooling* e das habilidades para os setores de atividade econômica no Brasil, de modo que seja possível ter uma análise de concentração ocupacional intrasetorial, possibilitando uma investigação mais detalhada. Uma análise mais detalhada pode ser feita, ainda, considerando as macrorregiões brasileiras, uma vez que as diferenças regionais brasileiras são intensas, desse modo, justifica-se avaliar como essas peculiaridades podem ser relevantes para explicar a concentração ocupacional.

3 ESPECIALIZAÇÃO OCUPACIONAL CATEGORIZADA POR HABILIDADES: PRÊMIO SALARIAL E DECOMPOSIÇÃO SALARIAL

3.1 INTRODUÇÃO

O estudo da desigualdade salarial é fundamental na compreensão do bem-estar econômico e distribuição de renda. Essa dinâmica econômica é reflexo do relacionamento intrínseco entre o mercado de trabalho e o mercado de bens. Nesse sentido, a temática é bastante explorada na literatura econômica (HOFFMANN, 2001; BARROS *et al.*, 2006; SABOIA, 2007; COELHO; CORSEUIL, 2002; FREGUGLIA; MENESES-FILHO; SOUZA, 2007; CACCIAMALI; TATEI; ROSALINO, 2009; MORICONI *et al.*, 2009). No Brasil, dadas as discrepâncias econômicas, sociais e culturais entre as macrorregiões, a análise da relação entre salários e espaço vem ganhado destaque, principalmente, no contexto de aglomeração de trabalhadores e firmas. Essa análise permite a formulação de políticas públicas com o intuito de mitigar diferenciais salariais que impeçam o desenvolvimento regional, em especial quando esses diferencias não são explicados pelas características produtivas dos trabalhadores e firmas (EHRL; MONASTERIO, 2020; BARUFI; HADDAD; NIJKAMP, 2016; NEVES JUNIOR, 2018; SILVA; AZZONI, 2021; SILVA, 2019).

Dentro desta perspectiva, diversas pesquisas buscam explicar os prêmios salariais nas cidades e regiões metropolitanas, com base nos efeitos da densidade, controlando as características não observadas dos indivíduos, das firmas ou de ambos (COMBES; DURANTON; GOBILLON, 2008; BARUFI; HADDAD; NIJKAMP, 2016; SILVA; AZZONI, 2021). A densidade tem sido uma variável de fundamental importância para a análise dos salários, em especial no contexto das economias urbanas. Entretanto, formas mais precisas de avaliar a aglomeração dos trabalhadores tem sido objeto de estudo na literatura econômica, no contexto das economias de localização, em especial no que se refere a aglomeração industrial. Todavia, poucas pesquisas se direcionam à análise da aglomeração ocupacional na determinação dos salários. Goldman, Klier e Walstrum (2019) enfatizam que há uma literatura antiga que avalia como a aglomeração afeta as decisões de localização das indústrias, todavia, só mais recentemente estuda-se como as economias de aglomeração afetam as decisões de localização dos trabalhadores.

Considerando a força das externalidades na determinação dos salários, Rossi-Hansberg, Sarte e Schwartzman (2019) indicam que essas externalidades de produção dependem da composição

ocupacional nas cidades, indicando que trabalhadores com maiores habilidades cognitivas necessitam das interações com trabalhadores mais capacitados, ao mesmo tempo em que as propiciam. Essa dinâmica no mercado de trabalho pode determinar os padrões de aglomeração ocupacional no espaço. Além disso, Rossi-Hansberg, Sarte e Schwartzman (2019) indicam que, em geral, as empresas acabam por se especializarem onde há mais intensidade de trabalhadores habilidosos. Neste sentido, percebe-se que a demanda de trabalho local acaba por ser influenciada pela oferta de trabalhadores habilidosos inicialmente estabelecidos. Esses trabalhadores acabam atraindo ainda mais trabalhadores com seus perfis de habilidades, intensificando o *learning*, um importante microfundamento da aglomeração, e favorecendo a determinação dos salários para grupos ocupacionais.

As evidências iniciais dessa pesquisa, a partir dos dados do Relatório Anual de Informações Sociais (RAIS) do Ministério da Economia (ME), indicam que no Brasil há uma relação positiva entre salários e densidade. Em 2018, observa-se que a elasticidade da densidade sobre os salários é de 3,9%. Todavia, essa medida não permite tecer comparações entre o efeito nos salários locais de acordo com o perfil de habilidades dos trabalhadores. Um avanço para a literatura no Brasil seria avaliar como a especialização categorizada por habilidades pode impactar nos salários e qual o papel da densidade nesse processo.

No Brasil, a distribuição dos trabalhadores entre as categorias de habilidades cognitivas, sociais e manuais é bastante desigual. Em 2018, verificou-se que existia um grande número de trabalhadores sociais e manuais no total de trabalhadores do mercado de trabalho formal: os trabalhadores intensivos em habilidades sociais tem uma participação de 46,9% e os trabalhadores intensivos em habilidades manuais tem uma participação de 43,5%. No entanto, há um pequeno número de trabalhadores intensivos em habilidades cognitivas, cuja participação no total de trabalhadores do mercado de trabalho formal brasileiro é de 9,6%, indicando que a estrutura produtiva do Brasil é pouco complexa em atividades cognitivas. Cabe ainda ressaltar que, do total de trabalhadores intensivos em habilidades cognitivas, a maioria tem seus postos de trabalho na macrorregião Sudeste, com um percentual de 41,67%, seguindo-se pelas macrorregiões Sul (21,72%), Nordeste (18,73%), Centro-Oeste (13,31%) e Norte (4,56%). Desse modo, percebe-se que os trabalhadores por perfis de habilidade se distribuem de forma desigual no território nacional e essa desigualdade pode ser ainda mais intensa quando se considera unidades regionais menores, como os Arranjos Populacionais brasileiros.

A distribuição espacial dos trabalhadores por habilidade impacta diretamente nas medidas de localização espacial, em especial, a especialização dos trabalhadores. Nesse contexto, é fundamental analisar os efeitos da especialização ocupacional categorizada por habilidades. Ressalta-se que, independentemente da participação de trabalhadores categorizados por habilidades no Brasil, o nível de especialização capta de fato os efeitos mais intensos de externalidades. Nesse sentido, deve-se investigar não só a participação dos trabalhadores intensivos em habilidades (cognitivas, sociais e manuais) e o prêmio salarial individual dessas habilidades, mas também o prêmio médio agregado da especialização de habilidades.

Para avaliar a aglomeração ocupacional na determinação salarial considera-se medidas de especialização. Wheaton e Lewis (2002) consideram que são os mercados locais ocupacionais e industriais que possibilitam retornos crescentes de escala, podendo determinar positivamente os salários dos trabalhadores, ou seja, quando há mais trabalhadores em uma indústria ou ocupação os efeitos intrínsecos da aglomeração, como as externalidades de conhecimento e ideias, permitem maiores retornos. Neste sentido, a relação convencional entre aglomeração e salários, identificada a partir de medidas simples de aglomeração, como área metropolitana e densidade, é pouco explicativa.

Rossi-Hansberg, Sarte e Schwartzman (2019) verificaram que os salários dos trabalhadores em ocupações específicas, condicionadas às suas habilidades na composição do trabalho total das cidades, dependem das externalidades entre as ocupações e entre os trabalhadores locais. Diante disto, seguindo a perspectiva de Wheaton e Lewis (2002) e Rossi-Hansberg, Sarte e Schwartzman (2019), busca-se aplicar para o Brasil uma análise mais direta do impacto da escala de mercado e externalidades sobre os salários dos trabalhadores em ocupações categorizadas por nível de habilidade cognitiva, social e manual. Logo, se tem uma análise mais precisa de como a aglomeração ocupacional pode determinar os salários, sendo um avanço em relação aos trabalhos já realizados no Brasil que utilizam tamanho da cidade e urbanização como medidas de aglomeração.

No Brasil, Neves Junior (2018) observa um maior prêmio salarial para as habilidades cognitivas, assim quanto maior o número de trabalhadores cognitivos no país maiores são as vantagens produtivas e de distribuição de renda. Desse modo, considerando que os dados da RAIS de 2018 mostram que o Brasil tem um pequeno percentual de trabalhadores intensivos em habilidades cognitivas e que a maioria destes estão localizados na macrorregião Sudeste, o

Brasil teria poucas vantagens econômicas e uma clara desigualdade regional. No entanto, ao considerar medidas espaciais que permitem inferir de forma mais eficiente os efeitos das externalidades sobre os salários e a produção, torna-se possível examinar a especialização ocupacional categorizada por habilidades para traçar um perfil do mercado de trabalho brasileiro e tecer maiores considerações acerca de suas vantagens, bem como elaborar políticas públicas para intensificá-las.

Neste sentido, esta pesquisa teve como objetivo principal mensurar como a especialização ocupacional, categorizada pelas habilidades requeridas em cada ocupação, pode determinar os salários. Isto é, estimou-se o prêmio salarial da especialização por habilidades classificadas em cognitivas, sociais e manuais para o mercado de trabalho brasileiro no período compreendido entre 2009 e 2018, controlando a endogeneidade causada pela causalidade reversa da escolha local e a endogeneidade causada pela heterogeneidade não observada de trabalhadores e firmas. Adicionalmente, decompõem-se os prêmios salariais da especialização ocupacional, categorizada por habilidades, em fatores associados à aglomeração e às características locais, utilizando como base o modelo de Combes, Duranton e Gobillon (2008).

Para tanto, foram estabelecidas medidas de especialização ocupacional no Brasil a fim de verificar como estas medidas determinam os diferenciais salariais. Assim, considerando as 2.557 ocupações delimitadas pela Classificação Brasileira de Ocupações (CBO-2002) foi possível categorizar as ocupações pela intensidade das habilidades requeridas a partir da métrica de habilidades de Maciente (2013) e aprimorada por Neves Júnior (2018), estabelecendo um indicador de especialização nas ocupações com habilidade cognitivas, sociais e manuais. Desse modo, foi possível avaliar suas peculiaridades na determinação dos salários por *clusters* de habilidades, considerando como unidade espacial de análise os Arranjos Populacionais brasileiros estabelecidos pelo IBGE (2016). Uma medida de especialização ocupacional categorizada por habilidades cognitivas, sociais e manuais permite compreender melhor o padrão de aglomeração dos trabalhadores e as externalidades geradas, pois espera-se que o *learning* dependa da composição ocupacional das cidades.

Entretanto, a análise do prêmio salarial da especialização ocupacional categorizada por habilidades apresenta alguns problemas estatísticos de endogenidade que merecem atenção. Roy (1951), Heckman, Stixrud e Urzua (2006), Acemoglu e Autor (2011) e David e Deming (2017) expressam a relação entre ocupações, salários e habilidades e reforçam o problema de

endogenidade da escolha ocupacional. Todavia, como as ocupações foram categorizadas pelas habilidades dos trabalhadores que são mais intensas para o desempenho da ocupação, dado que a variável de interesse, nesse modelo, é a especialização ocupacional no local de emprego e não a ocupação individual, elimina-se o problema da endogeneidade da escolha ocupacional individual. Além disso, a especialização ocupacional foi definida no nível dos Arranjos Populacionais e não dos indivíduos. No entanto, ainda é possível identificar um problema de endogeneidade se os trabalhadores com determinado tipo de habilidade decidirem se localizar nas áreas com maior especialização ocupacional naquele tipo.

Goldman, Klier e Walstrum (2019) e Combes, Duranton e Gobillon (2008) corroboram a principal conclusão do modelo de transmissão de habilidades de Duranton e Puga (2004), que conclui que existe um estado estacionário onde todos os trabalhadores habilidosos optam por residir nas cidades a fim de transmitirem as suas habilidades adquiridas, principalmente naquelas onde a densidade, representada pela proporção de trabalhadores não habilidosos, for maior. Dessa forma, identificou-se um problema de endogeneidade nesta pesquisa, uma vez que buscou-se verificar os diferenciais salariais a partir da especialização dos Arranjos Populacionais nas ocupações categorizadas por habilidade (cognitivas, sociais e manuais). Para contornar este problema utilizou-se variáveis instrumentais para lidar com o problema de endogeneidade existente e obter parâmetros consistentes. Além disso, foi possível lidar com o problema incluindo os efeitos fixos de individuais (habilidades) e firmas (produtividade) na equação salarial.

Abowd, Kramarz e Margolis (1999) desenvolveram um modelo de decomposição estatística dos salários, denominado de modelo AKM, onde mostram a importância de incluir o efeito das características não observadas dos trabalhadores e firmas para explicar os salários. O trabalho inicial dos autores buscou decompor o diferencial salarial setorial na França considerando a heterogeneidade não observada dos indivíduos e firmas. Na literatura brasileira, Silva (2017) aplicou esta metodologia para o mercado de trabalho brasileiro na busca por explicar o *sorting* no mercado de trabalho e observou que existe uma maior contribuição dos efeitos fixos individuais dos trabalhadores para explicar os salários, o que também está de acordo com as evidências apontadas por Abowd, Kramarz e Margolis (1999). Todavia, Silva (2017) aplicou a metodologia AKM em um modelo da Economia Urbana, ao incluir os efeitos da aglomeração na decomposição, fazendo uso de variáveis instrumentais para a densidade e da iluminação noturna para mensurar a área urbana.

Rossi-Hansberg, Sarte e Schwartzman (2019) indicam que a análise de múltiplas indústrias e ligações de insumo-produto são fundamentais para determinar a demanda por diferentes ocupações, estabelecendo assim a intensidade de cada ocupação nas indústrias locais. Além disso, Goldman, Klier e Walstrum (2019), ao avaliarem a aglomeração de ocupações intrasetoriais, evidenciam a importância em se considerar características das firmas e indivíduos conjuntamente para definir os padrões de aglomeração no espaço. Os mesmos utilizam dados ao nível de trabalhador e estabelecimentos, avaliando conjuntamente indústria, ocupação e local de trabalho. Diante destas considerações fica evidente que a especialização ocupacional capta não só variáveis relacionadas ao trabalhador, mas também relacionadas às firmas. Assim, considera-se fundamental aplicar a análise de decomposição salarial desenvolvida por Abowd, Kramarz e Margolis (1999) e Combes, Duranton e Gobillon (2008) a fim de verificar o papel efeito fixo da firma e do indivíduo na determinação do prêmio salarial da especialização ocupacional.

No Brasil, Ehrl e Monasterio (2020) buscaram avaliar como a concentração espacial de habilidades determina o nível de salários urbanos a partir da classificação de cinco dimensões de habilidades. Os autores ainda corrigem o problema da causalidade reversa a partir do uso do instrumento *shift-share* de Bartik. Todavia, é importante ressaltar a contribuição desse ensaio para a literatura brasileira, uma vez que se estima o prêmio salarial da especialização ocupacional categorizada por habilidades e a medida de especialização de habilidades é uma variável mais completa para avaliar os impactos da aglomeração. Enquanto a medida de concentração de Ehrl e Monasterio (2020) detecta a quantidade média de trabalho com determinadas habilidades no município, as medidas de especialização permitem ver a proporção de trabalhadores em cada categoria de habilidade dentro da unidade regional, ou seja, essa pesquisa avalia o quão especializado é o Arranjo Populacional em determinada habilidade, permitindo apreender melhor os efeitos de externalidades.

Além disso, essa pesquisa traz como contribuição empírica o controle da endogeneidade na determinação do prêmio de habilidades, considerando não só instrumento *Bartik shift* ocupacional, mas também variáveis instrumentais históricas e a heterogeneidade não observada de trabalhadores e firmas. Outra importante contribuição é a decomposição dos salários locais devido aos efeitos brutos e líquidos da especialização de habilidades, onde é possível avaliar a relação entre densidade populacional e especialização de habilidades no mercado de trabalho brasileiro.

Esta pesquisa está dividida em cinco seções, além desta introdução. Na seção seguinte, é apresentada uma estrutura teórica que permite compreender a relação entre habilidades e ocupação na determinação dos salários urbanos. Além disso, evidências empíricas são apontadas para a literatura internacional e nacional. Na seção 3, apresenta-se a estratégia empírica desenvolvida e a estratégia de identificação, dividida em duas partes: 1) a metodologia utilizada para determinar o prêmio salarial da especialização ocupacional, considerando o modelo de variáveis instrumentais e o painel de dados, para o controle da endogeneidade resultante da escolha locacional e da heterogeneidade não observada, respectivamente, e 2) a metodologia desenvolvida por Abowd, Kramarz e Margolis (1999) de decomposição salarial. Além disso, apresentam-se os dados utilizados e o tratamento empírico.

Na seção 4 é feita uma análise do mercado de trabalho formal brasileiro para o período de 2009-2018, considerando o contexto da especialização ocupacional categorizada por habilidades. Os resultados e discussões serão apresentados na seção 5, essa é dividida em: Prêmio salarial da especialização ocupacional categorizada por habilidades (modelos de variáveis instrumentais e de painel); e Decomposição do prêmio salarial da especialização ocupacional categorizada por habilidades. Por fim, a seção 6 apresenta as considerações finais desse ensaio.

3.2 HABILIDADES E OCUPAÇÃO NA DETERMINAÇÃO DOS SALÁRIOS URBANOS

Esta seção busca demonstrar como as habilidades dos trabalhadores impactam na dinâmica do mercado de trabalho, alterando a composição do emprego e salários, uma vez que é indissociável a análise entre habilidades e ocupações (MACIENTE, 2013; ACEMOGLU; AUTOR, 2011). Como esta pesquisa visa verificar como a especialização ocupacional impacta sobre os salários é de fundamental importância considerar as habilidades desempenhadas pelos trabalhadores uma vez que a determinação ocupacional e até mesmo a concentração e especialização ocupacional, devido aos efeitos de *sharing (labor pooling)*, *matching e learning*, são influenciadas pelas habilidades dos trabalhadores.

Vale ressaltar que, os modelos teóricos aqui apresentados servem como respaldo para avaliar os problemas empíricos enfrentados nessa pesquisa, demostrando a importância de considerar as habilidades dos trabalhadores como determinantes dos salários locais e da especialização (Modelo Ricardiano) e como o problema da causalidade reversa pode surgir, no contexto da determinação dos salários locais, em função da especialização (Modelo de difusão de

conhecimentos nas cidades). Isto é, as abordagens teóricas servem como base para o entendimento das relações empíricas presentes neste estudo e não são usadas para fins de replicação. Não obstante, se faz necessário algumas considerações teóricas gerais sobre a relação entre habilidades, ocupações e salários espaciais.

Como estabelecido pelo modelo de Roy (1951), trabalhadores com habilidades específicas acabam se direcionando para determinadas ocupações, existindo uma endogeneidade da escolha ocupacional. Além disso, trabalhadores em determinadas ocupações, que apresentam um rol de habilidades similares, tendem a se aglomerarem a fim de obter aumento de produtividade dados os efeitos de *sharing (labor pooling), matching e learning*. Logo, as habilidades, além de determinarem a escolha ocupacional, determinam a concentração e especialização dos trabalhadores em ocupações similares sendo essa uma variável fundamental para explicar os salários urbanos bem como suas alterações. David e Deming (2017), Acemoglu e Autor (2011), Heckman, Stixrud e Urzua (2006) evidenciam o papel das habilidades dos trabalhadores na dinâmica econômica do mercado de trabalho.

Heckman, Stixrud e Urzua (2006) consideram que há endogeneidade entre as habilidades cognitivas e não cognitivas e a escolaridade, enfatizando que a escolaridade é uma variável de escolha que explica os salários, mas, pode ser determinada pelas habilidades. Ao mesmo tempo, as habilidades cognitivas influenciam diretamente os salários. Os autores evidenciam que as habilidades cognitivas e não cognitivas têm o mesmo efeito sobre a escolha ocupacional. No que se refere aos salários, embora as habilidades cognitivas expliquem mais a variação dos salários, os efeitos das habilidades não cognitivas também são fortes nesta determinação. Quando se considera o emprego e a experiência, o gradiente de habilidades não cognitivas é maior do que o das habilidades cognitivas. Assim, os resultados dos autores apontam para a importância das habilidades não cognitivas, o que contradiz a literatura do comportamento humano de Herrnstein e Murray (1994) e Jensen (1998), que estabelecem a supremacia das habilidades cognitivas na explicação dos fenômenos econômicos e sociais no mercado de trabalho.

A base de dados proposta por Maciente (2013), que foi utilizada para a determinação dos salários nesta pesquisa, engloba um vetor de 21 fatores de habilidades, conforme mencionado no Ensaio anterior. Estas habilidades englobam aspectos cognitivos e não cognitivos, em um maior grau de desagregação. Conforme mencionado por Heckman, Stixrud e Urzua (2006) é

evidente o papel das habilidades para a escolha ocupacional e, por sua vez, para a concentração ocupacional nos centros aglomerativos. Dessa forma, uma hipótese desta pesquisa é que as habilidades determinam diretamente e indiretamente, via determinação da especialização ocupacional, os salários dos trabalhadores.

Considerando o papel crescente das habilidades sociais no desempenho do mercado de trabalho, David e Deming (2017) buscam explicar qual o impacto das habilidades sociais nos níveis de emprego e salários. Os autores indicam que as habilidades sociais têm certas peculiaridades que as diferem das habilidades cognitivas. Assim, ao longo da sua pesquisa, investigam a relação e a diferença nos salários e empregos a partir dessas classificações de habilidades. Esta ideia surge das constatações empíricas apontadas por Heckman (1995) e intensificadas por Heckman, Stixrud e Urzua (2006) de que as habilidades cognitivas explicam apenas uma parte das variações e comportamentos no mercado de trabalho, chamando atenção para o papel das habilidades não cognitivas, o que inclui as habilidades sociais e muitas outras dimensões de habilidades. Assim, ao observarem que as mudanças tecnológicas modificaram e vem modificando a estrutura de empregos, tarefas e ocupações no mercado de trabalho, David e Deming (2017) enfatizam que a interação social e humana é muito difícil de ser substituída e, neste sentido, indicam que empregos com altos níveis salariais exigem cada vez mais habilidades sociais, estas, por sua vez, podem ser complementares às habilidades cognitivas, intensificando ainda mais sua importância na determinação dos salários altamente remunerados.

David e Deming (2017) consideram o modelo teórico apresentado por Acemoglu e Autor (2011) como um elo fundamental para avaliar as habilidades sociais, a distribuição e complexidade de tarefas e a variação da produtividade dos trabalhadores. O modelo de Acemoglu e Autor (2011) diferencia habilidades e tarefas e os trabalhadores são alocados em ocupações que lhe proporcionem vantagens comparativas e que aumentam sua produtividade. Tal como Acemoglu e Autor (2011) e Heckman, Stixrud e Urzua (2006), indicam que as habilidades podem vir a determinar a escolha ocupacional dos trabalhadores, ao afirmarem que trabalhadores com maiores habilidades sociais estão mais propensos a trabalharem em ocupações menos rotineiras e intensivas em habilidades sociais, e essa escolha determinaria o retorno salarial dos trabalhadores.

Acemoglu e Autor (2011) evidenciam a tendência recente de demanda e oferta de habilidades no mercado de trabalho, enfatizando que nos EUA verifica-se um prêmio salarial às habilidades

capturadas, inicialmente, pelo retorno salarial dos trabalhadores com nível superior de escolaridade. Neste sentido, observam que, apesar do elevado aumento da oferta de trabalhadores com nível superior, a demanda por estes também sofreu aumento considerável o que acabaria por explicar parte do retorno salarial positivo para trabalhadores qualificados. No Brasil, esta tendência também é observada e diversos trabalhos evidenciam um diferencial salarial positivo e crescente para trabalhadores de ensino superior.

Rodrigues, Vieira e Freguglia (2015), considerando a existência de aumento de desigualdade entre trabalhadores de alta qualificação, evidenciado no trabalho de Tavares e Menezes-Filho (2013), buscam explicar a evolução dessa desigualdade considerando uma abordagem ocupacional. Avaliando o índice de Gini por nível de escolaridade, os autores perceberam um aumento na desigualdade entre os trabalhadores do ensino superior, tendência totalmente contrária para os demais grupos de escolaridade. Assim, indicam que esta conjuntura pode ser explicada pela heterogeneidade de habilidades entre esses trabalhadores ou mudanças na estrutura das ocupações causadas pelo processo tecnológico.

Acemoglu e Autor (2011) apresentam uma modelagem teórica e empírica que permite estabelecer a relação entre habilidades, ocupação e tecnologias e seus impactos sobre os salários. Os mesmos apresentam, inicialmente, o modelo canônico, que reflete uma organização sobre a vasta literatura que trata de habilidades e desigualdade salarial. Posteriormente, dadas as suas limitações, apresentam um modelo mais completo e que se adequa mais a realidade: o Modelo Ricardiano do Mercado de Trabalho.

3.2.1 O Modelo Ricardiano do Mercado de Trabalho

O modelo canônico considera diferentes ocupações e habilidades, sendo estas imperfeitamente substituíveis. É fundamental ressaltar que, neste contexto, não há distinção entre habilidades e ocupações, sendo essa uma limitação do modelo. Considerando dois tipos de habilidades, alta (H) e baixa (L), cada tipo de trabalhador que possui alta habilidades desempenha tarefas afins, que são diferentes das tarefas desempenhadas pelos trabalhadores de baixa habilidade. Assim, tem-se que L é a oferta de trabalhadores de baixa habilidade e H é a oferta de trabalhadores de alta habilidade (ACEMOGLU; AUTOR, 2011):

$$L = \int_{i \in \mathcal{L}} l_i d_i \text{ e } H = \int_{i \in \mathcal{H}} h_i d_i \tag{3.1}$$

Em que $\mathcal{L}$ e $\mathcal{H}$ indicam o conjunto de trabalhadores *i* de baixa habilidade e alta habilidade, respectivamente, que possuem uma determinada distribuição de unidades de eficiência (l_ie h_i). Considerando o lado da produção, considera-se a seguinte função de produção com elasticidade de substituição constante:

$$Y = [(A_L L)^{\frac{\sigma-1}{\sigma}} + (A_H H)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \quad (3.2)$$

Em que A_L e A_H representam o fator tecnológico e $\sigma \epsilon [0, \infty)$ representa a elasticidade de substituição entre os trabalhadores de alta e baixa habilidade. Os trabalhadores são substitutos perfeitos quando $\sigma > 1$ e complementares perfeitos se $\sigma < 1$.

Acemoglu e Autor (2011) consideram a interpretação deste termo fundamental para verificar as consequências no mercado de trabalho no que se refere à distribuição de rendimentos. Considerando a hipótese de concorrência perfeita, os salários dos trabalhadores representam o produto marginal dos tipos de trabalhadores. Logo, diferenciando a função de produção se obtém:

$$w_L = \frac{\partial Y}{\partial L} = A_L^{\frac{\sigma-1}{\sigma}} [A_L^{\frac{\sigma-1}{\sigma}} + A_H^{\frac{\sigma-1}{\sigma}} (H/L)^{\frac{\sigma-1}{\sigma}}]^{\frac{1}{\sigma-1}}$$

$$w_H = \frac{\partial Y}{\partial H} = A_H^{\frac{\sigma-1}{\sigma}} [A_L^{\frac{\sigma-1}{\sigma}} (H/L)^{-\frac{\sigma-1}{\sigma}} + A_H^{\frac{\sigma-1}{\sigma}}]^{\frac{1}{\sigma-1}}$$

Em que, estes representam o salário unitário, enquanto o salário do trabalhador *i* é dado pela multiplicação deste salário unitário pelo seu grau de eficiência. Verifica-se que $\partial w_L/(\frac{\partial H}{L}) > 0$, o que indica que a medida que aumenta a oferta de trabalhadores qualificados os salários dos trabalhadores de baixa habilidade aumenta, isso ocorre dado o efeito de substituição existente entre os trabalhadores. Neste sentido, verifica-se ainda, de modo inverso, que $\partial w_H/(\frac{\partial H}{L}) < 0$, ou seja, se a oferta de trabalhadores qualificados aumenta, considerando a demanda fixa, os seus salários devem diminuir. Observa-se, ainda, que $\partial w_L/\partial A_L > 0$, $\partial w_L/\partial A_H > 0$, $\partial w_H/\partial A_L > 0$ e $\partial w_H/\partial A_H > 0$, isso indica que incrementos tecnológicos para qualquer tipo de trabalhadores fazem com que o salário de ambos aumentem. Acemoglu e Autor (2011)

asseguram que o modelo canônico não permite explicar decréscimos nos salários em situações de aumento de oferta tecnológica.

A análise sobre os diferenciais salariais, ou o prêmio de habilidade, pode ser feita combinando as equações de salários marginais, pela seguinte relação w_H/w_L e obtendo-se o seu logaritmo:

$$ln\omega = \frac{\sigma-1}{\sigma}\ln\left(\frac{A_H}{A_L}\right) - \frac{1}{\sigma}\ln(\frac{H}{L}) \tag{3.3}$$

ou seja, o prêmio de habilidade depende dos níveis tecnológicos por tipo de habilidade e da oferta relativa de trabalhadores. Assim, considerando um determinado nível tecnológico, um aumento relativo de trabalhadores habilidosos provoca uma diminuição no prêmio de habilidade por um fator $1/\sigma$, que representa o inverso da elasticidade de substituição entre trabalhadores de alta e baixa habilidade.

Acemoglu e Autor (2011) evidenciam para os EUA um aumento na oferta de trabalhadores de alta habilidade, todavia, não se verificou uma diminuição no prêmio salarial dos trabalhadores mais habilidosos, conforme já mencionado. Esta tendência por sua vez não representa uma negação empírica do modelo canônico, dado que ao avaliar a equação acima, acredita-se, que simultaneamente, há um aumento na oferta de trabalhadores habilidosos e existiram mudanças nos padrões tecnológicos que modificaram a composição da demanda por trabalho, aumentando a demanda por habilidade. Os autores, assim, evidenciam o sucesso do modelo canônico, tanto teoricamente como empiricamente, todavia não deixam de apontar suas limitações, principalmente dado o atual desenvolvimento econômico e tecnológico.

A principal limitação conceitual apontada por Acemoglu e Autor (2011) se refere ao fato de que, no modelo, as ocupações (tarefas) e as habilidades são tratadas como equivalentes. Essa hipótese deve ser relaxada uma vez que, conforme indicado pelos autores, uma ocupação (tarefa) é a aplicação do trabalho que permite produzir um bem ou serviço, já a habilidade se refere capacidade (destreza, determinação, facilidade) de execução da tarefa. Uma ocupação pode exigir um rol de habilidades diversas em maior ou menor intensidade, logo um trabalhador com determinado nível de habilidade pode desempenhar várias tarefas. Dada a dinâmica atual do mercado de trabalho envolvida em uma trajetória de crescente desenvolvimento tecnológico, tratar ocupações e habilidades como fatores distintos se adequa mais a realidade econômica.

Vale mencionar que, a diferença entre ocupações e habilidades está de acordo com a abordagem de Maciente (2013), isto é, as diferentes habilidades que os trabalhadores podem possuir são usadas em maior ou menor grau em diversas ocupações.

Uma segunda limitação reside na forma como a tecnologia é tratada. O modelo canônico considera a tecnologia como exógena e, além disso, existe um viés de habilidade, ou seja, as novas tecnologias são voltadas para trabalhadores com altas habilidades. A quarta revolução industrial pela qual o mundo passa, caminha para uma tendência de substituição de cérebros, isto é, tecnologia de engenharia genética e neurotecnologias. Talvez essa limitação do modelo canônico não seja tão forte no Brasil, pois o desenvolvimento tecnológico é menos acentuado do que nos EUA e em outros países mais desenvolvidos. No Brasil, de acordo com Maciente, Rauen e Kubota (2019), a quarta revolução industrial caminha a passos muito lentos quando se compara com outros países como os EUA.

Dadas essas limitações, Acemoglu e Autor (2011) avaliam em detalhes a estrutura do mercado de trabalho dos EUA e União Europeia e percebem algumas tendências como: declínio nos salários dos trabalhadores de baixa qualificação; o avanço tecnológico alterou a estrutura ocupacional do mercado de trabalho, substituindo determinadas ocupações; e alterações sistemáticas na composição do emprego entre ocupações, com aumento de emprego em ocupações com alta escolaridade. Dessa forma, os autores indicam que essa polarização reflete mudanças na alocação de habilidades entre as ocupações no mercado de trabalho.

Para estabelecer um modelo que incorpore melhor estas tendências, um arcabouço teórico denominado de Modelo Ricardiano do Mercado de Trabalho é proposto de forma a estabelecer a relação entre habilidades, tecnologia, ocupações e salários. Esse nome é proposto uma vez que o modelo estabelece uma análise de vantagem comparativa entre os diferentes tipos de trabalhadores, dado suas habilidades. O novo modelo incorpora de forma explícita a diferença entre habilidades e tarefas, além de avaliar a polarização ocupacional considerando pelo menos três tipos de habilidades: baixa, média e alta.

O modelo estabelece uma função de produção Cobb-Douglas, onde apenas um bem final é produzido a partir de uma combinação de um *continuum* de tarefas ou ocupações:

$$lnY = \int_0^1 \ln y(i)di \tag{3.4}$$

Em que $y(i)$ é o nível de produção que corresponde a ocupação i. Conforme mencionado, o fator de produção trabalho, classifica os trabalhadores entre: baixa, média e alta habilidade, e a tecnologia incorporada no processo produtivo também apresenta a mesma classificação. Assim, cada ocupação ou tarefa apresentará uma função de produção definida como:

$$y(i) = A_L\alpha_L(i)l(i) + A_M\alpha_M(i)m(i) + A_H\alpha_H(i)h(i) + A_K\alpha_K(i)k(i) \tag{3.5}$$

ou seja, uma ocupação pode ser desempenhada com um conjunto diverso de habilidades, além de um determinado nível de capital empregado. A produtividade dos trabalhadores de baixa, média e alta habilidade na ocupação i é definida por $\alpha_L(i)$, $\alpha_M(i)$ e $\alpha_H(i)$, respectivamente. As variáveis $l(i)$, $m(i)$ e $h(i)$ representam a quantidade de trabalhadores de baixa, média e alta habilidade, respectivamente, que desempenham a ocupação i.

Convém ressaltar que, apesar da ocupação poder ser desempenhada por trabalhadores de qualquer tipo de habilidade, é evidente que algum tipo pode apresentar vantagem comparativa, capturada pela magnitude do coeficiente alfa. Assim, cada ocupação ou tarefa pode ter como fator de produção mais significante na produção um dos diferentes tipos de habilidades. Logo, cada ocupação apresenta uma estrutura produtiva diferente, trazendo um caráter mais real para o modelo.

A fim de deixar mais clara a ideia de vantagem comparativa, adota-se a hipótese de que a relação entre a produtividade de dois tipos de trabalhadores nos limites mínimos e máximos, $\alpha_L(i)/\alpha_M(i)$ e $\alpha_M(i)/\alpha_H(i)$, é continuamente diferenciável e estritamente decrescente. O que tem como implicação que os índices mais elevados correspondem a tarefas mais complexas nas quais os trabalhadores de alta habilidade são melhores do que os trabalhadores de habilidade média e os trabalhadores de habilidade média são melhores do que os trabalhadores de baixa qualificação (ACEMOGLU; AUTOR, 2011).

Acemoglu e Autor (2011) estabelecem as condições para um equilíbrio, inicialmente sem considerar o nível de capital da economia, de modo que as máquinas não podem substituir os trabalhadores. Esse equilíbrio tem como base teórica a hipótese de vantagem comparativa:

considerando um conjunto de ocupações/tarefas I_L e I_H, verifica-se que, se determinada ocupação tiver vantagem comparativa com a utilização de trabalhadores de baixa habilidade, $i < I_L$, esta será executada por trabalhadores de baixa habilidade; o mesmo acontece quando considera-se ocupações onde $i > I_H$, as quais, dada sua vantagem comparativa na utilização de trabalhadores de alta habilidade, serão desempenhadas por esses; por fim, as ocupações intermediarias médias $i \in (I_L, I_H)$, consideradas tarefas rotineiras e administrativas, devem utilizar trabalhadores que tenham habilidades médias. Dessa forma, os trabalhadores vão se direcionando para as tarefas em que apresentam vantagem comparativa. Esse argumento embasa o método de classificação das ocupações brasileiras nessa pesquisa: intensivas em habilidades cognitivas, intensivas em habilidades sociais e intensivas em habilidades manuais, conforme será explicado na próxima seção.

A determinação final do equilíbrio depende ainda dos preços dos fatores de produção, w_L, w_M e w_H, onde os salários dos tipos de trabalhadores são diferentes, de modo que quanto maior o nível de habilidade maior o salário, sendo esses determinados numa situação de concorrência perfeita. Considerando um conjunto de ocupações I_L, onde os custos de produção utilizando trabalhadores de baixa e média habilidade são os mesmos e a hipótese de vantagem comparativa é válida, ou seja, $\alpha_L(i)/\alpha_M(i)$ é estritamente decrescente, os custos com mão de obra para executar uma tarefa do tipo $i < I_L$ serão menores se forem utilizados trabalhadores com baixa habilidade. Na situação inversa, $i > I_L$, a vantagem comparativa estabelece uma condição ótima com a utilização de trabalhadores de habilidade média. A mesma análise pode ser aplicada para o caso de I_H.

Dessa forma, percebe-se que a minimização dos custos implica que, considerando os diferentes níveis de salários para cada tipo de habilidade, a alocação ótima dos trabalhadores entre as ocupações obedecerá ao critério de vantagem comparativa, sendo que o conjunto de ocupações será dividido em três conjuntos convexos: 1. Ocupações desempenhadas por trabalhadores de baixa habilidade; 2. Ocupações desempenhadas por trabalhadores de média habilidade e 3. Ocupações desempenhadas por trabalhadores de alta habilidade. Diante destas considerações, o equilíbrio que pode ser resumido no Lema 1 (ACEMOGLU; AUTOR, 2011, p. 1122, "tradução nossa"): *Em qualquer equilíbrio existem* I_L *e* I_H *tais que* $0 < I_L < I_H < 1$ *e para qualquer* $i < I_L, m(i) = h(i) = 0$, *para qualquer* $i \in (I_L, I_H), l(i) = h(i) = 0$, *e para qualquer* $i > I_H, l(i) = m(i) = 0$.

Como já mencionado, esse lema é constituído a partir da hipótese de vantagem comparativa, que deixa clara a possibilidade de substituição das habilidades entre as ocupações, ideia fundamental do modelo segundo Acemoglu e Autor (2011). De forma geral, verifica-se que os agentes dessa economia (empresas e trabalhadores) definem de forma ótima a alocação de ocupações para cada grupo de habilidade. Essa ideia encontra-se relacionada com o Modelo de Roy (1951), onde a ideia central que rege a sua fundamentação é a de que os indivíduos se autosselecionam nas ocupações de acordo com suas respectivas habilidades. Isso ocorre porque cada ocupação exige inúmeras habilidades dos trabalhadores, todavia, algumas habilidades são mais apropriadas para determinadas atividades do que para outras, dessa forma, os trabalhadores direcionam-se para os setores onde suas habilidades proporcionariam maior produtividade ou são melhores aproveitadas, resultando em maior rendimento[18]. Desse modo, um agente só escolhe uma ocupação se seus ganhos forem maiores nessa ocupação, o que sugere relações e resultados endógenos no mercado de trabalho.

A oferta de trabalhadores habilidosos e a alocação de ocupações determinará o diferencial salarial de equilíbrio entre os trabalhadores. Conforme já mencionado, considerando um mercado competitivo, o salário para cada nível de habilidade é o mesmo independentemente da tarefa desempenhada. Logo, o produto marginal para cada grupo de habilidade também será o mesmo:

$$w_L = p(i)A_L\alpha_L(i)\ para\ todo\ i < I_L$$
$$w_M = p(i)A_M\alpha_M(i)\ para\ todo\ I_L < i < I_H$$
$$w_H = p(i)A_H\alpha_H(i)\ para\ todo\ i > I_H$$

Neste sentido, as diferenças de produtividade entre os trabalhadores de mesmas habilidades desempenhando tarefas diferentes, são obtidas pelas diferenças salariais dos mesmos, ou seja, pelo preço dessas duas tarefas. Sendo o salário dado, tem-se que, ao desempenhar a tarefa i ou i' o salário é o mesmo:

[18] O trabalho original desenvolvido por Roy (1951) considerava a alocação de trabalhadores entre os setores da caça e pesca. Mais tarde Lee (1978) adapta este modelo de modo a investigar os efeitos da sindicalização sobre a estrutura salarial, averiguando a autosseleção para participar de um sindicato. Seguindo o mesmo modelo, Borjas (1987) avalia a autosseleção entre imigrantes, considerando que os trabalhadores migram de um país para outro avaliando seus ganhos médios, ou seja, a partir do confronto de suas habilidades e das atividades tipicamente predominantes em cada país, os indivíduos decidiriam por aquele onde os ganhos provenientes do seu trabalho fossem maiores.

$$w_L = p(i')A_L\alpha_L(i')$$
$$w_L = p(i)A_L\alpha_L(i)$$

Logo:

$$P_L \equiv p(i')A_L\alpha_L(i') = p(i)A_L\alpha_L(i) \quad (3.6)$$

Em que P_L representa o índice de preço das tarefas desempenhadas pelos trabalhadores de baixa habilidade. O mesmo pode ser aplicado para os demais tipos de habilidades:

$$P_M \equiv p(i')A_M\alpha_M(i') = p(i)A_M\alpha_M(i) \quad (3.7)$$
$$P_H \equiv p(i')A_H\alpha_H(i') = p(i)A_H\alpha_H(i) \quad (3.8)$$

Considerando a condição de compensação do mercado, dada por:

$$\int_0^1 l(i)di \leq L\,, \int_0^1 m(i)di \leq M \;\; e \;\; \int_0^1 h(i)di \leq H$$

Obtém-se:

$$l(i) = \frac{L}{I_L} \; para \; todo \; i < I_L$$
$$m(i) = \frac{M}{I_H - I_L} \; para \; todo \; I_H > \; i > I_L$$
$$h(i) = \frac{H}{1 - I_H} \; para \; todo \; i > I_H$$

Com isso, é possível estabelecer a relação dos índices de preços de habilidades das tarefas desempenhados pelos trabalhadores. Considerando os casos extremos, a relação entre trabalhadores de alta e média habilidade é dada por:

$$\frac{P_H}{P_M} = \left(\frac{A_H H}{1-I_H}\right)^{-1}\left(\frac{A_M M}{I_H - I_L}\right) \quad (3.9)$$

A relação entre média e baixa habilidade é dada por:

$$\frac{P_M}{P_L} = \left(\frac{A_M M}{I_H - I_L}\right)^{-1}\left(\frac{A_L L}{I_L}\right) \quad (3.10)$$

A relação entre os índices de preços permite determinar os salários e captar as desigualdades salariais existentes entre os grupos de habilidades. Como em qualquer modelo de concorrência perfeita, os salários para cada tipo de trabalhador são determinados pelo valor do respectivo produto marginal. Desse modo, os salários para cada tipo de habilidade são expressos da seguinte forma:

$$w_L = P_L A_L$$
$$w_M = P_M A_M$$
$$w_H = P_H A_H$$

A relação entre os níveis salarias expressam a desigualdade salarial existente entre os trabalhadores de diferentes perfis de habilidades. Considerando as equações de preços relativos pode-se expressar a relação entre os salários de alta e média habilidade (ou o diferencial salarial dos trabalhadores de alta e média habilidade como)[19]:

$$\frac{w_H}{w_M} = \left(\frac{1-I_H}{I_H-I_L}\right)\left(\frac{H}{M}\right)^{-1} \tag{3.11}$$

O diferencial salarial entre trabalhadores de média e baixa habilidade é dado por:

$$\frac{w_M}{w_L} = \left(\frac{I_H-I_L}{I_L}\right)\left(\frac{M}{L}\right)^{-1} \tag{3.12}$$

Como é possível verificar, os salários relativos dependem da oferta de trabalhadores de baixa, média e alta habilidade e das ocupações ou tarefas desempenhadas por grupos extremos de habilidade, alta e baixa. Assim, verifica-se que a estrutura ocupacional dos trabalhadores e sua oferta determinam os diferenciais salariais dos trabalhadores categorizados por nível de habilidade. E esse é justamente o objetivo desta pesquisa, verificar como a especialização ocupacional categorizada por habilidades (que representam a oferta de trabalho de habilidades) determina os salários. Esse modelo deixa clara a relação entre ocupações, salários e habilidades. Além disso, expressa que existe um diferencial salarial quando se fala em ocupações que se

[19] Considerações adicionais sobre como se chegar as equações, que expressam a diferença salarial entre os trabalhadores de diferentes tipos de habilidade, podem ser verificadas no Anexo.

deve às habilidades dos trabalhadores, sendo essas diferentes do nível de escolaridade e da ocupação desempenhada, conforme hipótese fundamental do modelo.

Entretanto para completar essa abordagem teórica é fundamental incluir na análise teórica um modelo que capta a transmissão de habilidades num contexto urbano, considerando as características das cidades, uma vez que a medida de especialização proposta nesta pesquisa considera as diferenças locais no espaço urbano. Para tanto, torna-se fundamental avaliar os microfundamentos da aglomeração no que se refere ao *learning*, modelo este proposto por Duranton e Puga (2004) que avalia a transmissão de habilidades em um contexto de especialização urbana, objetivo da próxima seção.

Entretanto, é fundamental, antes de finalizar essa seção, ressaltar que as pesquisas de Roy (1951), Heckman, Stixrud e Urzua (2006), Acemoglu e Autor (2011) e David e Deming (2017) expressam a relação entre ocupações, salários e habilidades. Desse modo, ao se tratar sobre diferenciais salarias que categorizem o trabalhador por ocupações, em especial quando se trata de concentração e especialização ocupacional, é fundamental considerar as habilidades dos trabalhadores, pois os retornos às habilidades são comprovados na literatura empírica. Devido a isso, considera-se fundamental, ao buscar explicar como a especialização ocupacional determina os salários, incluir variáveis que captem as habilidades dos trabalhadores. Para suprir essa necessidade, um vetor de habilidades obtido a partir da base de dados de Maciente (2013) foi considerado na modelagem empírica, a fim de categorizar essas ocupações por tipo de habilidades (cognitiva, social e manual) e evitar possíveis vieses de habilidade. Assim, como as ocupações foram categorizadas pelas habilidades dos trabalhadores que são mais intensas para o desempenho da ocupação, elimina-se o problema da endogeneidade da escolha ocupacional, dado que a variável de interesse nesse modelo, especialização ocupacional, categoriza os trabalhadores pela intensidade de habilidades requeridas nas ocupações, em cada área.

Vale mencionar, ainda, a existência de endogeneidade presente no modelo, uma vez que se constatou que habilidades determinam a escolha ocupacional e, aliadas às forças aglomerativas, essas habilidades estabelecem condições propícias para a concentração ocupacional, determinando os salários dos trabalhadores. Todavia, convém indicar que, no que se refere a endogeneidade da escolha ocupacional, dado o objetivo e a estratégia empírica desta pesquisa, este problema não interfere diretamente na determinação dos parâmetros, uma vez que não foi

utilizada a variável ocupação, mas um indicador agregado de especialização ocupacional, considerando os Arranjos Populacionais do Brasil, onde as ocupações foram ainda categorizadas por habilidades (cognitivas, sociais e manuais), ou seja, a variável principal de análise avalia a escolha locacional dos indivíduos a partir das suas habilidades. Adicionalmente, Combes, Duranton e Gobillon (2008) indicam que existe um crescimento mais forte nos salários nas grandes cidades cuja causa principal é a concentração mais rápida e intensa de habilidades nesses grandes centros, essa evidência está de acordo com a principal conclusão teórica do modelo de Duranton e Puga (2004) que avalia a transmissão de habilidades e o equilíbrio estático dos trabalhadores habilidosos e não habilidosos no espaço urbano. Diante deste cenário, formas para corrigir essa endogeneidade serão propostas posteriormente.

3.2.2 O Modelo de difusão de conhecimentos nas cidades (Transmissão de habilidades)

O modelo de Duranton e Puga (2004) mostra como a geografia desempenha um papel crucial na aquisição de habilidades. Dessa forma, é possível verificar o papel das cidades e sua densidade na determinação das habilidades dos indivíduos, a partir dos mecanismos de *sharing (labor pooling), learning e matching* que se tornam mais intensos em grandes centros e atraem mais trabalhadores habilidosos, intensificando a concentração ocupacional na tentativa de obter melhores resultados salariais. Os autores estabelecem uma ideia básica de que as cidades são um ambiente perfeito para transmissão de conhecimento e habilidades, pois a proximidade entre indivíduos facilita a troca de informações e habilidades. Duranton e Puga (2004) indicam que quanto maiores as cidades, maiores as possibilidades de difusão de habilidades e, em um equilíbrio estático, as maiores cidades acabam atraindo indivíduos mais habilidosos, pois assim estes maximizam sua função de bem-estar, refletida em seu consumo.

A formulação do modelo considera a ideia de gerações sobrepostas, onde indivíduos vivem por dois períodos: tem-se indivíduos jovens no primeiro período e velhos no segundo. Logo, os indivíduos buscam maximizar o seu consumo esperado ao longo de sua vida. Considera-se que todos os indivíduos no primeiro período são inexperientes, assim possuem baixos níveis de habilidades, e que essas habilidades determinam o nível de produtividade dos indivíduos, e por sua vez, seus salários. A produtividade de um indivíduo sem habilidades é $\underline{\beta}$ e a produtividade de um indivíduo com habilidades é $\overline{\beta}$, como condição básica tem-se que $\overline{\beta} > \underline{\beta}$.

Os trabalhadores inexperientes podem optar por se tornar habilidosos e isso vai determinar a escolha do local onde irão residir no início e no final da sua vida. É interessante, ainda, mencionar que morar na cidade é uma condição necessária para a aquisição de habilidades, todavia, não é uma condição suficiente. Os autores ainda estabelecem que a probabilidade de se tornar hábil na cidade *i* depende da quantidade de trabalhadores qualificados lá existentes, logo, depende de uma distribuição de probabilidade $f(N_i^S)$, onde N_i^S é o número de trabalhadores qualificados na cidade *i*, com $f' > 0$ e $f'' < 0$.

Como hipótese adicional os autores estabelecem que o único benefício de morar nas cidades é a possibilidade de se tornarem mais habilidosos. Além disso, morar nas cidades tem custos de alugueis e transporte, τN_i, o que não ocorre no interior, dessa forma, os indivíduos enfrentam um *trade-off*. Além disso, os indivíduos qualificados podem optar por permanecer nas cidades, pois os mesmos recebem uma remuneração por transferirem suas habilidades aos trabalhadores jovens. Duranton e Puga (2004) indicam que a decisão de onde morar quando jovem reflete um investimento arriscado em capital humano, pois apesar dos custos de se viver na cidade, o simples fato de viver em uma cidade não implica que o indivíduo conseguirá absorver habilidades. Se o indivíduo se tornar habilidoso na cidade, ele aufere um valor endógeno, V_i. Esse excedente, por sua vez, é dividido igualmente ao longo do ciclo de vida, $V_i/2$.

Assim, de modo geral, o modelo de transmissão de habilidades de Duranton e Puga (2004) conclui que existe um estado estacionário onde todos os trabalhadores jovens optam por morar nas cidades para tentarem adquirir habilidades, podendo ter sucesso ou não, e todos os trabalhadores habilidosos optam por residir nas cidades a fim de transmitir as suas habilidades adquiridas, principalmente naquelas onde a densidade, representada pela proporção de trabalhadores não habilidosos, for maior.

Desse modo, intensifica-se o que já foi mencionado na seção anterior, onde Combes, Duranton e Gobillon (2008) afirmam que a existência de trabalhadores habilidosos nos grandes centros é mais intensa. Dessa forma, identifica-se um problema de endogeneidade no modelo estabelecido nesta pesquisa, uma vez que busca-se verificar os diferenciais salariais que captem a especialização nos Arranjos Populacionais nas ocupações categorizadas por habilidade (cognitivas, sociais e manuais). Essa endogeneidade resulta de choques locais que elevam/reduzem os salários em determinadas ocupações e atraem/expulsam mais trabalhadores

dessas ocupações categorizadas por habilidades para estes locais, uma vez que essa medida de especialização ocupacional capta a escolha locacional dos indivíduos a partir das suas habilidades. As formas utilizadas para a correção dessa endogeneidade serão apresentadas na seção de metodologia.

3.2.3 Evidências empíricas

Os modelos teóricos apresentados mostraram a importância de avaliar os determinantes dos salários considerando a ocupação dos trabalhadores e suas habilidades. Além disso, demonstram os problemas empíricos de causalidade reversa no uso das habilidades para explicar os salários espacialmente. Na literatura empírica poucos trabalhos se concentram na relação salários, habilidades e espaço, dada a dificuldade de se tratar com variáveis que possam vir a capturar as habilidades.

3.2.3.1 Evidências empíricas internacionais

Na literatura internacional, as pesquisas de Wheaton e Lewis (2002) e Rossi-Hansberg, Sarte e Schwartzman (2019) avaliam as externalidades causadas nos salários dos trabalhadores dada a concentração ocupacional, mas somente a pesquisa de Rossi-Hansberg, Sarte e Schwartzman (2019) concentra-se no papel das habilidades para a determinação desta concentração.

Rossi-Hansberg, Sarte e Schwartzman (2019) consideram como um insumo valioso aqueles trabalhadores que desenvolvem ocupações que exigem habilidades cognitivas não rotineiras, sendo estes fundamentais para o bem-estar da economia. Além disso, a concentração destes trabalhadores intensifica as externalidades positivas nos salários, uma vez que, a produtividade marginal dos trabalhadores aumenta, dadas as possibilidades de aprendizado e colaboração. Neste sentido, os autores questionam como a alocação desses trabalhadores pode aumentar o bem-estar geral da economia. Assim, o estudo tem como objetivo principal avaliar como a concentração ocupacional dos trabalhadores categorizados em habilidades cognitivas não rotineiras e habilidades não cognitivas pode determinar os salários e, dada a ideia de externalidade, como essa pode determinar a alocação dos indivíduos no espaço. Adicionalmente, dada uma alocação ótima, definida a partir de um modelo de equilíbrio espacial com múltiplas indústrias, os autores buscam entender e projetar políticas espaciais para melhorar a alocação ocupacional dos trabalhadores dos EUA.

Considerando trabalhadores dos EUA para os anos de 2011 e 2015 e usando as Tabelas de Uso do *Bureau of Economic Analysis* (BEA), Rossi-Hansberg, Sarte e Schwartzman (2019) ao calcularem o salário médio dos trabalhadores em cada cidade controlando pelas características observadas dos trabalhadores, verificaram que os salários das ocupações cognitivas não rotineiras, em relação aos demais (ocupações não cognitivas), aumentam com o número de trabalhadores em ocupações cognitivas na composição do emprego total, o que corrobora a ideia de externalidades positivas sobre os salários dada a concentração. Assim, as diferenças nos salários nas cidades dependem da composição do emprego ocupacional. Mesmo considerando a importância da demanda de trabalhadores, representada pelo *mix* de indústrias das diferentes cidades dos EUA, os autores verificaram que as parcelas salariais diferem em mais de um para um em relação às diferenças no *mix* industrial. Logo, os autores indicam que a composição industrial não pode ser o único determinante da demanda por trabalho e dos salários entre as cidades.

No que se refere às externalidades de produção por tipo de trabalhador, constatou-se que a produtividade aumenta dada a presença ou proximidade de trabalhadores do mesmo grupo ocupacional, um aumento no logaritmo da concentração ocupacional aumenta os salários dos trabalhadores com habilidades cognitivas não rotineiras em 86,3% e dos trabalhadores com habilidades não cognitivas em 57,5%. Todavia, o estudo indicou que estas estimativas são tendenciosas na medida em que os trabalhadores decidem onde morar. Assim, os trabalhadores com mais habilidades cognitivas não rotineiras podem se mover para as cidades nas quais acreditam que seriam mais produtivos.

Assim, os autores decidem controlar as regressões salariais por características locais geográficas e históricas que determinam as amenidades locais, essas variáveis explicam porque os trabalhadores, de modo geral, podem se direcionar para determinadas cidades. Adicionalmente, buscaram contornar a endogeneidade do modelo, isto é, trabalhadores mais habilidosos tendem a se direcionar para as cidades maiores que possam aumentar a sua produtividade dada a concentração ocupacional favorável. Para tanto, uma estratégia de variável instrumental foi utilizada, de modo a isolar a simultaneidade residual entre a variação na produtividade exógena e a alocação do trabalho.

Os autores, seguindo a literatura padrão, usaram as seguintes variáveis instrumentais: população do século anterior, de acordo com Ciccone e Hall (1996), a migração do século anterior e a

presença de áreas destinadas a faculdades, de acordo com Card (2001) e Moretti (2004). Os instrumentos utilizados se mostraram fortes. Um estimador GMM continuamente atualizado (GMM-CUE) também foi adotado, considerando os mesmos instrumentos, e as estimativas indicaram que o diferencial salarial calculado, sem considerar os instrumentos, foram subestimados, já que uma variação de uma unidade no log da concentração ocupacional aumenta os salários dos trabalhadores com habilidades cognitivas não rotineiras em 126,5% e dos trabalhadores com habilidades não cognitivas em 93,8%. Dado este cenário, os autores têm como conclusão principal que a segregação ocupacional é benéfica para a economia, indicando assim que os esforços para que essa segregação não aconteça, em termos de política espacial, se configuram em um erro.

Wheaton e Lewis (2002) também encontraram evidências de que a concentração e especialização ocupacional e industrial tem efeitos benéficos sobre os salários dos trabalhadores. Os autores indicam que, quando há mais trabalhadores em uma indústria ou ocupação, dentro de uma região específica, a correspondência e o aprendizado entre os trabalhadores aumentam, gerando produtividade e inovação. Os autores indicam que em situações de concentração/especialização os trabalhadores “mergulham” em uma própria piscina, onde por exemplo, programadores de computador têm maiores ganhos em capital humano e produtividade quando estão mergulhados em âmbitos que concentrem suas próprias tarefas.

Os autores propõem um novo recurso de pesquisa com o uso de variáveis de concentração e especialização: concentração de ocupação/indústria é o número de trabalhadores de uma área metropolitana e ocupação/indústria dividido pelo total de trabalhadores nacionais. A especialização de ocupação/indústria é o número de trabalhadores de uma área metropolitana e ocupação dividido pelo total de trabalhadores da região metropolitana. Usando uma amostra de uso público de 5% do Censo dos Estados Unidos de 1990, os autores consideram 220 áreas metropolitanas, 424 categorias ocupacionais e 77 categorias de indústria. O modelo estimado em mínimos quadrados ordinários inclui, além das variáveis principais de análise: especialização e concentração de ocupação e indústria (emprego e estabelecimento), variáveis de controle relacionadas às características específicas dos indivíduos.

Wheaton e Lewis (2002) ressaltam, ainda, o papel dos custos de vida nas diferentes regiões metropolitanas consideradas. Para tanto, incluem efeitos fixos estruturais ao nível de regiões

metropolitanas, indústria e ocupação a fim de tentar minimizar os efeitos de custos de vida sobre os salários. Verificou-se que a especialização ocupacional apresentou variação de 0 a 7%, para as 220 regiões metropolitanas e 424 ocupações, e este intervalo provoca um aumento de 23% nos salários. Quanto a especialização industrial, houve um intervalo de variação de 0 a 68%, entre as 220 regiões metropolitanas e 77 indústrias, o que provoca um aumento de 30% nos salários dos trabalhadores. No que se refere a concentração, os valores foram menores, todavia significativos: a concentração ocupacional apresentou uma variação entre 0 e 46%, provocando um aumento de 12% nos salários e a concentração industrial apresentou uma variação entre 0 e 41%, afetando positivamente os salários em torno de 16%.

Wheaton e Lewis (2002) também enfatizam que a aglomeração ocupacional é favorável para a economia, tal qual enfatizado por Rossi-Hansberg, Sarte e Schwartzman (2019). Wheaton e Lewis (2002) questionam, diante dos seus resultados, por que não há concentração e especialização completas. Os autores respondem essa questão indicando que existem outras considerações de produção que são importantes tal como acesso aos mercados, matérias-primas ou recursos educacionais.

Mion e Naticchioni (2009) analisam como as economias de aglomeração, medidas por variáveis espaciais como densidade, potencial de mercado e especialização industrial, podem determinar os salários na Itália. Todavia, os autores chamam atenção que a maior parte das pesquisas nesse ramo ignoram a natureza das externalidades oriundas dessa aglomeração, cuja fonte principal é a interação entre trabalhadores e firmas. Nesse sentido, os autores buscam controlar a heterogeneidade não observada dos trabalhadores e firmas, uma vez que, tanto trabalhadores como empresas apresentam localizações desiguais no território nacional. No que se refere aos trabalhadores, os autores indicam três razões para esses estarem distribuídos localmente de forma desigual: as cidades maiores oferecem melhor infraestrutura e bens de consumos de lazer, maiores retornos à educação e permitem um maior acúmulo de capital humano. Já as empresas se localizam desproporcionalmente, pois cidades maiores melhoram a correspondência entre trabalhadores e firmas no mercado, aumentando a produtividade.

A fim de contornar esse problema de endogeneidade local de trabalhadores e firma, os autores utilizam um painel de dados para a Itália para o período de 1991-1998. A fonte de dados principal para a elaboração do painel foram os dados fornecidos pelo INPS (*Italian Social Security Institute*). A equação salarial estimada possui, além das variáveis que captam a

aglomeração (densidade, potencial de mercado e especialização industrial), características observadas de indivíduos e firmas, especialmente o tamanho da firma. Várias especificações empíricas foram estimadas pelos autores, a fim de verificar como as variáveis de aglomeração se comportavam: (1) *pooled* de dados, (2) efeitos fixos dos trabalhadores, (3) com a variável tamanho da firma e efeitos fixos dos trabalhadores, (4) efeitos fixos de trabalhadores e firmas, (5) efeitos fixos de trabalhadores e firmas (com uma amostra diferente e um modelo de dois estágios), (6) com a variável tamanho da firma e efeitos fixos dos trabalhadores (uso de IV para contornar a endogeneidade da escolha local) e (7) com tamanho da firma e efeitos fixos dos trabalhadores e firmas (com uma amostra diferente e maior). As especificações que controlam os efeitos fixos de trabalhadores e firmas, simultaneamente, utilizaram o modelo desenvolvido por Abowd e outros (1999). Os autores verificaram que a magnitude dos coeficientes de aglomeração espacial sofre uma diminuição quando a equação de salários inclui como controle o tamanho da firma, justificando considerar a variável como determinante para a heterogeneidade das empresas no espaço. Todavia, os autores evidenciam que essa variável capta apenas uma parte dessa heterogeneidade justificando o controle da heterogeneidade não observada da firma, a partir da estratégia desenvolvida por Abowd e outros (1999).

Mion e Naticchioni (2009) chamam a atenção para o fato de que os efeitos fixos das firmas podem representar não apenas a produtividade das empresas, mas também o poder sindical das empresas, e nesse último caso não há relação com questões espaciais. Desse modo, consideram que a especificação (3) e, por consequência sua extensão, a especificação (6) apresentam as estimativas mais confiáveis. Considerando a especificação base (o *pooled* de dados), verifica-se que todas as variáveis espaciais foram significantes ao nível de 1%. Na especificação (3), após controlar pela heterogeneidade da empresa (tamanho da firma) e heterogeneidade não observada dos trabalhadores, a especialização passa a ser insignificante (antes com um valor de 0,55%), o coeficiente de densidade diminui de 2,21% para 0,56% e o coeficiente de potencial de mercado diminui de 10,88% para 4,53%. Na especificação (6), após o controle da endogeneidade da escolha local, os autores também observaram uma alteração na magnitude das estimativas que captam a aglomeração espacial. O coeficiente de especialização continua insignificante, a densidade diminui para 0,2% (significante ao nível de 10%) e o potencial de mercado passa para 4,64% (significante ao nível de 1%).

Ao classificar as províncias italianas em províncias de baixa e de alta densidade, os autores puderam calcular as habilidades médias nessas províncias, dado os efeitos fixos individuais

obtidos a partir da especificação (6). Assim, verificaram que a diferença de habilidades médias entre as províncias de alta densidade e baixa densidade corresponde a uma magnitude de 9,43% nos salários médios. Ainda considerando as diferenças nas densidades das províncias, os autores avaliaram como o *matching* varia com relação a densidade espacial. Para tanto, os autores indicam que a correlação entre as habilidades médias e o tamanho da empresa em determinado mercado local capta o grau de *matching*, isto é, a correspondência entre os atributos gerais dos indivíduos e as empresas. O autor define atributos gerais como aqueles relacionados aos efeitos fixos dos indivíduos que, dada a especificação empírica da equação de salários dos autores, reflete a produtividade relacionada ao capital humano geral, como capacidade e educação.

Mion e Naticchioni (2009) observaram uma relação negativa entre o *matching* e a densidade nas províncias da Itália, a correlação entre as duas variáveis foi de -0,39% (significante ao nível de 1%). Os autores atribuem essa relação ao fato de que existem diferenças entre as competências e habilidades gerais e específicas dos trabalhadores: quando se considera os atributos gerais, quanto maior o mercado menor a probabilidade de encontrar correspondência entre trabalhadores e empresas. Essa situação não aconteceria quando se considera atributos específicos que são pouco frequentes entre os trabalhadores e mais valiosos para as empresas. Assim, quando os atributos específicos são levados em consideração, a relação entre *matching* e densidade pode ser positiva. A relação negativa é robusta para o caso em que se considera os atributos/habilidades gerais para a determinação do *matching*, condição testada na pesquisa.

Apesar de não considerarem a especialização ou concentração ocupacional na sua pesquisa Combes, Duranton e Gobillon (2008) indicam a importância de se considerar as habilidades dos trabalhadores na determinação dos salários locais. A composição das habilidades dos trabalhadores, onde determinados trabalhadores se direcionariam e se classificariam em determinadas áreas de emprego, obedeceria à própria dinâmica de distribuição das indústrias no espaço, onde algumas delas seriam mais intensivas em mão de obra com maior habilidade. Adicionalmente, os autores indicam que é fundamental considerar, ainda, na determinação dos salários locais, as dotações locais ou naturais e as interações entre trabalhadores e empresas. No que se refere as dotações locais ou naturais, os autores indicam que trabalhadores em algumas áreas que apresentam características geográficas e localização favorável poderiam apresentar um maior produto marginal. Quanto às interações entre trabalhadores e empresas, os autores utilizam as argumentações teóricas de Marshall, onde os efeitos de *matching*, *sharing* e *learnig* provocariam um aumento da produtividade dos trabalhadores como resultado do

tamanho do mercado (economias de urbanização) ou concentração geográfica industrial (economias de localização).

Dessa forma, utilizando um painel de dados para os trabalhadores franceses, os autores indicam que a principal contribuição da pesquisa é utilizar essas três vertentes (habilidades, dotações e interações) em uma estrutura unificada a fim de explicar os salários locais. Ao trabalhar com um painel de dados os autores controlam a heterogeneidade dos trabalhadores, dos locais e das indústrias, sugerindo que as abordagens anteriormente adotadas, não dão a ênfase necessária à classificação dos trabalhadores por área, sendo este um problema grave de variável omitida que deve ser corrigido, objetivo da pesquisa dos mesmos.

A fim de obter um painel de dados Combes, Duranton e Gobillon (2008) utilizam os registros da *Déclarations Annuelles des Données Sociales* (DADS), considerando os funcionários franceses que trabalham nos setores da indústria e serviços, para os anos de 1976, 1980, 1984, 1988, 1992 e 1996. Para definir o local de trabalho, considerou-se as áreas de emprego denominadas de "zonas de emprego", onde na França apresentam-se em um total de 341. Os autores realizaram o estudo em duas etapas. Inicialmente, estimaram um painel de dados a fim de determinar os salários dos trabalhadores considerando as características observadas dos mesmos, as características locais das indústrias e suas variáveis de interação, além da inclusão do efeito fixo dos trabalhadores, indústrias e locais. Assim, os efeitos fixos por área/ano são interpretados como índices salariais locais após o controle das características observadas e não observadas dos trabalhadores e dos efeitos da indústria. No segundo estágio da pesquisa busca-se explicar os efeitos fixos locais, isto é, detectar a importância das dotações locais e interações entre indústrias, ou seja, dos índices de concentração industrial, na determinação dos salários locais (efeitos fixos de uma área em determinado ano).

No primeiro estágio, Combes, Duranton e Gobillon (2008) corroboram sua hipótese principal. Os resultados estimados indicam que as habilidades dos trabalhadores são de suma importância para determinação dos salários e apresentam um papel superior ao ambiente local e a indústria na determinação salarial. Os efeitos fixos dos trabalhadores apresentaram uma correlação com os salários de 0,78, superior a qualquer outra variável incluída no modelo, observou-se ainda que as variáveis que captam as interações entre as indústrias têm pouco poder explicativo. No segundo estágio, incluiu-se como variáveis independentes os efeitos fixos de tempo, um vetor

de dotações locais, um vetor de interações locais entre as indústrias (logaritmo da densidade do emprego local, logaritmo da área e o inverso do índice de Herfindahl).

Os autores ainda indicam que o uso de dotações locais pode provocar endogeneidade no modelo, uma vez que algumas características locais provavelmente são endógenas aos salários locais, assim consideram apenas quatro variáveis de dotações exógenas (costa marítima, montanhas, lagos e "patrimônio cultural ou arquitetônico extraordinário") e as seguintes variáveis instrumentais: densidade logarítmica da população urbana em 1831, 1861, 1891 e 1921, o potencial do mercado com dados da população em 1831 e um índice de periferia (a distância média logarítmica de todas as outras áreas de emprego). Os autores justificam o uso das duas primeiras variáveis instrumentais indicando que aglomeração do passado não está relacionada com os choques de produtividade atuais.

Verificou-se, pelos resultados econométricos, que a densidade do emprego tem maior destaque na explicação dos efeitos fixos locais (apresentando uma correlação de 0,84), sendo seguida pelo potencial de mercado e a área terrestre. Ao considerar o modelo instrumentalizado verificou-se que o uso das variáveis instrumentais se justifica, a principal alteração foi uma mudança notável no coeficiente de densidade que ficou em torno de 3%, abaixo da literatura padrão (no intervalo de 4%-8%). Além disso, também houve uma queda significativa no coeficiente de potencial de mercado e o coeficiente de área terrestre se tornou insignificante. Os autores justificam esse resultado dado o fracasso das pesquisas anteriores em controlar a heterogeneidade individual não observada.

Florida e outros (2011) também avaliam a relação entre habilidades, cidades e salários para a economia dos EUA, todavia, considerando uma abordagem mais direta para habilidades dos trabalhadores. Categorizando as habilidades em analítica, inteligência social e habilidades físicas, os autores buscam verificar como essas habilidades determinam os salários regionais. Florida e outros (2011) evidenciam a importância das habilidades para os indivíduos e as cidades, pois essas habilidades contribuem para aumentar os salários individuais e, por sua vez, os salários regionais, o que pode determinar o crescimento econômico de uma região. Consideram ainda que a aglomeração de pessoas qualificadas nas grandes cidades se intensifica cada vez mais e as habilidades tendem a aumentar cada vez mais o tamanho das cidades. Nesse sentido, os autores buscam verificar se diferentes tipos de habilidade impactam de forma distinta na economia, tendo como hipótese principal que a distribuição de habilidades é

diferente entres as cidades e que maiores cidades têm maior número de trabalhadores com habilidades analíticas e de inteligência social, e essas habilidades impactam positivamente nos salários regionais quando comparadas às habilidades físicas.

Desse modo, Florida e outros (2011) utilizam uma modelagem econométrica que tem como variável principal as habilidades, em três dimensões: habilidades analíticas, inteligência social e habilidades físicas. As variáveis foram incluídas em modelos separados, usando as seguintes variáveis de controle: densidade, população e porcentagem de trabalhadores com ensino superior. Essas variáveis são obtidas para 361 regiões metropolitanas dos EUA para os anos de 1999 e 2008, a fim de comparar e verificar a tendência comportamental ao longo desse período. Os dados foram obtidos a partir das *Occupational Employment Statistics* (OES), provenientes da *Bureau of Labor Statistics*. Para obtenção do percentual das habilidades ocupacionais por região, os autores utilizam as informações disponíveis na O*NET. Como utilizaram os dados da O*NET de 2007 e tendo em vista as atualizações da base e sua não adequação à series temporais, pois os valores da escala de classificação obtidos se mantêm constantes ao longo do tempo, os autores criaram uma variável definida por score de habilidades regionais. Essa variável é o percentual das habilidades ocupacionais ponderada pelo emprego ocupacional em cada região.

Florida e outros (2011) observaram que há uma correlação positiva entre salários regionais e habilidades analíticas, o mesmo ocorre com as habilidades de inteligência social. Todavia, as habilidades físicas apresentam uma correlação negativa com os salários regionais, que se intensifica entre os períodos considerados, mais do que quando se considera as habilidades analíticas e de inteligência social. No que se refere ao tamanho das regiões, verificou-se uma correlação negativa com habilidades físicas, que também se intensifica ao longo do período considerado, já as habilidades analíticas e de inteligência social apresentaram uma correlação positiva com o tamanho das regiões, isto é, cidades maiores tem maior número de trabalhadores com habilidades analíticas e de inteligência social quando comparadas às habilidades físicas.

Os resultados das estimações corroboram as análises de correlações. Verificou-se, para o ano de 2008, que a elasticidade das habilidades analíticas sobre os salários regionais e a elasticidade das habilidades de inteligência social sobre os salários regionais são positivas, 0,582 e 0,5, respectivamente. Por sua vez, a elasticidade das habilidades físicas sobre os salários regionais é negativa (-0,314). Percebeu-se um aumento considerável para todas as elasticidades de

habilidades quando comparadas ao período de 1999. Desse modo, os autores enfatizam a importância das habilidades para a definição dos salários regionais, principalmente quando comparadas com a escolaridade dos trabalhadores, pois a elasticidade da parcela da população com ensino superior sobre os salários regionais, apesar de positiva, apresentou menor magnitude (em torno de 0,1 para todas as especificações) e se manteve estável ao longo do período considerado.

Considerando os microfundamentos da aglomeração adotados por Duranton e Puga (2004), *sharing*, *matching* e *learning*, Bacolod, Blum e Strange (2009) analisam como esses podem vir a determinar os salários correspondentes às habilidades cognitivas, interpessoais e físicas. Os autores acreditam que tanto para as habilidades cognitivas quanto para as interpessoais é mais fácil ver os canais pelos quais são afetadas, isto é, trabalhadores intensivos em habilidades cognitivas e interpessoais aprendem facilmente com os outros, além disso, os mesmo podem ser beneficiados pelas interações no mercado de trabalho aglomerado, tanto no que se refere ao *matching* quanto ao *sharing*, uma vez que esses são especializados. Esse contexto tende a aumentar o preço dessas habilidades nas grandes cidades.

Os autores chamam atenção para as habilidades interpessoais, pois indicam que o *learning*, *sharing* e *matching* são inerentes a interação, logo as habilidades interpessoais refletem perfeitamente o fenômeno de interações nas cidades. No que se refere às habilidades físicas, os autores têm uma visão diferente. Apesar de ser possível verificar como os trabalhadores intensivos em habilidades físicas podem ser beneficiados pelo *sharing* e *matching*, as habilidades físicas podem não ser facilmente beneficiadas pelo *learning*, isto é, o processo de aprendizagem não ocorre de forma natural como nas habilidades cognitivas e interpessoais. Dessa forma, os autores esperam que a aglomeração não intensifique o aumento no preço das habilidades físicas.

Bacolod, Blum e Strange (2009) também buscaram avaliar o prêmio salarial para as habilidades cognitivas, interpessoais e físicas para a economia dos EUA. Além disso, os autores avaliam a distribuição espacial das habilidades. Para a categorização das habilidades, os autores, tal como Florida e outros (2010), usaram os dados da O*NET e a definição de ocupações do DOT, enquanto para a obtenção das outras variáveis usaram dados do Censo dos EUA e também da NLSY. No que se refere à alocação de habilidades entre as cidades, considerando quatro classes de cidades (pequenas, médias, grandes e muito grandes), os autores observaram que existe uma

relação positiva entre tamanho das cidades e as habilidades, porém essa relação é mais fraca quando comparada à distribuição espacial pela ótica da educação, indústria ou ocupação. Assim, verifica-se que nas grandes cidades há maior habilidade cognitiva, porém, a diferença não pode ser considerada suficiente para qualifica-las como intensivas em habilidades cognitivas.

A especificação econométrica base do trabalho de Bacolod, Blum e Strange (2009) considera os dados do Censo dos EUA, a partir de um modelo hedônico do impacto da urbanização nos preços das habilidades dos trabalhadores. Nessa especificação tem-se um *pooled* de dados cuja variável dependente é o logaritmo do salário do trabalhador *i* na ocupação *j* e na área metropolitana *s*. Variáveis de controle são introduzidas na análise tal como: escolaridade, raça, sexo, estado civil, idade e idade ao quadrado. A variável independente principal é o vetor que identifica a habilidade (cognitiva, interpessoal e física) do trabalhador *i* na ocupação *j* que podem variar por *s* áreas metropolitanas, incluindo uma variável de interação entre população das áreas metropolitanas e habilidades. Os autores observaram que o efeito das habilidades cognitivas, interpessoais e físicas isoladamente sobre os salários foram positivas, todavia, a variável de interação entre habilidade física e logaritmo da população foi negativa (-0,1036), o que não ocorre com as variáveis de interação para as habilidades cognitivas (0,1224) e interpessoais (0,0294). Assim, verifica-se que o aumento dos salários associado à aglomeração é maior para trabalhadores com habilidades cognitivas e interpessoais. Na especificação que inclui conjuntamente todas as habilidades os principais resultados permanecem.

Bacolod, Blum e Strange (2009) ainda consideram possíveis vieses de estimação nos seus resultados, pois acreditam que, além de erros de medição nas habilidades estarem relacionados com o tamanho da cidade, a heterogeneidade não observada dos trabalhadores pode impactar nos salários regionais. Para contornar esses vieses, os autores utilizam a base NLSY79 que permitem capturar a heterogeneidade não observada dos trabalhadores e áreas metropolitanas, capturando o efeito fixo de indivíduos interagido com as áreas metropolitanas. Adicionalmente, os autores inserem no modelo as variáveis AFQT, uma medida de capacidade cognitiva, e o Índice Rotter, uma medida de capacidades pessoais, ambas presentes na NLSY79.

Verificou-se que as variáveis AFQT e Índice Rotter apresentam um efeito positivo sobre os salários, todavia, o efeito das capacidades cognitivas é maior. Quanto aos resultados que captam os efeitos fixos individuais interagidos com as áreas metropolitanas, evidenciou-se que o

prêmio salarial para as habilidades cognitivas e interpessoais é maior nas grandes cidades. Já as habilidades físicas apresentam um prêmio salarial negativo nas grandes cidades. Esses resultados corroboram os resultados obtidos a partir do Censo dos EUA, logo todas as especificações testadas pelos autores indicam que as habilidades cognitivas e interpessoais têm preços maiores nas grandes cidades, o inverso ocorre com as habilidades físicas. Assim, os autores comprovam que a heterogeneidade não observada não modifica as suas principais evidências. Desse modo, percebe-se que os resultados de Bacolod, Blum e Strange (2009) se aproximam daqueles obtidos por Florida e outros (2011).

Na pesquisa de Guvenen e outros (2020), tal como nas pesquisas de Bacolod, Blum e Strange (2009) e Florida e outros (2011), utilizou-se uma categorização dos trabalhadores por habilidade, a partir dos dados da O*NET e NLSY79, para avaliar a determinação dos salários, todavia, a categorização feita pelos autores consiste em: habilidades matemáticas, habilidades verbais (ambas dentro do que se considera habilidades cognitivas) e habilidades sociais. Além disso, os autores analisam o papel das habilidades considerando uma problemática diferente da abordada por Bacolod, Blum e Strange (2009) e Florida e outros (2011), pois não consideram variáveis espaciais, mas a incompatibilidade de habilidades multidimensionais.

Guvenen e outros (2020) elaboram uma medida empírica que estabelece a qualidade da correspondência entre o portfólio de habilidades dos trabalhadores e o portfólio de habilidades exigida para a realização das ocupações, podendo assim mensurar a incompatibilidade de habilidades. Essa incompatibilidade pode ser positiva, onde os trabalhadores são superqualificados, e negativa, onde os trabalhadores são subqualificados. A métrica de incompatibilidade consiste em identificar o portfólio de habilidades dos trabalhadores a partir das variáveis ASVAB (que apresenta scores de habilidades cognitivas), Escala do Lócus de Controle de Rotter e a Escala de Auto-Estima de Rosenberg (que apresentam scores de habilidades não cognitivas, que os autores classificam como habilidades sociais) presentes na NLSY79. Já o portfólio de habilidades requeridas em cada ocupação é identificado a partir das informações disponíveis na O * NET, onde para cada ocupação há 277 descrições ou características que são utilizadas para execução das mesmas. Dada a classificação de habilidades matemáticas, verbais ou sociais, os autores utilizaram 26 descrições da O*NET.

Os resultados obtidos pelos autores apontam que a incompatibilidade em qualquer uma das formas inseridas na equação é significativa e impacta negativamente sobre os salários.

Verificou-se que a interação entre incompatibilidade e posse ocupacional é significativa e que no decil superior (trabalhador com maior incompatibilidade) o salário é 6,8% menor, quando comparado com o decil inferior. Considerando a incompatibilidade cumulativa esse valor passa a ser 9,4%. Além disso, realizando uma análise contrafactual, os autores evidenciam que a eliminação de toda incompatibilidade aumentaria os salários médios em 11%. Ao analisar as incompatibilidades pela ótica das dimensões de habilidades desagregadas, verificou-se que o efeito negativo sobre os salários permanece, todavia, em qualquer variável de incompatibilidade (atual, interação e cumulativa) e em qualquer especificação, nas habilidades sociais os efeitos sobre os salários são menores quando se considera as habilidades cognitivas (matemática ou verbal).

3.2.3.2 Evidências empíricas nacionais

Na literatura empírica nacional, as pesquisas direcionadas para avaliar a especialização ocupacional e o papel das habilidades na determinação dos salários são ainda mais escassas. Pode-se citar as pesquisas de Ehrl e Monasterio (2020), na análise da especialização ocupacional por habilidades, Barufi, Haddad e Nijkamp (2016) e Neves Júnior (2018), na análise da determinação das habilidades sobre os salários locais.

Ehrl e Monasterio (2020) enfatizam a importância em diferenciar as habilidades de acordo com suas repercussões no mercado de trabalho. Os mesmos indicam que as habilidades rotineiras estão associadas a pouca interação entre os trabalhadores, sendo obtidas a partir de treinamento, as habilidades analíticas estão associadas a resolução de problemas, necessitando de aprendizado constate e as habilidades interpessoais estão associadas a maiores desempenhos de interação, especialização e troca de tarefas. Desse modo, os autores têm como hipótese principal que trabalhadores intensivos em habilidades interpessoais e analíticas produzem maiores externalidades de produtividade, independentemente do tamanho das cidades em que estão inseridos.

A fim de avaliar como a concentração espacial de habilidades determina o nível de salários urbanos no Brasil, Ehrl e Monasterio (2020) classificam cinco dimensões de habilidades: analíticas, interpessoais, cognitivas de rotina, manuais de rotina e habilidades manuais não rotineiras. Os autores aprofundam a pesquisa indicando os vieses e buscando corrigi-los. Citam a casualidade reversa, onde regiões com maiores salários atraem trabalhadores mais

qualificados, além da heterogeneidade não observada dos trabalhadores na determinação dos salários e erros de medição na medida de concentração de habilidades. Esse erro resulta da definição da variável que é obtida a partir da média municipal das pontuações de habilidades dos trabalhadores, essa pontuação é alcançada a partir da soma dos elementos selecionados da O*NET de acordo com cada dimensão de habilidade. Nesse sentido, como a base da O*NET é criada para as ocupações nos EUA, erros de medição são possíveis de acontecer. Embora, os autores indiquem um consenso na literatura (HANDEL, 2012; LEWANDOWSKI, 2017) de que há consistência substancial nas pontuações das habilidades da O*NET podendo ser empregada em outros países, sem mudanças significativas na classificação e identificação das habilidades nas ocupações.

Ehrl e Monasterio (2020) utilizaram um modelo de dois estágios cuja variável instrumental principal é um *shift-share* que permite corrigir o problema de causalidade reversa. Os vetores de concentração de habilidades são ponderados pela variável *shift-share,* que é definida pela multiplicação da participação regional das ocupações em 1994 e da taxa de crescimento nacional das ocupações entre 1994 e 2010 (ano de referência). Essa variável reflete o impacto de choques de demanda nos municípios. Utilizando um painel obtido a partir dos dados da RAIS, para o período entre 2007 e 2013, a modelagem empírica dos autores busca explicar os salários, em nível regional, capturados pelo efeito fixo da região e obtidos no primeiro estágio do modelo. O painel busca explicar os salários, considerando as características observadas dos indivíduos e os efeitos fixos dos indivíduos, setores, anos e região.

Os autores ainda aplicaram o mesmo modelo base para o Censo brasileiro de 2010, a fim de considerar o mercado de trabalho informal. Comparando o modelo de OLS com o IV, os autores verificaram que, nas estimativas com o uso de variáveis instrumentais, a magnitude dos efeitos das habilidades sobre os salários aumenta. A concentração de habilidades que apresenta coeficientes positivos foi: analítica, interpessoal e cognitiva de rotina; já a concentração de habilidade manual de rotina e de habilidades manual não rotineiras apresenta efeito negativo. Em todos os modelos estimados, a concentração de habilidades analíticas ganha destaque. Ehrl e Monasterio (2020) estabelecem, ainda, que o aumento na magnitude dos coeficientes indica que o viés devido ao erro de medição domina no problema de pesquisa apresentado. No que se refere aos resultados com os dados do Censo, todas as concentrações de habilidade se tornaram significativas e os sinais permanecem os mesmos.

Utilizando também uma abordagem econométrica de dois estágios, Barufi, Haddad e Nijkamp (2016) buscam estudar os diferenciais salariais industriais locais para o Brasil. Os autores também dão ênfase fundamental às habilidades dos trabalhadores na determinação salarial e utilizam como base de dados as categorias de habilidades de Maciente (2013) e os dados da RAIS para o período de 2004, 2008 e 2012. Os setores industriais são categorizados em: S1 - Manufatura de baixa tecnologia; S2 - Manufatura de tecnologia de ponta; S3 - Manufatura de alta tecnologia; S4 - Serviços com menos conhecimento e S5 - Serviços intensivos em conhecimento. Todas as análises econométricas são realizadas para as cinco classificações separadamente, a fim de comparar o efeito das externalidades de aglomeração e do *mix* industrial nos diferentes setores dada a sua intensidade tecnológica.

Na estratégia de primeiro estágio, Barufi, Haddad e Nijkamp (2016) consideraram como variáveis para explicar os salários: idade, educação, fatores de habilidades, tamanho da firma e uma *dummy* para cada combinação de região, setor e ano. Essa *dummy* será considerada o resíduo espacial ou salário espacial que será utilizada como variável dependente no segundo estágio. Como variáveis explicativas no segundo estágio tem-se medidas associadas com externalidades de aglomeração, *mix* industrial e características geográficas: logaritmo da densidade, área, medidas de especialização, diversidade e competição, *dummies* de setor, anuais e macrorregionais e características geográficas (altitude e distância da linha do Equador).

Barufi, Haddad e Nijkamp (2016) justificam o uso de dois estágios, pois o termo de erro do primeiro estágio pode estar correlacionado com características individuais de determinados setores e regiões, uma vez que trabalhadores próximos e que trabalham no mesmo setor acabam compartilhando de características que não são determinadas pelos controles individuais. Além disso, os autores ainda utilizam variáveis instrumentais para captar melhor o efeito das variáveis do segundo estágio na determinação do salário espacial. Os instrumentos utilizados são: densidade populacional em 1940; distância para a costa; e *dummies* para a produção de cana e café e exploração de ouro durante o período colonial.

Os resultados do primeiro estágio, após o controle da heterogeneidade não observada dos indivíduos, indicam um papel importante das habilidades. Apesar de algumas categorias perderem a significância, os coeficientes permanecem positivos e significativos para habilidades cognitivas. Além disso, verifica-se valores crescentes para os coeficientes que captam tamanho das empresas. No segundo estágio, após o controle da heterogeneidade não

observada dos indivíduos, verificou-se os coeficientes de densidade permaneceram positivos e significativos para todos os setores, mas de acordo com a literatura os autores observaram que esse coeficiente de urbanização diminui significativamente após o controle dos efeitos fixos individuais. Quanto as medidas de *mix* industrial, existiram mudanças significativas. No que se refere às medidas de diversificação, estas apresentaram um sinal positivo, mas não significativo para o setor S3. O grau de especialização tornou-se não significativo para os setores S2, S3 e S5. Assim, os autores mostram a importância do controle da heterogeneidade não observada na determinação do salário espacial, já que a não inclusão desse controle pode conduzir a estimativas inconsistentes. Por fim, no que se refere ao modelo de variáveis instrumentais, as principais conclusões parecem se confirmar, dado que os coeficientes de densidade se mostram positivos e significativos.

No Brasil, o trabalho de Neves Júnior (2018) busca avaliar a determinação dos salários a partir da categorização dos trabalhadores em habilidades cognitivas, sociais e motoras, avaliando ainda como o tamanho das cidades pode impactar nesses salários. Para a classificação de cada dimensão de habilidades o autor utilizou variáveis da O*NET, aplicando a técnica de fator comum, tal como elaborado por Maciente (2013). O autor utiliza os microdados da RAIS para o período compreendido entre 2003 e 2013 e inclui os municípios brasileiros que pertençam a definição do IBGE (2015) de Arranjos Populacionais e municípios isolados, compondo um total de 374 áreas de mercado de trabalho local, formadas por 1976 municípios. Os Arranjos Populacionais são classificados para capturar a ideia de escala urbana: as aglomerações pequenas têm população menor que 100.000 habitantes, as médias têm população entre 100.000 e 750.00 habitantes, as grandes têm população entre 750.000 e 2.500.000 habitantes e as muito grandes têm população superior a 2.500.000 habitantes.

A fim de captar o *sorting* espacial dos trabalhadores, o autor elabora uma estratégia empírica que capte a heterogeneidade não observada dos trabalhadores a partir de um painel de dados. As principais variáveis para explicar os salários são um vetor de habilidades, um vetor de tamanho do urbano e um vetor de interação entre habilidades e tamanho urbano. A regressão capta ainda os efeitos fixos da área de mercado local e das firmas, uma vez o as características inerentes das áreas de mercado podem estar relacionadas com o tamanho urbano e a maior produtividade nos grandes centros pode ser determinada pela presença das melhores firmas; além disso, inclui-se na regressão efeitos combinados (indivíduo-área, indivíduo-firma, área-firma e indivíduo-firma-área).

Após testar diversas especificações para a estratégia empírica da pesquisa, Neves Júnior (2018) observou que o resultado principal da sua pesquisa é aquele que capta o efeito fixo individual, indicando que a heterogeneidade individual é mais importante do que a heterogeneidade da firma ou do mercado local para analisar as diferenças salarias das habilidades frente ao tamanho urbano. Nesse sentido, avaliando o modelo principal, o autor demostra que existe uma relação positiva entre salários e tamanho dos centros urbanos para as habilidades cognitivas e sociais, todavia, quando se considera as habilidades motoras nas aglomerações muito grandes há perdas salariais. A relação positiva e crescente com relação ao tamanho dos centros é mais clara para as habilidades cognitivas, onde nos centros médios o diferencial salarial é 0,001, nos centros grandes é 0,035 e nos centros muito grandes chega a 0,066, no que se refere as habilidades sociais, esse padrão crescente não é verificado para os centros muito grandes, cujo coeficiente foi de 0,042, enquanto nos centro grandes foi de 0,117, isto é, o desempenho das habilidades sociais é melhor em centros grandes e o desempenho das habilidades cognitivas se destaca nos centros muito grandes.

Quanto às habilidades motoras, o diferencial salarial permanece praticamente o mesmo em médios e grandes centros (0,042 e 0,045, respectivamente), todavia, conforme já destacado, em centros muito grandes esse diferencial se torna negativo, -0,028. Quando o autor, considera todos os efeitos fixos e efeitos fixos combinados, a magnitude de todos os diferenciais diminuem e alguns perdem significância, no que se refere às habilidades cognitivas a significância permanece somente para os centros muito grandes (0,025), para as habilidades sociais tem-se diferenciais significantes para centros grandes (0,051) e muito grandes (-0,039), todavia, para centros muitos grandes têm-se perdas salariais. Já para as habilidades motoras, as perdas salariais para centro muito grandes tornam-se insignificantes. Entretanto, o Neves Júnior (2018) indica que a interpretação desses resultados merece cautela, uma vez que a inclusão de tantos controles de heterogeneidade não observada combinados vai além das características de *sorting* de trabalhadores e firmas.

Considerando as evidências empíricas abordadas nessa seção, percebe-se que são poucos os estudos que tem como foco a análise da relação entre salários, habilidades e espaço. Na literatura brasileira, a maioria dos trabalhos analisa o prêmio salarial das habilidades no espaço, considerando o tamanho das cidades ou centros urbanos (NEVES JÚNIOR, 2018, BARUFI; HADDAD; NIJKAMP, 2016, EHRL; MONASTERIO, 2020). Apenas a pesquisa de Ehrl e Monasterio (2020) considera uma medida de aglomeração de habilidades. Conforme verificado,

os autores avaliam a determinação dos salários urbanos a partir da concentração espacial de habilidades (média de trabalhadores com determinadas habilidades no município) e corrigem o problema da causalidade reversa a partir do uso do instrumento *shift-share* de Bartik. Nesse sentido, é fundamental indicar que, a principal contribuição dessa pesquisa para a literatura brasileira é obter o prêmio salarial da especialização ocupacional categorizada por habilidades, ou seja, verifica-se o quão especializado é o Arranjo Populacional em determinada habilidade, permitindo apreender melhor os efeitos de externalidades.

Além disso, essa pesquisa traz como contribuição empírica o controle da endogeneidade na determinação do prêmio de habilidades, considerando não só o instrumento *Bartik shift* ocupacional (como feito por Ehrl e Monasterio (2020)), mas também variáveis instrumentais históricas (metodologia aplicada por Rossi-Hansberg, Sarte e Schwartzman (2019) para os EUA) e a heterogeneidade não observada de trabalhadores e firmas. Como contribuição à literatura internacional, essa pesquisa realiza uma decomposição dos salários locais devido aos efeitos brutos e líquidos da especialização de habilidades, onde é possível avaliar a relação entre densidade populacional e especialização de habilidades no mercado de trabalho brasileiro.

3.3 ESTRATÉGIA EMPÍRICA

Esta seção busca indicar e explicar a abordagem empírica utilizada a fim de medir como a especialização ocupacional, categorizando os trabalhadores a partir das suas habilidades (cognitivas, social e manual), pode determinar os salários no Brasil, tendo como foco o controle da endogeneidade da escolha locacional dos trabalhadores a partir de instrumentos adequados. Além disso, estabelece-se uma estratégia empírica para verificar como o prêmio salarial da especialização ocupacional pode ser modificado quando se controla a endogeneidade causada pela heterogeneidade não observada dos trabalhadores e firmas. Para a decomposição salarial, utiliza-se uma metodologia baseada na proposta de Abowd, Kramarz e Margolis (1999) e Combes, Duranton e Gobillon (2008), a fim de determinar como a densidade populacional afeta os diferenciais de salários em cada tipo de especialização ocupacional categorizada por habilidades.

3.3.1 Especialização ocupacional

Ao determinar o prêmio salarial da especialização ocupacional categorizada por tipos de habilidades mais intensamente desenvolvidas pelos trabalhadores, habilidades essas captadas pela metodologia de Maciente (2013), tem-se uma avaliação mais precisa de como o *learning*, o *sharing* (*labor pooling)* e o *matching* podem determinar os salários no mercado de trabalho. Conforme já mencionado, os trabalhadores tendem a manter uma proximidade com trabalhadores que ocupam atividades similares a fim de aumentarem a sua produtividade. Uma medida de especialização ocupacional categorizada por habilidades cognitivas, sociais e manuais permite compreender melhor o padrão de aglomeração dos trabalhadores e as externalidades geradas, pois espera-se que os efeitos de *learning, sharing* (*labor pooling)* e *matching* dependam da composição ocupacional das cidades.

A categorização em habilidades cognitivas, sociais e manuais nessa pesquisa segue o desenvolvimento das pesquisas em torno da temática de aglomeração, salários e habilidades. Como visto na seção de literatura é comum a dicotomia em habilidades cognitivas e não cognitivas. A pesquisa de Roy (1950) já indicava a importância em se considerar habilidades multidimensionais, no sentido de que os trabalhadores se direcionariam para aquelas ocupações que se adequassem mais as suas habilidades. Heckman, Stixrud e Urzua (2006) rompem com a ideia padrão de que apenas as habilidades cognitivas, muitas vezes associadas em nível superior de escolaridade, são importantes para explicar o prêmio salarial dos trabalhadores. Os autores enfatizam a importância das habilidades cognitivas e não cognitivas na determinação da escolaridade, salários, emprego, experiência e escolha ocupacional, mostrando que as habilidades não cognitivas são tão importantes quanto as cognitivas em análises de mercado de trabalho. Acemoglu e Autor (2011) apresentam uma modelagem teórica e empírica que permite analisar como as habilidades, em uma vertente multidimensional, podem impactar na ocupação e salários dos trabalhadores. Com o avanço da literatura tornou-se comum a análise das habilidades categorizadas em cognitivas e não cognitivas, entre as habilidades não cognitivas ainda existe uma subdivisão importante: habilidades sociais e habilidades manuais (FLORIDA *et al.*, 2011; BACOLOD; BLUM; STRANGE, 2009; GUVENEN *et al.*, 2020; ROSSI-HANSBERG; SARTE; SCHWARTZMAN, 2019; NEVES JÚNIOR, 2018).

As habilidades sociais ganham cada vez mais destaque dada a evolução e importância do setor de serviços e comércio, pois mesmo com a evolução tecnológica, as ocupações que são

intensivas em habilidades sociais tenderão a se manter, pois a interação humana não é facilmente substituída por máquinas ou insumos alternativos (MACIENTE; RAUEN; KUBOTA, 2019; DAVID; DEMING, 2017). Neste sentido, a categorização das habilidades, tal como apresentada nessa pesquisa, está de acordo com a evolução dos estudos na literatura.

Seguindo a metodologia de Wheaton e Lewis (2002) e Rossi-Hansberg, Sarte e Schwartzman (2019), a especialização ocupacional categorizada por habilidades é definida para o Brasil da seguinte forma:

$$EO_h = {}^{L_{hat}}/_{T_{at}} \tag{3.13}$$

Em que,

$T_{a,t}$ é o emprego total no Arranjo Populacional (a) no período (t)

$L_{h,a,t}$ é o emprego ocupacional categorizado pela habilidade (h) no Arranjo Populacional (a) no período (t)

h são as categorias de habilidades ocupacionais (h= 3, habilidades cognitivas, habilidades sociais e habilidades manuais)

a é o número dos Arranjos Populacionais (376 unidades geográficas)

Para a construção da variável "Grau de especialização ocupacional" a pesquisa fez uso da quantidade de trabalhadores em cada ocupação, de acordo com a CBO-2002. Para tanto, utilizou-se os microdados da RAIS-ME para o período de 2009 a 2018. Com isso, a partir da categorização de Neves Júnior (2018), cada ocupação foi classificada em três categorias: intensivas em habilidades cognitivas, intensivas em habilidades sociais e intensivas em habilidades manuais. Após a classificação das 2.632 ocupações disponíveis na base de dados de Maciente (2013) e dada a definição do IBGE (2019) para as 376 unidades geográficas que englobam os Arranjos Populacionais e municípios isolados, que são definidos como médias e grandes concentrações, pode-se então calcular o grau de especialização ocupacional para cada uma das três classificações de habilidades gerando assim um vetor de 3 variáveis, a partir da divisão entre o número de trabalhadores da categoria ocupacional categorizada pela habilidade (h) no Arranjo Populacional (a) no período (t) e o número de trabalhadores totais do Arranjo Populacional (a) no período (t). Convém mencionar que, como se trata de uma medida de participação de emprego essa pode variar entre 0 e 100%, para cada categoria de habilidades.

É necessário, ainda, esclarecer como a classificação entre habilidades cognitivas, sociais e manuais foi estabelecida. A base de dados de Maciente (2013) obteve 21 fatores de habilidades, a saber: Cognitivas; Manutenção e operação; Assistência; Gerenciais; Design e engenharia; Transporte; Artísticas; Precisão e automação; Trabalho supervisionado; Ensino e ciências sociais; Força física; Telecomunicação; Independência; Ciências naturais; Atenção; Experiência no trabalho; Gestão de conflitos; Trabalho em grupo; Vendas; Monitoramento e conformidade e Clericais. A fim de analisar como as habilidades dos trabalhadores determinam o prêmio salarial urbano, utilizando a base de dados de Maciente (2013), Neves Júnior (2018) criou três dimensões de habilidades: cognitivas, sociais e motoras. O autor observou que as habilidades estão relacionadas positivamente com os salários urbanos. Com maior destaque para as habilidades cognitivas e sociais.

Pensando na robustez do seu modelo, Neves Júnior (2018) realizou algumas métricas alternativas para captar as habilidades cognitivas, sociais e motoras. Uma das alternativas foi utilizar diretamente as variáveis criadas por Maciente (2013). Desse modo, dos 21 fatores de habilidades obtidos pelo autor, e já listados acima, foram escolhidos, da ampla variedade dos fatores, aqueles que mais se aproximavam das habilidades cognitivas, sociais e motoras. Para representar as habilidades cognitivas o autor utilizou o fator cognitivo (fator 1 de Maciente (2013)), para as habilidades sociais o autor utilizou o fator trabalho em equipe (fator 18 de Maciente (2013)) e para as habilidades motoras o autor selecionou o fator manutenção e operação (fator 2 de Maciente (2013). De acordo com Neves Júnior (2018), todas as alternativas testadas apresentaram resultados do *ranking* de habilidades dentro das ocupações muito próximos, existindo apenas alterações marginais, não existindo mudanças relevantes.

Desse modo, considerando que o objetivo principal desta pesquisa é determinar o prêmio salarial da especialização ocupacional categorizada pelas habilidades cognitivas, sociais e motoras, deve-se inicialmente categorizar as ocupações de acordo com as habilidades já mencionadas. Essa pesquisa utiliza a alternativa estratégica abordada por Neves Júnior (2018), isto é, se fez uso da base original de Maciente (2013), utilizando os fatores de habilidade cognitiva (fator 1), trabalho em equipe (fator 18) e manutenção e operação (fator 2) para representar habilidades cognitivas, sociais e motoras, respectivamente. Optou-se por utilizar essa estratégia pois é a mais robusta no sentido de que todas as 263 variáveis presentes na ONET foram utilizadas. A base de dados final dos 21 fatores é apresentada da seguinte forma: as linhas

(2.632) representam as ocupações e as colunas (21) representam os fatores de habilidades extraídos da análise fatorial[20].

Assim, para cada ocupação verifica-se a intensidade do fator de habilidade que esta requer, neste caso percebe-se que em cada ocupação pode-se requerer todos os fatores de habilidades, mas com intensidades diferentes. O valor deste fator para cada ocupação pode ser positivo, zero ou negativo, um valor positivo indica que essa habilidade contribui positivamente para o desempenho da ocupação, um valor de zero indica que esta habilidade não interfere no desempenho da ocupação e um valor negativo indica que esta habilidade interfere negativamente para o desempenho da ocupação, isto é, não é necessária para o bom desempenho da ocupação. Dessa forma, quanto maior o valor do fator de habilidade mais importante este é para o desempenho da ocupação.

Assim, nesta pesquisa, a fim de categorizar as ocupações entre intensivas em habilidades, cognitivas, sociais ou manuais, cria-se uma medida de robustez, avaliando, inicialmente, se o fator individual para cada ocupação é maior do que a mediana de uso desse fator entre todas as ocupações. Neste caso se o fator individual para uma ocupação for maior que a sua mediana pode-se defini-la como intensiva naquele fator. Todavia, como essa condição, de fator individual ser maior que a mediana, pode ocorrer para as três dimensões de habilidades numa mesma ocupação, um segundo critério foi adotado, avaliando e comparando os fatores individuais por ocupação. Dessa forma, tem-se que:

1. Se, para uma ocupação, o fator individual for maior do que a mediana de todas as ocupações apenas para uma dimensão de habilidade (entre as três analisadas), essa habilidade define a sua intensidade, ou seja, a ocupação é intensiva nessa habilidade, assumindo que o fator individual das demais dimensões de habilidades é menor do que a mediana;
2. Se, para uma ocupação, o fator individual não for maior do que a mediana em qualquer dimensão de habilidade, foi utilizada a regra de avaliar qual fator de habilidade individual é maior para cada ocupação, o maior fator define se a ocupação é intensiva em habilidades cognitivas, sociais ou manuais;

[20] Essas linhas incluem as ocupações militares que, conforme já mencionado, não foram consideradas nesta pesquisa, isto é, foram desconsideradas da base de dados de Maciente (2013): as primeiras 19 linhas correspondentes ao primeiro grupo ocupacional 0 (Forças Armadas, Policiais e Bombeiros Militares).

3. Se, para uma ocupação, o fator individual for maior que a mediana em mais de uma dimensão de habilidade, foi utilizada a regra de avaliar qual fator de habilidade individual é maior para cada ocupação, o maior fator define se a ocupação é intensiva em habilidades cognitivas, sociais ou manuais.

Isto é, nas condições 2 e 3, verifica-se para cada ocupação, qual dos três vetores de habilidades considerados (fator 1, fator 18 ou fator 2) apresenta maior valor, e assim, pôde-se classificar a ocupação. Por exemplo, um Bioengenheiro apresenta um valor de 1,866663 para o fator 1, um valor de -1,37003 para o fator 18 e um valor de 0,978275 para o fator 2, assim, classifica-se essa ocupação como intensiva em habilidades cognitivas. Percebe-se a adequação da base de dados de Maciente (2013), pois um Bioengenheiro pertence a uma área da engenharia voltada para biologia e saúde, logo, é uma ocupação que exige muito raciocínio e conhecimento teórico onde as habilidades cognitivas são fundamentais. Para essa ocupação, são requisitadas, ainda, as habilidades manuais, pois há a necessidade de realizar experimentos, todavia, é uma ocupação onde não se faz necessário nenhuma interação social. Outro exemplo é um Pintor de letreiro, este apresenta um valor de -0,96167 para o fator 1, um valor de 0,700887 para o fator 2 e um valor de -0,76976 para o fator 18, neste sentido, pode-se classificar essa ocupação como intensiva em habilidades manuais, ou seja, um pintor requer pouca habilidade social e cognitiva (essa ainda em menor grau). Já um Agente de microcrédito apresenta um valor de -0,27787 para o fator 1, um valor de -1,57963 para o fator 2 e um valor de 1,401233 para o fator 18, logo essa ocupação pode ser classificada como intensiva em habilidades sociais, pois um Agente de microcrédito tem como função principal o atendimento ao público a fim de realizar empréstimos. Neste sentido, é de fundamental importância o contanto interpessoal para esta ocupação, sendo de pouca importância as habilidades motoras.

3.3.2 Modelagem empírica e prêmio salarial da especialização ocupacional

Utilizando os microdados da RAIS-ME para o período de 2009 a 2018, estima-se uma equação de salários, a fim de determinar o prêmio salarial da especialização ocupacional categorizada por habilidades. Foram realizados alguns exercícios empíricos, onde no primeiro estima-se um modelo de *pooling* de dados, de modo a contornar o problema de endogeneidade da escolha locacional, por meio do uso de variáveis instrumentais. Posteriormente, um modelo com dados em painel foi estimado, controlando a endogeneidade de variável omitida, dado a heterogeneidade não observada dos trabalhadores e firmas.

A equação salarial do modelo completo, que foi estimada em várias especificações, é dada por:

$$\ln(w_{ijt}) = \beta_0 + \beta_{1h}\ln(EO_{h,a,t}) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{jt}} + \boldsymbol{T_t} + \varepsilon_{ijt} \quad (3.14)$$

Em que, $logw_{ijt}$ é o logaritmo natural do salário-hora real do indivíduo *i* na firma *j* no período *t*, $\boldsymbol{X_{it}}$ é um vetor de características socioeconômicas observáveis, que inclui variáveis de controle individuais, tais como idade, idade ao quadrado, experiência, experiência ao quadrado e variáveis *dummies* de gênero (sem o efeito fixo individual) e o grau de instrução, $\boldsymbol{E_{jt}}$ é um vetor de características do emprego do trabalhador, mais especificamente, *dummies* de setor de atividade econômica de acordo com a CNAE e *dummies* de tamanho do estabelecimento, $\ln(EO_{h,a,t})$ é a variável de interesse que indica o logaritmo natural do grau de especialização ocupacional no Arranjo populacional *a*, e h são as categorias de habilidades ocupacionais (cognitivas, sociais e manuais) de acordo com a classificação de habilidades estabelecida por Maciente (2013) e categorizada por Neves Júnior (2018), $\boldsymbol{T_t}$ é um vetor de *dummies* de ano e β são os vetores de parâmetros a serem estimados.

Convém ressaltar que é estimada uma equação salarial, tal como (3.14), para cada categoria de habilidades h (cognitivas, sociais e manuais). Obtendo-se, assim, três modelagens empíricas diferentes. Alguns exercícios empíricos são feitos de modo a obter maiores informações acerca da aglomeração de trabalhadores categorizados por habilidades. Assim, a equação (3.14) é estimada substituindo o grau de especialização ocupacional categorizada por habilidades pelo indicador de aglomeração categorizado por habilidades ($L_{h,a,t}$), que reflete o número de trabalhadores em cada Arranjo Populacional categorizado pelas suas habilidades intensivas.

3.3.2.1 Estratégia de identificação

Conforme já mencionado na seção anterior, o modelo de Combes, Duranton e Gobillon (2010) evidenciam que é mais intensa a presença de trabalhadores habilidosos nos grandes centros. Logo, no modelo proposto acima há um problema de endogeneidade que se refere a escolha locacional do indivíduo: busca-se determinar o prêmio de especialização nos Arranjos Populacionais do Brasil, onde a variável principal de interesse é a especialização ocupacional categorizada por habilidades nos Arranjos Populacionais brasileiros. Como a variável é obtida a partir de uma unidade geográfica local, observa-se que trabalhadores mais habilidosos

poderiam se direcionar para determinados centros, ou seja, a endogeneidade resulta de choques locais que atraem/expulsam mais trabalhadores dessas ocupações categorizadas por habilidades para estes locais.

A estratégia de identificação deste trabalho consiste no uso de instrumentos históricos e de choques exógenos de demanda por trabalho para controlar a endogeneidade e a causalidade reversa mencionadas, além de usar um modelo longitudinal que possibilita controlar o viés de heterogeneidade não observada de indivíduos e firmas.

No que se refere aos instrumentos históricos, para o controle da endogeneidade da escolha local, baseado na metodologia de Rossi-Hansberg, Sarte e Schwartzman (2019), um modelo de variáveis instrumentais de dois estágios foi considerado. Os autores utilizaram três variáveis instrumentais para a escolha locacional na determinação dos diferenciais de salários dos trabalhadores, dado a composição do mercado de trabalho entre trabalhadores com habilidades cognitivas não rotineiras e trabalhadores com habilidades não cognitivas nos centros urbanos: população do século anterior, de acordo com Ciccone e Hall (1996); a migração do século anterior e a presença de áreas destinadas a faculdades, de acordo com Card (2001) e Moretti (2004), respectivamente. Nesta pesquisa, utiliza-se também de fontes históricas como variáveis instrumentais, mas de acordo com a disponibilidade dos dados no Brasil. Dessa forma, duas variáveis instrumentais foram utilizadas: população de 1920, 1940 e 1991, a partir dos microdados do Censo de Demográfico disponibilizados pelo IBGE e IPEADATA[21] e o Número de Cursos de Graduação Presenciais em 1991 e 2010 disponibilizado pelo Instituto Nacional de Estudos e Pesquisa (Inep) na Sinopse Estatística da Educação Superior, com base nos resultados do Censo da Educação Superior[22].

Seguindo as justificativas de Hansberg, Sarte e Schwartzman (2019) e Ciccone e Hall (1996) uma série histórica de população passada permite captar fontes originais de aglomeração que estão atreladas aos padrões populacionais históricos, dadas as preferências dos trabalhadores, e neste caso capta as preferências que não estejam atreladas a produtividade atual. Neste sentido, a população brasileira de 1920, 1940 e 1991 permite verificar esses padrões populacionais passados que determinam a escolha locacional do indivíduo.

[21] Disponível em: https://www.ibge.gov.br/estatisticas/sociais/populacao/25089-censo-1991-6.html?edicao=25090&t=publicacoes e no IPEADATA.
[22] Disponível em: http://portal.inep.gov.br/web/guest/sinopses-estatisticas-da-educacao-superior

No que se refere a variável histórica “Número de Cursos de Graduação Presenciais em 1991 e 2010”, baseando-se nas justificativas de Hansberg, Sarte e Schwartzman (2019) e Moretti (2004), compreende-se que a presença de faculdades em determinadas cidades pode estar relacionada com a composição de trabalhadores que residem nestas cidades, atraindo indivíduos mais habilidosos e captando padrões passados de aglomeração, impactando diretamente na escolha local dos indivíduos e influenciando nas chances de emprego.

Quanto ao instrumento que capta choques de demanda por trabalho, o controle da endogeneidade foi feito com base no *Shiftshare instrument*, comumente utilizado na literatura para controle de endogeneidade nos salários ou outras variáveis típicas do mercado de trabalho (BARTIK, 1991; SILVA; AZZONI, 2021; EHRL; MONASTERIO, 2020). O *Shiftshare instrument* foi formalmente especificado e desenvolvido na pesquisa de Bartik (1991), o que possibilitou denominar a variável de choques de demanda que representam previsões de crescimento, sendo então usado como uma variável instrumental para indicadores do mercado de trabalho (*Bartik Instrumental Variable*). A pesquisa original de Bartik (1991) objetiva mostrar que as políticas locais de desenvolvimento econômico, entendidas como aquelas de assistência direta às empresas, podem afetar o crescimento das regiões (estados ou áreas metropolitanas) e o desenvolvimento econômico da economia nacional. O autor usa como variável principal o crescimento do emprego para explicar variáveis que representam a estrutura do mercado de trabalho, tais como: desemprego, força de trabalho, horas semanais, preços da habitação e salários reais. Dentro desta perspectiva, Bartik (1991) tem como preocupação básica a ligação intrínseca entre os efeitos de curto e longo prazo no mercado de trabalho, isto é, como o crescimento do emprego local provoca mudanças nessas variáveis econômicas no curto prazo e como essas mudanças se manifestam no longo prazo.

Bartik (1991) considera 25 regiões metropolitanas dos EUA para um período compreendido entre 1972 e 1986, permitindo avaliar o efeito da mudança anual em cada região metropolitana. Dessa forma, o autor possibilita que o modelo esclareça como as variações no crescimento do emprego local em relação à média nacional está relacionada às variações da variável dependente local em relação à média nacional. Todavia, conforme já mencionado, o autor tem como foco principal avaliar como políticas públicas de desenvolvimento econômico podem impactar na economia local, e essas políticas são aquelas direcionadas às empresas, isto é, que buscam aumentar a demanda local de trabalho. Além disso, sua variável independente principal é o

crescimento do emprego local, e esse crescimento pode ser ocasionado tanto por efeitos da oferta de trabalho como por efeitos da demanda por trabalho. Nesse sentido, as regressões podem estar enviesadas se não for possível separar os efeitos da oferta de trabalho e de terra dos efeitos da demanda de trabalho.

Para contornar o problema de endogeneidade, Bartik (1991) cria um instrumento de choques de demanda. O autor enfatiza a importância em se considerar variáveis que captem como a demanda por trabalho muda nas regiões metropolitanas, mas considerando constante a oferta de trabalho, cujo efeito estaria no termo de erro. Para criar efeitos de choques de demanda, o autor considera os efeitos de crescimento nacional, de *shift* (mudança) e de *share* (participação). O componente de crescimento nacional calcula o crescimento como se todos os setores locais tivessem um crescimento tal como a média nacional, o componente *shift* calcula o crescimento extra considerando que os setores locais crescem diferente da média nacional dos setores e o componente *share* calcula e compara o crescimento extra dos setores locais em relação à média nacional[23]. As expressões para esses efeitos de crescimento nacional, de *share* e de *shift* são representadas, respectivamente, como se segue (BARTIK, 1991):

$$G_{mt} = \sum_{j} R_{jmb} \cdot \frac{(R_t - R_{nt-1})}{R_{nb}}$$

$$G_{mt} = \sum_{j} R_{jmb} \cdot \left[\frac{\left(R_{jnt} - R_{jnt-1}\right)}{R_{jnb}} - \frac{(R_{nt} - R_{nt-1})}{R_{nb}} \right]$$

$$G_{mt} = \sum_{j} R_{jmb} \cdot \frac{\left(R_{jnt} - R_{jnt-1}\right)}{R_{jnb}}$$

Em que R é o número de trabalhadores, isto é, o nível de emprego, o subscrito j representa o setor econômico, o subscrito m representa a região metropolitana, o subscrito n representa o país, o subscrito t representa o ano e o subscrito b representa o ano base considerado na pesquisa.

[23] Todavia, Bartik (1991) considera como instrumentos na sua pesquisa o componente de crescimento nacional e o componente *share.*

É interessante mencionar que Bartik (1991) considera importante a análise dos salários pela ótica local e ocupacional. O autor usa esses instrumentos ao analisar como o crescimento local afeta os salários reais especificados por local e ocupação, mantendo constante as habilidades individuais. Bartik (1991) verifica que choques de emprego têm efeitos significativos no longo prazo sobre a estrutura ocupacional, mas não sobre os salários. O autor indica que no curto prazo um aumento no emprego faz com que indivíduos de menor escolaridade ou menor leque de habilidades consigam adquirir empregos mais complexos, o que não aconteceria se não houvesse um choque de demanda positivo. Dessa forma, os trabalhadores aumentam o seu capital humano, o que permite que os mesmos possam permanecer nesses empregos após a minimização do choque inicial. Assim, no longo prazo, a estrutura ocupacional é alterada. No que se refere aos salários, como os empregadores preenchem suas novas vagas com indivíduos menos qualificados, a pressão sobre os salários pode ser evitada. O autor ainda enfatiza que mesmo que esse aumento salarial fosse inevitável para algumas ocupações, esse aumento não seria permanente pois não houve nenhuma mudança no estoque de capital humano desses trabalhadores e, além disso, a migração ajustaria a oferta de trabalho, diminuindo os salários aos patamares anteriores ao choque.

Essa pesquisa avalia a determinação do salário real categorizado por habilidades e localização, sendo uma medida mais precisa de salário real local, usada por Bartik (1991). Busca-se verificar como a especialização ocupacional categorizada por habilidades determinam os salários. Como já mencionado há endogeneidade no modelo, que foi corrigida a partir de variáveis instrumentais. A análise acima justifica o uso do instrumento Bartik (1991) e permite formular a estratégia de identificação, pois como estabelecido pelo autor, choques de emprego (demanda) modificam a estrutura ocupacional, e por sua vez a especialização ocupacional, mas não os salários ocupacionais, o que assegura a eficácia do instrumento para essa pesquisa. Ehrl e Monasterio (2020) e Acemoglu e Autor (2011) estabelecem que o crescimento do emprego nos países entre as ocupações se dá por múltiplos fatores relacionados a mudanças tecnológicas e mudanças na estrutura produtiva setorial dos países. Nesse sentido, Ehrl e Monasterio (2020) indicam que as mudanças tecnológicas que alteram a estrutura de emprego do país são diferentes e independentes dos fatores que determinam os salários.

Nessa pesquisa, o instrumento de choques de demanda por trabalho de Bartik (1991) foi utilizado em nível de ocupação, não de indústria como o autor originalmente construiu. A análise em nível ocupacional se ajusta mais aos objetivos propostos nesta pesquisa e, além

disso, considera-se que choques de demanda mais específicos e precisos podem ser captados pelas ocupações. Assim, uma análise mais minuciosa de choques de demanda no mercado de trabalho pode ser feita, pois cada ocupação está atrelada a determinado produto ou serviço, diferentemente de setores que tem um nível de agregação considerável. A análise em nível de ocupação para o instrumento Bartik (1991) também foi feita por Ehrl e Monasterio (2020).

Tal como as pesquisas de Ehrl e Monasterio (2020) e Silva (2017), esse estudo fez uso do instrumento *Bartik shift*, uma vez que alguns autores como Borusyak e outros (2018) indicam que o componente *share* pode ser endógeno. Além disso, justifica-se o uso do *Bartik shift* pois o instrumento capta mudanças que podem representar melhor choques de demandas locais baseados no crescimento nacional das ocupações de forma simples.

Com base em uma análise de estudos empíricos que utilizam o instrumento originalmente desenvolvido por Bartik (1991), Silva (2017) levanta uma limitação do instrumento: os choques locais podem influenciar ou determinar os choques de demanda nacional, se essas áreas locais possuírem uma grande parcela da produção de determinado setor. De acordo com a autora, essa condição pode ocorrer quando o emprego de um setor for muito concentrado em uma localidade, o que comprometeria a exogeneidade do instrumento. Autor e Duggan (2003) propõem que no cálculo do crescimento do emprego nacional seja excluído o emprego da respectiva localidade. Silva (2017) ainda alerta que se existir dependência produtiva entre as localidades, essa estratégia pode deixar resíduos no crescimento nacional de um crescimento local. Todavia, a autora indica que esse problema ocorre quando a dimensão local é pequena, por exemplo, cidades. A autora, ao utilizar uma dimensão espacial maior, os Arranjos Populacionais brasileiros, indica que a estratégia de Autor e Duggan (2003) é suficiente para garantir exogeneidade ao instrumento. Como a pesquisa desenvolvida nesse ensaio também faz uso da dimensão espacial de Arranjos Populacionais brasileiros, o instrumento utilizado nesse estudo é definido como se segue:

$$BARTIK_{a,t}^{h} = \sum_{o} E_{o,a,t_0} \cdot \frac{\left(E_{o,(n-a),t} - E_{o,(n-a),t_0}\right)}{E_{o,(n-a),t_0}} \tag{3.15}$$

Em que, E_{o,a,t_0} é o nível de emprego na ocupação o no arranjo populacional a para o ano base t_0 (1994) multiplicado pela taxa de crescimento nacional do emprego (excluindo o emprego do respectivo Arranjo populacional $(n-a)$) na ocupação o para cada ano da base de dados (2009

à 2018) e o ano base (1994): $\left(E_{o,(n-a),t} - E_{o,(n-a),t_0}\right)/E_{o,(n-a),t_0}$. O subscrito h indica as categorias de habilidades ocupacionais (h=3, cognitivas, sociais e manuais). Assim, tem-se um instrumento *Bartik shift* ocupacional para cada categoria de habilidades que foi utilizado para corrigir a endogeneidade na equação salarial da especialização ocupacional categorizada pelas habilidades respectivas.

O ano base de 1994 foi considerado porque representa uma defasagem mínima de 15 anos em relação ao primeiro ano da base de dados (2009), isso permite maior credibilidade e exogeneidade ao instrumento em relação aos choques locais recentes. Conforme já mencionado, os trabalhadores são categorizados por ocupações, nesse sentido se fez uso da CBO-2002, classificação de ocupações mais atualizada. Todavia, na base de dados da RAIS de 1994 a classificação de ocupações apresentada era a de 1994. Para criar a taxa de crescimento do emprego para o período compreendido entre 2009 e 2018, considerando como ano base 1994, foi necessário realizar uma conversão da CBO de 1994 para a CBO de 2002 para os dados de 1994, para tanto utilizou-se as Tábuas de conversão da Classificação Brasileira de Ocupações disponível no site do Ministério do Trabalho[24].

3.3.2.2 Métodos de Estimação

Desse modo, considerando a estratégia empírica de *pooling* de dados para os anos de 2009 a 2018, o modelo de dois estágios foi utilizado. Considerando um modelo de regressão múltipla com K variáveis, onde tem-se os vetores $Kx1$, deve existir um vetor de instrumentos z de dimensão $rx1$, de modo que $r \geq K$, que satisfaça as seguintes propriedades (CAMERON; TRIVEDI, 2005):

1. z não é correlacionado com o erro u;
2. z é correlacionado com o vetor regressor x (endógeno);
3. z é fortemente correlacionado, ao invés de fracamente correlacionado, com o vetor regressor x.

[24] Para maiores informações acesse: https://www.mtecbo.gov.br/cbosite/pages/tabua/FiltroConversao_CBO2002_CBO94_CIUO88.jsf

Dessa forma, estabelece-se que, se as propriedades 1 e 2 são necessárias para a consistência do modelo, a terceira propriedade fortalece a segunda em uma condição de amostra finita. Adicionalmente, indica-se que se a primeira condição falhar o instrumento é inválido, se a segunda condição falhar o instrumento é irrelevante e se a terceira condição falhar o instrumento é dito como fraco. Tem-se ainda a condição de ordem que estipula que o número de instrumentos deve ser pelo menos igual ao número de variáveis endógenas, logo $r \geq K$. Se $r = K$, o modelo é identificado e se $r > K$ o modelo é sobreidentificado.

No modelo empírico, considera-se que a variável que apresenta endogeneidade de escolha locacional é a especialização ocupacional categorizada por habilidades, EO_{hat}. Conforme já justificado acima, percebe-se que os instrumentos População brasileira de 1920, 1940 e 1991, o Número de Cursos de Graduação Presenciais em 1991 e 2010 e o *Bartik shift* ocupacional não determinam os salários atuais dos indivíduos, mas determinam sua escolha locacional atual interferindo diretamente na variável de especialização ocupacional. Além disso, verifica-se que o modelo é sobreidentificado para as variáveis históricas e identificado para a variável *Bartik shift* ocupacional. Desse modo, pode-se indicar que, de acordo com as propriedades econométricas, as variáveis instrumentais podem conduzir a estimativas mais consistentes.

Os estimadores de Mínimos Quadrados de dois Estágios são definidos da seguinte maneira, (CAMERON; TRIVEDI, 2005):

1. No primeiro estágio, tem-se a regressão de MQO da variável endógena com a variável instrumental como variável explicativa:

$$X = Z\theta + \psi$$

   Em que, o estimador será $\hat{X} = Z\hat{\theta} = Z(Z'Z)^{-1}Z'X$

2. No segundo estágio, tem-se a regressão de MQO da variável dependente com a endógena instrumentalizada:

$$y = \hat{X}\beta + u$$

O estimador $\hat{\beta}$ então será dado por: $\hat{\beta} = [\hat{X}'\hat{X}]^{-1}\hat{X}'y$. Substituindo $\hat{X}$ do primeiro estágio na equação anterior, tem-se:

$$\hat{\beta} = [X'Z(Z'Z)^{-1}Z'X]^{-1}[X'Z(Z'Z)^{-1}Z'y]$$

Este será um estimador consistente. Empiricamente, o modelo desta pesquisa pode ser representado em dois estágios. Inicialmente considerando como instrumentos as variáveis históricas, tem-se no primeiro estágio:

$$\ln(EO_{h,a,t}) = \theta_0 + \boldsymbol{\theta_1 POP_a} + \boldsymbol{\theta_2 NCGP_a} + \boldsymbol{\theta_3 X_{it}} + \boldsymbol{\theta_4 E_{jt}} + \boldsymbol{T_t} + \psi_{a,t} \quad (3.16)$$

Em que, $\ln(EO_{h,a,t})$ é a variável dependente endogenamente identificada, correspondendo ao indicador de especialização ocupacional, categorizado por habilidades. Como *h*=3 (habilidades cognitivas, sociais e manuais), tem-se três equações no primeiro estágio. As variáveis instrumentais são: POP_a, a população de 1920, 1940 e 1991, e $NCGP_a$, o número de Cursos de Graduação Presenciais em 1991 e 2010. No segundo estágio, tem-se:

$$\ln(w_{ijt}) = \beta_0 + \beta_{1h}\ln(\widehat{EO}_{h,a,t}) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{jt}} + \boldsymbol{T_t} + \varepsilon_{ijt} \quad (3.17)$$

Em que, $logw_{ijt}$ é o logaritmo natural do salário-hora real do indivíduo *i* na firma *j* no período *t*, $\boldsymbol{X_{it}}$ é um vetor de características socioeconômicas observáveis, que inclui variáveis de controle individuais, tais como idade, idade ao quadrado, experiência, experiência ao quadrado e variáveis *dummies* de gênero (sem o efeito fixo individual) e o grau de instrução, $\boldsymbol{E_{jt}}$ é um vetor de características do emprego do trabalhador, mais especificamente, *dummies* de setor de atividade econômica de acordo com a CNAE e *dummies* de tamanho do estabelecimento, $\ln(\widehat{EO}_{hat})$ é a variável de interesse instrumentalizada, que indica o logaritmo natural do grau de especialização ocupacional a partir das variáveis instrumentais adotadas, $\boldsymbol{T_t}$ são *dummies* de ano e β são os vetores de parâmetros a serem estimados de modo consistente.

Como teste de robustez, considerando o controle da endogeneidade a partir do instrumento *Bartik shift* ocupacional no primeiro estágio, tem-se:

$$\ln(EO_{h,a,t}) = \vartheta_0 + \vartheta_1 BARTIK_{h,a,t} + \boldsymbol{\vartheta_2 X_{it}} + \boldsymbol{\vartheta_3 E_{jt}} + \boldsymbol{T_t} + \epsilon_{a,t} \quad (3.18)$$

Em que, $\ln(EO_{h,a,t})$ é a variável dependente endogenamente identificada, correspondendo ao indicador de especialização ocupacional, categorizado por habilidades. Como *h*=3 (habilidades

cognitivas, sociais e manuais), tem-se três equações no primeiro estágio. A variável instrumental, conforme já mencionado, é o *Bartik shift* ocupacional. No segundo estágio, tem-se:

$$\ln(w_{ijt}) = \beta_0 + \beta_{1h}\ln(\widehat{EO}_{h,a,t}) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{jt}} + \boldsymbol{T_t} + \varepsilon_{ijt} \quad (3.19)$$

Em que, $logw_{ijt}$ é o logaritmo natural do salário-hora real do indivíduo *i* na firma *j* no período *t*, $\boldsymbol{X_{it}}$ é um vetor de características socioeconômicas observáveis, que inclui variáveis de controle individuais, tais como idade, idade ao quadrado, experiência, experiência ao quadrado e variáveis *dummies* de gênero (sem o efeito fixo individual) e o grau de instrução, $\boldsymbol{E_{jt}}$ é um vetor de características do emprego do trabalhador, mais especificamente, *dummies* de setor de atividade econômica de acordo com a CNAE e *dummies* de tamanho do estabelecimento, $\ln(\widehat{EO}_{hat})$ é a variável de interesse instrumentalizada, que indica o logaritmo natural do grau de especialização ocupacional a partir das variáveis instrumentais adotadas, $\boldsymbol{T_t}$ são *dummies* de ano e β são os vetores de parâmetros a serem estimados de modo consistente. Novamente, ressalta-se que as equações salariais (3.17) e (3.19) são estimadas para cada categoria de habilidades h (cognitivas, sociais e manuais), obtendo-se, assim, três modelagens empíricas diferentes para cada equação.

A estimação dos diferenciais salariais da especialização ocupacional categorizada pelas habilidades, controlada pelos dois grupos de variáveis instrumentais apresentadas nessa pesquisa, é uma contribuição à literatura, pois permite verificar como choques de oferta (variáveis históricas) e choques de demanda (*Bartik shift*) impactam na determinação dos salários locais por habilidades. Bartik (1991) enfatiza a importância em se diferenciar choques de oferta e demanda e como esses choques impactam no mercado de trabalho. De acordo como o autor, choques de oferta aumentam a população, podendo ter efeitos negativos sobre os salários reais, ou seja, estes diminuem após um choque de oferta, e podem ter diversas causas, inclusive melhoria das amenidades locais. Já os choques de demanda, ao aumentar a lucratividade local, seriam refletidos no mercado de trabalho com aumentos dos salários reais. Bartik (1991) indica que o crescimento geral do emprego se deve tanto a choques de oferta como a choques de demanda, e que o efeito positivo sobre os salários reais é mais acentuado quando se considera choques de demanda, ou seja, após o controle da endogeneidade no seu modelo obtém-se um aumento das estimativas de salários.

Desse modo, as variáveis históricas dessa pesquisa podem representar choques de oferta pois, em seu conjunto, representam a população defasada e o número de cursos presenciais defasados, que estariam associados às amenidades locais. Dada a metodologia de Bartik (1991), que cria um instrumento com o intuito de separar os efeitos de oferta dos efeitos de demanda, essa pesquisa busca controlar não apenas choques de oferta (variáveis históricas), mas também choques de demanda (*Bartik shift*). Assim, diante do exposto no parágrafo anterior, espera-se que os choques de demanda tenham um impacto positivo maior sobre os salários do que os choques de oferta, tal como preconizado por Bartik (1991). É importante mencionar a importância de utilizar variáveis defasadas, pois as modificações no mercado de trabalho se dão de forma lenta, principalmente, devido ao processo migratório que impacta na oferta de trabalho, nas mudanças estruturais relacionadas ao rol de habilidades dos trabalhadores de determinada localidade e, como consequência, na adequação à demanda de trabalho dessa localidade. Desse modo, utilizar variáveis históricas ou defasadas permite captar melhor os efeitos de curto e longo prazo, evidência levantada na pesquisa de Bartik (1991).

A segunda estratégia de controle da endogeneidade, dado a heterogeneidade não observada de trabalhadores e firma, é a utilização de um painel de dados. Pode-se verificar como as características não observadas dos indivíduos, das firmas e de seu efeito de par combinado (trabalhadores e firmas, definindo uma *proxy* de *matching*) podem impactar no prêmio salarial da especialização ocupacional. Considerando a modelagem em painel, foi estimado o seguinte modelo de regressão:

$$\ln(w_{ijt}) = \beta_0 + \beta_{1h}\ln\left(EO_{h,a,t}\right) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{jt}} + \emptyset_i + \boldsymbol{T_t} + \varepsilon_{ijt} \quad (3.20)$$

$$\ln(w_{ijt}) = \beta_0 + \beta_{1h}\ln\left(EO_{h,a,t}\right) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{jt}} + \omega_j + \boldsymbol{T_t} + \varepsilon_{ijt} \quad (3.21)$$

$$\ln(w_{ijt}) = \beta_0 + \beta_{1h}\ln\left(EO_{h,a,t}\right) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{jt}} + \varphi_{ij} + \boldsymbol{T_t} + \varepsilon_{ijt} \quad (3.22)$$

Em que, $\ln(w_{ijt})$ é o logaritmo natural do salário-hora real do indivíduo *i* na firma *j* no período *t*, $\boldsymbol{X_{it}}$ é um vetor de características socioeconômicas observáveis, que inclui variáveis de controle individuais, tais como experiência, experiência ao quadrado e variáveis *dummies* de grau de instrução, $\boldsymbol{E_{jt}}$ é um vetor de características do emprego do trabalhador, mais especificamente, *dummies* de setor de atividade econômica de acordo com a CNAE e *dummies* de tamanho do estabelecimento, $\ln(EO_{hnt})$ é a variável de interesse, que indica o logaritmo natural do grau de especialização ocupacional no Arranjo populacional *n*, h=3 (habilidades

cognitivas, sociais e manuais), de acordo com a classificação de habilidades estabelecida por Maciente (2013) e categorizada por Neves Júnior (2018), β são os vetores de parâmetros a serem estimados, $\emptyset_i$ é o efeito fixo do trabalhador, ω_j é o efeito fixo da firma, φ_{ij} é o o efeito fixo de trabalhador e firma, $\boldsymbol{T_t}$ são as *dummies* de tempo e ε_{it} é o erro estocástico. As equações (3.20), (3.21) e (3.22) foram estimadas separadamente para cada grupo de habilidade.

Convém ressaltar que o efeito fixo de trabalhador e firma (φ_{ij}) é um efeito fixo de par, isto é, tem-se um identificador para cada combinação de trabalhador e firma. Neste caso, se o trabalhador muda de firma um novo par (trabalhador/firma) será obtido. Desse ponto em diante, o efeito fixo de trabalhador e firma será chamado de *matching*. Além disso, é importante ressaltar que, espera-se um pequeno impacto das características não observadas individuais no prêmio da especialização ocupacional uma vez que esta já está categorizada por habilidades e, como é comum na literatura, as características não observadas dos indivíduos representam as suas habilidades inatas.

Todavia, considerando que, para categorizar as ocupações por habilidade considera-se, além das habilidades inatas, as habilidades adquiridas no exercício da ocupação pelo próprio *learning*. Portanto, dependendo da composição de habilidades (inatas ou adquiridas), a mudança nos coeficientes que captam o prêmio salarial da especialização pode ser expressiva. As características não observadas das firmas são fundamentais na análise e podem impactar de forma significativa, uma vez que, para a compreensão dos padrões de aglomeração no espaço, em especial análises de aglomeração ocupacional, o papel das firmas é fundamental. Esse papel pode estar atrelado com as ligações de insumo-produto. Nesse sentido, não considerar a heterogeneidade não observada das firmas traria um problema grave de variável omitida (ROSSI-HANSBERG; SARTE; SCHWARTZMAN, 2019; GOLDMAN, KLIER; WALSTRUM, 2019; COMBES; DURANTON; GOBILLON, 2008).

3.3.3 Decomposição salarial

Abowd, Kramarz e Margolis (1999), ao estudar os diferenciais salariais interindustriais, enfatizaram que os trabalhos já realizados anteriormente (DICKENS; KATZ, 1987; KRUEGER; SUMMERS, 1987; MURPHY; TOPEL, 1987; GIBBONS; KATZ, 1992) não conseguiram estimar com precisão esses diferenciais, pois os mesmos dependem simultaneamente das características não observadas dos indivíduos e das empresas, e nenhum

trabalho até então havia conseguido controlar essas heterogeneidades não observadas, dadas as limitações empíricas, além de falta de microdados que captassem as características de empresas e trabalhadores. Com o avanço desses microdados, em especial na França, onde os autores realizaram seu estudo, a evolução econométrica ainda apresentava limitações. Neste ponto reside a principal contribuição dos autores, não só para a análise dos diferenciais de salários interindustriais, mas para a Economia do Trabalho como um todo. Abowd, Kramarz e Margolis (1999) desenvolveram a primeira análise estatística de heterogeneidade simultânea considerando o nível individual e da firma na determinação e decomposição dos salários. O método de decomposição permite indicar se os diferenciais salariais se devem mais aos efeitos não observados das firmas ou dos indivíduos.

Esta técnica só é possível a partir de dados longitudinais que permitem a interação entre o método dos mínimos quadrados e dados em painel, controlando a heterogeneidade observada e não observada dos trabalhadores e das firmas. O método de decomposição salarial se aproxima da estimação de mínimos quadrados, mas permite que as heterogeneidades de indivíduos e empresas sejam correlacionadas. Dessa correlação, surge uma exigência estatística básica do modelo onde deve existir mobilidade dos indivíduos entre as firmas, sendo esta uma condição necessária para a identificação estatística do modelo, pois só assim é possível captar o efeito indivíduo e firma separadamente (ABOWD; KRAMARZ; MARGOLIS,1991).

Conforme já mencionada a metodologia desenvolvida pelos autores supracitados tem reflexos na Economia do Trabalho e Regional como um todo. Dado que esse ensaio tem como objetivo verificar a existência de prêmios salariais para a especialização ocupacional categorizada pelas habilidades dos trabalhadores, é importante realizar uma análise de decomposição a fim de detectar os efeitos da heterogeneidade não observada dos indivíduos e das firmas nos níveis de salários que se devem à especialização ocupacional, pois, como enfatizado Rossi-Hansberg, Sarte e Schwartzman (2019), a análise de múltiplas indústrias e ligações de insumo-produto é fundamental para determinar a demanda por diferentes ocupações, estabelecendo assim a intensidade de cada ocupação nas indústrias locais.

Além disso, Goldman, Klier e Walstrum (2019), ao avaliar a aglomeração de ocupações intrasetoriais, evidenciam a importância em se considerar as características das firmas e indivíduos conjuntamente para definir os padrões de aglomeração no espaço. Os mesmos utilizam dados ao nível de trabalhador e estabelecimentos, avaliando conjuntamente indústria,

ocupação e local de trabalho. Diante destas considerações fica evidente que a especialização ocupacional capta não só variáveis relacionadas ao trabalhador, mas também relacionadas às firmas. No modelo empírico desenvolvido neste ensaio considera-se, além de características dos indivíduos, características das firmas, assim, é relevante a análise da decomposição salarial na determinação dos prêmios salariais da especialização ocupacional.

Abowd, Kramarz e Margolis (1999) iniciam a derivação da sua estratégia empírica de decomposição salarial considerando as consequências da heterogeneidade de indivíduos e firmas na determinação dos salários, de modo geral. O estudo indica que esta heterogeneidade tem motivado diversas pesquisas, tanto no que se refere a diferenças permanentes entre indivíduos, quanto no que se refere a diferenças permanentes entre firmas. Os autores sintetizam estes modelos considerando a seguinte equação de salários:

$$y = X\beta + D\theta + F\psi + \varepsilon \tag{3.23}$$

Em que, y é o vetor $(N^* x\ 1)$ de salários anuais (em desvio da média), X é uma matriz $(N^* x\ P)$ de características observáveis em relação aos desvios médios, variáveis no tempo, D é uma matriz $(N^* x\ N)$ de identificadores de indivíduos, F é uma matriz $(N^* x\ mJ)$ de identificadores do efeito total da firma, em que i trabalha na data t, ε é o vetor de resíduos, β, θ e ψ são os parâmetros do modelo (Observe que $N^* = NT$). Ao apresentarem esse modelo, Abowd, Kramarz e Margolis (1999) enfatizam que a literatura, em geral, sempre aborda a análise dos diferenciais de salários considerando um dos efeitos ausentes, seja de indivíduos ou de firmas, modelos que os autores consideram inconsistentes pois não se adequam à realidade empírica, gerando ambiguidades nas interpretações dos parâmetros. Assim, enfatizam que a omissão ou agregação de um ou mais efeitos pode alterar a interpretação dos outros.

Com isso, Abowd, Kramarz e Margolis (1999) mostram estatisticamente como é diferente a estimação e interpretação dos efeitos quando a equação de salários não inclui o efeito puro de indivíduos e quando não inclui o efeito puro da firma, situações em que não é possível definir com exatidão o efeito de indivíduo e o efeito da firma, pois ambos se confundem. Assim, mais uma vez, ressaltam a importância de identificar com mais precisão os efeitos puros do indivíduo e da firma, indicando a contribuição do método de decomposição salarial que propõem.

Dessa forma, Abowd, Kramarz e Margolis (1999) propõem um método de estimação condicional, indicando que a denominação condicional se deve à relação de técnicas computacionais de modelos lineares e dados em painel, que permitem estimações com efeitos fixos. O método de estimação condicional permite, obtendo estimadores que sejam o mais semelhante possível com o de mínimos quadrados, identificar separadamente o efeito de indivíduo e firma, realizando uma decomposição salarial. Para tanto, algumas suposições de ortogonalidade adicionais ao modelo de mínimos quadrados devem ser consideradas. Essas suposição origina-se da interação de ordem superior entre as características observáveis, a heterogeneidade não observada de indivíduos e a heterogeneidade das firmas, que representam a correlação entre as matrizes X, D e F.

Abowd, Kramarz e Margolis (1999) testaram e avaliaram três técnicas de estimação do modelo geral para o método de estimação condicional, a saber o método de "ordem independente", o método de "ordem dependente", com efeitos de indivíduos em primeiro lugar e "ordem dependente", com efeitos de firmas em primeiro lugar. Esses métodos são aplicados em dois estágios. No método de ordem independente, não importa em qual ordem o efeito (indivíduo ou firma) é estimado primeiro, os resultados são iguais independente da ordem da estimação. Já no método de ordem dependente, os resultados diferem, dependendo de qual efeito é usado primeiro, isso acontece por que os resíduos da primeira etapa são usados para calcular as estimativas da segunda etapa. Adicionalmente, os autores indicam que, em todos os métodos e modelos estimados, o efeito de indivíduo é estatisticamente mais significante do que o de firma. O modelo AKM (1999) é considerado e adaptado para a proposta empírica desse ensaio, buscando controlar as heterogeneidades não observadas de trabalhadores e firmas na estimação da equação de salários (1º estágio da decomposição do prêmio de especialização).

Assim, a fim de avaliar a decomposição do prêmio salarial da especialização ocupacional categorizada por habilidades, utilizou-se a abordagem de estimação em dois estágios baseada em Combes, Duranton e Gobillon (2008). Inicialmente, estima-se uma equação de salário para cada habilidade, em *Pooled* de dados:

$$\ln(w_{it}) = \beta_0 + \sum_a \beta_{1\,at}^{\,h} \ln(EO_{h,a,t}) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{it}} + \boldsymbol{T_t} + \varepsilon_{it} \qquad (3.24)$$

Para captar os efeitos da especialização de habilidades sobre os salários dos trabalhadores foram incluídos os efeitos fixos de cada Arranjo Populacional $a = \{1, \dots, 375\}$. Se o efeito fixo do

Arranjo Populacional for positivo e estatisticamente significante, o vetor $\beta_{1_{at}}^{h}$ será a magnitude do prêmio da especialização ocupacional categorizada por habilidades. Como a equação salarial (3.24) é estimada para cada categoria de habilidades h (cognitivas, sociais e manuais), obtém-se três modelagens empíricas. Assim, $\beta_{1_{at}}^{h=cog}$ reflete o prêmio salarial da especialização ocupacional para habilidades cognitivas nos Arranjos Populacionais, $\beta_{1_{at}}^{h=soc}$ reflete o prêmio salarial da especialização ocupacional para habilidades sociais nos Arranjos Populacionais e $\beta_{1_{at}}^{h=man}$ reflete o prêmio salarial da especialização ocupacional para habilidades manuais nos Arranjos Populacionais. A especificação das demais variáveis é a mesma da equação (3.14). Nesta metodologia, a fim de captar os efeitos puros (ou brutos) do prêmio salarial, não há controle das heterogeneidades não observadas dos trabalhadores.

No segundo estágio aplica-se a decomposição salarial. Nesta etapa, utiliza-se os parâmetros $\beta_{1_{at}}^{h}$ estimados em (3.24) como variáveis dependentes. Dessa forma, obteve-setrês equações, uma para cada categorização de habilidades (cognitivas, sociais e manuais). Seguindo a contribuição de Silva (2017), que estimou a decomposição do prêmio salarial urbano para o Brasil, também baseada em Combes, Duranton e Gobillon (2008), utilizou-se como variável independente principal a densidade no emprego, de fundamental importância na economia urbana para captar os efeitos da aglomeração. Além disso, como os coeficientes decompostos neste ensaio refletem uma medida de especialização no espaço econômico, a variável se justifica no modelo aqui proposto. Assim a decomposição, para cada habilidade, no segundo estágio obedece à seguinte equação empírica:

$$\beta_{1_{at}}^{h} = \delta_0 + \delta_1 \ln(DE_{at}) + \delta_2 \ln(\acute{a}rea_{at}) + \boldsymbol{\delta_3 CA_{at}} + \boldsymbol{T_t} + \epsilon_{at} \qquad (3.25)$$

Em que, $\beta_{1_{at}}^{h=cog}$ reflete o prêmio salarial médio da especialização ocupacional para habilidades cognitivas nos Arranjos Populacionais, $\beta_{1_{at}}^{h=soc}$ reflete o prêmio salarial médio da especialização ocupacional para habilidades sociais nos Arranjos Populacionais, $\beta_{1_{at}}^{h=man}$ reflete o prêmio salarial médio da especialização ocupacional para habilidades manuais nos Arranjos Populacionais, DE_{at} é a densidade no arranjo *a* no período *t*, $\acute{a}rea_{at}$ é a área do arranjo no período *t* e $\boldsymbol{CA_{at}}$ é um vetor de características observáveis do arranjo, que inclui variáveis *dummies* de Arranjos Fronteiriço ou Turístico, e um conjunto de *dummies* de macrorregião a qual pertence o arranjo, e ϵ_{at} é o termo de erro.

Assim, considerando a literatura de Economia Urbana, foi verificado se o tamanho dos mercados de trabalho importa para o prêmio salarial da especialização ocupacional, através da densidade do emprego. Verificou-se como os efeitos fixos dos indivíduos e o efeito fixo das firmas determinam a decomposição dos efeitos da especialização de habilidades sobre os salários, analisando ainda se a decomposição difere quando se considera diferentes categoriais de habilidades. Dessa forma, o primeiro estágio com o controle da heterogeneidade não observada de trabalhadores e firmas, permite obter os efeitos líquidos da especialização sobre os salários. O efeito bruto para cada grupo de habilidade pôde ser captado pela equação (3.24), o efeito líquido foi analisado em três situações: com o controle de efeitos fixos do trabalhador (*i*); com o controle de efeitos fixos das firmas (*j*) e com o controle de efeitos fixos do trabalhador e das firmas (*matching*). No geral, tem-se quatro especificações para o primeiro estágio:

Especificação (1) – *Pooled* de dados:

$$\ln(w_{it}) = \beta_0 + \sum\nolimits_a \beta_{1_{at}}^{h} \ln(EO_{hat}) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{it}} + \boldsymbol{T_t} + \varepsilon_{it}$$

Especificação (2) – Painel com o controle de efeitos fixos do trabalhador (*i*):

$$\ln(w_{it}) = \beta_0 + \sum\nolimits_a \beta_{1_{at}}^{h} \ln(EO_{hat}) + \boldsymbol{\beta_2} X_{it} + \boldsymbol{\beta_3 E_{it}} + \boldsymbol{T_t} + \emptyset_i + \varepsilon_{it}$$

Especificação (3) – Painel com o controle de efeitos fixos da firma (*j*):

$$\ln(w_{it}) = \beta_0 + \sum\nolimits_a \beta_{1_{at}}^{h} \ln(EO_{hat}) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{it}} + \boldsymbol{T_t} + \omega_j + \varepsilon_{it}$$

Especificação (4) – Painel com o controle de efeitos fixos do trabalhador e firma (*ij*):

$$\ln(w_{it}) = \beta_0 + \sum\nolimits_a \beta_{1_{at}}^{h} \ln(EO_{hat}) + \boldsymbol{\beta_2 X_{it}} + \boldsymbol{\beta_3 E_{it}} + \boldsymbol{T_t} + \varphi_{ij} + \varepsilon_{it}$$

As especificações (2), (3) e (4), do primeiro estágio do modelo de decomposição, foram estimadas com o método de Efeitos Fixos, tradicional para modelos lineares de dados em painel. O método assume a existência de heterogeneidades não observadas, que são correlacionadas com uma ou mais variáveis do modelo, e garante a estimação consistente dos parâmetros após o controle do viés de variável omitida (CAMERON; TRIVEDI, 2005). O segundo estágio seguiu a dinâmica já apresentada na equação (3.25). Todavia, o problema de causalidade reversa também está presente na relação entre salários locais e densidade. Para contornar o problema,

considerando que a endogeneidade pode gerar estimativas tendenciosas, variáveis instrumentais foram utilizadas para determinação da relação entre densidade e efeitos da especialização de habilidades sobre os salários. Utilizou-se instrumentos históricos (população de 1940 e população de 1991) para a obtenção do vetor de densidade corrigido do viés de endogeneidade (estágio intermediário de controle de endogenidade).

3.3.4 Dados

Esta seção tem por objetivo apresentar e descrever as bases de dados utilizadas para determinação do prêmio salarial da especialização ocupacional categorizada pelas habilidades dos trabalhadores. As duas principais bases utilizadas foram o Relatório Anual de Informações Sociais (RAIS) e a base de dados de Maciente (2013), além disso utilizou-se séries históricas do Censo Demográfico Brasileiro disponibilizado pelo IBGE e do Censo da Educação Superior disponibilizado pelo Instituto Nacional de Estudos e Pesquisa (Inep). A apresentação das bases de dados da RAIS-ME e da base de Maciente (2013) já foi realizada no ensaio anterior, de modo que nessa seção é feita uma breve discussão acerca do tratamento da base de dados para essa pesquisa.

No que se refere a RAIS, foi necessário organizar os microdados em nível de indivíduos (PIS) e estabelecimentos (CNPJ). Essa organização foi feita anualmente e, após isso, gerou-se o painel de dados não balanceado. Considerando que a base de dados da RAIS pode apresentar inconsistências e erros de preenchimento, a limpeza da base foi realizada anualmente. Inicialmente, considerando apenas os trabalhadores com vínculos ativos, dado que a dinâmica de determinação local e salarial dos trabalhadores do setor público obedece à critérios peculiares, considerou-se apenas os trabalhadores do setor privado. Além disso, como é comum em análises do mercado de trabalho, considerou-se somente a população em idade ativa, isto é, trabalhadores entre 18 e 65 anos. Adicionalmente, para uma análise mais consistente do mercado de trabalho, no que se refere as condições em que este é estabelecido, exclui-se os trabalhadores com carga horária semanal menor que 20 horas, com tempo no emprego igual a zero meses e aqueles que recebem menos do que o salário mínimo vigente no respectivo ano. Exclui-se da base, ainda, os vínculos que apresentavam PIS inexistente ou igual a zero, remuneração inexistente ou igual a zero, municípios e ocupações sem registro, dado que estas variáveis são fundamentais para a obtenção dos resultados dessa pesquisa.

Além disso, para os trabalhadores que possuíam no mesmo ano vínculos (PIS) multiplicados, isto é, que possuíam mais de um vínculo e inconsistências, foram realizados procedimentos de ajuste para a seleção do vínculo principal. Inicialmente, foram excluídos aqueles trabalhadores que possuíam mais de 5 vínculos, pois certamente tratava-se de erro de preenchimentos, posteriormente exclui-se aqueles trabalhadores com gêneros diferentes entre os vínculos (PIS) repetidos. Além disso, manteve-se os vínculos com mais horas contratadas e com data de admissão mais antiga. Após essas exclusões, objetivando manter apenas um vínculo por trabalhador (PIS), exclui-se aleatoriamente os PIS repetidos na base. É fundamental ainda excluir aqueles vínculos cujo CNPJ ou CEI dos estabelecimentos não possuíam registro (eram inexistentes ou igual a zero), uma vez que essa variável é o identificador do estabelecimento.

Conforme já mencionado, a análise da especialização ocupacional nesta pesquisa considera como subunidades regionais os Arranjos Populacionais, dada a sua importância na dinâmica econômica do país. Nesse sentido, exclui-se aqueles vínculos que estão localizados em municípios que não participam de nenhum Arranjo Populacional ou que não sejam caracterizados como Média e Grande Concentração.

No que se refere ao identificador da firma, como a estratégia de identificação desse estudo requer mobilidade de trabalhadores entre firmas e Arranjos Populacionais e mobilidade de firmas entre Arranjos Populacionais ao longo do painel de dados, alguns filtros foram feitos na variável que identifica a firma: CNPJ. Algumas inconsistências estavam presentes na variável CNPJ, a saber: alguns estabelecimentos tinham localizações e/ou setores de atividades distintos no mesmo ano, todavia apresentavam o mesmo CNPJ. Nessas situações, optou-se pelo uso do identificador da firma a partir do Cadastro Específico do Instituto Nacional do Seguro Social (CEI) e não do Cadastro Nacional de Pessoas Jurídicas (CNPJ), como comumente é feito.

Esse critério segue a proposta de Silva (2017), que tem a RAIS como a base de dados principal na sua pesquisa, e opta por utilizar o CEI em substituição ao CNPJ, isso porque a autora verificou que no período considerado na sua pesquisa (2002-2014) alguns estabelecimentos apresentavam o mesmo CNPJ mesmo estando em unidades territoriais diferentes no mesmo ano, logo, não se tratava de um processo migratório dos estabelecimentos, e que o mesmo CNPJ estava registrado em setores de atividades distintos no mesmo ano. Assim, a autora concluiu que essas inconsistências poderiam ser causadas, além de erros de na declaração, pelo fato de algumas empresas filiais apresentarem na declaração da RAIS o mesmo CNPJ da empresa

matriz. Sendo observado a mesma tendência para o período compreendido nesta pesquisa, optou-se por considerar como identificador dos estabelecimentos o CEI, nesses casos específicos. Para casos em que a inconsistência ainda permanecia, os CNPJ foram excluídos, a fim de permitir verificar a mobilidade das firmas sem equívocos. Após essas correções, obteve-se um painel não balanceado em nível de trabalhadores e estabelecimentos. O painel apresentou 168.408.851 observações, referentes a 46.782.167 trabalhadores formais e 3.897.242 estabelecimentos, no período de 2009 a 2018.

Após a classificação dos trabalhadores de acordo com a intensidade de habilidades utilizadas em cada ocupação, considerando a base de Maciente (2013), para facilitar alguns exercícios empíricos, o painel total pode ser subdividido em: Trabalhadores intensivos em habilidades cognitivas (16.227.793 observações, referentes a 5.727.045 trabalhadores formais e 960.586 estabelecimentos), Trabalhadores intensivos em habilidades sociais (79.373.026 observações, referentes a 25.738.055 trabalhadores formais e 3.208.045 estabelecimentos) e Trabalhadores intensivos em habilidades manuais (72.808.032 observações, referentes a 23.505.302 trabalhadores formais e 2.293.037 estabelecimentos).

Convém ainda mencionar, que esta pesquisa é um excelente aparato para analisar o mercado de trabalho formal nos Arranjos Populacionais brasileiros, uma vez que não gerou uma amostra aleatória na primeira etapa deste estudo. Posteriormente, a fim de permitir viabilidade econométrica para a aplicação metodológica da decomposição salarial dos efeitos de especialização de habilidades, foram geradas amostras de 1% aleatorizadas e estratificadas por Arranjos Populacionais, para cada categoria de habilidades.

O Quadro 3.1 descreve as variáveis utilizadas nas estratégias empíricas deste ensaio. A variável que permite captar o prêmio salarial da especialização ocupacional categorizada por habilidades é o logaritmo do salário-hora real. Inicialmente, obtém-se o salário-hora, a partir da divisão entre variável “remuneração média do trabalhador” e horas semanais, obtida a partir da variável “quantidade de horas contratadas”, ambas presentes na RAIS. O salário-hora é então deflacionado pelo Índice de Preços ao Consumidor – Amplo (IPCA), fornecido pelo IBGE para o período considerado, sendo estabelecido como período base desta pesquisa dezembro de 2018. Finalmente é obtido o logaritmo natural do salário-hora real.

A variável independente principal no modelo de prêmios salarial está dentro da categoria de variáveis de aglomeração (espaciais): a especialização ocupacional, que já foi explicada no início da seção. Além da variável de especialização, tem-se, ainda, como variáveis espaciais a densidade populacional e a área do Arranjo Populacional, utilizadas como variáveis independentes na análise de decomposição salarial. A densidade é obtida a partir da divisão entre a população de um arranjo e sua área, ambas as variáveis são disponibilizadas pelo IBGE em nível de municípios. Para a pesquisa foi feita a compatibilização para a unidade regional de Arranjos Populacionais.

As variáveis de controle captam características dos trabalhadores, das firmas e dos Arranjos Populacionais. No que se refere às características dos trabalhadores, tem-se a idade, a idade ao quadrado, a experiência, a experiência ao quadrado, o gênero e o grau de instrução. Todas são disponibilizadas pela RAIS. A variável idade é obtida em anos, onde considerou-se trabalhadores que tenham entre 18 e 65 anos, conforme já mencionado. A variável experiência é obtida a partir da variável "Tempo no emprego" dada em meses e a variável gênero é uma *dummy* que identifica o sexo do trabalhador. A variável grau de instrução é obtida em ciclos de estudo, a partir da variável "grau de instrução" presente na RAIS, que indica o grau de escolaridade mais alto do trabalhador. Para a inclusão dos ciclos de escolaridade na base, criou-se um conjunto de variáveis *dummies* para cada ciclo apresentado no Quadro 3.1.

Quadro 3. 1 - Descrição das Variáveis

Variável	Descrição	Fonte
Logaritmo natural do salário-hora real	Logaritmo da remuneração média do trabalhador deflacionada pelo Índice de Preços ao Consumidor – Amplo (IPCA)	RAIS e IBGE
Logaritmo natural do Grau de especialização ocupacional	Logaritmo da divisão entre o emprego ocupacional categorizado pela habilidade (h) no Arranjo Populacional (*a*) no período (t) e o emprego total no Arranjo Populacional (*a*) no período (t)	RAIS e Maciente (2013)
Idade	Idade em anos	RAIS
Idade ao quadrado	Idade em anos elevada ao quadrado	RAIS
Experiência	Tempo no emprego (em meses)	RAIS
Experiência ao quadrado	Tempo no emprego (em meses) elevado ao quadrado	RAIS
Gênero	Variável *dummy* que estabelece 1 para masculino e 0 para feminino (Categoria base – feminino)	RAIS
Grau de instrução	Conjunto de variáveis *dummies* que identifica o nível de escolaridade do trabalhador, a saber: Analfabeto, Fundamental Incompleto, Fundamental Completo, Médio Incompleto, Médio Completo, Superior Incompleto e Superior Completo (Categoria base – analfabeto)	RAIS
Setor de atividade econômica	Conjunto de variáveis *dummies* que identifica o setor de atividade econômica do trabalhador de acordo com a CNAE 2.0 (Categoria base – setor agropecuário)	RAIS
Tamanho do Estabelecimento	Conjunto de variáveis *dummies* que identifica o tamanho do estabelecimento a partir da quantidade de trabalhadores contratados, dado o seguinte intervalo: 1 a 19 – micro; 20 a 99 – pequeno; 100 a 499 – médio e mais de 500 – grande (Categoria base – micro)	RAIS
População de 1920, 1940 e 1991	Série histórica da população brasileira por Arranjo Populacional para os anos de 1920, 1940 e 1991	Censo Demográfico
Cursos de Graduação Presenciais em 1991 e 2010	Série histórica para a quantidade de cursos de graduação presenciais por Arranjo Populacional para os anos de 1991 e 2010	Censo da Educação Superior
Logaritmo da Densidade Populacional	Logaritmo da divisão entre a população do Arranjo Populacional (*a*) no período (t) pela área desse Arranjo Populacional (*a*) no período (t)	IBGE
Logaritmo da Área	Logaritmo da área do Arranjo Populacional (*a*) no período (t)	IBGE

(continua)

(conclusão)

Variável	Descrição	Fonte
Fronteiriço	Variável *dummy* que estabelece 1 para Arranjo Populacional Fronteiriço e 0 caso contrário	IBGE (2016)
Turístico	Variável *dummy* que estabelece 1 para Arranjo Populacional Turístico e 0 caso contrário	IBGE (2016)
Macrorregião	Conjunto de variáveis *dummies* que identifica a macrorregião que o Arranjo Populacional (*a*) pertence, a saber: Norte, Nordeste, Sudeste, Sul e Centro-Oeste (Categoria base – Sudeste)	IBGE (2016)

Fonte: Elaboração própria (2021)

No que se refere as características das firmas, tem-se as características do setor da atividade econômica e do tamanho do estabelecimento, disponibilizados pela RAIS. O setor de atividade ao qual o estabelecimento pertence é obtido a partir da CNAE 2.0, a partir da qual criou-se um conjunto de variáveis *dummies* que identifica o setor de atividade econômica, correspondente. A variável tamanho do estabelecimento é classificada de acordo com a quantidade de trabalhadores contratados dos estabelecimentos, conforme classificação apresentada no Quadro 3.1. A partir dessa classificação criou-se um conjunto de variáveis *dummies*.

No que se refere às características dos Arranjos Populacionais tem-se a classificação entre Arranjo Populacional Fronteiriço e Turístico, além, da determinação de qual macrorregião o Arranjo Populacional pertence. Esses dados são disponibilizados pelo IBGE (2016). Os Arranjos Fronteiriço são aqueles que estão na fronteira internacional do Brasil (obedecendo o critério de contiguidade espacial). Para identifica-los na base de dados, cria-se uma variável *dummy* indicando aqueles que estão na fronteira. Os Arranjos Populacionais Turísticos são aqueles que tem localização litorânea e suas principais atividades são turísticas. Para identifica-los na base de dados, cria-se uma variável *dummy* indicando aqueles que são turísticos A macrorregião a qual pertence os Arranjos Populacionais é definida na base a partir de um conjunto de variáveis *dummies* de macrorregião.

No que se refere às variáveis instrumentais históricas, pode-se citar: População de 1920, 1940 e 1991 e Cursos de Graduação Presenciais em 1991 e 2010. Para as variáveis instrumentais históricas foi necessário realizar a equiparação dos dados do Censo Demográfico e do Censo da Educação Superior às Áreas Mínimas Comparáveis (AMCs). Essa estratégia se faz necessária, pois como são usadas variáveis defasadas de 1920, 1940 e 1991 podem existir incompatibilidades nas dimensões territoriais e unidades espaciais do Brasil.

Os dados do IBGE (2011) registram o rápido crescimento do número de municípios brasileiros e de sua população residente: em 1920 o número de municípios era aproximadamente 1.302, passando para 1.574, em 1940, e chegando a 4.491, em 1991. Em 2010, o Brasil computava 5.565 municípios (IBGE, 2011). Verifica-se que, ao longo do tempo, não apenas novos municípios foram formados, mas Estados Federais. Essa dinâmica deve-se a separações, disputas políticas e anexação. Desse modo, Erhl (2017) argumenta que estudos temporais de painéis de dados que incluam variáveis regionais merecem atenção, sendo necessário encontrar unidades espaciais consistentes no tempo. Esse cenário motivou a agregação de municípios em áreas geográficas mais abrangentes, denominadas de Áreas Mínimas Comparáveis (AMCs).

Para o período de 1991-2000, o IPEA disponibiliza 4.267 AMCs. Essa pesquisa usa a agregação desenvolvida por Reis et. a. (2011). Assim, após obter as AMCs para os períodos indicados (1920, 1940, 1991), foi possível, então, fazer a compatibilização dos municípios em Arranjos Populacionais, unidade regional considerada nessa pesquisa. Dessa forma, foi possível obter a população brasileira por município e, posteriormente, por Arranjo Populacional, para os anos de 1920, 1940 e 1991 e a quantidade de cursos de graduação presenciais por município e, posteriormente, por Arranjo Populacional para os anos de 1991 e 2010.

3.4 ESPECIALIZAÇÃO OCUPACIONAL CATEGORIZADA POR HABILIDADES: O MERCADO DE TRABALHO FORMAL BRASILEIRO (2009-2018)

O mercado de trabalho formal brasileiro[25] apresentou, entre o período de 2009 e 2018, uma certa estabilidade, o que já era esperado, dado que as maiores oscilações no emprego no Brasil estão concentradas no mercado informal[26]. O número de trabalhadores formais considerados nesta pesquisa oscilou entre 29.277.891 (2018) e 36.944.478 (2014) e entre 2009 e 2018 houve uma variação de -3,22% no total de trabalhadores formais (RAIS, ME).

O Brasil apresenta um dinamismo urbano intenso e uma estrutura urbana altamente concentrada. A maior parte dos trabalhadores formais brasileiros tem os seus postos de trabalho localizados em Arranjos Populacionais. De acordo com os dados do IBGE, o Brasil tem um

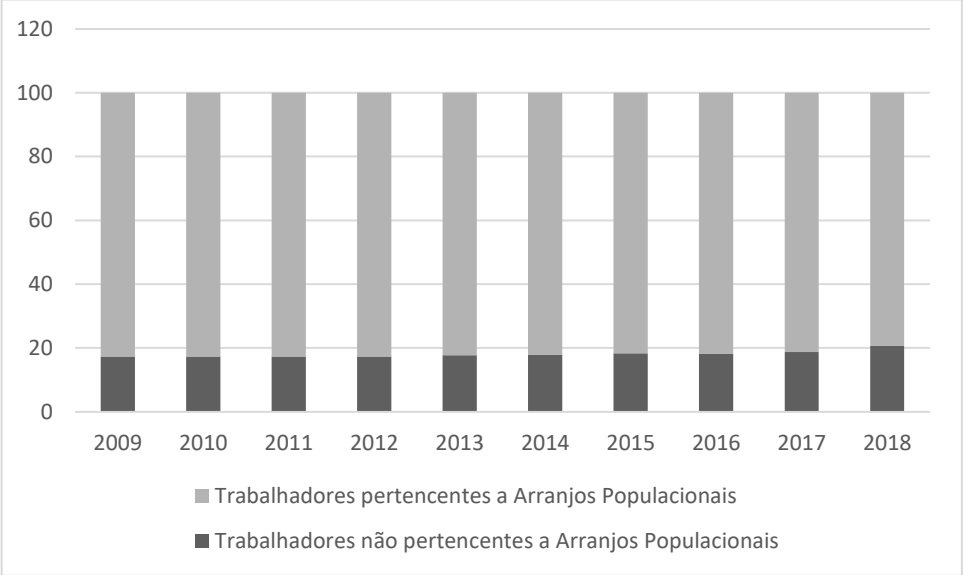
[25] Considera-se apenas os trabalhadores do setor privado com idades entre 18 e 65 anos, carga horária semanal maior do que 20 horas semanais e remuneração positiva acima do salário mínimo vigente em cada ano.

[26] É importante salientar que apesar da limitação da base da RAIS em analisar a dinâmica real do mercado de trabalho brasileiro, uma vez que considera apenas o mercado de trabalho formal, verifica-se que os dados considerados nessa pesquisa apresentam boa representatividade. Para o ano de 2018, por exemplo, os dados dessa pesquisa representam 31,57% da força de trabalho do Brasil, visto que, de acordo com os microdados da PNAD de 2018, tem-se um total de 92.736.067 trabalhadores na força de trabalho.

total de 294 Arranjos Populacionais formados por 953 municípios. Além disso, verifica-se que há 82 municípios isolados (médios e grandes) denominados de Concentrações Urbanas, totalizando 376 unidades regionais analisadas nessa pesquisa, que são chamadas genericamente de Arranjos Populacionais. Do total desses Arranjos Populacionais, 26 são classificados como grandes, 158 como médios e 192 como pequenos (IBGE, 2016). Em 2009, 82,7% dos trabalhadores formais brasileiros estavam alocados em Arranjos Populacionais. Percebe-se que, ao longo dos anos, essa participação manteve-se praticamente estável. Em 2018, esse percentual foi de 79,2% (FIGURA 3.1). Considerando os trabalhadores que estão alocados em Arranjos Populacionais, verifica-se, ainda, que a proporção de trabalhadores alocados em grandes Arranjos Populacionais é maior (67,5% em 2009 e 64,4% em 2018) (FIGURA 3.2).

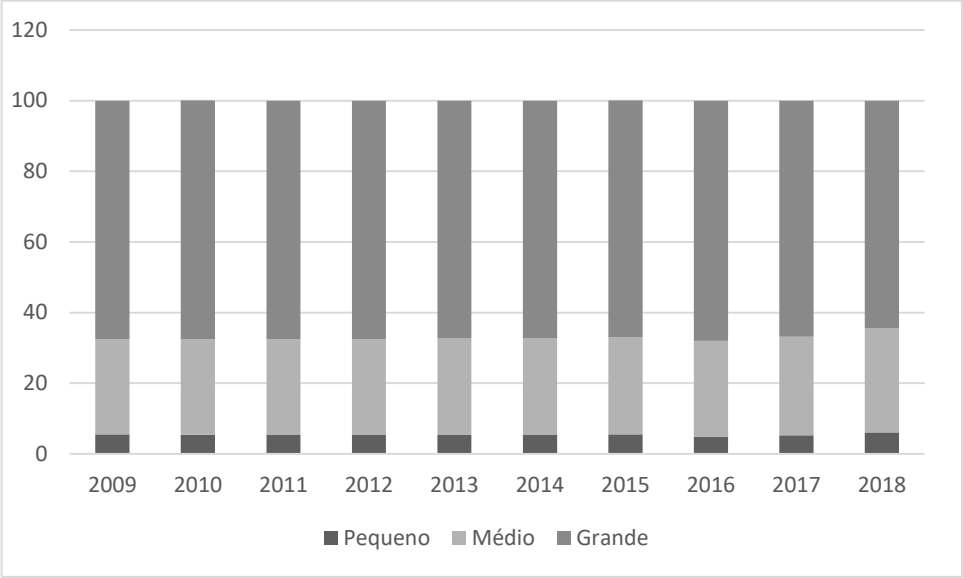
O mercado de trabalho formal brasileiro se destaca em habilidades sociais e manuais, em detrimento das habilidades cognitivas. No período de tempo analisado, apenas 9,6% dos trabalhadores formais que trabalham em Arranjos Populacionais são categorizados em ocupações intensivas em habilidades cognitivas. No que se refere aos trabalhadores cujas ocupações são intensivas em habilidades sociais e manuais esse percentual é de 47,1% e 43,2% (FIGURA 3.3).

Figura 3. 1 – Parcela dos trabalhadores formais alocados em Arranjos Populacionais



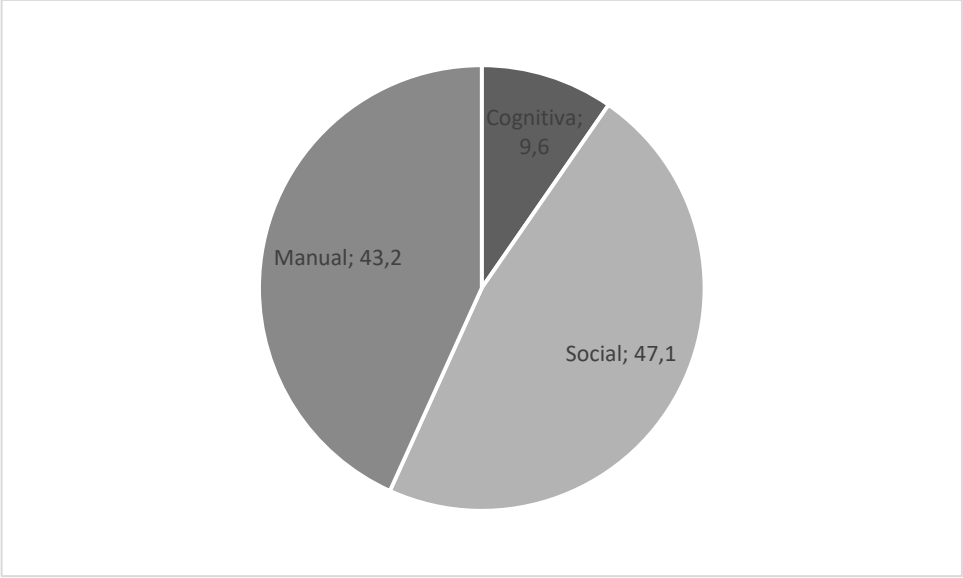
Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores do setor privado.

Figura 3. 2 – Composição dos trabalhadores formais por tamanho dos Arranjos Populacionais



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores do setor privado A classificação do tamanho dos Arranjos Populacionais brasileiros obedece aos critérios definidos pelo IBGE (2016): são considerados pequenos aqueles que tem uma população menor ou igual a 100 mil habitantes, médios aqueles que tem uma população entre 100 e 750 mil habitantes e grandes aqueles que tem uma população maior que 750 mil habitantes.

Figura 3. 3 – Composição dos trabalhadores formais por categorias de habilidades alocados em Arranjos Populacionais



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

Essa realidade é uma condição fundamental para definir a estrutura salarial brasileira. Os brasileiros recebem um salário-hora médio de 5,67 para o período compreendido entre 2009 e 2018. Quando se considera apenas os trabalhadores cujas ocupações são intensivas em habilidades cognitivas o salário-hora médio está muito acima da média nacional (11,93), o que não ocorre para os trabalhadores cujas ocupações são intensivas em habilidades sociais ou

manuais. Convém mencionar que os trabalhadores intensivos em habilidades manuais recebem a menor remuneração média (4,7), estando também atrás dos trabalhadores intensivos em habilidades sociais cuja remuneração média é 5,27. Como a participação dos trabalhadores cujas ocupações são intensivas em habilidades cognitivas é baixa, a média salarial brasileira declina (TABELA 3.1). Essa evidência está de acordo com os resultados de Neves Júnior (2018) que verificou que no Brasil, no período de 2003 a 2013, os trabalhadores cognitivos eram os que apresentavam maiores salários em detrimento dos trabalhadores sociais e manuais, esses últimos com a menor média salarial.

Verifica-se que os trabalhadores categorizados por habilidades apresentam muitas diferenças socioeconômicas e regionais. No Brasil, as regiões Sudeste e Nordeste concentram o maior número de trabalhadores, sendo assim caracterizadas como os maiores mercados de trabalho formais do país. No que se refere à distribuição regional dos trabalhadores, do total de trabalhadores formais do Brasil, 55,7% estão localizados na região Sudeste, sendo seguida pelas regiões Nordeste (16,4%), Sul (16%), Centro-Oeste (7,6%) e Norte (4,2%). Esse cenário é o mesmo quando se considera os grupos de trabalhadores isoladamente, ou seja, ao analisar a distribuição por características do emprego, com pequenas mudanças percentuais que não alteram o *ranking* de emprego entre as macrorregiões.

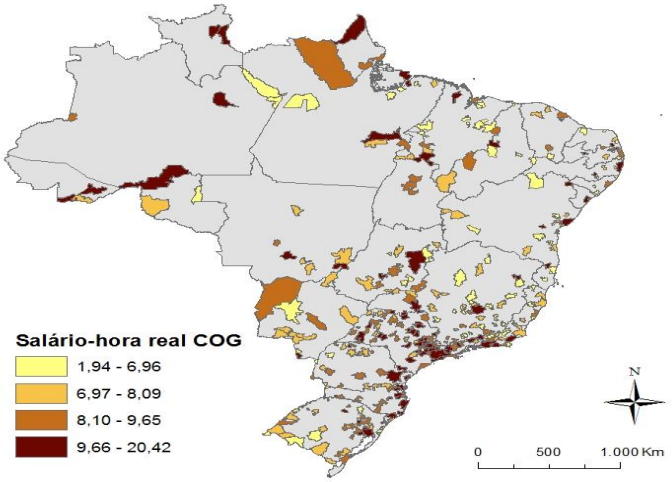
Verifica-se que dos trabalhadores alocados em Arranjos Populacionais a maioria deles estão localizados em grandes Arranjos Populacionais, conforme já mencionado (FIGURA 3.2). Essa composição não se altera quando se considera uma análise por grupo de habilidades. Todavia, é importante destacar que, para trabalhadores intensivos em habilidades cognitivas o percentual de trabalhadores localizados em grandes Arranjos Populacionais é 71,8%, um pouco superior à média nacional (67,4%) (TABELA 3.1). Uma análise mais precisa da estrutura salarial brasileira categorizada por habilidades requer uma abordagem espacial, considerando os Arranjos Populacionais. A Figura 3.4 apresenta a distribuição espacial do salário-hora real dos trabalhadores categorizados por habilidades para os Arranjos Populacionais.

A distribuição espacial do salário-hora real (médio) dos trabalhadores por categorias de habilidades mostra que, além dos trabalhadores intensivos em habilidades manuais apresentarem um menor salário médio por Arranjo Populacional, os maiores salários para esse perfil de trabalhadores se concentram nos Arranjos Populacionais pertencentes às regiões Sul, Sudeste e Centro-Oeste (FIGURA 3.4). A distribuição espacial do salário-hora real (médio)

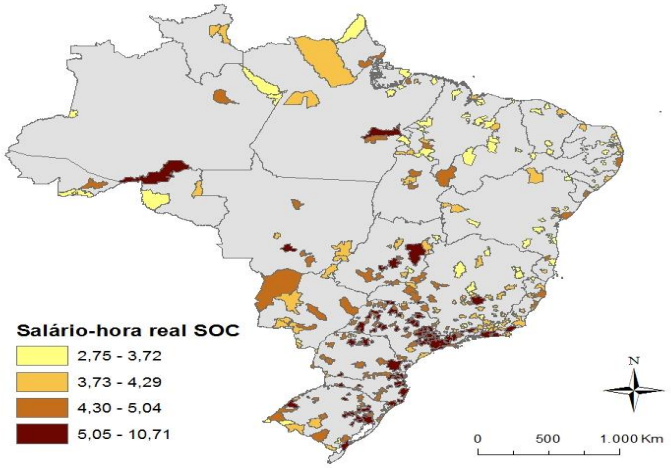
para os trabalhadores intensivos em habilidades sociais apresenta um comportamento semelhante, onde os maiores salários se concentram nos Arranjos Populacionais pertencentes às regiões Sul, Sudeste e Centro-Oeste. No que se refere aos trabalhadores intensivos em habilidades cognitivas, que possuem os maiores salários médios por Arranjo Populacional, pode-se indicar que a distribuição espacial do salário-hora real é mais uniforme do que as demais categoriais de habilidades, o que indica que esse perfil de trabalhadores tem melhores salários independente da região em que se encontram (FIGURA 3.4).

Figura 3. 4 – Distribuição espacial do salário-hora real por habilidades nos Arranjos Populacionais

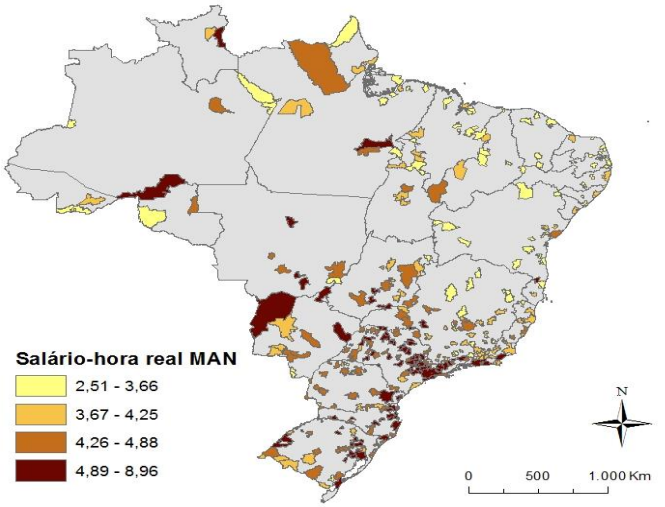
Painel 1: Habilidades Cognitivas



Painel 2: Habilidades Sociais



Painel 3: Habilidades Manuais



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)

No que se refere às características dos trabalhadores, percebe-se que a idade média dos trabalhadores do grupo de habilidade cognitiva (35,5) está bem próxima da média nacional (35,4), enquanto a idade média dos trabalhadores do grupo de habilidade social (34,4) está abaixo da média nacional e a idade média dos trabalhadores do grupo de habilidade manual (36,4) está acima (TABELA 3.1). Os trabalhadores do grupo de habilidades cognitivas são mais experientes no Brasil, o tempo médio no emprego para esses trabalhadores é de 48,9 meses, indicando menos volatilidade no mercado de trabalho para esses profissionais, já os trabalhadores do grupo de habilidades manuais são os que permanecem menos tempo no emprego (37,2), quando comparados com a média nacional (39,5). No que se refere ao sexo dos trabalhadores no Brasil, verifica-se que o sexo masculino tem maior participação (61,8%), sendo essa participação ainda mais intensificada quando se considera somente os trabalhadores do grupo de habilidades manuais (78,8%). Todavia, percebe-se que as mulheres são mais atuantes no grupo de habilidades cognitivas, cuja participação é 52,6%, e no grupo de habilidades sociais, cuja participação é 50,8%.

É fundamental examinar a composição dos grupos de trabalhadores categorizados por habilidades por nível de escolaridade, dado que nesta pesquisa a escolaridade não é vista como sinônimo de habilidades. Apesar da escolaridade representar uma credencial para ingressar em determinadas ocupações, ela não determina o rol de habilidades desempenhada pelo trabalhador em determinada ocupação. A análise do nível de escolaridade mostra um perfil mais favorável para os trabalhadores do grupo de habilidades cognitivas. Considerando o mercado de trabalho brasileiro total, verifica-se que a maioria dos trabalhadores tem ensino médio completo (51%). Ao analisar somente os trabalhadores do grupo de habilidades cognitivas verifica-se que a maior parte dos trabalhadores tem ensino superior completo (48%). Todavia, mesmo considerando trabalhadores intensivos em habilidades cognitivas, a participação de trabalhadores com ensino fundamental incompleto e completo é de 4,7%. Nos grupos de trabalhadores com habilidades manuais a participação de trabalhadores com ensino superior completo é muito baixa (1,1%), estando muito abaixo da média nacional (TABELA 3.1). Esse cenário indica que existe uma relação entre escolaridade e habilidades cognitivas, todavia, essa correspondência não é sempre satisfeita.

O emprego no Brasil destaca-se especialmente nos setores de Serviços e Comércio. Verifica-se que 48,3% dos trabalhadores brasileiros mantém vínculos empregatícios no setor de Serviços, sendo seguido pelo setor de Comércio (24,9%). Mesmo considerando uma análise decomposta

por grupo de habilidades dos trabalhadores, essa tendência se mantém. Todavia, a composição setorial sofre alterações significativas. Do total de trabalhadores com habilidades cognitivas, 71,1% destes estão alocados no setor de Serviços, já quando se considera o grupo de trabalhadores com habilidades manuais esse percentual diminui para 34,7%. Um setor que ganha destaque quando se considera trabalhadores com habilidades manuais é o da Transformação, cuja participação chega a 27,6%. Para os trabalhadores do grupo de habilidades sociais, destaca-se o setor de Serviços (56,1%) e Comércio (31,4%), cujas participações estão acima da média nacional. No Brasil, verifica-se que a maior parte dos trabalhadores estão empregadas em micro (44,5%) e pequenas (26,4%) empresas. No que se refere aos trabalhadores do grupo de habilidades cognitivas essa tendência se mantém, entretanto, a parcela de trabalhadores em grandes estabelecimentos aumenta quando comparada com a média nacional (TABELA 3.1).

Tabela 3. 1 - Características dos trabalhadores alocados em Arranjos Populacionais

Variável	Cognitiva	Social	Manual	Total
Características do trabalhador				
Salário-hora médio	11,93	5,27	4,7	5,67
Idade média	35,47	34,4	36,4	35,37
Tempo médio no emprego	48,87	39,78	37,22	39,55
Sexo				
Masculino	47,36%	49,2%	78,82%	61,83%
Feminino	52,64%	50,8%	21,18%	38,17%
Escolaridade				
Analfabeto	0%	0,13%	0,58%	0,31%
Fundamental Incompleto	1,55%	7,85%	21,2%	13,01%
Fundamental Completo	3,11%	11,2%	19,15%	13,86%
Médio Incompleto	2,47%	7,81%	10,47%	8,44%
Médio Completo	37,73%	57,83%	46,45%	50,97%
Superior Incompleto	7,09%	5,46%	1,05%	3,71%
Superior Completo	48,05%	9,73%	1,11%	9,69%
Características do emprego				
Setor				
Agropecuária	0,24%	0,22%	1,02%	0,57%
Extração Mineral	0,22%	0,15%	0,38%	0,26%
Transformação	8,69%	8,72%	27,63%	16,89%
Construção Civil	5,13%	3,37%	16,18%	9,08%
Comércio	14,55%	31,44%	20,06%	24,89%
Serviços	71,12%	56,06%	34,72%	48,28%
Serviços Domésticos	0%	0,01%	0,01%	0,01%
Organismos Internacionais	0,04%	0,02%	0,01%	0,02%
Tamanho do Estabelecimento				
Micro	32,93%	52,19%	38,61%	44,46%
Pequeno	27,13%	24,88%	27,92%	26,41%
Médio	20,02%	12,25%	18,46%	15,68%
Grande	19,93%	10,67%	15,02%	13,44%
Características regionais				
Macrorregião				
Norte	3,66%	4,18%	4,35%	4,2%
Nordeste	15,4%	17,01%	16,07%	16,45%
Sudeste	56,37%	56,13%	55,1%	55,71%
Sul	15,17%	14,91%	17,39%	16%
Centro-Oeste	9,39%	7,77%	7,09%	7,63%
Tamanho do Arranjo populacional				
Pequeno	3,92%	4,7%	6,27%	5,3%
Médio	24,3%	25,87%	29,63%	27,34%
Grande	71,77%	69,43%	64,11%	67,36%

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)

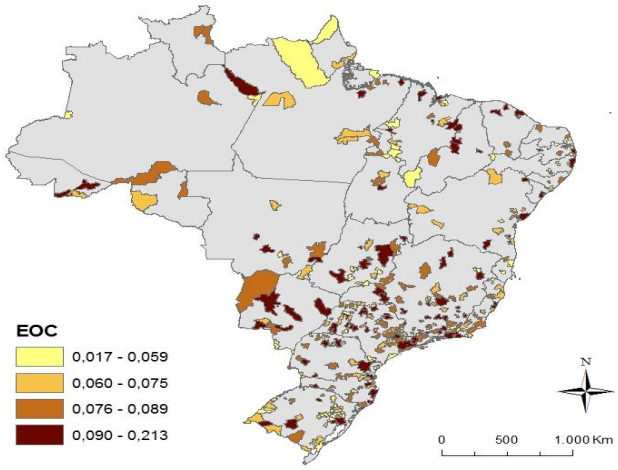
No Brasil, a especialização ocupacional categorizada por habilidades nos Arranjos Populacionais é maior para as habilidades manuais e sociais, em detrimento das habilidades cognitivas. A especialização manual e social no Brasil tem uma média de 0,4782 e 0,4475, respectivamente, enquanto a especialização cognitiva tem uma média de 0,0757. Essa tendência pode ser explicada pela pouca expressividade dos trabalhadores cognitivos nos Brasil (conforme verificado na FIGURA 3.3), indicando maior dispersão desses trabalhadores entre os Arranjos Populacionais brasileiros, impactando na determinação do índice de especialização ocupacional.

Já os trabalhadores cujas ocupações são intensivas em habilidades sociais têm a maior participação no mercado de trabalho formal brasileiro, nesse sentido verifica-se maior concentração desses trabalhadores no território nacional, impactando positivamente na determinação da especialização. Conforme verificado no Ensaio 1 dessa pesquisa (TABELA 2.6 e TABELA 2.8), a ocupação "Vendedor de comércio varejista" é a ocupação que apresenta maior concentração ocupacional no Brasil, o que corrobora as evidências aqui discutidas, uma vez que essa ocupação é categorizada por ser intensiva em habilidades sociais e é uma das mais típicas desse grupo de habilidade.

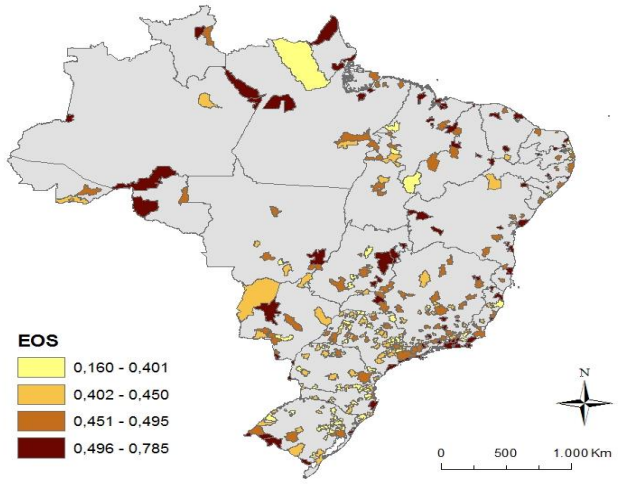
A Figura 3.5 apresenta a distribuição espacial da especialização ocupacional categorizada por habilidades para os Arranjos Populacionais. De modo geral, verifica-se que os índices de especialização cognitiva são bem menores quando comparados com os índices de especialização social e manual. Destaca-se que existe uma relação entre os Arranjos Populacionais especializados em habilidades sociais e manuais. Em sua maioria, os Arranjos Populacionais que têm menor especialização em habilidades sociais apresentam maior especialização em habilidades manuais e vice-versa. Dessa forma, há uma correlação negativa muito intensa entre especialização social e manual (QUADRO 3.2). Verifica-se que os maiores índices de especialização cognitiva estão agrupados nos Arranjos Populacionais pertencentes as regiões Sul e Sudeste, regiões com os maiores Arranjos Populacionais.

Figura 3. 5 – Distribuição espacial da especialização ocupacional categorizada por habilidades nos Arranjos Populacionais

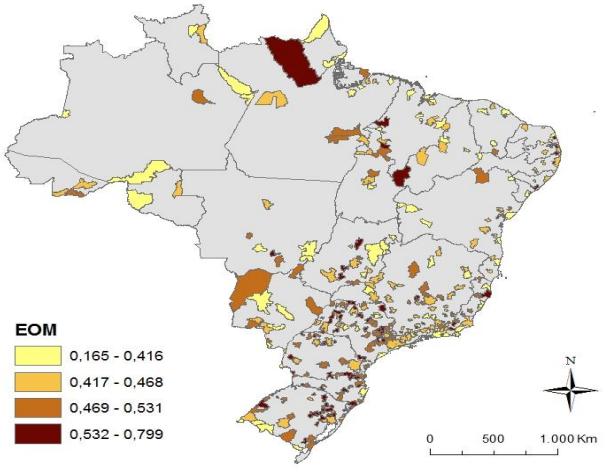
Painel 1: Habilidades Cognitivas



Painel 2: Habilidades Sociais



Painel 3: Habilidades Manuais



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)

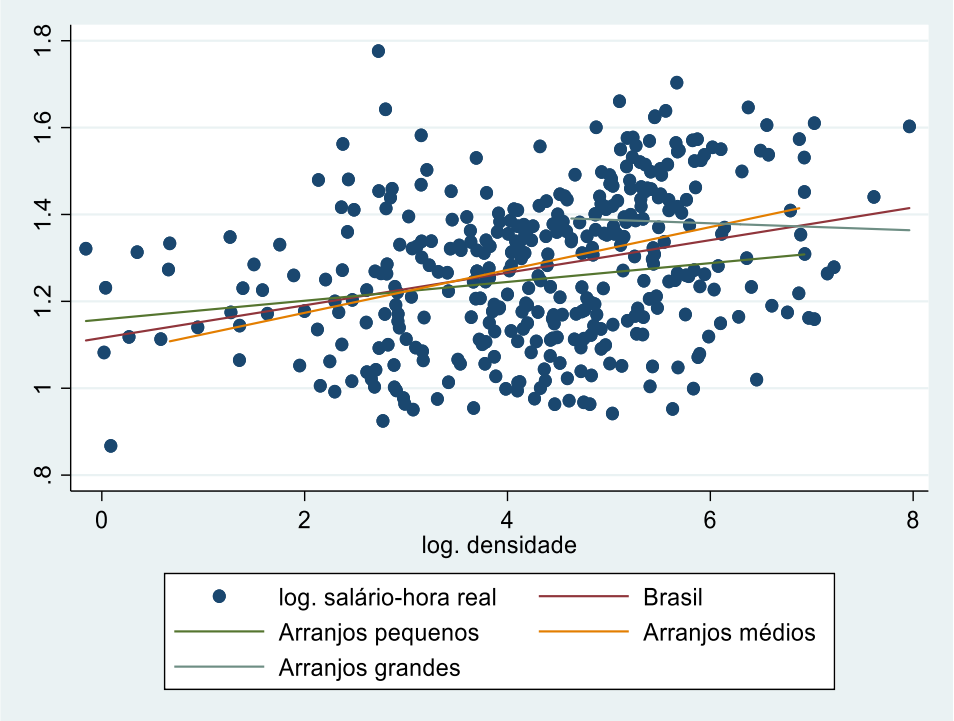
Quadro 3. 2 - Correlação entre as variáveis de especialização ocupacional categorizada por habilidades nos Arranjos Populacionais brasileiros

Correlação	EOC	EOS	EOM
EOC	1		
EOS	0,1975	1	
EOM	-0,4418	-0,9563	1

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)

A relação entre salário e densidade é positiva, conforme já mencionado na literatura. Essa evidência é mantida no Brasil quando se considera somente Arranjos Populacionais, cujas dinâmicas populacionais e urbanas são, de certa forma, semelhantes. No Brasil, a relação entre salário médio e densidade média nos Arranjos Populacionais é positiva no período de tempo considerado (2009-2018). A estimativa da elasticidade do salário-hora real em relação a densidade é de 0,0375%. Quando se considera a relação para Arranjos Populacionais pequenos (0,0216) e médios (0,0495) a relação se mantém positiva e significante, com maior elasticidade nos salários para Arranjos Populacionais médios. Todavia, quando se considera Arranjos grandes a relação é negativa (-0,008) e insignificante (FIGURA 3.6).

Figura 3. 6 – Salário-hora real e densidade nos Arranjos Populacionais (2009-2018)



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Salário-hora real deflacionado pelo IPCA. Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

Todavia, pode-se mencionar maior sensibilidade dos salários à especialização ocupacional categorizada por habilidades nos Arranjos Populacionais brasileiros. Além disso, destaca-se

que a relação positiva não é sempre verificada no Brasil, quando se considera especialização ocupacional categorizada por habilidades. Verifica-se que a especialização ocupacional de habilidades cognitivas e manuais tem uma relação positiva no salário-hora real do respectivo grupo de habilidade, a estimativa da elasticidade é de 0,1265 e 0,3587, respectivamente. Para as habilidades sociais foi observada uma relação negativa entre salário-hora real dos trabalhadores intensivos em habilidades sociais e a especialização social, a estimativa da elasticidade foi de -0,3273 (FIGURA 3.7).

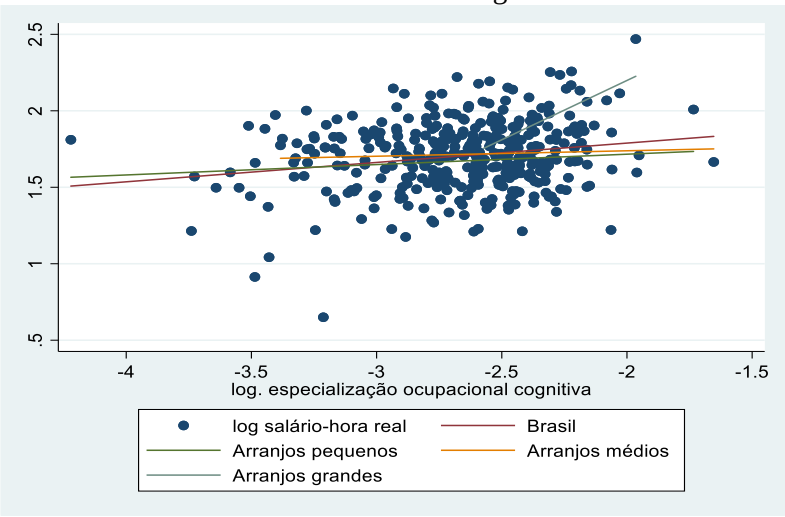
Uma análise mais minuciosa por tamanhos dos Arranjos Populacionais indica que a sensibilidade dos salários à especialização ocupacional cognitiva torna-se insignificante para Arranjos Populacionais pequenos e médios. Todavia, quando se considera Arranjos Populacionais grandes a relação é positiva (0,7759) e bem superior à média nacional, sugerindo que a especialização cognitiva provoca efeitos de externalidades maiores nos grandes Arranjos (FIGURA 3.7)

No que se refere a sensibilidade dos salários à especialização ocupacional manual, verifica-se que, ao considerar a relação para Arranjos Populacionais pequenos (0,3352) e médios (0,5866) a relação se mantém positiva e significante, mostrando uma relação crescente dado o tamanho dos centros. Todavia, quando se considera os Arranjos Populacionais grandes a relação é insignificante, esse fato pode sugerir que a especialização manual não é bem aproveitada nos grandes Arranjos (FIGURA 3.7)

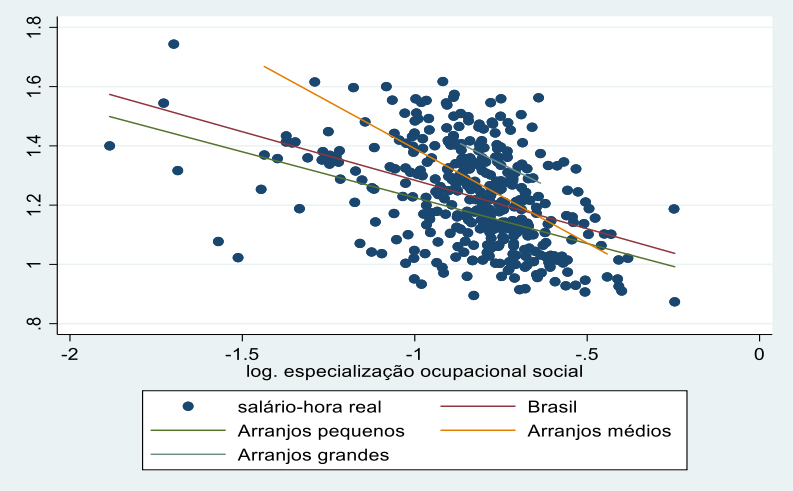
Considerando a sensibilidade dos salários à especialização ocupacional social, que em média apresentou uma relação negativa para o Brasil, observa-se que ao considerar a relação para Arranjos Populacionais pequenos (-0,3093) e médios (-0,6363) a relação se mantém negativa e significante, indicando que quanto maior o tamanho do Arranjo Populacional maior o impacto sobre os salários, exceto quando se considera os Arranjos Populacionais grandes, cuja relação torna-se insignificante (FIGURA 3.7)

Figura 3. 7 – Salário-hora real e especialização ocupacional categorizada por habilidades nos Arranjos Populacionais (2009-2018)

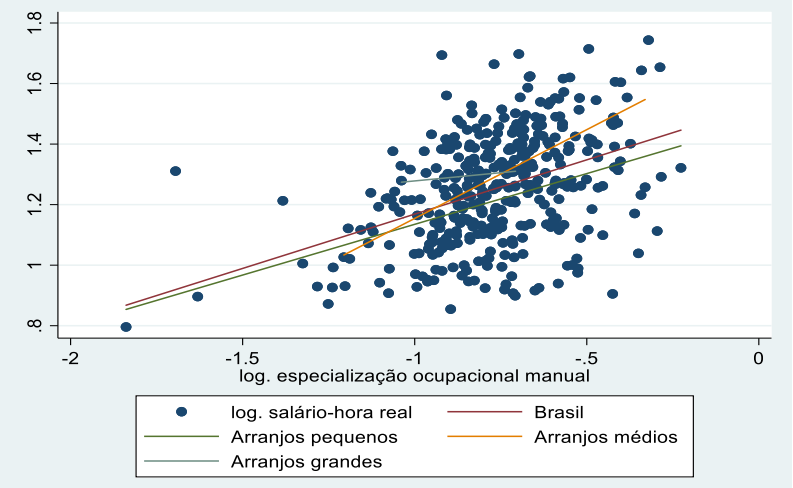
Painel 1: Habilidades Cognitivas



Painel 2: Habilidades Sociais



Painel 3: Habilidades Manuais



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Salário-hora real deflacionado pelo IPCA. Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

Para compreender a sensibilidade dos salários à especialização ocupacional média no Brasil, é importante mencionar que a especialização ocupacional categorizada por habilidades pode provocar efeitos no mercado de trabalho a partir de dois canais: competição e externalidades. No que se refere ao canal da competição, uma maior especialização local implica maior oferta de trabalhadores para determinada ocupação, implicando em uma diminuição dos salários. O canal da externalidade, por sua vez, gera um efeito positivo nos salários, pois quanto maior a especialização ocupacional maiores os efeitos do *learning*, possibilitando um maior rol de habilidades para esses trabalhadores. Nesse sentido, verifica-se que os dois canais têm efeitos opostos nos salários.

Considerando a elasticidade negativa entre salários e especialização social, compreende-se que o efeito competição supera o efeito de externalidade, o que não ocorre para a especialização cognitiva e manual. No que se refere à relação entre salários e especialização cognitiva é possível verificar que os efeitos de *learning* são intensos e a baixa oferta desses trabalhadores, que representam menos de 10% no mercado formal brasileiro (FIGURA 3.3), diminui o efeito de concorrência da especialização. Os trabalhadores categorizados em habilidades sociais são os maiores no Brasil e apresentam uma especialização média elevada, além disso, estão entre as ocupações mais concentradas em especial no comércio e serviços, conforme já mencionado, o que explicaria o efeito concorrência elevado. E, neste caso, os efeitos positivos da externalidade não superariam os efeitos negativos da competição.

Entretanto, para o caso dos trabalhadores intensivos em habilidades manuais, que também representam uma parcela considerável dos trabalhadores brasileiros e apresentam uma elevada especialização média, percebe-se uma elasticidade positiva. Isso pode ser explicado pelo fato do Brasil possuir uma estrutura produtiva de baixa complexidade. Dessa forma, a concorrência para trabalhadores intensivos em habilidades manuais é menor, isso faz com que esses trabalhadores sejam melhor absorvidos pelo mercado de trabalho. Além disso, considerando que o perfil dos trabalhadores sociais e manuais são parecidos, em especial no que se refere a escolaridade, onde a maioria dos trabalhadores para ambas as habilidades tem ensino médio completo (TABELA 3.1), os trabalhadores sociais e manuais podem ser, em algum grau, substitutos para algumas ocupações. Desse modo, espera-se que a concorrência no mercado manual seja menor dada a atratividade das ocupações no mercado social, uma vez que tem um maior salário médio e exige menos força física. Esse efeito concorrência menor para as

habilidades manuais, em comparação com as habilidades sociais, explicaria o efeito negativo da elasticidade entre salário e especialização social.

Na Economia Regional e Urbana difere-se economias de urbanização e economias de localização. Enquanto economias de urbanização é o agrupamento de estabelecimentos com funções de produção independentes e, por sua vez, trabalhadores com perfis diferentes, podendo ser representadas pela densidade do emprego, as economias de localização é o agrupamento de estabelecimentos ou trabalhadores com o mesmo perfil, indústria ou ocupação e pode ser representada por índices de concentração ou especialização. Essa pesquisa centra-se na análise de economias de localização, todavia, economia de localização e urbanização são importantes para explicar o mercado de trabalho e a determinação de salários. Dessa forma, é importante avaliar a relação entre densidade e especialização ocupacional categorizada por habilidades, isto é, se uma maior urbanização (densidade) explicaria uma maior especialização.

Os dados sugerem que no Brasil, no período considerado, a densidade está pouco relacionada com a especialização. Convém ainda mencionar que a densidade não é necessariamente uma condição *a priori* à especialização. É comum verificar que a partir da especialização industrial surge as condições propícias para o aumento da urbanização. Essa é uma visão comum dos teóricos da localização, onde uma indústria motriz intensificaria a atividade econômica regional provocando uma maior aglomeração e, por sua vez, uma maior densidade populacional (PERROUX, 1955; HIRSCHMAN, 1977). Ehn (2009) ainda ressalta que existe uma relação entre especialização, tecnologia e densidade populacional, entretanto essa relação ainda não é clara, indicando a necessidade de mais estudos teóricos e empíricos para compreender a relação. A codependência entre as variáveis especialização e densidade ainda é questionada por alguns autores. A depender das circunstâncias regionais e institucionais e do período analisado, Ehn (2009) indica que a densidade pode levar a uma maior especialização, todavia, em algumas situações a especialização pode levar a uma maior densidade, ou ainda, ressalta, a existência de um ciclo dinâmico entre as variáveis.

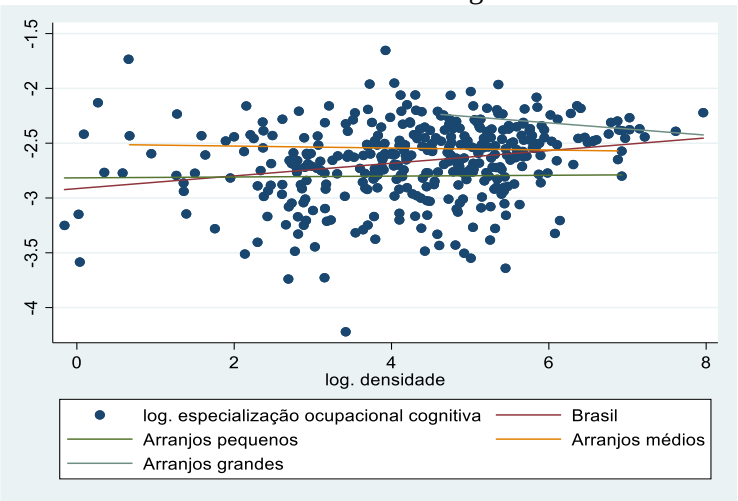
No mercado de trabalho brasileiro não é possível captar uma relação clara entre densidade e especialização ocupacional categorizada por habilidades, a partir da correlação entre as variáveis. Apenas a estimativa da elasticidade da densidade sobre a especialização cognitiva foi significante, todavia, muito baixa, em torno de 0,0578. A relação, porém, é insignificante quando avaliada por tamanhos dos Arranjos Populacionais. Além disso, tanto para a

especialização social como para a especialização manual as estimativas das elasticidades foram insignificantes (FIGURA 3.8).

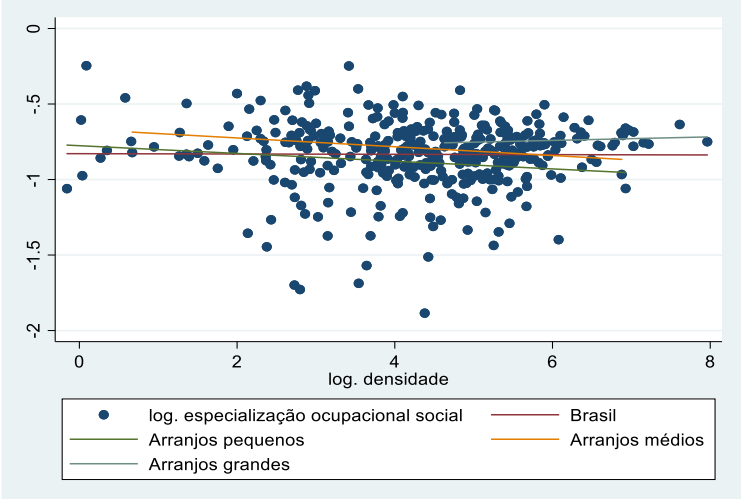
Convém ressaltar que outras características e variáveis ocupacionais do mercado de trabalho podem impactar nas relações expostas nas Figuras 3.7 e 3.8, essas são análises brutas que servem apenas como evidências empíricas iniciais. Diante disso, se faz necessário considerar e controlar essas características e variáveis a fim de obter resultados mais robustos. Essa estratégia empírica será realizada na próxima seção.

Figura 3. 8 – Especialização ocupacional categorizada por habilidades e densidade nos Arranjos Populacionais (2009-2018)

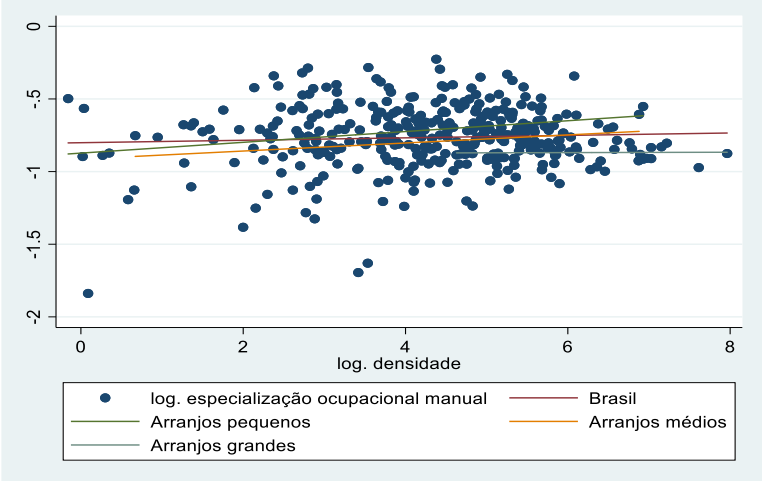
Painel 1: Habilidades Cognitivas



Painel 2: Habilidades Sociais



Painel 3: Habilidades Manuais



Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

3.5 RESULTADOS

Essa seção apresenta os resultados obtidos a partir das estratégias econométricas traçadas para obter evidências acerca do prêmio da especialização ocupacional categorizada por habilidades no Brasil. Considerando a estratégia de identificação fundamental deste estudo, serão apresentados os resultados dos modelos que lidam com a presença das endogeneidades existentes por meio de variáveis instrumentais e do controle das heterogeneidades não observáveis dos trabalhadores e firmas. Posteriormente, avalia-se a decomposição dos prêmios de especialização por habilidades.

3.5.1 Prêmio salarial da especialização ocupacional categorizada por habilidades

A estratégia de identificação desse estudo requer mobilidade dos trabalhadores entre firmas e Arranjos Populacionais e a mobilidade das firmas entre os Arranjos Populacionais ao longo do painel. A Tabela 3.2 mostra a mobilidade dos trabalhadores entre firmas e dos trabalhadores e firmas entre os Arranjos Populacionais. Observa-se que 8,1% dos trabalhadores com habilidades manuais mudam de Arranjos Populacionais, essa mobilidade é menor para os trabalhadores com habilidades cognitivas (4,4%) e habilidades sociais (5,2%). No que se refere à mobilidade dos trabalhadores entre as firmas, essa é bem mais expressiva, 31,23% dos trabalhadores com habilidades manuais mudam de firma, esse percentual é um pouco menor quando se compara com os trabalhadores de habilidades sociais (28,1%) e habilidades cognitivas (17,1%).

De modo geral, verifica-se que os trabalhadores intensivos em habilidades cognitivas têm os menores incentivos a mudar de emprego, os altos níveis de experiência (meses no emprego) já eram indícios desse fato. Esse cenário pode ser explicado pelos altos salários para esses grupos de trabalhadores, diminuindo os incentivos econômicos para mudar de emprego. A mobilidade das firmas entre Arranjos populacionais é menor do que a mobilidade dos indivíduos, fato que já era esperado, dado que a decisão de localização das empresas implica consideravelmente nos seus lucros e depende do tipo de atividade desempenhada. Considerando a base de dados dos trabalhadores intensivos em habilidades sociais, do total de firmas consideradas, 0,81% mudam de Arranjo Populacional. A mobilidade das firmas entre os Arranjos Populacionais também acontece para as bases com trabalhadores intensivos em habilidades cognitivas e manuais, mas com um menor percentual, 0,78% e 0,71%, respectivamente. Desse modo, percebe-se que as condições de identificação são satisfeitas para esse trabalho.

Tabela 3. 2 - Mobilidade de trabalhadores e firmas

Variável	Cognitiva	Social	Manual
	Observações	**Observações**	**Observações**
Mobilidade do trabalhador (i)			
Trabalhador muda de Arranjo Populacional	253.734	1.333.770	1.910.900
	4,43%	5,18%	8,13%
Trabalhador muda de firma	977.465	7.238.992	7.348.094
	17,07%	28,13%	31,26%
Total de trabalhadores	5.727.045	25.738.055	23.505.302
	100%	100%	100%
Mobilidade da firma (j)			
Firma muda de Arranjo Populacional	7.467	25.894	16.244
	0,78%	0,81%	0,71%
Total de firmas	960.586	3.208.045	2.293.037
	100%	100%	100%
Combinações de trabalhador e firma	6.961.754	25.665.735	25.256.703
	100%	100%	100%
Total de observações	16.227.793	72.808.032	79.373.026
	100%	100%	100%

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se apenas os trabalhadores em Arranjos Populacionais e do setor privado.

3.5.1.1 Prêmio salarial da especialização e aglomeração por habilidades: Modelo de variáveis instrumentais

No que se refere as variáveis instrumentais utilizadas nessa seção, variáveis históricas e *Bartik* ocupacional, é fundamental estabelecer que se considera como modelo mais adequado, o *benchmark*, aquele que corrige a endogeneidade a partir das variáveis históricas. Os modelos que utilizam como instrumento o *Bartik* ocupacional constituem-se testes de robustez. Isso porque o modelo com variáveis instrumentais históricas, além de permitir captar a dinâmica real do mercado de trabalho que acontece tanto devido a choques de oferta como choques de demanda, tem um maior número de instrumentos. Esse fato, torna o modelo sobreidentificado, que tem uma maior validade na literatura econométrica.

A Tabela 3.3 apresenta as estimativas do prêmio salarial da especialização ocupacional categorizada por habilidade no Brasil para o período de análise (2009-2018). Cada coluna da Tabela 3.3 representa uma regressão econométrica com a variável independente principal, a especialização ocupacional, que intitula a coluna (Cognitiva, Social e Manual), e as variáveis de controle, como: sexo, idade, idade ao quadrado, tempo no emprego, tempo no emprego ao quadrado, *dummies* de escolaridade, *dummies* de tamanho do estabelecimento, *dummies* de

setor de atividade econômica e *dummies* de ano. Para fins de ilustração só são inseridos na Tabela 3.3 os prêmios salariais da especialização ocupacional categorizada por habilidades. As Tabelas 3.4, 3.5 e 3.6, apresentam a mesma estrutura[27]. Os modelos completos com todas as estimativas das variáveis de controle estão no Apêndice B.

Considerando o modelo de *Pooled* OLS, os prêmios salariais da especialização das habilidades cognitiva, social e manual foram estatisticamente significantes ao nível de 1% e os modelos apresentaram um bom ajuste global, com um valor R^2 Ajustado de 0,5853, 0,6160 e 0,6362, respectivamente. Convém ainda mencionar que a maioria das estimativas das variáveis de controle foram estatisticamente significantes (TABELAS do APÊNDICE B).

O prêmio salarial da especialização de habilidades cognitivas e manuais, em nível de Arranjos Populacionais brasileiros, foi positivo. A especialização em habilidades cognitivas proporciona o maior prêmio salarial. Os resultados indicam que o aumento da especialização cognitiva em 1% provoca um aumento dos salários-hora reais dos trabalhadores intensivos em habilidades cognitivas em 0,37%, já para os trabalhadores intensivos em habilidades manuais esse aumento é de 0,13%. No que se refere às habilidades sociais, tal como na análise da elasticidade entre salário-hora e especialização média por Arranjos Populacionais apresentada na seção anterior (FIGURA 3.7), observou-se que o prêmio salarial da especialização social é negativo, indicando que um aumento de 1% na especialização social provoca uma diminuição de 0,21% no salário-hora real. Esse prêmio negativo pode ser explicado pelo efeito competição superar o efeito de externalidade no mercado de trabalhadores intensivos em habilidades sociais, pressionando os salários para baixo, conforme já mencionado.

[27] Nos modelos que consideram estimativas através de painel de dados de Efeitos Fixos não se considerou como variáveis de controle sexo e idade, pois são variáveis fixas ou que possuem variação constante no tempo.

Tabela 3. 3 - Prêmio salarial da especialização ocupacional por habilidades entre os grupos de trabalhadores (Cognitiva, Social e Manual): POLS e 2SLS

Variável Dependente: Logaritmo dos salários reais por Habilidade			
	***Pooled* OLS**		
Variável	**Cognitiva**	**Social**	**Manual**
$ln(EO_h)$	0,37042***	-0,20805***	0,13268***
	(0,00142)	(0,00087)	(0,00062)
Observações (NxT)	16.227.793	79.373.026	72.808.032
R^2 Ajustado	0,5853	0,6160	0,6362
Prob>F	0,0000	0,0000	0,0000
2SLS (IV Históricas) – 2º estágio			
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	1,01305***	1,0277***	-0,45785***
	(0,002749)	(0,003136)	(0,00156)
Observações (NxT)	15.003.790	72.837.470	65.503.892
R^2 Ajustado	0.5759	0,5851	0,6192
Prob>F	0.0000	0,0000	0,0000
2SLS (IV Históricas) – 1º estágio			
Instrumentos:			
log. da população de 1920	-0,0138***	0,04126***	-0,0402***
	(0,00015)	(0,00005)	(0,00005)
log. da população de 1940	-0,0421***	-0,0423***	0,05411***
	(0,00016)	(0,00005)	(0,00007)
log. da população de 1991	-0,0585***	0,02786***	-0,023***
	(0,0004)	(0,00011)	(0,00013)
log. do NCGP de 1991	0,0774***	-0,0078***	-0,0113***
	(0,00029)	(0,00009)	(0,00011)
log. do NCGP de 2010	0,101***	0,0013***	-0,0195***
	(0,00043)	(0,00013)	(0,00015)
2SLS (IV Bartik) – 2º estágio			
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	2,40958***	-0,09791***	0,1755***
	(0,02693)	(0,00316)	(0,00172)
Observações (NxT)	16.186.500	79.323.717	71.596.428
R^2 Ajustado	0,3745	0,6157	0,6346
Prob>F	0,0000	0,0000	0,0000
2SLS (IV Bartik) – 1º estágio			
Instrumentos:			
log. bartik	0,03712***	0,1962***	-0,0886***
	(0,00031)	(0,00027)	(0,00007)

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Entretanto, ao se considerar o controle da endogeneidade a partir das variáveis instrumentais históricas, modelo 2SLS (IV Históricas), verifica-se uma alteração na magnitude e/ou sinal do prêmio salarial para os trabalhadores intensivos em habilidades cognitivas, sociais e manuais, permanecendo significativos ao nível de 1%. As variáveis instrumentais foram todas significantes ao nível de 1%, para os três tipos de habilidades (TABELA 3.3). Essas variáveis

controlam a endogeneidade da escolha local, captando choques locais que atraem ou expulsam trabalhadores de determinadas ocupações e Arranjos Populacionais. Essa dinâmica pode estar subestimando ou superestimando os salários locais e, por sua vez, impactando no prêmio salarial da especialização ocupacional.

Os resultados do modelo 2SLS (IV Históricas), quando comparado ao modelo base (*Pooled* OLS), mostram um aumento do prêmio salarial para os trabalhadores intensivos em habilidades cognitivas e sociais, tornando o prêmio salarial para habilidades sociais positivo, e uma diminuição para os trabalhadores intensivos em habilidades manuais, tornando negativo o prêmio salarial para as habilidades manuais. Apenas o prêmio salarial para a especialização em habilidades cognitivas permaneceu com o mesmo sinal, porém verifica-se um considerável aumento desse prêmio salarial, quando comparado com o modelo base *Pooled* OLS, esse prêmio salarial é de 1,01305. No que se refere ao prêmio salarial das habilidades sociais e manuais verifica-se uma mudança de sinal, o prêmio salarial da especialização social é de 1,0277, enquanto o prêmio da especialização manual apresentou uma diminuição, apresentando uma estimativa de -0,45785.

Convém mencionar que, as variáveis instrumentais captam a mobilidade dos trabalhadores entre os Arranjos Populacionais sem considerar a distinção dos trabalhadores por categorias de habilidades. Dessa forma, a atração dos Arranjos Populacionais vai definir as condições iniciais de oferta e demanda nos mesmos. A indicação é que Arranjos Populacionais que tenham um salário médio maior receberão um maior número de trabalhadores independentemente das habilidades em que são mais intensivas. As descritivas mostraram que os trabalhadores em habilidades cognitivas têm menor mobilidade. No entanto, os Arranjos Populacionais intensivos em habilidades cognitivas, por exemplo, recebem trabalhadores mais habilidosos em todas as esferas, assim arranjos que são intensivos em habilidades cognitivas e sociais (que tem um salário médio maior) receberão um maior número de trabalhadores impactando na oferta local.

De modo geral, verifica-se que os movimentos populacionais e produtivos brasileiros ao longo das décadas passadas parece ter favorecido a especialização ocupacional dos trabalhadores cognitivos e sociais em detrimento dos trabalhadores manuais. É importante mencionar que, a relação dos fluxos populacionais passados não mostra uma tendência clara na determinação da especialização ocupacional categorizada por habilidades, entretanto, quando se considera o

número de cursos de graduação é visível uma relação. Para trabalhadores em habilidades cognitivas o número de cursos impacta positivamente na especialização, para as habilidades sociais o efeito é muito baixo mas vem crescendo, em um período anterior era negativo e no período seguinte tornou-se positivo, já para as habilidades manuais esse efeito é sempre negativo (TABELA 3.3). Isto é, uma maior oferta de educação nos Arranjos Populacionais atrai mais os trabalhadores em habilidades cognitivas e sociais, aumentando a especialização e, por sua vez, os salários.

Verifica-se que o prêmio salarial da especialização cognitiva e social estavam subestimados no modelo *Pooled* OLS, isso pode ser explicado pelo efeito de externalidade. Isso se deve por que o fluxo mais intenso de trabalhadores habilidosos nos Arranjos Populacionais possibilita aumentar a especialização. Assim, quanto mais trabalhadores habilidosos nesses Arranjos Populacionais maiores são os efeitos de *sharing, matching* e *learning* sobre os salários, aumentando-os. Entretanto, para os trabalhadores intensivos em habilidades manuais, os fluxos populacionais fazem com que os efeitos da externalidade sejam menores, ou seja, os efeitos de *sharing, matching* e *learning* sobre os salários não são intensificados. O resultado indica que o prêmio salarial da especialização manual estaria superestimado no *Pooled* OLS.

Essa evidência é corroborada pelo trabalho de Bacolod, Blum e Strange (2009) que indicam que a aglomeração não intensificaria o aumento dos salários dos trabalhadores com habilidades físicas (manuais). Os autores indicam que, diferente do que acontece para os trabalhadores com habilidades cognitivas e interpessoais (sociais), para os trabalhadores com habilidades físicas o processo de aprendizagem não ocorre de forma natural. Isto é, apesar de ser possível verificar como os trabalhadores intensivos em habilidades físicas podem ser beneficiados pelo *sharing* e *matching*, as habilidades físicas podem não ser facilmente beneficiadas pelo *learning.*

Considerando os dados para os trabalhadores das cidades dos EUA, no período entre 2011 e 2015, e separando os trabalhadores em dois grupos de habilidades, cognitivas não rotineiras e habilidades não cognitivas, Rossi-Hansberg, Sarte e Schwartzman (2019) verificaram que o prêmio salarial da especialização de habilidades para os dois grupos de trabalhadores era positivo. Todavia, com maior efeito para as habilidades cognitivas não rotineiras. Após o controle da endogenenidade da escolha local, através de variáveis instrumentais como a população do século anterior, a migração do século anterior e a presença de áreas destinadas a faculdades, os autores observaram que as estimativas estavam subestimadas. As estimativas

para habilidades cognitivas não rotineiras aumentaram, de 0,885 para 1,265, e a mesma tendência foi observada para as estimativas não cognitivas, cujo coeficiente era de 0,714 e passou para 0,938. Os autores também evidenciaram os efeitos de externalidades para esse aumento.

Convém mencionar que, a estratégia empírica proposta neste ensaio busca avaliar a especialização das habilidades com um maior nível de desagregação, representado pelas atividades não cognitivas, sociais e manuais. Essa desagregação permite uma análise da especialização com maior rigor, considerando as peculiaridades e diferenças entre as classificações de habilidades sociais e manuais, conforme já mencionado. No Brasil, verificou-se que a endogeneidade da escolha local também é importante na determinação do prêmio salarial da especialização categorizada por habilidades. E que as estimativas são, em sua maioria, subestimadas, com exceção para a especialização manual, cuja estimativa foi superestimada. Essa distinção se dá devido à própria categorização das habilidades nessa pesquisa, que permite analisar mais detalhadamente os efeitos de competição e externalidade para cada classificação mais desagregada de habilidades não cognitivas.

Além disso, a dinâmica populacional passada e a estrutura do mercado trabalho nos EUA e no Brasil são bastante diferentes, devendo destacar que a estrutura salarial nos EUA é menos desigual, com tendências de aumento nos salários nas últimas décadas e o avanço tecnológico intenso que modificam a estrutura de demanda por trabalho (ACEMOGLU; AUTOR, 2011). Conforme visto no Modelo Ricardiano do Mercado de Trabalho de Acemoglu e Autor (2011) um maior avanço tecnológico permite aumentar os salários para qualquer grupo de habilidades dos trabalhadores. Assim, considerado altas tecnologias nos EUA, também voltadas para ocupações de habilidades não cognitivas (manuais), as externalidades são mais intensas, aumentando os salários dada a aglomeração e a especialização/concentração ocupacionais, realidade diferente no Brasil. Essa tendência talvez explique o fato das estimativas de prêmio da especialização manual no Brasil estarem superestimadas no modelo base (*Pooled* OLS).

Como teste de robustez, obteve-se o modelo 2SLS (IV Bartik), também exposto na Tabela 3.4. Os resultados indicam um bom ajuste do modelo, com significância de 1% para o instrumento Bartik em todas as especificações. Verificou-se que, as estimativas do prêmio salarial da especialização (cognitiva, social e manual) permaneceram estatisticamente significantes ao nível de 1% e com o mesmo sinal, quando comparadas com o modelo base (*Pooled* OLS). Após

o controle da endogeneidade observou-se um aumento nos coeficientes estimados, com maior destaque para o prêmio salarial da especialização cognitiva. O prêmio salarial é positivo para a especialização cognitiva e manual, 2,40958 e 0,1755, respectivamente. No que se refere a especialização social obteve-se um coeficiente negativo de -0,09791 (TABELA 3.3).

Percebe-se um efeito positivo da externalidade na especialização categorizada por habilidades, todavia, não foi suficiente para permitir um prêmio salarial positivo para as habilidades sociais. O instrumento Bartik capta como choques de demanda podem interferir na estrutura ocupacional dos Arranjos Populacionais brasileiros, considerando como o crescimento do emprego local provoca mudanças nessas variáveis econômicas no curto prazo e como essas mudanças se manifestam no longo prazo. Os choques de demanda captados pelo instrumento Bartik indicam que esses choques foram positivos para a composição de habilidades dos Arranjos Populacionais, provocando um efeito positivo sobre os salários da especialização ocupacional em relação ao *Pooled* OLS, aumentando o prêmio salarial da especialização ocupacional categorizadas pelas habilidades.

Ehrl e Monasterio (2020) corrigiram o problema da endogeneidade da escolha locacional na determinação dos salários, dada a concentração espacial de habilidades no Brasil, através do uso do instrumento Bartik. Utilizando a base de dados da RAIS para o período de 2007 e 2013 e classificando cinco dimensões de habilidades, os autores verificaram que as estimativas iniciais estavam subestimadas, tal como foi observado para essa pesquisa. Os autores atribuíram esse aumento a erros de medição. Após a correção da endogeneidade observou-se que a concentração de habilidades analíticas, interpessoais e cognitiva de rotina apresentavam um efeito positivo sobre os salários, enquanto as habilidades manuais de rotina e de habilidades manuais não rotineiras apresentaram efeito negativo.

É interessante reforçar que, o instrumento de Bartik *shift* busca isolar os choques de demanda dos choques de oferta no mercado de trabalho, o que não ocorre com as variáveis instrumentais históricas. Nesse sentido, uma vez que o Bartik isola os efeitos de demanda é comum observar alterações mais brandas, todavia, significantes, nos coeficientes estimados. Verificou-se um aumento dos coeficientes em todas as categorias de habilidade, mas a relação entre salário e especialização não muda em comparação ao modelo base (*Pooled* OLS). Já para as variáveis históricas que captam alterações de oferta, mas não de forma isolada, percebe-se maior volatilidade, inclusive mudança de sinal. Além disso, as variáveis históricas possuem longas

defasagens temporais, enquanto o Bartik capta choques de demanda comparando o crescimento regional atual e a composição ocupacional do passado, uma vez que a variável representa o emprego atual estimado no Arranjo Populacional, dado o crescimento com relação ao ano base de 1991, de modo a preservar a exogeneidade da variável. Diante disso, o efeito da variável Bartik capta choques de demanda mais contemporâneos, provocando menores alterações, quando comparado com as variáveis históricas.

A especialização não somente implica em melhor produtividade para o grupo de trabalhador especializado em determinada habilidade, mas no conjunto geral de trabalhadores no mercado de trabalho. Observou-se que a especialização cognitiva sempre provoca um aumento nos salários de todos os trabalhadores, antes e após o controle da endogeneidade da escolha local pelos instrumentos históricos e o instrumento Bartik (TABELA B.7 do APÊNDICE B). Para a especialização em habilidades sociais há um efeito positivo sobre os salários gerais no modelo 2SLS (IV Históricas), já no que se refere à especialização manual o efeito positivo é visto no modelo *Pooled* OLS e no modelo 2SLS (IV Bartik). Verifica-se que, o comportamento das elasticidades entre a especialização e salários gerais, frente às estratégias de controle da endogeneidade, nos dois modelos de variáveis instrumentais, segue aquele que foi visto para a elasticidade entre especialização e salários por habilidades (TABELA 3.3). Esse resultado ressalta a importância da especialização na produtividade geral da economia, provocando um maior bem-estar econômico, em especial quando se considera a especialização cognitiva.

É importante mencionar que, o controle da endogeneidade da escolha local, que resulta da existência de trabalhadores habilidosos em direção a determinados Arranjos Populacionais, não provoca, necessariamente, uma redução nos prêmios salariais estimados. É comum acontecer essa redução quando se considera o prêmio salarial da escolaridade ou da ocupação, ou até mesmo de habilidades locais, por exemplo, pois os prêmios salariais poderiam estar sendo superestimados dado o efeito da habilidade, isto é, o maior número de trabalhadores habilidosos em determinados locais estaria inflando os salários locais. Essa redução não acontece com o prêmio salarial da especialização nessa pesquisa, pois quanto maior o número de trabalhadores habilidosos maiores os efeitos da externalidade (*sharing*, *learning* e *matching*), uma vez que o principal canal para se avaliar os efeitos da especialização é pelo efeito externalidade.

No entanto, ao analisar o prêmio salarial da escolaridade e do setor de atividade (variáveis de controle) pode-se verificar uma redução. As *dummies* de escolaridade e setor de atividade, após

o controle da endogeneidade local, tanto no modelo 2SLS (IV Históricas) quanto no modelo 2SLS (IV Bartik), para as três categorias de habilidades, apresentaram, em geral, uma diminuição nas estimativas, com maiores destaques quando se considera o grupo de trabalhadores intensivos em habilidades cognitivas (TABELAS B.1 a B.3 do APÊNDICE B). Esse resultado corrobora a ideia de que a endogeneidade da escolha local tem um efeito diferente quando se considera uma análise mais apurada da especialização de habilidades.

Como exercício empírico adicional, que permite complementar a análise da dinâmica do mercado de trabalho, obtém-se o prêmio da aglomeração por habilidades nos Arranjos Populacionais brasileiros. A Tabela 3.4 mostra o prêmio salarial da aglomeração por categorias de habilidades cognitiva, social e manual. A aglomeração nesta pesquisa é mensurada a partir do número de trabalhadores nos Arranjos Populacionais classificados por habilidades. De modo geral, verifica-se um que aglomeração tem efeito positivo para todas as classificações de habilidade, com maior destaque para a aglomeração dos trabalhadores intensivos em habilidades cognitivas. Considerando o modelo *Pooled* OLS, o prêmio da aglomeração cognitiva é de 0,0621, sendo seguido pelo prêmio da especialização social (0,0329) e o prêmio da especialização manual (0,0189), onde todas as estimativas foram significantes ao nível de 1%. Todavia, é fundamental mencionar que os efeitos da especialização são mais fortes do que o da aglomeração, conforme visto anteriormente na Tabela 3.3.

Ao controlar a endogeneidade da escolha local através de variáveis instrumentais históricas, observou-se um bom ajuste do modelo 2SLS (IV Históricas), com todas as variáveis instrumentais significantes ao nível de 1%. Os resultados mostram pouca alteração nos prêmios salarias da aglomeração categorizadas por habilidades, houve um pequeno aumento dos prêmios salariais estimados (TABELA 3.4). Esse aumento é um indício de que a externalidade impacta positivamente no prêmio da aglomeração, todavia, com menor intensidade quando comparado ao prêmio da especialização.

Tabela 3. 4 - Prêmio salarial da aglomeração por habilidades entre os grupos de trabalhadores (Cognitiva, Social e Manual): POLS e 2SLS

Variável Dependente: Logaritmo dos salários reais por Habilidade			
	***Pooled* OLS**		
Variável	**Cognitiva**	**Social**	**Manual**
$ln(L_h)$	0,0621***	0,03295***	0,01891***
	(0,00016)	(0,000052)	(0,000047)
Observações (NxT)	16.227.793	79.373.026	72.808.032
R^2 Ajustado	0,5930	0,6228	0,6387
Prob>F	0,0000	0,0000	0,0000
2SLS (IV Históricas) – 2º estágio			
Variável	Cognitiva	Social	Manual
$ln(L_h)$	0,06494***	0,03424***	0,01895***
	(0,00018)	(0,00006)	(0,00005)
Observações (NxT)	15.003.790	72.837.470	65.503.892
R^2 Ajustado	0,5913	0,6206	0,6356
Prob>F	0,0000	0,0000	0,0000
2SLS (IV Históricas) – 1º estágio			
Instrumentos:			
log. da população de 1920	-0,06303***	-0,05828***	-0,05539***
	(0,00026)	(0,00012)	(0,00014)
log. da população de 1940	0,03474***	0,0272***	0,0205***
	(0,00027)	(0,00014)	(0,00016)
log. da população de 1991	0,655***	0,6818***	0,6954***
	(0,00067)	(0,00031)	(0,00032)
log. do NCGP de 1991	-0,17731***	-0,18437***	-0,1753***
	(0,00062)	(0,00028)	(0,00029)
log do NCGP de 2010	0,5991***	0,58115***	0,55283***
	(0,00083)	(0,00037)	(0,00038)
2SLS (IV Bartik) – 2º estágio			
Variável	Cognitiva	Social	Manual
$ln(L_h)$	0,086***	-0,0044***	-0,01631***
	(0,00069)	(0,000)	(0,00016)
Observações (NxT)	16.186.500	79.323.717	71.596.428
R^2 Ajustado	0,5909	0,6128	0,6255
Prob>F	0,0000	0,0000	0,0000
2SLS (IV Bartik) – 1º estágio			
Instrumentos:			
Log. bartik	1,0402***	4,3625***	0,95397***
	(0,00288)	(0,0044)	(0,00109)

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Considerando o modelo 2SLS (IV Bartik), verificou-se que tanto os instrumentos como os prêmios salariais da aglomeração foram significantes em 1%. Os choques de demanda provocaram diminuições para o prêmio salarial da aglomeração dos trabalhadores intensivos em habilidades sociais e manuais, tornando ambos negativos, -0,0044 e -0,01631, respectivamente. Percebe-se que, os efeitos da demanda podem ter provocado uma redução nos

salários para os grupos menos qualificados, em especial quando se isola os efeitos da oferta. Já considerando os trabalhadores intensivos em habilidades cognitivas, os choques de demanda positivos provocam aumento no prêmio salarial da aglomeração (0,086), indicando uma posição mais favorável destes trabalhadores no mercado de trabalho.

No que se refere ao efeito da aglomeração sobre os salários gerais dos trabalhadores, se observa uma tendência semelhante ao efeito da aglomeração sobre os salários categorizados pelas habilidades dos trabalhadores. Uma maior aglomeração aumenta ou salários em 0,02898. Esse valor sofre pouca alteração quando se considera o controle da endogeneidade da escolha local a partir das variáveis históricas (0,03012), entretanto considerando os choques de demanda, verifica-se que a estimativa do modelo base (*Pooled* OLS) estava superestimada, provocando uma diminuição na estimativa do prêmio salarial da aglomeração (-0,027238) (TABELA B.8 do APÊNDICE B).

É importante mencionar que, enquanto os choques de demanda provocaram aumentos nos prêmios salariais da especialização categorizada por habilidades (TABELA 3.3), esses mesmos choques de demanda provocaram diminuições nos prêmios salariais das aglomerações (TABELA 3.4). Esse resultado se deve ao fato de que, apesar da especialização e da aglomeração captarem dinâmicas do mercado de trabalho, essas medidas têm propósitos diferentes. Enquanto a aglomeração é um indicador da organização dos mercados de trabalhos mais ou menos intensos e remete a medidas de diversificação das atividades produtivas, a especialização capta a estrutura produtiva das regiões, isto é, no que as regiões se destacam para gerar produção e qual a sua vantagem comparativa. Desse modo, é valido esperar que os choques de demanda sejam intrinsecamente relacionados com a especialização local, isto é, quanto mais especializado o local maior serão as políticas públicas ou as ações das empresas para aumentar essa especialização, gerando um ciclo vicioso. A aglomeração é uma medida mais geral e não tem uma relação tão clara com os choques de demanda.

Considerando que na dinâmica real do mercado de trabalho acontece tanto choques de oferta como choques de demanda, a presente pesquisa considera como modelo mais adequado, o *benchmark*, aquele que corrige a endogeneidade a partir das variáveis históricas que captam, além de choques de oferta, as amenidades e choques de demanda, conforme já mencionado. Essas variáveis possibilitam analisar melhor a dinâmica do mercado de trabalho, enquanto o instrumento de Bartik acaba isolando os choques de oferta. Além disso, é importante mencionar

que o modelo com variáveis instrumentais históricas tem um maior número de instrumentos, o que torna o modelo sobreidentificado, que tem uma maior validade na literatura econométrica. O instrumento de Bartik é interessante para verificar como o mercado de trabalho se comporta frente à choques específicos de demanda, todavia, a dinâmica do mercado de trabalho é muito mais complexa e envolve além de choques de demanda os choques de oferta.

3.5.1.2 Prêmio salarial da especialização e aglomeração por habilidades: Modelo de dados em painel

A estratégia empírica de painel de dados, com o controle da heterogeneidade não observada dos trabalhadores e firmas e do seu efeito combinado (*matching*), para a determinação do prêmio salarial da especialização ocupacional categorizada pelas habilidades pode ser visualizada na Tabela 3.5. De modo geral, os resultados mostram uma diminuição nas estimativas do prêmio salarial da especialização cognitiva, social e manual, mantendo-se significantes ao nível de 1% em todas as especificações (TABELA 3.5). Nesse sentido, a heterogeneidade não observada dos trabalhadores e firmas, bem como seu efeito par combinado e/ou *matching*, é importante na determinação dos prêmios salariais da especialização.

Considerando o modelo base (*Pooled* OLS), cujas estimativas já foram explicadas na Tabela 3.3, verifica-se uma diminuição no prêmio salarial da especialização pelo controle da heterogeneidade não observada dos trabalhadores para as habilidades cognitivas e manuais. O prêmio da especialização cognitiva manteve-se positivo, apesar da diminuição, e significante ao nível de 1% com um coeficiente de 0,0626, enquanto a diminuição do prêmio da especialização manual, provocou um diferencial salarial negativo de -0,00483, também significante ao nível de 1%. Todavia, o controle da heterogeneidade não observada dos trabalhadores provocou um aumento no prêmio da especialização social, tornando-o positivo (0,01967). O que sugere que sugere que o prêmio da especialização social está negativamente atrelado aos atributos não observáveis dos indivíduos e, após controlar esses atributos, torna-se positivo. Todavia, não se pode indicar que o prêmio salarial da especialização é anulado ao considerar esses atributos, ou seja, após controlar as heterogeneidades não observadas, os trabalhadores em habilidade sociais apresentaram um prêmio de especialização.

Tabela 3. 5 - Prêmio salarial da especialização ocupacional por habilidades entre os grupos de trabalhadores (Cognitiva, Social e Manual): Efeitos Fixos

Variável Dependente: Logaritmo dos salários reais por Habilidade			
		***Pooled* OLS**	
Variável	**Cognitiva**	**Social**	**Manual**
$ln(EO_h)$	0,37042***	-0,20805***	0,13268***
	(0,00142)	(0,00087)	(0,00062)
Observações (NxT)	16.227.793	79.373.026	72.808.032
R^2 Ajustado	0,5853	0,6160	0,6362
Prob>F	0,0000	0,0000	0,0000
	Efeito Fixo do trabalhador		
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	0,0626 ***	0,01967***	-0,00483***
	(0,00105)	(0,00078)	(0,00047)
Observações (NxT)	16.227.793	79.373.026	72.808.032
N. de grupos	5.727.045	25.738.055	23.505.302
R^2 (overall)	0,3182	0,4780	0,5271
R^2 (within)	0,8354	0,8689	0,8829
R^2 (between)	0,2782	0,3946	0,4348
Prob>F	0,0000	0,0000	0,0000
Hausman χ^2	2.313.767,24	6.533.692,75	2.530.729,89
Prob>χ^2	0,0000	0,0000	0,0000
	Efeito Fixo da firma		
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	-0,04402***	-0,07361***	0,04089***
	(0,00638)	(0,00288)	(0,00313)
Observações (NxT)	16.227.793	79.373.026	72.808.032
N. de grupos	960.586	3.208.045	2.293.037
R^2 (overall)	0,4655	0,5495	0,5528
R^2 (within)	0,5418	0,6521	0,6985
R^2 (between)	0,4608	0,5441	0,5862
Prob>F	0,0000	0,0000	0,0000
Hausman χ^2	88.187,37	253.687,30	121.874,00
Prob>χ^2	0,0000	0,0000	0,0000
	Efeito Fixo de trabalhador e firma		
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	0,00706***	-0,09359***	0,05219***
	(0,00073)	(0,001)	(0,00067)
Observações (NxT)	16.227.793	79.373.026	72.808.032
N. de grupos	6.961.754	25.665.735	25.256.703
R^2 (overall)	0,2659	0,5567	0,5528
R^2 (within)	0,8832	0,8181	0,8361
R^2 (between)	0,2469	0,4844	0,4719
Prob>F	0,0000	0,0000	0,0000
Hausman χ^2	2.199.490,44	1.501.570,15	671.495,12
Prob>χ^2	0,0000	0,0000	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

É importante mencionar que apesar da especialização ser categorizada pelas habilidades, essas são extraídas e classificadas a partir do rol de tarefas desempenhadas em cada ocupação, essas habilidades podem ser adquiridas pela escolaridade e experiências. Todavia, além desse rol de habilidades, os trabalhadores possuem habilidades inatas não observadas, como determinação, motivação, sagacidade, entre outras, que podem impactar na sua produtividade. Esse fato justifica a análise de heterogeneidade não observada dos trabalhadores em uma modelagem empírica de análise do prêmio da especialização de habilidades.

A diminuição observada para o prêmio da especialização cognitiva e manual deve-se ao fato de que a heterogeneidade não observada dos trabalhadores que capta as características não observadas, tais como habilidades inatas, podem impactar na sua produtividade. Assim, nesses casos os ganhos sobre os salários estariam superestimados. Já para a especialização social, verificou-se que as habilidades inatas dos trabalhadores estariam subestimando os salários dos trabalhadores intensivos em habilidades sociais, isso sugere que os atributos não observados dos trabalhadores nessa categoria estariam negativamente relacionados com o prêmio salarial da especialização. Uma explicação possível para esse resultado é que os trabalhadores em ocupações sociais possuem atributos não observados que não favorecem o aproveitamento das externalidades positivas da especialização.

Após o controle da heterogeneidade não observada das firmas, que capta a produtividade das empresas, todos os coeficientes são significativos ao nível de 1%. Além disso, os resultados mostram uma diminuição nesses coeficientes quando comparados com os do modelo base (*Pooled* OLS). O maior prêmio da especialização ocupacional, após o controle da heterogeneidade da firma, foi para as habilidades manuais (0,04089) e o menor para as sociais (-0,07361). No que se refere às habilidades cognitivas, essa diminuição no coeficiente foi tão intensa que tornou o prêmio negativo (-0,04402). Esse cenário se deve ao fato de que as firmas que empregam trabalhadores intensivos em habilidades cognitivas atuam em atividades mais intensivas nesse ramo e, em geral, apresentam elevada produtividade quando comparadas com as demais. Portanto, ao descontar o efeito associado à maior produtividade das firmas, o prêmio da especialização cognitiva desaparece.

Desse modo, verifica-se que, além dos trabalhadores intensivos em habilidades cognitivas serem os que detêm maiores habilidades inatas que interferem positivamente na sua produtividade, esses são empregados em firmas mais eficientes. A melhor correspondência de

emprego e qualificação, isto é, a possibilidade de firmas absorverem os melhores trabalhadores dado a sua especificidade técnica, é estabelecida como o *matching* no mercado de trabalho. Esse *matching* pode ser captado pelo modelo Efeito Fixo de trabalhador e firma. Os resultados indicam que o *matching* é importante na determinação dos prêmios salariais da especialização ocupacional categorizada por habilidades, pois todas as estimativas se mostram significantes ao nível de 1%. No entanto, é importante mencionar que o controle da heterogeneidade trabalhador-firma não é suficiente para eliminar o prêmio da especialização, mesmo após o controle os diferenciais salariais permanecem. Indicando que de fato existe um prêmio para a especialização ocupacional categorizada por habilidades para o Brasil, com exceção do grupo de habilidades sociais que já não apresentavam prêmio no modelo base (Pooled OLS).

Todos os prêmios salariais sofreram uma diminuição na sua magnitude quando comparados ao modelo base (*Pooled* OLS). Esses resultados indicam que as principais fontes de ganhos de especialização podem ser atribuídas às características não observados dos trabalhadores e firmas (*matching*). Destaca-se que, a maior diminuição foi no prêmio da especialização cognitiva que passa a ser 0,00706, indicando que a correspondência trabalhador-ocupação é mais intensa para habilidades cognitivas, refletindo um bom ajuste entre a oferta e demanda de trabalho para habilidades cognitivas.

Esse resultado, corrobora as evidências obtidas por Silva (2019) que buscou observar a probabilidade de determinação de *matching* para trabalhadores com ensino superior nos grandes centros urbanos. O autor observou que a probabilidade de *matching* aumenta com a maior especialização/concentração de trabalhadores mais habilidosos. Considerando que os trabalhadores intensivos em habilidades cognitivas no Brasil têm, em sua maioria, ensino superior completo (TABELA 3.1), essas evidências demonstram uma relação positiva entre prêmio salarial e *matching.* E, ao controlar os efeitos fixos das firmas e trabalhadores, o prêmio salarial da especialização das habilidades tende a diminuir, uma vez que estavam superestimados pelo efeito de *matching*, em especial para aqueles intensivos em habilidades cognitivas.

Neves Júnior (2018) e Silva (2017) também encontraram evidências de que a heterogeneidade não observada dos trabalhadores e firmas, bem como o seu efeito combinado, são fundamentais para explicar a determinação dos diferenciais salariais das habilidades nas cidades e os diferenciais salarias de localização no Brasil, respectivamente. Neves Júnior (2018) indica que

as grandes cidades atraem trabalhadores mais habilidosos e firmas mais produtivas, nesse sentido, o efeito combinado da heterogeneidade não observada leva a reduções na magnitude do prêmio salarial urbano das habilidades.

Mion e Naticchioni (2009), ao analisar os efeitos de variáveis espaciais, como densidade, potencial de mercado e especialização industrial, sobre os salários, verificaram que as estimativas dos efeitos caem acentuadamente quando se considera a heterogeneidade de firmas e trabalhadores. Os autores utilizam como estratégia empírica o modelo desenvolvido por Abowd e outros (1999) e identificam que o coeficiente de densidade e especialização diminuem e deixam de ser significantes, enquanto o coeficiente de potencial de mercado diminui para 2,17%, com um nível de significância de 1%. Os autores ainda estabelecem que o *sorting* não explica na sua totalidade as externalidades de mercado, uma vez que sua estimativa ainda foi positiva. O mesmo se observa para essa pesquisa, pois apesar da queda nas estimativas do prêmio da especialização ocupacional categorizada por habilidades, essas ainda se mantêm significantes e diferentes de zero, isto é, o diferencial salarial dado a especialização de habilidades permanece para a economia brasileira mesmo após o controle da heterogeneidade não observada das firmas e trabalhadores.

Conforme já mencionado, os efeitos da especialização sobre os salários específicos por grupo de habilidades são maiores que os efeitos da especialização sobre os salários gerais dos trabalhadores. Com o controle da heterogeneidade não observada dos trabalhadores, firmas e *matching* sobre o prêmio salarial da especialização (nos salários gerais) verifica-se um comportamento parecido com o observado para o prêmio da especialização sobre o respectivo grupo de habilidade. Em todas as especificações, as estimativas foram significantes ao nível de 1% e observou-se, para cada especificação, que a direção das mudanças nas estimativas foram as mesmas do modelo sobre os salários específicos, todavia, com uma menor sensibilidade. Nesse sentido, verifica-se que a heterogeneidade dos trabalhadores, firmas e o *matching* alteram a determinação do prêmio salarial da especialização sobre os salários gerais, todavia, com menor intensidade (TABELA B.9 do APÊNDICE B). Isto é, as habilidades inatas dos trabalhadores, a produtividade das firmas e a melhor correspondência entre trabalhador-firma interferem de forma mais forte em grupos específicos de análise, pois conseguem isolar melhor os efeitos e a direção dos mesmos nos salários específicos.

A Tabela 3.6 mostra o prêmio salarial da aglomeração por categorias de habilidades cognitiva, social e manual, considerando como critério para o controle da endogeneidade os efeitos fixos dos trabalhadores, firmas e o seu efeito combinado (*matching*). Como já mencionado, no modelo de *Pooled* OLS, aglomeração tem efeito positivo para todas as classificações de habilidade.

Ao controlar a heterogeneidade dos trabalhadores, verifica-se que os prêmios salariais da aglomeração diminuíram para todas as categorias de habilidades, mantendo-se ainda positivos e significativos ao nível de 1%. Quando se controla pela heterogeneidade não observada das firmas, apenas o prêmio da especialização social manteve-se significativo, todavia tornou-se negativo (-0,0024). Esses resultados indicam que a produtividade das firmas, isoladamente, apresenta forte contribuição para a determinação do prêmio da aglomeração categorizada por habilidades cognitivas e manuais, uma vez que o seu controle elimina a presença do prêmio. O efeito negativo para as habilidades sociais pode indicar que a aglomeração de trabalhadores nesse grupo está associada a um maior efeito competição, após controlar o prêmio associado às firmas.

Todavia, ao considerar o modelo Efeito Fixo de trabalhador e firma, que controla a heterogeneidade não observada da combinação de trabalhadores e firmas, o prêmio da aglomeração volta a ser significativo para todas as categorias de habilidades. Para as habilidades sociais e manuais esse prêmio é positivo e significativo ao nível de 1%, 0,02836 e 0,02028, respectivamente. Já para as habilidades cognitivas a estimativa é quase nula (-0,0007) e significante ao nível de 10%. Desse modo, verifica-se que quando comparado ao modelo de *Pooled* OLS, o *matching* provoca uma grande diminuição para o prêmio da aglomeração de habilidades cognitivas e sociais, em maior magnitude para as habilidades cognitivas, enquanto para as habilidades manuais há um pequeno aumento. Isso indica que o *matching* trabalhador-firma entre os trabalhadores categorizados como intensivos em habilidades manuais age negativamente sobre os salários ao se considerar os efeitos da aglomeração, uma vez que após o seu controle o prêmio torna-se maior.

Tabela 3. 6 - Prêmio salarial da aglomeração entre os grupos de trabalhadores (Cognitiva, Social e Manual): Efeitos Fixos

Variável Dependente: Logaritmo dos salários reais por Habilidade			
		***Pooled* OLS**	
Variável	**Cognitiva**	**Social**	**Manual**
$ln(L_h)$	0,0621***	0,03295****	0,01891***
	(0,00016)	(0,000052)	(0,000047)
Observações (NxT)	16.227.793	79.373.026	72.808.032
R^2 Ajustado	0,5930	0,6228	0,6387
Prob>F	0,0000	0,0000	0,0000
	Efeito Fixo do trabalhador		
Variável	Cognitiva	Social	Manual
$ln(L_h)$	0,02865***	0,0141***	0,01113***
	(0,00038)	(0,00012)	(0,00008)
Observações (NxT)	16.227.793	79.373.026	72.808.032
N. de grupos	5.727.045	25.738.055	23.505.302
R^2 (overall)	0,3324	0,4851	0,5305
R^2 (within)	0,8357	0,8690	0,8831
R^2 (between)	0,2934	0,4031	0,4389
Prob>F	0,0000	0,0000	0,0000
Hausman χ^2	2.193.869,09	6.394.107,99	2.431.664,70
Prob>χ^2	0,0000	0,0000	0,0000
	Efeito Fixo da firma		
Variável	Cognitiva	Social	Manual
$ln(L_h)$	0,00224	-0,002403**	0,00015
	(0,00218)	(0,00098)	(0,00094)
Observações (NxT)	16.227.793	79.373.026	72.808.032
N. de grupos	960.586	3.208.045	2.293.037
R^2 (overall)	0,4700	0,5472	0,5524
R^2 (within)	0,5417	0,6521	0,6984
R^2 (between)	0,4653	0,5401	0,5848
Prob>F	0,0000	0,0000	0,0000
Hausman χ^2	74.566,75	266.169,54	113.652,88
Prob>χ^2	0,0000	0,0000	0,0000
	Efeito Fixo de trabalhador e firma		
Variável	Cognitiva	Social	Manual
$ln(L_h)$	-0,0007*	0,02836 ***	0,02028***
	(0,00043)	(0,00009)	(0,00008)
Observações (NxT)	16.227.793	79.373.026	72.808.032
N. de grupos	6.961.754	25.665.735	25.256.703
R^2 (overall)	0,2649	0,5653	0,5562
R^2 (within)	0,8832	0,8200	0,8373
R^2 (between)	0,2459	0,4938	0,4751
Prob>F	0,0000	0.0000	0,0000
Hausman χ^2	2.315.308,44	1.483.176,49	651.190,24
Prob>χ^2	0,0000	0,0000	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Silva e Azzoni (2021) verifica como a heterogeneidade não observada pode impactar no prêmio da aglomeração dos Arranjos Populacionais no Brasil, isto é, nos efeitos de localização. A autora mostra que há grande interferência dos atributos não observados dos trabalhadores e firmas, todavia, menciona que os resultados sugerem maior contribuição dos atributos dos trabalhadores do que dos atributos das firmas. Tal evidência não foi observada para essa pesquisa, na avaliação do prêmio da aglomeração de habilidades através do controle da heterogeneidade não observada. No entanto, vale ressaltar que, a aglomeração nessa pesquisa é definida pelo número de trabalhadores em cada grupo de habilidades, diferentemente da aglomeração mensurada no estudo de Silva e Azzoni (2021) que considera o tamanho geral do mercado de trabalho. Considerando o prêmio da aglomeração nos salários gerais, a heterogeneidade não observada das firmas tem maior impacto (TABELA B.10 do APÊNDICE B).

No que se refere ao prêmio da aglomeração nos salários gerais verifica-se um prêmio de 0,02898 no modelo *Pooled* OLS, conforme já mencionado. Ao controlar pela heterogeneidade não observada dos trabalhadores e firmas isoladamente, os coeficientes mostram-se significantes ao nível de 1%, sofrendo uma diminuição considerável: passando para 0,01447 com o controle da heterogeneidade não observada dos trabalhadores, e para -0,002941 com o controle da heterogeneidade não observada das firmas. Nesse sentido, percebe-se que tanto as habilidades inatas dos trabalhadores como a produtividade e eficiência das firmas contribuem para aumentar os salários gerais locais. Vale mencionar o grande impacto da produtividade das firmas nos salários gerais, dado que o prêmio da aglomeração desaparece, se tornando negativo.

No entanto, ao considerar a modelagem mais completa que capta o *matching* no mercado de trabalho, verifica-se que o prêmio da aglomeração permanece quase o mesmo (uma leve diminuição) quando comparado com o modelo *Pooled* OLS, com um coeficiente de 0,02894, significante ao nível de 1% (Tabela B.10 do Apêndice B). Apesar de não haver alterações significativas no prêmio da aglomeração sobre os salários gerais, entre os modelos *Pooled*-OLS e aqueles que controlam a heterogeneidade da combinação de indivíduo e firma, a redução do prêmio no modelo que controla somente o efeito de firma é compensada quando se considera a melhor correspondência dos empregos, onde as firmas conseguem absorver os melhores trabalhadores.

De modo geral, verificou-se que em todos os modelos considerados a heterogeneidade não observada dos trabalhadores e firmas, bem como seu feito combinado, *matching*, são fundamentais para a determinação dos prêmios salariais da especialização ocupacional categorizada por habilidades e dos prêmios salariais da aglomeração. Na maioria dos testes empíricos, a ausência de controle para o *matching* no mercado de trabalho acaba superestimando os salários dos trabalhadores locais no Brasil.

Essa seção utilizou várias estratégias empíricas para o controle da endogeneidade do prêmio salarial da especialização ocupacional categorizada por habilidades, tanto variáveis instrumentais como painel de efeitos fixos. Nesse sentido, obteve-se uma rica análise dos efeitos da endogeneidade sobre a dinâmica do mercado de trabalho e, consequentemente, seus impactos no prêmio salarial da especialização categorizada por habilidades. A Tabela 3.7 resume os principais resultados encontrados na seção 3.5.1.

Tabela 3. 7 - Prêmio salarial da especialização ocupacional por habilidades entre os grupos de trabalhadores (Cognitiva, Social e Manual): Resumo

Variável Dependente: Logaritmo dos salários reais por Habilidade			
	Categorias de Habilidades		
Modelo	**Cognitiva**	**Social**	**Manual**
Pooled OLS	0,37042***	-0,20805***	0,13268***
2SLS (IV Históricas)	1,01305***	1,0277***	-0,45785***
Efeito Fixo do trabalhador	0,0626 ***	0,01967***	-0,00483***
Efeito Fixo da firma	-0,04402***	-0,07361***	0,04089***
Efeito Fixo de trabalhador e firma	0,00706***	-0,09359***	0,05219***

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Nota: * p < 0.10, ** p < 0.05, *** p < 0.01

3.5.2 Decomposição do prêmio salarial da especialização ocupacional categorizada por habilidades

Essa seção apresenta os resultados que decompõem os prêmios salariais da especialização ocupacional categorizada por habilidades. Todavia, antes analisar os resultados é importante mencionar que, devido a restrições computacionais a análise da decomposição do prêmio da especialização de habilidades foi feita considerando uma amostra aleatória de 1% das bases de dados da seção anterior. Com isso, se faz necessário analisar se as amostras obedecem às condições necessárias para a robustez do modelo de decomposição. Conforme já mencionado, a estratégia de identificação desse modelo requer mobilidade de trabalhadores entre firmas e Arranjos Populacionais e mobilidade de firmas entre Arranjos Populacionais ao longo do

painel. A Tabela 3.8 mostra a mobilidade dos trabalhadores entre firmas e a mobilidade das firmas e dos trabalhadores entre os Arranjos Populacionais para as amostras aleatórias obtidas.

Para os trabalhadores intensivos em habilidades cognitivas, a amostra conta com um total de 744.877 observações. Verifica-se que 7,12% dos trabalhadores mudaram de Arranjo Populacional, enquanto 27,82% mudaram de firma. As empresas que mudaram de Arranjo populacional representam um percentual de 0,65%. No que se refere aos trabalhadores intensivos em habilidades sociais, a amostra apresenta 3.830.690 observações, do total de trabalhadores na amostra 7,05% mudaram de Arranjo Populacional e 40,37% mudaram de firma. Considerando o total de firmas nessa amostra, 0,49% mudaram de Arranjo Populacional durante o período considerado. A amostra para os trabalhadores intensivos em habilidades manuais possui 3.536.201 observações, a mobilidade dos trabalhadores foi de 10,71% e 44,43% quando se considera Arranjo Populacional e firmas, respectivamente. No que se refere à mobilidade das firmas entre os Arranjos Populacionais para essa amostra, observa-se que 0,61% das firmas mudaram de local.

Tabela 3. 8 - Mobilidade de trabalhadores e firmas nas amostras

Variável	**Cognitiva**	**Social**	**Manual**
	Observações	**Observações**	**Observações**
Mobilidade do trabalhador (i)			
Trabalhador muda de Arranjo Populacional	11.342	54.870	76.459
	7,12%	7,05%	10,71%
Trabalhador muda de firma	44.302	314.078	317.064
	27,82%	40,37%	44,43%
Total de trabalhadores da amostra	159.291	777.960	713.566
	100%	100%	100%
Mobilidade da firma (j)			
Firma muda de Arranjo Populacional	689	3.134	2.913
	0,65%	0,49%	0,61%
Total de firmas da amostra	106.128	635.111	475.638
	100%	100%	100%
Combinações de trabalhador e firma	219.182	1.233.135	1.196.056
	100%	100%	100%
Total de observações da amostra	744.877	3.830.690	3.536.201
	100%	100%	100%

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Amostra aleatória de um 1% do painel inicial.

Verifica-se que o percentual de trabalhadores que mudaram de Arranjo Populacional e de firma nas amostras é maior que o observado na base total (TABELA 3.2), para todas as categorias de habilidades. No que se refere a mobilidade das empresas entre os Arranjos Populacionais, verifica-se um percentual ligeiramente menor, para todas as categorias de habilidades, quando comparado com a base total (TABELA 3.2). Desse modo, considera-se que as amostras são representativas e as condições de identificação são satisfeitas para as amostras obtidas.

A decomposição dos prêmios salariais da especialização ocupacional categorizada por habilidades é desenvolvida em dois estágios, seguindo a abordagem proposta em Combes, Duranton e Gobillon (2008) e o modelo AKM. Primeiro, se estima o prêmio salarial da especialização ocupacional categorizada por habilidades por Arranjo Populacional e ano, isto é, inclui-se o efeito fixo dos Arranjos Populacionais na análise, com as variáveis de controle já mencionadas. Nesse estágio, obtém-se o efeito da especialização sobre os salários nos Arranjos Populacionais brasileiros no período considerado (2009-2018). Em um segundo momento, a decomposição desses efeitos é feita, estimando o vetor do efeito da especialização sobre os salários contra as características observáveis dos Arranjos Populacionais. As variáveis de controle incluídas no segundo estágio seguem a abordagem de Silva (2017), baseada em Combes, Duranton e Gobillon (2008), que considera fundamental incluir variáveis da Economia Urbana na análise de decomposição de salários, como a densidade.

É importante frisar que, a densidade permite verificar a contribuição dos efeitos puros da urbanização sobre os prêmios salariais devidos à especialização de habilidades. Essa interpretação é de suma importância pois permite obter maiores evidências sobre a relação indissociável entre economias de localização (especialização/concentração) e economias de urbanização (densidade), dois ramos diferentes de economias de aglomeração que são analisadas, na maior parte dos estudos, separadamente na literatura econômica (EBERTS; MCMILLEN, 1999; GOLDMAN; KLIER; WALSTRUM, 2019; WHEATON; LEWIS, 2002). Desse modo, considerando que a estratégia empírica apresentada nessa pesquisa tem as habilidades (cognitivas, sociais e manuais) como fatores determinantes dessa especialização, os resultados dessa seção permitem compreender melhor a relação entre especialização e densidade e se essa relação pode ser alterada quando se considera os salários da especialização para cada tipo de habilidade.

A decomposição dos efeitos da especialização das habilidades sobre os salários é feita para cada grupo de habilidade separadamente: cognitiva (TABELA 3.9), social (TABELA 3.10) e manual (TABELA 3.11). Para cada grupo de habilidades foram estimadas quatro especificações para o primeiro estágio da decomposição: (1) *Pooled* de dados, (2) painel com o controle de efeitos fixos do trabalhador (*i*); (3) painel com o controle de efeitos fixos das firmas (*j*) e (4) painel com o controle de efeitos fixos do trabalhador e das firmas (*matching*) (*ij*). Cada especificação representa uma coluna que possibilita avaliar a decomposição dos efeitos da especialização sobre os salários brutos (especificação 1) e a decomposição dos efeitos da especialização sobre os salários líquidos (especificação 2, 3 e 4), após descontar as heterogeneidades não observadas.

Como já mencionado, variáveis que captam a aglomeração podem estar endogenamente determinadas. Como relatado para a variável de especialização ocupacional, o problema de causalidade reversa também está presente na relação entre salários locais e densidade. Desse modo, considerando que a endogeneidade pode gerar estimativas tendenciosas, para contornar o problema, variáveis instrumentais foram utilizadas para a determinação da relação entre densidade e efeitos da especialização de habilidades sobre os salários. Assim, o modelo de decomposição que, por natureza empírica, tem dois estágios, anteriormente mencionados, passa a apresentar três estágios nessa pesquisa. O primeiro estágio continua sendo a obtenção dos efeitos de especialização sobre o salário. Após essa etapa, foi feito o controle da endogeneidade da variável densidade a partir dos instrumentos históricos (população de 1940 e população de 1991)[28], obtendo o vetor de densidade corrigido (estágio intermediário de controle de endogenidade). Desse modo, pode-se então ir para o estágio final que foi a decomposição propriamente dita. As Tabelas 3.9, 3.10 e 3.11 mostram o estágio final de decomposição para cada grupo de habilidade após corrigida a endogeneidade.

De modo, geral, verifica-se nas Tabelas 3.9, 3.10 e 3.11 que os instrumentos foram significantes ao nível de 1%, justificando o modelo de variável instrumental para controle de endogeneidade. Enquanto a variável população de 1940 apresentou uma relação inversa com a densidade (em média -0,05), a variável população de 1991 apresentou uma relação positiva (em média 0,99). Poucas alterações na magnitude dos coeficientes e na sua significância foram observadas, quando se compara com o modelo que não controla a endogeneidade da densidade. Verifica-se

[28] A justificativa do uso desses instrumentos já foi feita na seção de metodologia.

que a maioria das estimativas estavam superestimadas. Os resultados sem o controle da endogeneidade estão nas Tabelas B.11, B.12 e B.13 do Apêndice B.

A Tabela 3.9 apresenta os resultados do modelo de decomposição dos efeitos da especialização cognitiva sobre os salários nos Arranjos Populacionais, utilizando os instrumentos de população defasada. A primeira coluna mostra a decomposição considerando os efeitos brutos da especialização cognitiva sobre os salários (primeiro estágio em *Pooled* de dados). Nessa especificação, verifica-se um bom ajuste do modelo onde o coeficiente de determinação é 0,6924. Os resultados indicam que há um efeito positivo sobre os salários médios da especialização cognitiva local se o arranjo for fronteiriço. Além disso, os arranjos das macrorregiões Norte e Nordeste apresentam efeitos positivos sobre os salários da especialização cognitiva, relativamente ao Sudeste. Essas características indicam que há vantagens produtivas para a especialização cognitiva, impactando positivamente nos salários locais.

Os resultados ainda indicam que quanto maior a área do arranjo menores os efeitos sobre os salários médios devidos à especialização cognitiva, isso indica que os efeitos de escala decorrentes de um aumento da área não são aproveitados quando se considera a especialização cognitiva. Isso se deve ao fato de que a especialização ocupacional categorizada por habilidades cognitivas não depende necessariamente de vastas extensões de território (ou população), mas de vantagens de capital humano específicas. Nesse sentido, como as regiões são economias abertas e condição fundamental para a especialização é a existência de mercado consumidor, a especialização de habilidades pode se intensificar mesmo em áreas menores uma vez que as exportações são factíveis.

No que se refere a variável de principal interesse, a densidade, uma descoberta interessante é obtida para o Brasil: a densidade apresenta um efeito negativo de 2,62% sobre os salários locais da especialização cognitiva. Apesar de ser notório na literatura a relação positiva entre densidade e salários (em geral), essa seção visa explorar o papel da densidade como um determinante dos efeitos da especialização sobre os salários. Nesse sentido, deve-se avaliar a relação entre especialização e densidade. A densidade é uma medida geográfica que expressa a população relativa, enquanto a especialização ocupacional categorizada por habilidades é uma medida econômica mais precisa para captar organização produtiva local, além de economias de aglomeração.

Tabela 3. 9 - Decomposição do efeito da especialização cognitiva sobre os salários

Variável Dependente: Efeito da especialização cognitiva sobre os salários locais				
2SLS (IV Históricas) – 2º estágio				
Variável	**(1)**	**(2)**	**(3)**	**(4)**
	EEC	**EEC(*i*)**	**EEC(*j*)**	**EEC(*ij*)**
Log. da densidade	-0,0262***	-0,01379***	-0,00091	0,00431
	(0,00532)	(0,0031)	(0,00378)	(0,00312)
Log. da área	-0,0409***	-0,01455***	0,007	0,00046
	(0,00971)	(0,00437)	(0,005)	(0,00489)
Fronteiriço	0,07012**	-0,02807	-0,0583*	-0,04201
	(0,03095)	(0,02657)	(0,03409)	(0,02698)
Turístico	-0,01422	-0,04028***	0,01088	-0,02042
	(0,01443)	(0,01385)	(0,01374)	(0,01253)
Norte	0,04499**	-0,00767	-0,04412*	0,0065
	(0,01754)	(0,0222)	(0,02320)	(0,01408)
Nordeste	0,1045***	0,0366***	0,0112	0,01423
	(0,01366)	(0,00673)	(0,01627)	(0,01123)
Sul	-0,03342**	0,00282	-0,01936	-0,0103
	(0,01667)	(0,00796)	(0,01407)	(0,01006)
Centro-Oeste	-0,06089	-0,03223*	-0,00048	0,00285
	(0,05881)	(0,01919)	(0,01211)	(0,01754)
Constante	0,5097***	0,3922***	0,02044	0,05519
	(0,06721)	(0,02598)	(0,0354)	(0,03513)
Dummies de ano	SIM	SIM	SIM	SIM
Observações	2.681	2.668	2.554	2.550
R^2 Ajustado	0,6924	0,5848	0,1785	0,2839
Prob>F	0,0000	0,0000	0,0000	0,0000
2SLS (IV Históricas) – 1º estágio				
Instrumentos:				
log. da população de 1940	-0,0503***	-0,0503***	-0,0502***	-0,0502***
	(0,01619)	(0,01619)	(0,01632)	(0,01632)
log. da população de 1991	0,99828***	0,9982***	0,9975***	0,9975***
	(0,01562)	(0,01562)	(0,01578)	(0,01578)

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01; Resultados assumem peso analítico proporcional ao número de trabalhadores em cada Arranjo Populacional.

Conforme visto na Figura 3.8 não existe uma relação clara entre especialização cognitiva e densidade nos Arranjos Populacionais brasileiros, desse modo, não é possível afirmar que uma maior densidade gera uma maior especialização e como consequência um aumento nos salários locais. Além disso, as medidas que captam economias de localização modificam os salários por dois canais que surgem a partir da interação entre empresas e trabalhadores: efeito competição e efeito externalidade. A densidade capta economias de urbanização que estão atreladas a aumentos da população, todavia, o aumento populacional pode provocar um efeito competição acentuado, exercendo impacto negativo sobre os salários. Assim, Arranjos Populacionais mais densos podem provocar um efeito competição superior ao efeito de externalidades sobre os

salários da especialização cognitiva, provocando uma diminuição do efeito da especialização de habilidades sobre os salários locais. Além disso, a diversificação produtiva inerente aos grandes centros urbanos pode limitar a especialização e atenuar as suas externalidades positivas.

Silva e Azzoni (2021) buscou avaliar a decomposição dos efeitos de localização sobre os salários, isto é, no seu primeiro estágio a equação de salários estimava os salários locais dos Arranjos Populacionais brasileiros. A autora decompõe os salários considerando, além das variáveis mencionadas nessa pesquisa, os efeitos fixos dos trabalhadores e indivíduos, uma vez que os salários locais obtidos no primeiro estágio desconsideravam os efeitos de firma e indivíduo. Nessa pesquisa, diferentemente, o efeito da especialização sobre os salários é obtido no primeiro estágio a partir de um *Pooled* de dados e de um painel de efeitos fixos de trabalhadores e firma, conforme já mencionado, obtendo os efeitos líquidos da especialização sobre os salários.

Considerando a especificação de Silva e Azzoni (2021), que não considera na decomposição os efeitos não observados de trabalhadores e firmas, sendo essa a especificação que mais se aproxima do resultado da coluna (1) - *Pooled* de dados desta pesquisa, verifica-se que há uma relação positiva entre densidade e efeitos de localização após o controle da endogeneidade da variável densidade pelo instrumento *Bartik* setorial: um efeito positivo de 2,49%. A relação positiva entre os efeitos da localização sobre os salários e a densidade é justificada na pesquisa da autora pois a densidade nada mais é que uma medida de aglomeração, assim quanto maior a densidade maior seria os salários locais dos Arranjos Populacionais brasileiros, pois há uma relação clara entre densidade e salários nas regiões metropolitanas. Uma relação positiva também foi observada nos trabalhos de Combes, Duranton e Gobillon (2008) para a França e Barufi, Haddad e Nijkamp (2016) para o Brasil.

Todavia, esta pesquisa contribui para a literatura pois realiza uma decomposição dos efeitos da especialização de habilidades sobre os salários, trazendo evidências novas da relação entre densidade e salários locais a partir da especialização de habilidades. A relação negativa obtida nessa pesquisa indica que os efeitos de localização sobre os salários (SILVA; AZZONI, 2021), os efeitos de áreas metropolitanas sobre os salários (COMBES; DURANTON; GOBILLON, 2008) e os efeitos setoriais sobre os salários (BARUFI; HADDAD; NIJKAMP, 2016) não captam todas as facetas que a aglomeração pode causar nos salários. Analisar os efeitos da especialização sobre os salários locais mostra a complexidade das externalidades da

aglomeração: a densidade impacta negativamente nos salários locais decorrentes da especialização, pois o efeito competição supera os efeitos externalidades. Esse efeito competição pode ser captado a partir da análise de especialização ocupacional, uma medida que, indiretamente, permite avaliar a relação entre trabalhadores e empresas locais.

Na especificação (2), onde se decompõe os efeitos da especialização cognitiva sobre os salários locais líquidos dos efeitos fixos dos trabalhadores (*i*), verifica-se que a variável densidade diminui sua magnitude, mas continua com um efeito negativo de 1,37%, isso indica que a heterogeneidade não observada dos trabalhadores é importante na determinação da relação entre a densidade e os salários locais da especialização cognitiva, o que reduz o poder da variável densidade na explicação desses salários. Isto é, a heterogeneidade não observada do trabalhador está correlacionada positivamente com a densidade. O mesmo pode ser verificado para a variável área cujo efeito negativo diminuiu de 4,09% para 1,45%. No que se refere às demais variáveis, observa-se que Arranjo Turístico passa a ser significativo com um efeito negativo sobre os salários (-0,04028), Arranjos pertencente ao Nordeste mantêm sua significância e sinal, diminuindo de magnitude (0,0366) e arranjos pertencentes ao Centro-Oeste passam a apresentar um menor prêmio de especialização relativo, significantes ao nível de 10% (0,03223).

A heterogeneidade da firma também se mostra muito importante na determinação dos efeitos da especialização cognitiva sobre os salários locais. Na especificação (3), onde se decompõe os efeitos da especialização sobre os salários locais líquidos dos efeitos fixos das firmas (*j*), as variáveis densidade e área deixam de ser significantes, indicando que a produtividade das firmas acaba anulando o efeito da densidade sobre os salários locais. Além disso, verifica-se pouco poder explicativo (R^2) nessa especificação, apenas as variáveis Fronteiriço e Norte são significantes ao nível de 10%.

Na especificação (4), onde se decompõe os efeitos da especialização sobre os salários locais líquidos dos efeitos fixos dos trabalhadores e firmas (*ij*), verifica-se a importância do *matching* na determinação dos salários locais da especialização cognitiva, pois todos os coeficientes passam a ser insignificantes e verifica-se que o modelo tem pouco poder explicativo (R^2). Desse modo, a densidade perde seu poder explicativo sobre os salários locais da especialização cognitiva quando se considera o *matching* no mercado de trabalho. Logo, a produtividade das

firmas e o *matching* tem maior poder explicativo do que a densidade populacional na determinação dos salários locais da especialização cognitiva.

Silva e Azzoni (2021) verificou que, na decomposição dos efeitos de localização no Brasil, os efeitos fixos individuais e de firmas foram componentes muito importantes, todavia, a inclusão destes efeitos não anulou a o efeito da densidade na explicação dos salários locais. Os resultados da autora levam-na a concluir que o componente mais importante da decomposição é o efeito fixo dos trabalhadores. Mion e Naticchioni (2009), ao realizam a decomposição dos salários locais na Itália, observam que o principal componente de variação nos salários são os efeitos fixos dos trabalhadores e o tamanho da firma, já as variáveis de efeito espacial, densidade e potencial de mercado, têm pequenos efeitos nos salários, com maior destaque para o potencial de mercado.

Nessa pesquisa, também é observado a importância da heterogeneidade não observada dos trabalhadores e das firmas como fonte de variação dos salários locais da especialização cognitiva, uma vez que na decomposição dos efeitos líquidos de especialização a variável densidade deixa de ser importante. Desse modo, observa-se que, na decomposição dos salários da especialização esses efeitos são tão importantes que suprimem o efeito puro da densidade.

Esse resultado pode ainda ser corroborado pelas evidências obtidas por Mion e Naticchioni (2009), que verificaram uma relação inversa entre *matching* e densidade. Considerando que a correlação entre habilidades dos trabalhadores e tamanho das firmas é uma medida de *matching*, os autores verificaram que nas províncias italianas, uma maior densidade está associada a um menor *matching*: uma correlação negativa de 0,39. Os autores mencionam que essa relação negativa pode ser explicada porque se considera os atributos gerais dos trabalhadores para obter a medida de correspondência. Nesse sentido, considerando os três canais dos efeitos da aglomeração sobre os salários (*sharing*, *matching* e *learning*) e a estratégia empírica de decomposição dos efeitos líquidos da especialização sobre os salários, que permite verificar a importância do *matching* (especificação (4)), pode-se argumentar que quanto maior a densidade menor o *matching* e, por sua vez, menores os efeitos positivos sobre os salários por esse canal dos efeitos da aglomeração. Logo, as evidências de Mion e Naticchioni (2009) permitem estabelecer que as relações entre medidas espaciais e salários apresentam ainda controvérsias, em especial, quando se considera na análise os múltiplos efeitos que a aglomeração pode causar

na economia. Assim, a relação positiva entre densidade e salário pode não estar presente em todas os contextos econômicos, como foi verificado nessa pesquisa.

A Tabela 3.10 apresenta os resultados do modelo de decomposição dos efeitos da especialização social sobre os salários nos Arranjos Populacionais, utilizando os instrumentos de população defasada. As evidências para a decomposição dos efeitos da especialização social sobre os salários são as mesmas verificadas para as habilidades cognitivas, em especial no papel da densidade sobre a decomposição. A primeira coluna mostra a decomposição considerando os efeitos brutos da especialização social sobre os salários (primeiro estágio em *Pooled* de dados). Verifica-se um bom ajuste do modelo (o coeficiente de determinação é 0,7614). Os resultados indicam um efeito positivo sobre os salários médios locais da especialização social se o Arranjo Populacional for fronteiriço.

Os arranjos das macrorregiões Norte e Nordeste apresentam efeitos positivos sobre os salários da especialização social, essas características indicam que há vantagens produtivas para a especialização social. No que se refere a área e a densidade o efeito é negativo, -0,03379 e -0,0506, respectivamente. Essas variáveis também apresentaram efeito negativo para a especialização cognitiva (TABELA 3.9), todavia, a magnitude desses efeitos para as habilidades sociais é um pouco maior. Esse fato pode ser explicado pelo efeito competição mais acentuado para as habilidades sociais.

Tabela 3. 10 - Decomposição do efeito da especialização social sobre os salários

Variável Dependente: Efeito da especialização social sobre os salários locais				
2SLS (IV Históricas) – 2º estágio				
Variável	**(1)**	**(2)**	**(3)**	**(4)**
	EES	**EES(*i*)**	**EES(*j*)**	**EES(*ij*)**
Log. da densidade	-0,0506***	-0,0239***	0,005	-0,00186
	(0,01388)	(0,00409)	(0,00962)	(0,00591)
Log. da área	-0,03379**	-0,0185***	-0,00218	0,00798
	(0,01612)	(0,00541)	(0,01275)	(0,00932)
Fronteiriço	0,0909**	0,01925	-0,01428	-0,01547
	(0,0424)	(0,0562)	(0,04429)	(0,05123)
Turístico	-0,02212	-0,05374**	-0,0022	0,00614
	(0,04544)	(0,02637)	(0,04388)	(0,02897)
Norte	0,1413***	0,02296	0,2121***	-0,01567
	(0,05081)	(0,02062)	(0,04462)	(0,03814)
Nordeste	0,368***	0,1502***	0,0712***	-0,0397***
	(0,03759)	(0,01513)	(0,02439)	(0,01341)
Sul	-0,1179***	-0,03307**	-0,0378	-0,0661**
	(0,02771)	(0,01399)	(0,03571)	(0,02646)
Centro-Oeste	0,00091	-0,02929**	-0,0299	-0,0425
	(0,03939)	(0,01392)	(0,05706)	(0,03101)
Constante	0,572***	0,1591***	0,0153	0,03935
	(0,15193)	(0,03883)	(0,08397)	(0,05544)
Dummies de ano	SIM	SIM	SIM	SIM
Observações	2.764	2.763	2.726	2.723
R^2 Ajustado	0,7614	0,6919	0,3391	0,1708
Prob>F	0,0000	0,0000	0,0000	0,0000
2SLS (IV Históricas) – 1º estágio				
Instrumentos:				
log. da população de 1940	-0,0489***	-0,0489***	-0,0488***	-0,0488***
	(0,01469)	(0,01469)	(0,01476)	(0,01475)
log. da população de 1991	0,9915***	0,9915***	0,9913***	0,99128***
	(0,01667)	(0,01667)	(0,0167)	(0,01676)

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01; Resultados assumem peso analítico proporcional ao número de trabalhadores em cada Arranjo Populacional.

Na especificação (2), onde se decompõe os efeitos da especialização sobre os salários locais líquidos dos efeitos fixos dos trabalhadores (*i*), verifica-se que a variável densidade diminui sua magnitude (-0,0239). Na especificação (3) e (4), a densidade passa a ser insignificante para explicar os efeitos da especialização social sobre os salários locais, líquidos dos efeitos fixos das firmas (*j*) e dos efeitos fixos dos trabalhadores e firmas (*ij*), respectivamente. Os fatos que podem explicar esses resultados já foram discutidos anteriormente.

A Tabela 3.11 apresenta os resultados do modelo de decomposição dos efeitos da especialização manual sobre os salários nos Arranjos Populacionais, utilizando os instrumentos de população

defasada. Os resultados para a decomposição indicam que os efeitos da especialização manual sobre os salários se comportam de maneira diferente da especialização cognitiva e social. Na especificação (1) somente as características dos arranjos e a macrorregião ao qual pertencem foram significantes para explicar o efeito bruto da especialização manual sobre os salários: Norte (0,2119), Nordeste (0,3803) e Centro-Oeste (0,08502) apresentam efeitos positivos sobre os salários da especialização manual. A principal variável de interesse, densidade, não foi significante em nenhuma das especificações, indicando que a densidade deixa de ser importante para explicar a decomposição da especialização manual.

Além disso, outra importante diferença para a especialização manual é o efeito da variável área que é positiva e significante nas especificações (2), (3) e (4). Isso indica que as economias de escala são mais importantes para explicar os efeitos da especialização manual sobre os salários, em detrimento da densidade populacional. Desse modo, verifica-se que, quanto maior a área do arranjo maiores serão os efeitos sobre os salários médios devidos à especialização manual. Os efeitos de escala decorrentes de um aumento da área são aproveitados na especialização manual. A especialização ocupacional categorizada por habilidades manuais pode depender de vastas extensões de território, o que pode ser típico das ocupações desempenhadas no setor agrícola ou mesmo em alguns segmentos industriais.

Na especificação (4), onde se decompõe os efeitos da especialização manual sobre os salários locais líquidos dos efeitos fixos dos trabalhadores e firmas (*ij*), somente a variável área e a macrorregião Nordeste foram significantes: a área exerce um efeito positivo de 2,65% e a macrorregião Nordeste exerce um efeito positivo de 5,42%, essas características indicam que há vantagens produtivas para a especialização manual.

Tabela 3. 11 - Decomposição do efeito da especialização manual sobre os salários

Variável Dependente: Efeito da especialização manual sobre os salários locais				
2SLS (IV Históricas) – 2º estágio				
Variável	**(1)**	**(2)**	**(3)**	**(4)**
	EEM	**EEM(*i*)**	**EEM(*j*)**	**EEM(*ij*)**
Log. da densidade	-0,01903	-0,007	0,00494	0,001
	(0,01214)	(0,00597)	(0,00899)	(0,00616)
Log. da área	0,02657	0,02329***	0,0249**	0,02655***
	(0,01665)	(0,00803)	(0,0127)	(0,00911)
Fronteiriço	0,09381	0,05778	0,09085	0,03376
	(0,05756)	(0,04252)	(0,0585)	(0,0517)
Turístico	0,07485	0,05691**	0,03188	0,01567
	(0,05169)	(0,02667)	(0,03059)	(0,03718)
Norte	0,2119***	0,06239	0,03398	-0,051**
	(0,07256)	(0,0482)	(0,0606)	(0,02308)
Nordeste	0,3803***	0,1979***	-0,0312	0,05417***
	(0,03669)	(0,02041)	(0,04375)	(0,0138)
Sul	-0,0995***	-0,02969	0,0037	-0,00056
	(0,03316)	(0,01931)	(0,0309)	(0,02808)
Centro-Oeste	0,08502**	0,03117*	0,05823**	0,0228
	(0,03722)	(0,01888)	(0,02874)	(0,03217)
Constante	-0,2974**	-0,30868***	-0,4963***	-0,35989
	(0,1287)	(0,052)	(0,0712)	(0,0533)
Dummies de ano	SIM	SIM	SIM	SIM
Observações	2.755	2.751	2.718	2.717
R^2 Ajustado	0,5265	0,4726	0,11	0,1226
Prob>F	0,0000	0,0000	0,0000	0,0000
2SLS (IV Históricas) – 1º estágio				
Instrumentos:				
log. da população de 1940	-0,0538***	-0,0538***	-0,0538***	-0,0538***
	(0,01575)	(0,01575)	(0,01577)	(0,01577)
log. da população de 1991	0,9989***	0,9989***	0,9987***	0,9987***
	(0,01685)	(0,01685)	(0,01686)	(0,01686)

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01; Resultados assumem peso analítico proporcional ao número de trabalhadores em cada Arranjo Populacional.

Portanto, a decomposição dos efeitos da especialização por habilidades é de suma importância, pois permite verificar as peculiaridades que não seriam observadas considerando a especialização ocupacional geral. Observou-se, ainda, que os efeitos da densidade merecem ser analisados com cautela quando se investiga a especialização ocupacional categorizada por habilidades.

Salienta-se a importância em comparar os resultados da decomposição dos efeitos da especialização por habilidades com os resultados obtidos na seção anterior. Constatou-se que as características não observadas dos trabalhadores e das firmas são fundamentais na relação

entre densidade e efeitos da especialização de habilidades sobre os salários locais, uma vez que quando se considera o efeito líquido da especialização sobre os salários locais a densidade deixa de ser importante, passando a ser estatisticamente insignificante. Essas evidências corroboram os resultados obtidos na seção anterior, onde verificou-se que a heterogeneidade não observada dos trabalhadores e firmas, bem como do seu efeito combinado, *matching*, são fundamentais para a determinação dos prêmios salariais da especialização ocupacional categorizada por habilidades no Brasil, onde na maior parte dos resultados os prêmios da especialização estavam superestimados. Isto é, ao isolar o efeito fixo dos trabalhadores e firmas, as variáveis capazes de mensurar a aglomeração e seus impactos nos salários perde magnitude, e no caso da densidade perde importância. Nesse sentido, destaca-se o papel da produtividade dos trabalhadores, da eficiência das firmas e do *matching* no mercado de trabalho brasileiro como essenciais para explicar a relação entre especialização de habilidades e salários locais.

3.6 CONSIDERAÇÕES FINAIS

O objetivo desse ensaio foi determinar o prêmio salarial da especialização categorizada por habilidades (cognitiva, social e manual), considerando o controle da endogeneidade da escolha local e da heterogeneidade não observada. A análise foi feita para o Brasil no período compreendido entre 2009 e 2018. Adicionalmente, realizou-se a decomposição do prêmio salarial da especialização por habilidades para entender a sua relação com a densidade, ou seja, de que forma a aglomeração afeta o prêmio dos salários locais considerando os efeitos da especialização de habilidades. Para tanto, considerou-se como unidades regionais para a especialização local os Arranjos Populacionais brasileiros e municípios isolados que se classificam como médias e grandes concentrações urbanas, delimitados pelo IBGE (2016), totalizando 376 unidades geográficas. Para a classificação dos trabalhadores em intensivos em habilidades cognitivas, sociais e manuais, avaliou-se as 2.557 ocupações da CBO (2002) de acordo com a base de dados de Maciente (2013).

Verificou-se que, os trabalhadores intensivos em habilidades cognitivas apresentam a maior média salarial (11,93) no Brasil, sendo seguido pelos trabalhadores sociais (5,27) e manuais (4,7). Esse fato pode ser explicado pelas características socioeconômicas dos trabalhadores das ocupações intensivas em habilidades cognitivas, que apresentam maior percentual com ensino superior e maior experiência. Todavia, é importante mencionar que, a análise do salário médio

não indica que o prêmio da especialização ocupacional categorizada por habilidades seguirá o mesmo padrão.

No modelo de *Pooled* de dados, apesar da especialização intensiva em habilidades cognitivas gerar um prêmio salarial positivo (0,37042), a especialização social apresenta um prêmio salarial negativo (-0,20805), isto é, mesmo apresentando uma média salarial maior do que a dos trabalhadores manuais, a especialização em habilidades sociais não gera um prêmio salarial positivo. Para os trabalhadores intensivos em habilidades manuais, o prêmio salarial foi positivo (0,13268). Esse contexto se deve ao fato de que a análise da especialização ocupacional se relaciona, intrinsecamente, com os efeitos da aglomeração, e os seus canais de *sharing*, *matching* e *learning* são fundamentais na determinação do diferencial de salários oriundos da especialização. Convém mencionar que, o resultado negativo para o prêmio da especialização social pode ser explicado pelo fato do efeito competição superar o efeito de externalidade para os trabalhadores nessa categoria de habilidade.

Nesse sentido, percebendo-se os complexos canais e efeitos que a aglomeração pode provocar na especialização e, por consequência nos salários, essa dinâmica mostra a urgência em analisar o prêmio da especialização categorizada por habilidades no Brasil: uma análise mais completa do que a análise do prêmio salarial das habilidades individuais dos trabalhadores. Essa estratégia empírica foi uma das contribuições dessa pesquisa. Outra importante contribuição desse ensaio foi a análise do prêmio da especialização considerando o controle da endogeneidade usando duas estratégias metodológicas. A primeira estratégia é o controle da endogeneidade da escolha local a partir da utilização de variáveis instrumentais (variáveis históricas e, como robustez, o *Bartik shift* ocupacional). A segunda estratégia é o controle da endogeneidade associada às heterogeneidades não observadas de trabalhadores e firmas a partir do uso de um painel de dados com a inclusão de efeitos fixos.

De modo geral, os resultados do prêmio de especialização com o controle da endogeneidade da escolha local pelo uso de variáveis instrumentais indicaram que as estimativas para o Brasil estavam subestimadas. Após o uso dos instrumentos os prêmios da especialização de habilidades aumentaram, devido aos efeitos intensos das externalidades como canal a explicar a aglomeração: um fluxo intenso de trabalhadores habilidosos nos Arranjos Populacionais estimula o aumento da especialização impactando positivamente nos efeitos de *sharing*, *matching* e *learning* sobre os salários.

No que se refere ao uso dos instrumentos históricos, o modelo 2SLS (IV Históricas) provocou um aumento no prêmio da especialização cognitiva (1,01305) e social (1,0277). Todavia, o prêmio da especialização manual desapareceu (-0,45785), indicando que os fluxos populacionais fazem com que os efeitos da externalidade sejam menores para esse grupo de trabalhadores, dado que, para trabalhadores com habilidades físicas, o processo de aprendizagem pode não ocorrer de maneira natural a partir das interações. Nesse cenário, o prêmio salarial estaria superestimado.

Ao realizar um exercício empírico que permite avaliar o prêmio da aglomeração por habilidades sobre os salários, observou-se que o prêmio da aglomeração é maior para as habilidades cognitivas, em todas os modelos estimados, o que reflete uma posição favorável dos trabalhadores intensivos em habilidades cognitivas. Esse resultado, além de possibilitar comparar os efeitos da aglomeração com os efeitos da especialização, justifica a importância em avaliar a especialização de habilidades, uma vez que os salários têm maior sensibilidade à especialização por habilidades, em detrimento da aglomeração por habilidades. Além disso, verificou-se que, após o controle da endogeneidade pelas variáveis instrumentais históricas, houve um pequeno aumento na magnitude das estimativas dos prêmios da aglomeração, indicando que a externalidade impacta positivamente no prêmio da aglomeração, todavia, com menor intensidade quando comparado ao prêmio da especialização. Essas evidências indicam a importância em estudar melhor a especialização ocupacional categorizada por habilidades dado que seus efeitos no mercado de trabalho são mais intensos do que os de aglomeração, em especial aqueles gerados através da externalidade: ao controlar a endogeneidade com IV histórica, foi encontrado um prêmio salarial de 1,013% e 1,028% para as especializações cognitiva e social e um prêmio de 0,065% e 0,034% para a aglomeração cognitiva e social.

No que se refere à determinação do prêmio salarial da especialização de habilidades com controle da endogeneidade a partir da heterogeneidade não observada dos trabalhadores e firmas, de modo geral, os resultados apontaram uma diminuição nas estimativas do prêmio salarial da especialização cognitiva, social e manual. Essa diminuição se deve ao fato de que a heterogeneidade não observada de trabalhadores e firmas superestimam os salários, mesmo em um contexto de especialização de habilidades, onde o canal da externalidade é determinante para os salários.

Verifica-se um grande impacto das características não observadas dos trabalhadores que, nesse contexto, onde as habilidades ocupacionais já foram controladas, representam as habilidades inatas dos indivíduos. Os prêmios da especialização cognitiva e manual apresentaram uma diminuição na sua magnitude, verificando-se um prêmio positivo para as habilidades cognitivas (0,0626) e negativo para as habilidades manuais (-0,00483). No entanto, o controle da heterogeneidade não observada dos trabalhadores provocou um aumento no prêmio da especialização social, tornando-o positivo (0,01967), indicando que os atributos não observados dos trabalhadores intensivos em habilidades sociais estariam negativamente relacionados com os prêmios de especialização. As heterogeneidades não observadas das firmas se mostraram fundamentais na análise da especialização de habilidades, impactando de forma significativa no prêmio da especialização, dado que os prêmios da especialização cognitiva, social e manual estavam superestimados. Isso se deve ao fato de que o padrão de especialização dos trabalhadores deve-se também às ligações de insumo-produto das empresas no espaço.

As características não observadas das empresas, ao representar a produtividade das mesmas, provocou uma intensa diminuição na magnitude do prêmio da especialização cognitiva, tornando esse prêmio negativo (-0,04402). Dessa forma, pode-se inferir que os trabalhadores intensivos em habilidades cognitivas são empregados em firmas de maior produtividade em relação aos demais. A melhor correspondência entre trabalhadores e firmas, analisada a partir do *matching* (controle da heterogeneidade dos trabalhadores e firmas), é fundamental na determinação do prêmio salarial da especialização de habilidades, todavia, o diferencial salarial da especialização permanece: especialização cognitiva (0,00706), social (-0,09359) e manual (0,05219). Visto que o prêmio da especialização cognitiva é o que apresentou a maior diminuição na sua magnitude após o controle da heterogeneidade não observada do *matching*, quando comparado ao modelo de *Pooled* OLS, ressalta-se que os trabalhadores intensivos em habilidades cognitivas são os que possuem as melhores habilidades inatas e mantêm vínculos empregatícios em firmas mais produtivas, estabelecendo uma melhor correspondência entre trabalhadores e empregadores.

De modo geral, comparando a abordagem de variáveis instrumentais e o painel de dados com efeitos fixos, percebe-se a complexidade na análise da especialização por habilidades no Brasil, pois enquanto o problema da escolha local acaba subestimando os prêmios salariais a heterogeneidade não observada faz o inverso, fatos que são justificados pela literatura empírica, dado que tanto as habilidades dos trabalhadores como a produtividade das firmas impactam

positivamente sobre os salários. Todavia, uma evidência inegável é que o prêmio salarial para a especialização de habilidades cognitivas é sempre positivo, independente do critério de controle de endogeneidade, logo, a especialização cognitiva sempre traz vantagem econômica.

Todavia, verificou-se que, em algumas especificações econométricas, a especialização social e a manual podem trazer um diferencial negativo (onde o efeito competição supera o efeito de externalidade), indo de encontro com a literatura econômica que, em geral, prevê vantagens para a especialização ocupacional e industrial. Os resultados dessa pesquisa auxiliam na compreensão do fenômeno da especialização ocupacional no Brasil, permitindo elaborar políticas públicas direcionadas a grupos de trabalhadores categorizados por habilidades, políticas que permitam que o efeito de externalidade tenha maior destaque suprimindo o efeito competição. Os governos podem incentivar a demanda desses trabalhadores em determinadas ramos da economia, em especial para as empresas de serviços e comércio, minimizado os efeitos de concorrência de modo a usufruir dos efeitos benéficos que a aglomeração pode causar através dos *spillovers* tecnológicos. Por outro lado, é possível que algumas categorias de habilidades não sejam beneficiadas pelas externalidades da especialização.

Na análise da decomposição dos salários locais, mensurados a partir do prêmio de especialização por habilidades, obteve-se uma evidência original nessa pesquisa: a densidade tem um impacto negativo na determinação dos prêmios salariais associados aos efeitos brutos da especialização cognitiva (-0,0262) e social (-0,0506). Evidências similares foram obtidas na determinação dos prêmios salariais associados aos efeitos líquidos da especialização cognitiva (-0,0138) e social (-0,0239), após o controle da heterogeneidade não observada dos trabalhadores, onde observou-se uma queda nos coeficientes. Essa tendência é explicada pela correlação positiva entre o efeito fixo do indivíduo e a densidade. A relação negativa obtida se deve ao fato de que, além da especialização de habilidades não ter uma relação clara com a densidade, os Arranjos Populacionais mais densos podem ter um efeito competição superior ao efeito de externalidade. Esse contexto, faz com que quanto maior a densidade menores serão os salários oriundos da especialização de habilidades.

É importante ainda destacar que a especialização cognitiva tende a ser maior em Arranjos Populacionais mais densos. Considerando a escala do grau de especialização cognitiva no Brasil (0,0171 e 0,2128), pode-se citar, como exemplo, o Arranjo Populacional de “São Paulo – SP ” e o de “Dourado – SP ”, onde o grau de especialização em habilidades cognitivas se destaca

como os maiores (0,11 e 0,2128, respectivamente) e possuem uma grande densidade média (2.874,89 habitantes/km² e 50.65914 hab/km², respectivamente). Já o Arranjo Populacional de "Internacional Chuy/Uruguai - Chuí/Brasil", com uma densidade média de 30,55 hab/km², apresenta o menor grau de especialização em habilidades cognitivas: 0,017. Esse cenário, pode sugerir que os centros menos densos devem pagar maiores salários para atraírem os trabalhadores que são intensivos em habilidades cognitivas ou, ainda, que o perfil do emprego baseado na estrutura produtiva local pode ser mais relevante para explicar a distribuição dos trabalhadores intensivos em habilidades cognitivas.

A habilidade social, diferente da cognitiva, não tem um padrão de especialização claro em maiores centros, todavia, dada a intensa oferta desses trabalhadores no Brasil, acredita-se que os Arranjos Populacionais menores não teriam a necessidade de ofertar incentivos salariais para atrair os trabalhadores que são intensivos em habilidades sociais. Não obstante, pode-se alegar que, o maior impacto negativo que a densidade provoca nos efeitos da especialização social (quando comparado com o da especialização cognitiva) sobre os salários é devido ao efeito de competição mais intenso nesse mercado. Todavia, quando se decompõe os efeitos da especialização cognitiva e social sobre os salários locais líquidos dos efeitos fixos dos trabalhadores e firmas, verifica-se que a densidade passa a ser insignificante, isto é, não explica os diferenciais salariais locais da especialização, indicando que o *matching* acaba anulando o efeito da densidade sobre os salários locais.

Quanto à decomposição dos salários devido a especialização manual, a densidade sempre foi insignificante na determinação dos prêmios salariais associados aos efeitos brutos e líquidos da especialização manual. Todavia, a área, ou extensão territorial do arranjo, tem um impacto positivo na determinação dos efeitos líquidos da especialização manual, isto é: as economias de escala são mais importantes do que a densidade. Considerando os efeitos da especialização sobre os salários locais líquidos dos efeitos fixos dos trabalhadores e firmas, a área exerce um efeito positivo de 2,65%.

De modo geral, os resultados indicam que o diferencial salarial da especialização por habilidades persiste mesmo após realizar os controles definidos na estratégia de identificação da pesquisa. Nesse contexto, percebe-se que o estudo do prêmio salarial da especialização categorizada por habilidades e a decomposição desses efeitos é de fundamental importância, pois permitem verificar as peculiaridades comportamentais de oferta e demanda para cada tipo

de mercado de habilidades. Desse modo, a compreensão dos efeitos de externalidades no mercado de habilidades possibilita uma visão mais ampla da dinâmica econômica entre salários e espaço, possibilitando elaborar políticas públicas mais eficazes e direcionadas. Assim, verifica-se que, apesar do pequeno número de trabalhadores intensivos em habilidades cognitivas no Brasil, a especialização destes sempre gera um prêmio positivo.

Além disso, verifica-se que a especialização cognitiva reflete positivamente não apenas na produtividade desse grupo de trabalhadores, mas no conjunto geral de todos os trabalhadores no mercado de trabalho. É interessante mencionar que, políticas públicas direcionadas para intensificar o desenvolvimento de habilidades cognitivas, aumentando a especialização e, por sua vez, os salários médios no Brasil, possibilitaria gerar uma melhor distribuição de renda. A medida simples de aglomeração, captada pela densidade, não garante ou explica os efeitos da especialização sobre os salários locais, isto é, compreende-se que a dinâmica do mercado por meio de uma maior urbanização não garante um aumento dos salários locais da especialização. Visto que o efeito fixo dos trabalhadores, o efeito fixo das firmas e o *matching* são os componentes de maior contribuição na decomposição dos salários da especialização ocupacional categorizada por habilidades, se faz necessário a elaboração e aplicação de políticas públicas que garantam maior produtividade dos trabalhadores e das firmas, bem como melhor correspondência entre os mesmos.

Por fim, extensões futuras desse trabalho envolvem avaliar os prêmios da especialização categorizada por habilidades, bem como sua decomposição, considerando uma maior desagregação, tanto no que se refere ao leque de habilidades quanto às unidades regionais. Essa maior desagregação possibilitaria uma análise mais detalhada, fato indispensável dada a vasta extensão territorial brasileira, permeada por intensas desigualdades regionais.

4 CONCLUSÃO

O objetivo central desta tese foi investigar os determinantes da aglomeração ocupacional e os efeitos da aglomeração urbana e ocupacional sobre os salários locais. Enfatiza-se que, o termo aglomeração envolve variáveis como espaço, trabalhadores e firmas. Dessa forma, trazendo conceitos da Economia do Trabalho e da Economia Regional, essa tese teve como principais objetos de análise as ocupações e habilidades dos trabalhadores. Além disso, essa análise torna-se mais relevante em um país com intensas desigualdades regionais como o Brasil. Os resultados evidenciaram que existe aglomeração ocupacional no mercado de trabalho formal Brasil e essa aglomeração é heterogênea regionalmente.

A concentração (ou aglomeração) ocupacional é mensurada a partir da participação dos trabalhadores de determinada ocupação e Arranjo Populacional no total de trabalhadores do Brasil, captada pelo índice de Ellison-Glaeser (1997) adaptado por Gabe e Abel (2009). Observou-se que há concentração ocupacional no Brasil no período de 2009 a 2018, com uma média de 0,0633[29]. No que se refere a especialização ocupacional categorizada por habilidades, mensurada a partir da participação dos trabalhadores de determinada habilidade e Arranjo Populacional no total de trabalhadores do respectivo Arranjo Populacional, verificou-se que as especializações manual e social no Brasil apresentaram uma média de 0,4782 e 0,4475, respectivamente, enquanto a especialização cognitiva apresentou uma média de 0,0757. Assim, observou-se que, além do Brasil ser pouco especializado em habilidades cognitivas, os maiores índices de especialização para essa categoria de habilidade estão agrupados nos Arranjos Populacionais pertencentes às regiões Sul e Sudeste, regiões com os maiores Arranjos Populacionais em tamanho populacional.

Essa dinâmica impacta diretamente nos salários locais. Verifica-se que, a especialização ocupacional de habilidades cognitivas tem uma relação positiva com o salário-hora real do respectivo grupo de habilidade e, quando se considera os Arranjos Populacionais grandes (população>750 mil habitantes), a relação positiva é intensificada. Logo, a especialização cognitiva provoca efeitos positivos de externalidades, que são maiores nos grandes Arranjos Populacionais. Essas evidências, acerca da concentração e especialização ocupacional no Brasil, e seus efeitos nos salários locais, em especial no que se refere à especialização

[29] Para o mercado de trabalho dos EUA, Gabe e Abel (2009) encontraram uma média de 0,1872, superior à verificada nessa pesquisa.

categorizada pelas habilidades, pode resultar em alguns questionamentos: de modo geral, a maior aglomeração ocupacional tem efeitos positivos no bem-estar da economia?

A literatura econômica aponta efeitos favoráveis para a aglomeração ocupacional. A explicação para os efeitos positivos, dados os efeitos de externalidade, é que a dinâmica econômica das respectivas unidades regionais é intensificada, provocando aumento nos salários locais e maior bem-estar. Os autores indicam, ainda, que os efeitos de externalidades podem beneficiar regiões menos aglomeradas ou menos especializadas a partir da transferência inter-regional dos efeitos positivos da aglomeração sobre o bem-estar econômico (ROSSI-HANSBERG; SARTE; SCHWARTZMAN, 2019; GABE; ABEL, 2009; WHEATON; LEWIS, 2002; PERROUX, 1955; HIRSCHMAN, 1977; MARSHALL, 1920). Todavia, convém mencionar que, essa transferência inter-regional aconteceria somente no longo prazo. Nesse cenário, deve-se ter cautela ao avaliar como a aglomeração ocupacional pode impactar no Brasil, um país com intensas desigualdades regionais, uma vez que a concentração e especialização ocupacional podem intensificar ou perpetuar as desigualdades regionais pré-existentes.

Dadas essas considerações, é importante reforçar as contribuições desta tese. No primeiro ensaio, a principal contribuição para a literatura foi avaliar a aglomeração ocupacional e seus determinantes considerando uma medida exógena de *labor pooling* que considera choques idiossincráticos e as habilidades dos trabalhadores. Já no segundo ensaio, a principal contribuição foi estimar o prêmio salarial da especialização ocupacional categorizada por habilidade e investigar a decomposição dos efeitos brutos e líquidos sobre os salários locais devido à especialização. Todavia, as contribuições dessa tese para o Brasil vão além das mencionadas, pois os resultados podem contribuir, mesmo que de forma limitada, para os estudos de desigualdades regionais no mercado de trabalho.

A análise do diferencial salarial estabelece como se dá a determinação dos salários, isto é, quais as variáveis que explicam os salários dos trabalhadores, tendo como foco a desigualdade interpessoal de renda. Enquanto o estudo da desigualdade regional refere-se ao problema da desigualdade da renda agregada devido à concentração da produção, ou seja, trata-se da desigualdade inter-regional da renda. Essa pesquisa avalia o prêmio salarial da especialização ocupacional, onde a variável espaço (Arranjo Populacional) é primordial na determinação desses prêmios. Essa tese avança na literatura, pois analisa não só a desigualdade interpessoal, mas a desigualdade entre as regiões a partir da variável de especialização. Logo, apesar da

análise de prêmios salariais, por si só, não ser suficiente para tecer conclusões acerca de desigualdades regionais, verifica-se que a análise do prêmio da especialização pode trazer indícios de que o Brasil tem uma composição e distribuição heterogênea de trabalhadores categorizados por habilidades no espaço. E, dada sua extensão territorial, surgem padrões espaciais extremamente desiguais. Esse estudo ratificou essas tendências.

Desse modo, infere-se que os prêmios salariais entre as unidades espaciais ao longo do país, que captam mudanças no padrão do emprego e da composição de habilidades, podem estar perpetuando ou intensificando as desigualdades regionais no país, uma vez que os resultados dessa pesquisa indicaram um prêmio salarial positivo para a especialização em habilidades cognitivas, em todas as modelagens empíricas estimadas, e a especialização cognitiva está concentrada em Arranjos Populacionais pertencentes às regiões Sul e Sudeste. Adicionalmente, ao avaliar as características dos trabalhadores categorizados por habilidades, pôde-se inferir que as desigualdades regionais no Brasil não são somente explicadas pela estrutura produtiva do país e sua concentração no espaço, mas pelas características dos trabalhadores, em especial as suas habilidades, que diferem entre as regiões. Dessa forma, pode-se indicar que as desigualdades regionais no Brasil também podem ser reflexo das desigualdades na composição dos trabalhadores.

Promover uma menor desigualdade regional não é uma tarefa fácil, principalmente em um país em desenvolvimento de grandes dimensões como o Brasil. A principal medida seria modificar a estrutura produtiva do país, promovendo a produção em setores e atividades mais complexos que exijam maiores habilidades cognitivas, permitindo o desenvolvimento dessas, em todas as regiões do país. Essa medida, por sua vez, constitui-se uma política pública de longo prazo. Não obstante, políticas públicas de desenvolvimento regional de curto prazo podem objetivar aumentar a competitividade dos setores intensivos em habilidades manuais no Brasil, permitindo maiores exportações. Além disso, políticas públicas de subsídios à indústria brasileira, intensiva em habilidades sociais e manuais, possibilitariam promover choques positivos na demanda por trabalho. As duas políticas promoveriam um cenário que intensificaria a demanda por trabalhadores manuais e sociais aumentando os salários locais dos mesmos e possibilitando uma melhor distribuição de renda e maior bem-estar. Por fim, considerando que a mobilidade de capital produtivo e financeiro é uma realidade, deve-se salientar que uma política pública de desenvolvimento regional eficaz deve incidir sobre os indivíduos, proporcionando às regiões mais pobres do Brasil um maior investimento em

educação de qualidade, considerando as novas demandas de habilidades, o que permitiria atrair investidores para essas regiões. Essa política permitiria, ao mesmo tempo, reduzir as desigualdades inter-regionais e interpessoais de renda no Brasil, gerando um efeito agregado positivo sobre o bem-estar do país.

REFERÊNCIAS

ABOWD, J. M.; KRAMARZ, F.; MARGOLIS, D. N. High Wage Workers and High Wage Firms. **Econometrica**, v. 67, n.2, p.251-333, mar. 1999.

ACEMOGLU, D.; AUTOR, D. Skills, tasks and technologies: implications for employment and earnings. *In:* ASHENFELTER, O.; CARD, D. (ed.). **Handbook of labor economics**, [S.l]: Elsevier, 2011. v. 4, p. 1043-1171.

ACEMOGLU, D. Training and innovation in an imperfect labour market. **The Review of Economic Studies**, v. 64, n.3, p.445-464, jul. 1997.

ALMEIDA, E. T.; ROCHA, R. M. Labor pooling as an agglomeration factor: Evidence from the Brazilian Northeast in the 2002–2014 period. **EconomiA,** n. 19, p. 236–250, 2018

ALKAY, E.; HEWINGS, G. J. The determinants of agglomeration for the manufacturing sector in the Istanbul metropolitan area. **The Annals of Regional Science**, Springer, v. 48, n. 1, p. 225–245, 2012.

AUTOR, D. H.; DUGGAN, M. G. The rise in the disability rolls and the decline in unemployment. **The Quarterly Journal of Economics**, v. 118, n.1, p. 157-206, 2003.

ANDRADE, P. H. P.; GONÇALVES, E.; FREGUGLIA, R. S. Concentração geográfica de ocupações por regiões metropolitanas brasileiras. **Revista Brasileira de Estudos Regionais e Urbanos (RBERU)**, v. 8, n. 2, p. 155-173, 2014.

BACOLOD, M.; BLUM, B. S.; STRANGE, W. C. Skills in the city. **Journal of Urban Economics**, v. 65, p. 136–153, 2009.

BANCO MUNDIAL. **Retomando o caminho para a inclusão, o crescimento e a sustentabilidade:** Brasil - diagnóstico sistemático de país. Disponível em: https:// documents1.worldbank.org/curated/pt/686871468197371171/pdf/101431-REVISED-PORTUGUESE-v2-SCD-Sumario-Executivo.pdf. Acesso em: 10 nov. 2021.

BARROS, R.; CARVALHO, M.; FRANCO, S.; MENDONÇA, R. Uma análise das principais causas da queda recente na desigualdade de renda brasileira. **Econômica: Revista do Programa de Pós-Graduação em Economia da UFF**, v. 8, n. 1, p. 117-147, 2006.

BARTIK, Timothy J**. Who benefits from state and local economic development policies?** Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, 1991. p. 354.

BARUFI, A. M. B.; HADDAD, E. A.; NIJKAMP, P. Industrial scope of agglomeration economies in Brazil. **The Annals of Regional Science**, v. 56, n. 3, p. 707-755, 2016.

BIDERMAN, C.; GUIMARÃES, N. Desigualdades, discriminação e políticas públicas: uma análise a partir de setores selecionados da atividade produtiva no Brasil. Comunicação, *In*: ENCONTRO NACIONAL DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS DO TRABALHO, n. 8, 2003, São Paulo. **Anais**[...]. São Paulo: ABET, 2003.

BLACKBURN, M. L.; NEUMARK, B. D. Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials. **Quarterly Journal of Economics**, v. 107, n. 4, p. 1421-1435, 1992.

BORJAS, G. J. Self-selection and the earnings of immigrants. *In*: ZIMMERMANN, K.; BAUER, T. (org.). **The economics of migration.** Northampton: Edward Elger, 1987. p. 332-54.

BORUSYAK, K.; HULL, P.; JARAVEL, X. **Quasi-experimental shift-share research designs.** [S.l]: NBER, 2018. (Working paper, n. 24997).

BURKI, A. A.; KHAN, M. A. **Agglomeration economies and their effects on technical inefficiency of manufacturing firms**. **LUMS Economics**, v. 1, n. 12, 2012.

CACCIAMALI, M. C.; TATEI, F.; ROSALINO, J. W. Estreitamento dos diferenciais de salários e aumento do grau de discriminação, limitações da mensuração padrão? **Planejamento e Políticas Pública**, v. 33, p. 195-222, 2009.

CAMERON, A. C.; TRIVEDI, P. K. **Microeconometrics:** methods and applications. New York: Cambridge University Press, 2005.

CARD, D. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. **Journal of Labor Economics,** v. 19, n.1, p. 22–64, 2001.

CICCONE, A.; HALL, R. E. Productivity and the density of economic activity. **The American Economic Review**, v. 86, n.1, p. 54-70, 1996.

COELHO, A. M.; CORSEUIL, C. H. **Diferenciais salariais no Brasil:** um breve panorama. Brasília: IPEA, 2002. (Texto para discussão, n. 898).

COMBES, P. P.; DURANTON, G.; GOBILLON, L. Spatial wage disparities: Sorting matters!. **Journal of Urban Economics**, v. 63, n. 2, p. 723-742, 2008.

COMBES, P.P.; DURANTON, G.; GOBILLON, L. The identification of agglomeration economies. **Journal of Economic Geography**, v. 11, p. 253–266, 2011.

COMBES, P. P.; GOBILLON, L. The empirics of agglomeration economies. *In*: DURANTON, G.; HENDERSON, J. V.; STRANGE, W. C. (eds.). **Handbook of Regional and Urban Economics.** Amsterdam: Elsevier, 2015. p. 247-348.

COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira**: Cana-de-açúcar. Brasília: Conab, 2018. v.5, p. 1-62, 2018.

CRUZ, B. O.; FURTADO, B. A.; MONASTERIO, L.; JÚNIOR RODRIGUES, W. **Economia regional e urbana:** teorias e métodos com ênfase no Brasil. Brasília: IPEA, 2011.

DAVID, J. D. The Growing importance of social skills in the labor market. **The Quarterly Journal of Economics**, v. 132, n. 4, p. 1593–1640, 2017.

DINIZ, M. B.; ARRAES, R. A. Desenvolvimento econômico e desigualdade de renda no Brasil. *In*: ENCONTRO REGIONAL EM ECONOMIA, n. 10, 2005, Fortaleza. **Anais** [...]. Fortaleza: Universidade Federal do Ceará, 2005.

DIXIT, A.; STIGLITZ, J. Monopolistic competition and optimum product diversity. **American Economic Review**, v. 67, p. 297-308, 1977.

DURANTON, G.; OVERMAN, H. G. Testing for localization using micro-geographic data. **The Review of Economic Studies**, v. 72, n. 4, p. 1077–1106, 2005.

DURANTON, G.; PUGA, D. Micro-foundations of urban agglomeration economies. *In*: HENDERSON, J. V.; THISSE, J.F. (eds.). **Handbook of regional and urban economics**. Amsterdam: Elsevier, 2004. v. 4, p. 2063-2117.

EBERTS, R. W.; McMILLEN, D. P. Agglomeration economies and urban public infrastructure. *In*: CHESHIRE, P.; MILLS, E. S. (eds.). **Handbook of regional and urban economics**. Amsterdam: Elsevier, 1999. p. 1455–1495.

EHN, M. **On the causes and effects of specialization:** a mathematical approach. Suécia: Västerås: Mälardalen University Press Licentiate, 2009. (Theses, n. 103).

EHRL, P. Minimum comparable areas for the period 1872-2010: an aggregation of Brazilian municipalities. **Estudos Econômicos**,v. 47, n.1, p. 215-229, 2017.

EHRL, P.; MONASTERIO, L. Spatial skill concentration agglomeration economies. **Journal Regional Science**, v. 61, p. 1–22, 2020.

ELLISON, G.; GLAESER, E. L. Geographic Concentration in US Manufacturing Industries: A Dartboard Approach. **Journal of Political Economy**, v. 105, n. 5, 1997.

ELLISON, G.; GLAESER, E. L. The geographic concentration of industry: Does natural advantage explain agglomeration? **The American Economic Review**, v. 89, n. 2, p. 311-316, 1999.

ELLISON, G.; GLAESER, E. L.; KERR, W. R. What causes industry agglomeration? Evidence from coagglomeration patterns. **The American Economic Review**, v. 100, n. 3, p. 1195–1213, 2010.

EVARINI, A.; ARAÚJO, E.; GARCIA, M. F. G.; SOUZA, S. C. I. Os diferenciais de salário no Brasil no período de 1999-2009: uma investigação empírica na perspectiva da teoria da segmentação. *In*: ENCONTRO NACIONAL DE ECONOMIA, 39., 2011, Foz do Iguaçu. **Anais** [...]. Foz do Iguaçu: ANPEC, 2011.

FERREIRA NETO, A. B.; FREGUGLIA, R. S.; FAJARDO, B. A. G. Diferenciais salariais para o setor cultural e ocupações artísticas no Brasil. **Economia Aplicada**, v. 16, n. 1, p. 49-76, 2012.

FLORIDA R.; MELLANDER, C.; STOLARICK, K.; ROSS, A. Cities, Skills and Wages. **Journal of Economic Geography**, v.12, 2011.

FREGUGLIA, R. S. F.; MENEZES-FILHO, N.; SOUZA, D. B. Diferenciais salariais inter-regionais, interindustriais e efeitos fixos individuais: uma análise a partir de Minas Gerais. **Estudos Econométricos**, São Paulo, v.37, n. 1, p. 129-150, 2007.

FUJITA, M.; KRUGMAN, P.; VENABLES, A. **The spatial economy:** cities, regions and international trade. Cambridge: Massachusetts Institute of Techhnology, 1999.

FUJITA, M.; THISSE, J. **Economics of agglomeration:** cities, industrial location, and regional growth. Cambridge: Cambridge University Press, 2002.

GABE, T. M.; ABEL, J. R. **Labor market pooling and occupational agglomeration**. New New York: Federal Reserve Bank of New York Staff Reports, n. 392, 2009. Disponível em: http://dx.doi.org/10.2139/ssrn.1473365. Acesso em: 13 mar. 2019.

GOLDMAN, B.; KLIER, T.; WALSTRUM, T. Within-industry agglomeration of occupations: Evidence from census microdata. **Journal Regional Science**, v. 59, n. 5, p. 1-21, 2019.

GREENE, W. H. **Econometric analysis.** 7. ed. New Jersey: Prentice Hall, 2008.

GUVENEN, F.; KURUSCU B.; TANAKA S.; WICZER, D. Multidimensional skill mismatch. **American Economic Journal: Macroeconomics**, v. 12, n. 1, p. 44-210, 2020.

HANDEL, M. J. **Trends in job skill demands in OECD countries.** Paris: OECD Social, Employment and Migration Working Papers, n. 143, 2012. Disponível em: https://doi.org/ 10.1787/5k8zk8pcq6td-en. Acesso em: 13 mar. 2019

HECKMAN, J. J. Lessons from the Bell Curve. **Journal of Political Economy**, v. 103, p. 1091–1120, 1995.

HECKMAN, J. J.; STIXRUD, J.; URZUA, S. The Effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. **Journal of Labor Economics**, v. 24, n. 3, p. 411-482, 2006.

HECKMAN, J. Sample selection bias as a specification error. **Econometrica**, v. 47, p.153–161, 1979.

HECKMAN, J. The common structure of statistical models of truncation, sample selection e limited dependent variables and a simple estimator for such models. *In*: BERG, S. V. (org). **The Annais of Economic and Social Measurement**. Massachusetts: NBER, 1976. p. 475-492.

HENDERSON, J. V. Community development: t1he effects of growth and uncer- tainty. **American Economic Review**, v. 70, p. 894-910, 1980.

HERRNSTEIN, R. J.; MURRAY, C. A. **The bell curve:** intelligence and class structure in American life. New York: Free Press, 1994.

HIRSCHMAN, A. O. Transmissão interregional e internacional do crescimento econômico. *In*: SCHWARTZMAN, J. (org). **Economia Regional:** textos escolhidos. Belo Horizonte: CEDEPLAR/UFMG, 1977. p. 35-52.

HOFFMANN, R. Distribuição de renda e crescimento econômico. **Estudos Avançados**, v. 15, n. 41, 2001.

IBGE. **Arranjos populacionais e concentrações urbanas do Brasil.** 2. ed. Rio de Janeiro, 2016. Disponível em: https://www.ibge.gov.br/apps/arranjos_populacionais/2015/pdf/publicacao.pdf. Acesso em: 14 jun. 2019.

IBGE. **Evolução da divisão territorial do Brasil 1872–2010.** Rio de Janeiro, 2011.

IBGE. **Pesquisa anual de comércio. Relatórios metodológicos.** 2. ed. Rio de Janeiro, 2000. v. 12. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv7052.pdf. Acesso em: 20 jun. 2020.

IBGE. **Séries históricas:** PIB a preços de mercado – Taxa acumulada em 4 trimestres (%). Disponível em: https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9300-contas-nacionais-trimestrais.html?=&t=series-historicas. Acesso em: 20 set. 2021.

JACINTO, P. A. Diferenciais de salários por gênero na indústria avícola da região Sul do Brasil: uma análise com micro dados. **Revista de Economia e Sociologia Rural**, Brasília, v. 43, n. 3, p. 529-555, 2005.

JENSEN, A. R. **The g factor:** the science of mental ability. Westport, CT: Praeger, 1998.

KRUGMAN, P. R. Increasing returns, monopolistic competition, and international trade. **Journal of International Economics**, v. 9, p. 469-479, 1979.

KRUGMAN, P. R. **Geography and trade.** Cambridge, MA: Mit Press, 1991.

KRUGMAN, P.; VENABLES, A. J. Globalization and the inequality of nations. **The Quarterly Journal of Economics**, v. 110, n. 4, 1995.

LEE, L. Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables. **International Economic Review**, v.19, p.415–433, 1978.

LEWANDOWSKI, P. The changing nature of jobs in central and eastern europe. **IZA World of Labor**, v. 351, p. 1–10, 2017.

MACIENTE, A. N. **A concentração do emprego no Brasil entre 1994 e 2005:** aspectos regionais e setoriais. [S.l]: IPEA, 2008.

MACIENTE, A. N. Uma análise setorial e regional das competências laborais no Brasil. *In*: NASCIMENTO, P. A. M. M. (org). **Radar**: tecnologia, produção e comércio exterior. Brasília: Ipea, 2012. v.23, p. 27-36.

MACIENTE, A. N.; RAUEN, C. V.; KUBOTA, L. C. Tecnologias digitais, habilidades ocupacionais e emprego formal no Brasil entre 2003 e 2017. **Mercado de Trabalho**: **Conjuntura e Análise**, v. 25, n. 66, p. 1-15, 2019.

MACIENTE, A. **The determinants of agglomeration in Brazil:** input-output, labor and knowledge externalities. 2013. Tese (Doutorado em Economia) — University of Illinois at Urbana-Champaign, 2013.

MAROCO, J. **Análise estatística com utilização do SPSS.** Lisboa: Edições Sílabo, 2007.

MARSHALL, A. **Principles of economics.** 8 ed. London: Macmillan, 1920.

MAUREL, F.; SÉDILLOT, B. A measure of the geographic concentration in french manufacturing industries. **Regional Science and Urban Economics**, v. 29, n. 5, p. 575–604, 1999.

MION, G.; NATICCHIONI, P. The spatial sorting and matching of skills and firms. **The Canadian Journal of Economics / Revue canadienne d'Economique**, v. 42, n. 1, p. 28-55, 2009.

MORETTI, E. Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. **Journal of Econometrics**, v. 121, n. 1-2, p. 175–212, 2004.

MORICONI, G. M.; MOURA NETO, J. S.; MARCONI, N.; ARVATE, P. R. Diferentes padrões de políticas salariais nos estados brasileiros: uma análise a partir do diferencial de salários público-privado. **Revista de Economia Política**, v. 29, n.3, p. 114-134, 2009.

BRASIL. Ministério da Economia. **Classificação Brasileira de Ocupações (CBO):** informações gerais. 2017. Disponível em: http://www.mtecbo.gov.br/cbosite/pages/informacoesGerais.jsf. Acesso em: 28 jun. 2019.

BRASIL. Ministério da Economia. **Manual de Orientação da Relação Anual de Informações Sociais (RAIS):** ano-base 2018. – Brasília: SPPE, DES, CGET, 2019.

MONTEIRO NETO, A. M. **Desigualdades regionais no Brasil:** características e tendências recentes. [S.l]:IPEA, 2014. (Boletim regional, urbano e ambiental, n. 9).

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **Marco da Programação no País (CPF) FAO para o Brasil 2013-2016**, 2013. Disponível em: http://www.fao.org/3/a-bp560o.pdf. Acesso em: 25 jun. 2020.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **FAOSTAT, Brazil**, 2018. Disponivel em: http://www.fao.org/faostat/en/#country/21. Acesso em: 25 jun. 2020.

O * NET RESOURSE CENTER. **Skills**, 2018. Disponível em: www.onetcenter.org/dictionary/23.3/excel/skills.html. Acesso em: 17 jun. 2019.

OVERMAN, H. G.; PUGA, D. Labor Pooling as a Source of Agglomeration: An Empirical Investigation. *In*: GLAESER, E. L. **Agglomeration economics**. [S.l.]: University of Chicago Press, 2010. p. 133-150.

PERROUX, F. O Conceito de pólo de crescimento. *In*: SCHWARTZMAN, J. (org.) **Economia regional:** textos escolhidos. Belo Horizonte: CEDEPLAR/UFMG, 1995. p. 145-152.

PUGA, D. The magnitude and causes of agglomeration economies. **Journal of Regional Science**, v. 50, n.1, p. 203-219, 2010.

REIS, E.; PIMENTEL, M.; ALVARENGA, A. I.; SANTOS M. C. H. Áreas mínimas comparáveis para os períodos intercensitários de 1872 a 2000. *In*: SIMPÓSIO BRASILEIRO DE CARTOGRAFIA HISTÓRICA, 1., 2011, Paraty, Rio de Janeiro. **Anais** [...]. Paraty: SBC, 2011.

RESENDE, M. Industrial coagglomeration: some state-level evidence for Brazil. **Nova Economia, SciELO Brasil**, v. 25, n. 1, p. 181–194, 2015.

RESENDE, M.; WYLLIE, R. Ubicación industrial y políticas de incentivos locales en Brasil: una investigación empírica. **Perspectivas**, v. 2, p. 9–24, 2004.

RODRIGUES, L. C.; VIEIRA, M. T.; FREGUGLIA, R. S. O papel das ocupações na desigualdade salarial brasileira: uma análise da evolução recente entre trabalhadores de alta qualificação. *In*: ENCONTRO NACIONAL DE ECONOMIA, 43., 2015, Florianópolis, Santa Catarina. **Anais** [...]. Florianópolis: ANPEC, 2015.

ROSENBERG, M. **Society and the adolescent self-image.** Princeton, NJ: Princeton University Press, 1965.

ROSENTHAL, S. S.; STRANGE, W. C. The determinants of agglomeration. **Journal of Urban Economics**, v. 50, n. 2, p. 191–229, 2001.

ROSSI-HANSBERG; SARTE; SCHWARTZMAN. **Cognitive hubs and spatial redistribution.** Richmond: Federal Reserve Bank of Richmond, 2019.

ROTTER, J. B. **Generalized expectancies for internal versus external control of reinforcement.** Washington, DC: American Psychological Association, 1966.

ROY, A. D. Some thoughts on the distribution of earnings. **Oxford Econ.**, v. 3, p. 145-146, 1951.

RUMBERGER, R. W.; LEVIN, H. M. Forecasting the impact of new technologies on the future job market. **Technological Forecasting and Social Change**, v. 27, n. 4, p. 399-417, 1985.

SABOIA, João. Efeitos do salário mínimo sobre a distribuição de renda no Brasil no período 1995/2005 – resultados de simulações. **Econômica**, v. 9, n. 2, p. 270-296, 2007.

SCHWAB, K. **The fourth industrial revolution.** New York: Crown Business, 2016.

SILVA, A. S. **Economias de aglomeração e *matching* de ocupação e qualificação no mercado de trabalho brasileiro.** 2019. Tese (Doutorado em Economia) - Universidade Federal de Bahia, Salvador¸2019. 153 p.

SILVA, D.; AZZONI, C. Worker and firm heterogeneity, agglomeration, and wages in Brazil. **Papers in Regional Science**, p. 1– 27, 2021. Disponível em: https://doi.org/ 10.1111/pirs.12637. Acesso em: 15 out. 2021.

SILVA, C. E. L.; HEWINGS, G. J. A decisão sobre investimento em capital humano em um arranjo produtivo local (APL): Uma abordagem teórica. **Revista Brasileira de Economia**, v. 64, n.1, p. 67-79, 2010.

SMITH, A. **An inquiry into the nature and causes of the wealth of nations.** London: Printed for W. Strahan, and T. Cadell., 1776.

TAVARES, P. A.; MENEZES-FILHO, N. A. Human capital and the recent fall of earnings inequality in Brazil. **Brazilian Review of Econometrics**, v. 31, n.2, p. 231–257, 2013.

VALE, G. M. V.; CASTRO, J. M. Clusters, arranjos produtivos locais, distritos industriais: reflexões sobre aglomerações produtivas. **Análise Econômica**, Porto Alegre, v. 28, n. 53, p. 81-97, 2010.

VENABLES, A. J. Equilibrium locations of vertically linked industries. **International Economic Review**, v. 37, n. 2, p. 341-359, 1996.

WHEATON, W. C.; LEWIS, M. Urban wages and labor market agglomeration. **Journal of Urban Economics**, v. 51, p. 542-562, 2002.

VIEIRA FILHO, J. E. R. **Expansão da fronteira agrícola no Brasil:** desafios e perspectivas. Brasília: Ipea, 2016. (Texto para discussão, n. 2223).

WILLIAMSON, O. E. **Markets and hierarchies:** analysis and antitrust implications. New York: The Free Press, 1975.

APÊNDICES

APÊNDICE A – Resultados complementares do Capítulo **2**

Tabela A.1 - Categorias de Habilidades Ocupacionais

Categorias	**Descrição das categorias de habilidades**
1	Habilidades cognitivas
2	Habilidades de manutenção e operação
3	Habilidades de assistência
4	Habilidades gerenciais
5	Habilidades de design e engenharia
6	Habilidades de transporte
7	Habilidades artísticas
8	Habilidade de precisão e automação
9	Habilidades de trabalho supervisionadas
10	Habilidades de ensino e ciências sociais
11	Força física
12	Habilidades de telecomunicação
13	Habilidades de independência
14	Habilidades em ciências naturais
15	Habilidades de atenção
16	Experiência no trabalho
17	Habilidades de gestão de conflitos
18	Habilidades de trabalho em grupo
19	Habilidades de vendas
20	Habilidades de monitoramento e conformidade
21	Habilidades clericais

Fonte: Maciente (2013)

Tabela A.2 – Estatística descritiva das variáveis normalizadas

Variável	Média	Desvio	Mínimo	Máximo
EGga	-7,31e-11	1	-0,656	58,04
Labor Pooling	0	1	-0,389	88,27
Habilidades cognitivas	1,21e-10	1	-28,17	7,574
Habilidades de manutenção e operação	7,82e-11	1	-31,04	15,27
Habilidades de assistência	-2,37e-10	1	-16,20	26,11
Habilidades gerenciais	-1,84e-10	1	-20,78	17,25
Habilidades de design e engenharia	-0	1	-25,17	16,49
Habilidades de transporte	-1,57e-10	1	-14,53	48,37
Habilidades artísticas	0	1	-32,50	9,116
Habilidade de precisão e automação	0	1	-44,70	28,97
Habilidades de trabalho supervisionadas	5,58e-11	1	-29,54	21,58
Habilidades de ensino e ciências sociais	1,57e-10	1	-22,94	13,53
Força física	8,57e-11	1	-10,05	30,34
Habilidades de telecomunicação	-0	1	-26,60	21,34
Habilidades de independência	6,11e-11	1	-25,12	21,29
Habilidades em ciências naturais	2,46e-10	1	-26,34	23,27
Habilidades de atenção	-8,71e-11	1	-39,85	25,65
Experiência no trabalho	0	1	-19,74	11,61
Habilidades de gestão de conflitos	1,22e-10	1	-21,51	35,34
Habilidades de trabalho em grupo	-7,54e-11	1	-58,19	88,46
Habilidades de vendas	6,31e-11	1	-11,38	51,15
Habilidades de monitoramento e conformidade	-1,40e-10	1	-21,23	16,30
Habilidades clericais	2,11e-10	1	-23,96	36,50
Participação de trabalhadores com ensino superior	5,98e-09	1	-0,765	1,831
Participação de trabalhadores no setor da agropecuária	8,49e-10	1	-0,240	7,802
Participação de trabalhadores no setor de extração mineral	-3,68e-09	1	-0,209	11,74
Média do tamanho do estabelecimento	-2,39e-09	1	-2,034	2,515
Experiência média	1,29e-10	1	-1,983	13,88

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Considera-se os trabalhadores em Arranjos Populacionais e do setor privado.

Tabela A.3 – Resultado completo da modelagem *Labor Pooling* ponderado pelas habilidades e determinantes da aglomeração ocupacional no Brasil: 2009 a 2018

Variável Dependente: EGga		
	***Pooled* OLS**	**Efeito Fixo (EF)**
Labor Pooling (LP)	0,0169***	0,0057***
	(0,00261)	(0,0014)
Vetor de Habilidades		
Cognitivas	-0,0925***	-0,3549***
	(0,0296)	(0,03501)
Manutenção e operação	-0,5496***	-0,1412***
	(0,02821)	(0,02825)
Assistência	-0,0156	-0,1471***
	(0,01165)	(0,0095)
Gerenciais	-0,355***	0,2783***
	(0,01429)	(0,02403)
Design e engenharia	0,1844***	-0,2447***
	(0,02594)	(0,0247)
Transporte	0,1682***	0,1065***
	(0,01736)	(0,01494)
Artísticas	0,2514***	0,3474***
	(0,02482)	(0,02789)
Precisão e automação	-0,1133***	-0,414***
	(0,01598)	(0,01334)
Trabalho supervisionadas	0,3205***	0,0209
	(0,01296)	(0,01717)
Ensino e ciências sociais	-0,1177***	0,0468**
	(0,01231)	(0,01979)
Força física	0,2574***	0,4305***
	(0,02233)	(0,02404)
Telecomunicação	-0,077***	-0,0265***
	(0,01134)	(0,00894)
Independência	0,2995***	0,1723***
	(0,01227)	(0,01635)
Ciências naturais	0,0819***	0,1216***
	(0,01511)	(0,01496)
Atenção	-0,0144	0,2247***
	(0,01891)	(0,01669)
Experiência no trabalho	0,2897***	-0,428***
	(0,01946)	(0,0226)
Gestão de conflitos	-0,0994***	-0,0054
	(0,01153)	(0,00882)
Trabalho em grupo	0,064***	-0,0214***
	(0,01255)	(0,00621)
Vendas	0,9406***	0,5609***
	(0,02589)	(0,03094)
Monitoramento e conformidade	0,3137***	-0,0442***
	(0,01193)	(0,01654)
Clericais	0,2607***	0,2547***
	(0,01735)	(0,01896)

(continua)

(continuação)

Variável Dependente: EGga		
	***Pooled* OLS**	**Efeito Fixo (EF)**
Variáveis Interação (*Labor Pooling* x Habilidades)		
LP x Cognitivas	-0,1707***	-0,0247
	(0,03317)	(0,02257)
LP x Manutenção e operação	0,6835***	0,0377*
	(0,02768)	(0,01926)
LP x Assistência	-0,0317**	0,0539***
	(0,01316)	(0,0086)
LP x Gerenciais	0,342***	-0,0194
	(0,01738)	(0,01187)
LP x Design e engenharia	-0,2017***	0,0718***
	(0,02657)	(0,01844)
LP x Transporte	-0,1052***	-0,021*
	(0,01704)	(0,0107)
LP x Artísticas	-0,316***	-0,1279***
	(0,02454)	(0,01678)
LP x Precisão e automação	-0,0276*	0,0107***
	(0,01471)	(0,01001)
LP x Trabalho supervisionadas	-0,2775***	-0,0654***
	(0,01433)	(0,00957)
LP x Ensino e ciências sociais	-0,0179	-0,0196**
	(0,01267)	(0,00957)
LP x Força física	-0,3508***	-0,1258***
	(0,03318)	(0,02475)
LP x Telecomunicação	-0,0177	-0,024**
	(0,01258)	(0,00949)
LP x Independência	-0,2215***	-0,0127
	(0,01215)	(0,01048)
LP x Ciências naturais	-0,0469***	0,002
	(0,01365)	(0,00864)
LP x Atenção	-0,0035	0,0731***
	(0,01995)	(0,01409)
LP x Experiência no trabalho	-0,3723***	-0,0149
	(0,01926)	(0,01432)
LP x Gestão de conflitos	0,0203	0,0354***
	(0,01465)	(0,00941)
LP x Trabalho em grupo	-0,0685***	0,0163**
	(0,01315)	(0,00647)
LP x Vendas	-0,2199***	-0,1416***
	(0,02466)	(0,01581)
LP x Monitoramento e conformidade	-0,2447***	-0,0272***
	(0,01296)	(0,00977)
LP x Clericais	-0,1324***	-0,1541***
	(0,01707)	(0,01189)

(continua)

(conclusão)

Variável Dependente: EGga		
	***Pooled* OLS**	**Efeito Fixo (EF)**
Variáveis de Controle		
P. de trabalhadores com ensino superior	-0,0093***	-0,0117
	(0,00245)	(0,00824)
P. de trabalhadores no setor da agropecuária	-0,0246***	-0,0054
	(0,00227)	(0,00376)
P. de trabalhadores no setor de extração mineral	-0,0091***	0,0008
	(0,00223)	(0,00361)
Média do tamanho do estabelecimento	0,0242***	0,0003
	(0,00251)	(0,00277)
Experiência média	0,003	-0,0003
	(0,00242)	(0,0023)
Constante	0,0456***	0,0542***
	(0,00683)	(0,00314)
Efeitos Temporais		
Dummies de ano	SIM	SIM
N. observações	21.804	21.804
R^2 global	0,8858	0,7097
R^2 within		0,6028
R^2 between		0,7162
Prob>F	0,0000	0,0000
Hausman		$\chi^2(56)=4509$
Prob > χ^2		0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: *** significativo a 1%; ** significativo a 5%; * significativo a 10%. Erro padrão entre parênteses.
Obs.: Todas as variáveis foram normalizadas com média 0 e desvio padrão de 1.

Tabela A.4 – Determinantes da aglomeração ocupacional no Brasil (2009 a 2018): *Labor Pooling* ponderado pelas habilidades e efeito marginal do *Labor Pooling* completo

Variável Dependente: EGga			
	***Pooled* OLS**		
Habilidades	γ_h **(Habilidades)**	μ_h **(LP x habilidades)**	$\delta + \mu_h$ $(\delta = \mathbf{0,0169})$
Cognitivas	-0,0925***	-0,1707***	-0,1538
Manutenção e operação	-0,5496***	0,6835***	0,7004
Assistência	-0,0156	-0,0317**	-0,0148
Gerenciais	-0,355***	0,342***	0,3589
Design e engenharia	0,1844***	-0,2017***	-0,1848
Transporte	0,1682***	-0,1052***	-0,0883
Artísticas	0,2514***	-0,316***	-0,2991
Precisão e automação	-0,1133***	-0,0276*	-0,0107
Trabalho supervisionadas	0,3205***	-0,2775***	-0,2606
Ensino e ciências sociais	-0,1177***	-0,0179	-
Força física	0,2574***	-0,3508***	-0,3339
Telecomunicação	-0,077***	-0,0177	-
Independência	0,2995***	-0,2215***	-0,2046
Ciências naturais	0,0819***	-0,0469***	-0,03
Atenção	-0,0144	-0,0035	-
Experiência no trabalho	0,2897***	-0,3723***	-0,3554
Gestão de conflitos	-0,0994***	0,0203	-
Trabalho em grupo	0,064***	-0,0685***	-0,0516
Vendas	0,9406***	-0,2199***	-0,203
Monitoramento e conformidade	0,3137***	-0,2447***	-0,2278
Clericais	0,2607***	-0,1324***	-0,1155

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: *** significativo a 1%; ** significativo a 5%; * significativo a 10%. Erro padrão entre parênteses.
Obs.: Todas as variáveis foram normalizadas com média 0 e desvio padrão de 1.

Tabela A.5 – Resultado em nível do *Labor Pooling* como determinante da aglomeração ocupacional no Brasil: 2009 a 2018

Variável Dependente: EGga		
	***Pooled* OLS**	**Efeito Aleatório (EA)**
	(I)	(I)
Labor Pooling (LP)	-0,0000053	-0,0000000089
	(0,000006)	(0,0000015)
Constante	0,0695***	0,068***
	(0,00195)	(0,00184)
Efeitos Temporais		
Dummies de ano	SIM	SIM
N. observações	21.804	21.804
R^2 global	0,0101	0,0100
R^2 within		0,1910
R^2 between		0,0000
Prob>F	0,0000	0,0000
Hausman		$\chi^2(9)= 1,95$
Prob > χ^2		0,9923

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: *** significativo a 1%; ** significativo a 5%; * significativo a 10%. Erro padrão entre parênteses.

Tabela A.6 – Resultado em nível dos determinantes da aglomeração ocupacional no Brasil: 2009 a 2018

Variável Dependente: EGga				
	***Pooled* OLS**		**Efeito Fixo (EF)**	
	(II)	(III)	(II)	(III)
Labor Pooling (LP)		-0,0000044*		0,00000032
		(0,0000024)		(0,000001)
Vetor de Habilidades				
Cognitivas	-6,7292***	-6,6064***	-6,1063***	-8,533***
	(0,26386)	(0,26031)	(0,73146)	(0,74022)
Manutenção e operação	2,6166***	2,658***	-0,9317***	-1,1797***
	(0,10085)	(0,10045)	(0,21222)	(0,20073)
Assistência	-0,4881***	-0,4126***	-0,8815***	-0,9266***
	(0,04079)	(0,03974)	(0,06058)	(0,05668)
Gerenciais	-0,6458***	-0,6094***	5,3919***	4,7005***
	(0,10267)	(0,10069)	(0,39172)	(0,39273)
Design e engenharia	-1,4136***	-1,3744***	-2,5405***	-2,578***
	(0,10803)	(0,10485)	(0,18869)	(0,19609)
Transporte	0,1155***	0,1439***	0,4819***	0,5069***
	(0,02728)	(0,02696)	(0,06086)	(0,05638)
Artísticas	-0,2142	-0,2099	3,1632***	5,2409***
	(0,20259)	(0,19872)	(0,45766)	(0,46031)
Precisão e automação	-0,4682***	-0,4277***	-1,7661***	-1,5286***
	(0,02218)	(0,02257)	(0,02781)	(0,02703)
Trabalho supervisionadas	0,7828***	0,7224***	-0,4688***	0,1217
	(0,0258)	(0,02533)	(0,07783)	(0,07637)
Ensino e ciências sociais	-2,0323***	-2,0323***	0,0527	0,3768
	(0,06955)	(0,06975)	(0,22706)	(0,23532)
Força física	1,2254***	1,2479***	6,4387***	6,4118***
	(0,12343)	(0,12192)	(0,27641)	(0,26645)
Telecomunicação	-0,8506***	-0,8197***	-0,4265***	-0,3731***
	(0,03508)	(0,03393)	(0,0583)	(0,05578)
Independência	1,3269***	1,3715***	1,189***	1,6707***
	(0,03236)	(0,03182)	(0,08754)	(0,08164)
Ciências naturais	1,0991***	1,1032***	2,1276***	2,1938***
	(0,09228)	(0,08993)	(0,18835)	(0,18575)
Atenção	0,1916***	0,2136***	1,8262***	1,6978***
	(0,03939)	(0,0385)	(0,07635)	(0,07221)
Experiência no trabalho	-3,0098***	-3,0497***	-12,122***	-11,198***
	(0,152)	(0,1499)	(0,35157)	(0,32979)
Gestão de conflitos	-0,4068***	-0,3729***	0,2252***	0,1255***
	(0,01992)	(0,01919)	(0,01891)	(0,01741)
Trabalho em grupo	-0,0015**	-0,0002	-0,0008***	-0,0004***
	(0,00034)	(0,00032)	(0,00016)	(0,00014)
Vendas	3,4475***	3,3453***	2,0207***	1,7227***
	(0,04877)	(0,04827)	(0,10788)	(0,1022)
Monitoramento e conformidade	1,6868***	1,6684***	-2,2889***	-1,0143***
	(0,07343)	(0,07372)	(0,20516)	(0,2069)

(continua)

(conclusão)

Variável Dependente: EGga				
	***Pooled* OLS**		**Efeito Fixo (EF)**	
	(II)	(III)	(II)	(III)
Clericais	0,1581***	0,1528***	0,2314***	0,071*
	(0,02115)	(0,02156)	(0,03655)	(0,03783)
Variáveis de Controle				
P. de trabalhadores com ensino superior		-0,0008		-0,0029
		(0,0007)		(0,0021)
P. de trabalhadores no setor da agropecuária		-0,0112***		-0,0044
		(0,00203)		(0,00297)
P. de trabalhadores no setor de extração mineral		-0,0097***		0,001
		(0,00297)		(0,00424)
Média do tamanho do estabelecimento		0,0043***		0,0001
		(0,00042)		(0,00041)
Experiência média		-4,33e-08		-2.37e-06
		(9,81e-06)		(8.26e-06)
Constante	0,0649***	0,0546***	0,0612***	0,0616***
	(0,0008)	(0,00121)	(0,00039)	(0,00118)
Efeitos Temporais				
Dummies de ano	SIM	SIM	SIM	SIM
N. observações	21.804	21.804	21.804	21.804
R^2 global	0,8369	0,8481	0,6575	0,7014
R^2 within			0,5554	0,5871
R^2 between			0,6653	0,7077
Prob>F	0,0000	0,0000	0,0000	0,0000
Hausman			$\chi^2(29)$= 4668,14	$\chi^2(24)$= 2928,77
Prob > χ^2			0,0000	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: *** significativo a 1%; ** significativo a 5%; * significativo a 10%. Erro padrão entre parênteses.

Tabela A.7 – Resultado completo em nível da modelagem *Labor Pooling* ponderado pelas habilidades e determinantes da aglomeração ocupacional no Brasil: 2009 a 2018

Variável Dependente: EGga		
	***Pooled* OLS**	**Efeito Fixo (EF)**
Labor Pooling (LP)	0,00001***	0,000005***
	(0,0000025)	(0,0000013)
Vetor de Habilidades		
Cognitivas	-2,3379***	-8,972***
	(0,7485)	(0,88515)
Manutenção e operação	-6,1014***	-1,5679***
	(0,31328)	(0,31368)
Assistência	-0,1308	-1,2289***
	(0,09734)	(0,07941)
Gerenciais	-6,2414***	4,893***
	(0,2513)	(0,42256)
Design e engenharia	2,1697***	-2,8796***
	(0,30527)	(0,29062)
Transporte	0,7875***	0,4988***
	(0,08131)	(0,06997)
Artísticas	5,4716***	7,561***
	(0,54014)	(0,60714)
Precisão e automação	-0,4344***	-1,5874***
	(0,06128)	(0,05117)
Trabalho supervisionadas	1,5144***	0,0987
	(0,06124)	(0,08115)
Ensino e ciências sociais	-1,6947***	0,6744**
	(0,17728)	(0,28485)
Força física	3,8342***	6,4117***
	(0,33269)	(0,3581)
Telecomunicação	-0,6708***	-0,2309***
	(0,0988)	(0,07791)
Independência	1,9709***	1,1339***
	(0,08078)	(0,10759)
Ciências naturais	1,2611***	1,8714***
	(0,23255)	(0,23022)
Atenção	-0,0844	1,3168***
	(0,11086)	(0,09785)
Experiência no trabalho	6,463***	-9,5498***
	(0,43422)	(0,50427)
Gestão de conflitos	-0,3271***	-0,0179
	(0,03794)	(0,02903)
Trabalho em grupo	0,0066***	-0,0022***
	(0,00129)	(0,00064)
Vendas	4,6114***	2,7497***
	(0,12692)	(0,15168)
Monitoramento e conformidade	4,8883***	-0,69***
	(0,18599)	(0,25785)

(continua)

(continuação)

Variável Dependente: EGga		
	***Pooled* OLS**	**Efeito Fixo (EF)**
Clericais	0,9291***	0,908***
	(0,06185)	(0,06757)
Variáveis Interação (*Labor Pooling* x Habilidades)		
LP x Cognitivas	-0,1081***	-0,0156
	(0,021)	(0,01429)
LP x Manutenção e operação	0,2173***	0,0119*
	(0,0088)	(0,00612)
LP x Assistência	-0,0061**	0,0103***
	(0,00251)	(0,00164)
LP x Gerenciais	0,1694***	-0,0096
	(0,0086)	(0,00588)
LP x Design e engenharia	-0,0609***	0,02167***
	(0,00802)	(0,00557)
LP x Transporte	-0,0141***	-0,0028*
	(0,00229)	(0,00145)
LP x Artísticas	-0,1931***	-0,0782***
	(0,01499)	(0,01025)
LP x Precisão e automação	-0,0031*	0,0012
	(0,00165)	(0,00113)
LP x Trabalho supervisionadas	-0,0333***	-0,0078***
	(0,00172)	(0,00114)
LP x Ensino e ciências sociais	-0,0069	-0,0076**
	(0,00493)	(0,00372)
LP x Força física	-0,0967***	-0,0347***
	(0,00914)	(0,00682)
LP x Telecomunicação	-0,0037	-0,005**
	(0,00263)	(0,00199)
LP x Independência	-0,0423***	-0,0024
	(0,00232)	(0,002)
LP x Ciências naturais	-0,0206***	0,0009
	(0,00597)	(0,00378)
LP x Atenção	-0,0005	0,0109***
	(0,003)	(0,00211)
LP x Experiência no trabalho	-0,2112***	-0,0085
	(0,01093)	(0,00813)
LP x Gestão de conflitos	0,0014	0,0024***
	(0,00101)	(0,00065)
LP x Trabalho em grupo	-0,0002***	0,00004**
	(0,00003)	(0,00001)
LP x Vendas	-0,0301***	-0,0194***
	(0,00338)	(0,00216)
LP x Monitoramento e conformidade	-0,0981***	-0,0109***
	(0,00519)	(0,00391)
LP x Clericais	-0,0128***	-0,0149***
	(0,00166)	(0,00115)

(continua)

(conclusão)

Variável Dependente: EGga		
	***Pooled* OLS**	**Efeito Fixo (EF)**
Variáveis de Controle		
P. de trabalhadores com ensino superior	-0,0023***	-0,0029
	(0,00061)	(0,00206)
P. de trabalhadores no setor da agropecuária	-0,0191***	-0,0042
	(0,00176)	(0,00292)
P. de trabalhadores no setor de extração mineral	-0,0105***	0,0009
	(0,00258)	(0,00416)
Média do tamanho do estabelecimento	0,0035***	0,00004
	(0,00036)	(0,0004)
Experiência média	0,00001	-1,11e-06
	(0,000008)	(8,11e-06)
Constante	0,0544***	0,0608***
	(0,00105)	(0,00116)
Efeitos Temporais		
Dummies de ano	SIM	SIM
N. observações	21.804	21.804
R^2 global	0,8858	0,7097
R^2 within		0,6028
R^2 between		0,7162
Prob>F	0,0000	0,0000
Hausman		$\chi^2(43)$= 3908,69
Prob > χ^2		0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: *** significativo a 1%; ** significativo a 5%; * significativo a 10%. Erro padrão entre parênteses.

APÊNDICE B – Resultados complementares do Capítulo 3

Tabela B.1 – Prêmio salarial da especialização ocupacional de habilidades cognitivas: POLS e 2SLS

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variáveis	***Pooled* OLS**	**2SLS (IV Históricas)**	**2SLS (IV Bartik)**
Variáveis explicativas			
$ln(EO_h)$	0,37042***	1,10305***	2,40958***
	(0,00143)	(0,00275)	(0,02693)
Características do trabalhador			
Idade	0,04852***	0,04943***	0,04621***
	(0,00019)	(0,000198)	(0,00023)
Idade²	-0,00046***	-0,00047***	-0,00043***
	(0,0000025)	(0,0000027)	(0,00000302)
Tempo no emprego	0,00257***	0,00265***	0,00311***
	(0,000012)	(0,000013)	(0,0000157)
Tempo no emprego²	-0,0000024***	-0,0000026***	-0,0000035***
	(0,00000043)	(0,000000044)	(0,00000049)
Masculino	0,26442***	0,02718***	0,27234***
	(0,00064)	(0,000657)	(0,00075)
Escolaridade (Categoria base - Analfabeto)			
Fundamental Incompleto	-0,12933	-0,16176*	-0,19984**
	(0,09181)	(0,09574)	(0,10029)
Fundamental Completo	-0,04784	-0,07603	-0,13967
	(0,09180)	(0,095739)	(0,10026)
Médio Incompleto	0,03589	-0,02108	-0,04717
	(0,09180)	(0,09574)	(0,10026)
Médio Completo	0,16925*	0,11603	0,96969
	(0,09179)	(0,09573)	(0,10024)
Superior Incompleto	0,48420***	0,42067***	0,36158***
	(0,09180)	(0,095735)	(0,10026)
Superior Completo	0,91339***	0,84723***	0,77818***
	(0,09179)	(0,095732)	(0,10026)
Características do emprego			
Tamanho do Estabelecimento (Categoria base - Micro)			
Pequeno	0,19852***	0,195298***	0,17257***
	(0,00067)	(0,0007056)	(0,00086)
Médio	0,33501***	0,33287***	0,29347***
	(0,00079)	(0,000829)	(0,00107)
Grande	0,48864***	0,457515***	0,29012***
	(0,000903)	(0,0009307)	(0,002802)

(continua)

(conclusão)

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	***Pooled* OLS**	**2SLS (IV Históricas)**	**2SLS (IV Bartik)**
Setor de Atividade (Categoria base - Agropecuária)			
Extração Mineral	0,88271***	0,93883***	0,85269***
	(0,01083)	(0,01207)	(0,01326)
Transformação	0,11858***	0,09068***	-0,07975***
	(0,00546)	(0,00656)	(0,00858)
Construção Civil	0,21127***	0,16646***	-0,06617***
	(0,005501)	(0,00659)	(0,00896)
Comércio	-0,04699***	-0,09513***	-0,33996***
	(0,00543)	(0,00653)	(0,008987)
Serviços	0,03861***	-0,01196*	-0,28164***
	(0,005402)	(0,00649)	(0,00913)
Serviços Domésticos	-0,13139***	-0,16513***	-0,40657***
	(0,03515)	(0,03175)	(0,35091)
Organismos Internacionais	0,33845***	0,20272***	-0,25683***
	(0,01626)	(0,01604)	(0,019901)
Dummies de ano	SIM	SIM	SIM
Constante	0,75312***	2,3726***	6,09196***
	(0,09202)	(0,09621)	(0,122902)
Observações (NxT)	16.227.793	15.003.790	16.186.500
R^2 Ajustado	0,5853	0,5759	0,3745
Prob>F	0,000	0,000	0,000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.2 – Prêmio salarial da especialização ocupacional de habilidades sociais: POLS e 2SLS

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	***Pooled* OLS**	**2SLS (IV Históricas)**	**2SLS (IV Bartik)**
Variáveis explicativas			
$ln(EO_h)$	-0,20806***	1,02775***	-0,09791***
	(0,00087)	(0,00314)	(0,00316)
Características do trabalhador			
Idade	0,02359***	0,02187***	0,02343***
	(0,000056)	(0,0000603)	(0,000056)
Idade²	-0,00024***	-0,00022***	-0,00024***
	(0,00000079)	(0,00000085)	(0,00000078)
Tempo no emprego	0,00278***	0,00286***	0,00279***
	(0,0000056)	(0,0000059)	(0,0000055)
Tempo no emprego²	-0,0000014***	-0,0000016***	-0,0000014***
	(0,000000025)	(0,000000025)	(0,000000024)
Masculino	0,15953***	0,16249***	0,15972***
	(0,000205)	(0,00022)	(0,000205)
Escolaridade (Categoria base - Analfabeto)			
Fundamental Incompleto	0,02596***	0,03004***	0,02676***
	(0,00234)	(0,00250)	(0,00234)
Fundamental Completo	0,05958***	0,06233***	0,06035***
	(0,00234)	(0,00249)	(0,00234)
Médio Incompleto	0,08422***	0,08997***	0,08525***
	(0,00234)	(0,00250)	(0,00234)
Médio Completo	0,16603***	0,16926***	0,16672***
	(0,00233)	(0,00249)	(0,00233)
Superior Incompleto	0,40421***	0,42917***	0,40618***
	(0,00237)	(0,00253)	(0,00237)
Superior Completo	0,77545***	0,79941***	0,77679***
	(0,00239)	(0,00254)	(0,00238)
Características do emprego			
Tamanho do Estabelecimento (Categoria base - Micro)			
Pequeno	0,11553***	0,10830***	0,11483***
	(0,00021)	(0,00023)	(0,00021)
Médio	0,19613***	0,19259***	0,19563***
	(0,00032)	(0,00034)	(0,00032)
Grande	0,20087***	0,18679***	0,19964***
	(0,00036)	(0,00038)	(0,00036)

(continua)

(conclusão)

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	***Pooled* OLS**	**2SLS (IV Históricas)**	**2SLS (IV Bartik)**
Setor de Atividade (Categoria base - Agropecuária)			
Extração Mineral	0,68311***	0,64298***	0,67465***
	(0,00578)	(0,00599)	(0,00573)
Transformação	0,14531***	0,16268***	0,14464***
	(0,00232)	(0,00269)	(0,00227)
Construção Civil	0,22042***	0,17879***	0,21374***
	(0,00236)	(0,00274)	(0,00232)
Comércio	-0,01866***	-0,06099***	-0,02503***
	(0,00229)	(0,00267)	(0,00225)
Serviços	-0,01741***	-0,07182***	-0,02506***
	(0,00229)	(0,00267)	(0,00225)
Serviços Domésticos	-0,18401***	-0,22478***	-0,19043***
	(0,00665)	(0,00751)	(0,00665)
Organismos Internacionais	0,41307***	0,31903***	0,40124***
	(0,00988)	(0,00983)	(0,00986)
Dummies de ano	SIM	SIM	SIM
Constante	0,14993***	1,14826***	0,24193***
	(0,00345)	(0,00448)	(0,00425)
Observações (NxT)	79.373.026	72.837.470	79.323.717
R^2 Ajustado	0,616	0,5851	0,6157
Prob>F	0,000	0,000	0,000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.3 – Prêmio salarial da especialização ocupacional de habilidades manuais: POLS e 2SLS

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	***Pooled* OLS**	**2SLS (IV Históricas)**	**2SLS (IV Bartik)**
Variáveis explicativas			
$ln(EO_h)$	0,13268***	-0,45786***	0,17553***
	(0,00063)	(0,00156)	(0,00172)
Características do trabalhador			
Idade	0,029509***	0,02849***	0,02945***
	(0,000048)	(0,000052)	(0,000048)
Idade²	-0,000317***	-0,000305***	-0,00032***
	(0,00000065)	(0,00000069)	(0,00000065)
Tempo no emprego	0,002497***	0,002497***	0,00248***
	(0,0000058)	(0,00000601)	(0,0000057)
Tempo no emprego²	-0,0000015***	-0,0000014***	-0,0000014***
	(0,000000029)	(0,000000029)	(0,000000027)
Masculino	0,33351***	0,33766***	0,33393***
	(0,000201)	(0,000218)	(0,000203)
Escolaridade (Categoria base - Analfabeto)			
Fundamental Incompleto	0,173806***	0,17616***	0,17259***
	(0,0010996)	(0,001176)	(0,00109)
Fundamental Completo	0,23928***	0,23878***	0,23757***
	(0,0011023)	(0,00119)	(0,001097)
Médio Incompleto	0,251189***	0,24818***	0,24962***
	(0,001115)	(0,00119)	(0,001109)
Médio Completo	0,33365***	0,33079***	0,33192***
	(0,001099)	(0,00117)	(0,00109)
Superior Incompleto	0,585979***	0,59937***	0,58491***
	(0,0015504)	(0,00165)	(0,00155)
Superior Completo	0,88469***	0,89350***	0,88421***
	(0,0018179)	(0,00190)	(0,00181)
Características do emprego			
Tamanho do Estabelecimento (Categoria base - Micro)			
Pequeno	0,08917***	0,08712***	0,090201***
	(0,000182)	(0,00020)	(0,00018)
Médio	0,160835***	0,15997***	0,16184***
	(0,00024)	(0,00026)	(0,00024)
Grande	0,22094***	0,21239***	0,22109***
	(0,000319)	(0,00033)	(0,000317)

(continua)

(conclusão)

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	***Pooled* OLS**	**2SLS (IV Históricas)**	**2SLS (IV Bartik)**
Setor de Atividade (Categoria base - Agropecuária)			
Extração Mineral	0,42909***	0,41016***	0,43942***
	(0,00261)	(0,00297)	(0,00263)
Transformação	0,17771***	0,17904***	0,18772***
	(0,00087)	(0,00107)	(0,00087)
Construção Civil	0,12322***	0,08075***	0,13165***
	(0,00087)	(0,00107)	(0,00089)
Comércio	0,07356***	0,03319***	0,08215***
	(0,00087)	(0,00107)	(0,00089)
Serviços	0,08965***	0,04022***	0,09844***
	(0,00087)	(0,001069)	(0,00089)
Serviços Domésticos	-0,06659***	-0,10651***	-0,05789***
	(0,007135)	(0,007406)	(0,00722)
Organismos Internacionais	0,413797***	0,29961***	0,42740***
	(0,011267)	(0,01126)	(0,01128)
Dummies de ano	SIM	SIM	SIM
Constante	-0,09385***	-0,52629***	-0,06557***
	(0,001689)	(0,00215)	(0,00201)
Observações (NxT)	72.808.032	65.503.892	71.596.428
R^2 Ajustado	0,6362	0,6192	0,6346
Prob>F	0,000	0,000	0,000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.4 – Prêmio salarial da especialização ocupacional de habilidades cognitivas: Efeitos Fixos

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	**EF (trabalhador)**	**EF (firma)**	**EF (trabalhador e firma)**
Variáveis explicativas			
$ln(EO_h)$	0,06265***	-0,04403***	0,00707***
	(0,00106)	(0,00639)	(0,00073)
Características do trabalhador			
Tempo no emprego	0,00127***	0,00340***	0,00123***
	(0,0000085)	(0,000053)	(0,000011)
Tempo no emprego²	-0,0000047***	-0,0000044***	-0,0000047
	(0,000000033)	(0,000000202)	(0,000000027)
Escolaridade (Categoria base - Analfabeto)			
Fundamental Incompleto	-0,01417	-0,21135***	-0,00973
	(0,02735)	(0,06498)	(0,02364)
Fundamental Completo	-0,01761	-0,15615**	-0,008701
	(0,02732)	(0,06589)	(0,02361)
Médio Incompleto	-0,03024	-0,13039**	-0,02112
	(0,02734)	(0,063602)	(0,02363)
Médio Completo	-0,02156	-0,06462	-0,01458
	(0,02729)	(0,06369)	(0,02358)
Superior Incompleto	-0,01332	0,05364	-0,02214
	(0,02731)	(0,06376)	(0,02359)
Superior Completo	0,10861***	0,46525***	0,06999***
	(0,02730)	(0,06404)	(0,02358)
Características do emprego			
Tamanho do Estabelecimento (Categoria base - Micro)			
Pequeno	0,06961***	0,01698***	0,02804***
	(0,00046)	(0,00207)	(0,000407)
Médio	0,14035***	0,02724***	0,05483***
	(0,00069)	(0,00432)	(0,00064)
Grande	0,19769***	0,03198***	0,05802***
	(0,00093)	(0,00844)	(0,00086)

(continua)

(conclusão)

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	**EF(trabalhador)**	**EF (firma)**	**EF (trabalhador e firma)**
Setor de Atividade (Categoria base - Agropecuária)			
Extração Mineral	0,21489***	-0,02224	-0,004305
	(0,01135)	(0,03258)	(0,00933)
Transformação	0,03621***	-0,00539	-0,01054**
	(0,00674)	(0,01807)	(0,00495)
Construção Civil	0,05899***	0,00213	-0,0061004
	(0,00684)	(0,01926)	(0,0050807)
Comércio	-0,01689**	-0,01287	-0,01905***
	(0,00671)	(0,01837)	(0,00488)
Serviços	0,01737***	0,00352	-0,00548
	(0,00668)	(0,01799)	(0,00485)
Serviços Domésticos	0,03755*	-0,03307	-0,002697
	(0,02236)	(0,03549)	(0,01302)
Organismos Internacionais	0,06721***	-0,02894	0,03836***
	(0,01193)	(0,02726)	(0,00746)
Dummies de ano	SIM	SIM	SIM
Constante	1,74339***	1,46458***	1,71309***
	(0,02823)	(0,06738)	(0,02413)
Observações (NxT)	16.227.793	16.227.793	16.227.793
N. de grupos	5.727.045	960.586	6.961.754
R^2 (overall)	0,318	0,4655	0,2659
R^2 (within)	0,8354	0,5418	0,8832
R^2 (between)	0,2782	0,4608	0,2469
Prob>F	0,000	0,000	0,000
Hausman χ^2	2313767,24	88187,37	2199490,44
Prob>χ^2	0,000	0,000	0,000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.5 – Prêmio salarial da especialização ocupacional de habilidades sociais: Efeitos Fixos

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	**EF (trabalhador)**	**EF (firma)**	**EF (trabalhador e firma)**
Variáveis explicativas			
$ln(EO_h)$	0,01967***	-0,07361***	-0,09359***
	(0,00078)	(0,00288)	(0,001008)
Características do trabalhador			
Tempo no emprego	0,001705***	0,00341***	0,00233***
	(0,0000031)	(0,000027)	(0,0000046)
Tempo no emprego²	-0,0000043***	-0,0000043***	-0,0000043***
	(0,000000014)	(0,00000015)	(0,000000026)
Escolaridade (Categoria base - Analfabeto)			
Fundamental Incompleto	-0,01027***	0,01725***	-0,01249***
	(0,00132)	(0,002104)	(0,00182)
Fundamental Completo	-0,01726***	0,05157***	-0,00574***
	(0,00132)	(0,00216)	(0,00183)
Médio Incompleto	-0,02383***	0,05061***	-0,00847***
	(0,00133)	(0,00216)	(0,00183)
Médio Completo	-0,00964***	0,10728***	0,02645***
	(0,00132)	(0,00220)	(0,00182)
Superior Incompleto	0,02716***	0,20376***	0,16193***
	(0,00137)	(0,00239)	(0,00187)
Superior Completo	0,10393***	0,52970***	0,41174***
	(0,00138)	(0,00312)	(0,00193)
Características do emprego			
Tamanho do Estabelecimento (Categoria base - Micro)			
Pequeno	0,04915***	0,01716***	0,07508***
	(0,00015)	(0,00046)	(0,00019)
Médio	0,08961***	0,01459***	0,14261***
	(0,00025)	(0,00116)	(0,00032)
Grande	0,10517***	0,00762***	0,16636***
	(0,00033)	(0,00282)	(0,00044)

(continua)

(conclusão)

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	**EF(trabalhador)**	**EF (firma)**	**EF (trabalhador e firma)**
Setor de Atividade (Categoria base - Agropecuária)			
Extração Mineral	0,18691***	0,05317	0,51449***
	(0,00465)	(0,03403)	(0,00806)
Transformação	0,02991***	0,01681***	0,04575***
	(0,00185)	(0,00535)	(0,00251)
Construção Civil	0,04628***	0,01822***	0,11372***
	(0,00191)	(0,00683)	(0,00258)
Comércio	-0,00832***	0,00991*	-0,03639***
	(0,00183)	(0,00525)	(0,00248)
Serviços	-0,00892***	0,01336**	-0,02405***
	(0,00183)	(0,00535)	(0,00248)
Serviços Domésticos	-0,03807***	0,01478	-0,0736***
	(0,005805)	(0,01319)	(0,00799)
Organismos Internacionais	-0,00905	-0,02639	0,19507***
	(0,00611)	(0,05141)	(0,01065)
Dummies de ano	SIM	SIM	SIM
Constante	1,10148***	0,91529***	0,96723***
	(0,002323)	(0,00619)	(0,00316)
Observações (NxT)	79.373.026	79.373.026	79.373.026
N. de grupos	25.738.055	3.208.045	25.665.735
R^2 (overall)	0,4780	0,5495	0,5567
R^2 (within)	0,8689	0,6521	0,8181
R^2 (between)	0,3946	0,5441	0,4844
Prob>F	0,000	0,000	0,000
Hausman χ^2	6533692,75	253687,30	1501570,15
Prob>χ^2	0,000	0,000	0,000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.6 – Prêmio salarial da especialização ocupacional de habilidades manuais: Efeitos Fixos

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	**EF (trabalhador)**	**EF (firma)**	**EF (trabalhador e firma)**
Variáveis explicativas			
$ln(EO_h)$	-0,00484***	0,04089***	0,05219***
	(0,00048)	(0,00313)	(0,00068)
Características do trabalhador			
Tempo no emprego	0,00147***	0,00275***	0,00186***
	(0,0000028)	(0,000036)	(0,0000043)
Tempo no emprego²	-0,000004***	-0,0000037***	-0,000004***
	(0,000000014)	(0,00000021)	(0,000000025)
Escolaridade (Categoria base - Analfabeto)			
Fundamental Incompleto	0,00618***	0,07294***	0,03231***
	(0,00072)	(0,00386)	(0,00106)
Fundamental Completo	0,00355***	0,10017***	0,05145***
	(0,00073)	(0,00394)	(0,00107)
Médio Incompleto	0,000299	0,09247***	0,05456***
	(0,00074)	(0,00395)	(0,00108)
Médio Completo	0,01044***	0,13321***	0,09891***
	(0,00073)	(0,00396)	(0,00107)
Superior Incompleto	0,03567***	0,24661***	0,24597***
	(0,00109)	(0,00441)	(0,00153)
Superior Completo	0,05639***	0,44465***	0,42061***
	(0,00115)	(0,00705)	(0,00194)
Características do emprego			
Tamanho do Estabelecimento (Categoria base - Micro)			
Pequeno	0,04793***	0,01259***	0,05166***
	(0,00013)	(0,00067)	(0,00017)
Médio	0,09409***	0,019309***	0,10184***
	(0,00019)	(0,00147)	(0,00026)
Grande	0,12844***	0,02364***	0,13911***
	(0,00026)	(0,00292)	(0,00038)

(continua)

(conclusão)

Variável Dependente: Logaritmo dos salários reais por Habilidade			
Variável	**EF(trabalhador)**	**EF (firma)**	**EF (trabalhador e firma)**
Setor de Atividade (Categoria base - Agropecuária)			
Extração Mineral	0,13284***	0,00139	0,34889***
	(0,00204)	(0,01498)	(0,00376)
Transformação	0,04385***	-0,00116	0,07464***
	(0,000903)	(0,00818)	(0,00117)
Construção Civil	0,03859***	0,0023	0,10253***
	(0,00092)	(0,008301)	(0,00118)
Comércio	0,00193**	-0,00479	0,03116***
	(0,000909)	(0,00799)	(0,00117)
Serviços	0,01009***	0,00372	0,02016***
	(0,000906)	(0,00803)	(0,00117)
Serviços Domésticos	-0,01191**	0,00312	-0,06838***
	(0,00501)	(0,01104)	(0,00764)
Organismos Internacionais	0,08221***	0,04336	0,22244***
	(0,00702)	(0,031108)	(0,01197)
Dummies de ano	SIM	SIM	SIM
Constante	1,03778***	1,03179***	0,99041***
	(0,001202)	(0,00809)	(0,00163)
Observações (NxT)	72.808.032	72.808.032	72.808.032
N. de grupos	23.505.302	2.293.037	25.256.703
R^2 (overall)	0,5271	0,5528	0,5528
R^2 (within)	0,8829	0,6985	0,8361
R^2 (between)	0,4348	0,5862	0,4719
Prob>F	0,000	0,000	0,000
Hausman χ^2	2530729,89	121874,00	671495,12
Prob> χ^2	0,000	0,000	0,000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.7 – Prêmio salarial da especialização ocupacional por habilidades: POLS e 2SLS

Variável Dependente: Logaritmo dos salários reais			
	***Pooled* OLS**		
Variável	**Cognitiva**	**Social**	**Manual**
$ln(EO_h)$	0,13839***	-0,1790***	0,03016***
	(0,00026)	(0,00054)	(0,00052)
Observações (NxT)	168.407.777	168.408.843	168.408.844
R^2 Ajustado	0,6241	0,6228	0,6219
Prob>F	0,0000	0,0000	0,0000
2SLS (IV Históricas) – 2º estágio			
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	0,472427***	0,907864***	-0,83779***
	(0,000618)	(0,00194)	(0,00137)
Observações (NxT)	153.345.152	153.345.152	153.345.152
R^2 Ajustado	0,6146	0,5964	0,6021
Prob>F	0,0000	0,0000	0,0000
2SLS (IV Históricas) – 1º estágio			
Instrumentos:			
log. da população de 1920	-0,01266***	0,04053***	-0,04257***
	(0,00005)	(0,00004)	(0,00003)
log. da população de 1940	-0,03857***	-0,0403***	0,05989***
	(0,00006)	(0,00004)	(0,00004)
log. da população de 1991	-0,0447***	0,02335***	-0,0261***
	(0,00015)	(0,00009)	(0,00008)
log. do NCGP de 1991	0,0631***	-0,00481***	-0,01063***
	(0,00009)	(0,00007)	(0,00007)
log. do NCGP de 2010	0,10573***	0,00628***	-0,01576***
	(0,00015)	(0,0001)	(0,0001)
2SLS (IV Bartik) – 2º estágio			
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	1,05064***	-0,08165***	0,03321***
	(0,00245)	(0,002067)	(0,00144)
Observações (NxT)	167.893.096	168.305.622	166.387.993
R^2 Ajustado	0,5323	0,6225	0,6211
Prob>F	0,0000	0.0000	0,0000
2SLS (IV Bartik) – 1º estágio			
Instrumentos:			
log. bartik	0,07922***	0,2116***	-0,0888***
	(0,00011)	(0,00021)	(0,00005)

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.8 – Prêmio salarial da aglomeração: POLS e 2SLS

Variável Dependente: Logaritmo dos salários reais	
***Pooled* OLS**	
$ln(L_h)$	0,02898***
	(0,000038)
Observações (NxT)	168.408.851
R^2 Ajustado	0,6276
Prob>F	0,0000
2SLS (IV Históricas) – 2º estágio	
$ln(L_h)$	0,03012***
	(0,00004)
Observações (NxT)	153.345.152
R^2 Ajustado	0,6262
Prob>F	0,0000
2SLS (IV Históricas) – 1º estágio	
Instrumentos:	
log. da população de 1920	-0,0576***
	(0,00009)
log. da população de 1940	0,02524***
	(0,000105)
log. da população de 1991	0,6854***
	(0,00022)
log. do NCGP de 1991	-0,1796***
	(0,0002)
log. do NCGP de 2010	0,5699***
	(0,00026)
2SLS (IV Bartik) – 2º estágio	
$ln(L_h)$	-0,027238***
	(0,000109)
Observações (NxT)	167.717.169
R^2 Ajustado	0,6060
Prob>F	0,0000
2SLS (IV Bartik) – 1º estágio	
Instrumentos:	
log. bartik	1,9842***
	(0,00123)

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.9 – Prêmio salarial da especialização ocupacional por habilidades: Efeitos Fixos

Variável Dependente: Logaritmo dos salários reais			
	***Pooled* OLS**		
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	0,13839***	-0,1790***	0,03016***
	(0,00026)	(0,00054)	(0,00052)
Observações (NxT)	168.407.777	168.408.843	168.408.844
R^2 Ajustado	0,6241	0,6228	0,6219
Prob>F	0,0000	0,0000	0,0000
	Efeito Fixo do trabalhador		
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	0,01984***	0,01421***	-0,03566***
	(0,000166)	(0,00042)	(0,0004)
Observações (NxT)	168.407.777	168.408.843	168.408.844
N. de grupos	46.781.964	46.782.167	46.782.163
R^2 (overall)	0,4623	0,4614	0,4615
R^2 (within)	0,8585	0,8585	0,8585
R^2 (between)	0,3565	0,3554	0,3555
Prob>F	0,0000	0,0000	0,0000
Hausman χ^2	16.800.000	16.800.000	16.600.000
Prob>χ^2	0,0000	0,0000	0,0000
	Efeito Fixo da firma		
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	-0,03446***	-0,06879***	0,07222 ***
	(0,001013)	(0,00238)	(0,00246)
Observações (NxT)	168.407.777	168.408.843	168.408.844
N. de grupos	3.897.191	3.897.242	3.897.239
R^2 (overall)	0,5498	0,5519	0,5513
R^2 (within)	0,6385	0,6385	0,6385
R^2 (between)	0,5755	0,5792	0,5784
Prob>F	0,0000	0,0000	0,0000
Hausman χ^2	290.798,93	255.250,86	240.935,37
Prob>χ^2	0,0000	0,0000	0,0000
	Efeito Fixo de trabalhador e firma		
Variável	Cognitiva	Social	Manual
$ln(EO_h)$	0,08306***	-0,09983***	-0,00141**
	(0,00026)	(0,00062)	(0,00058)
Observações (NxT)	168.407.777	168.408.843	168.408.844
N. de grupos	34.338.195	34.338.250	34.338.248
R^2 (overall)	0,5742	0,5730	0,5723
R^2 (within)	0,7359	0,7354	0,7352
R^2 (between)	0,4553	0,4540	0,4532
Prob>F	0,0000	0,0000	0,0000
Hausman χ^2	1.695.499,28	1.696.815,10	1.689.984,30
Prob>χ^2	0,0000	0,0000	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.10 – Prêmio salarial da aglomeração: Efeitos Fixos

Variável Dependente: Logaritmo dos salários reais	
***Pooled* OLS**	
$ln(L_h)$	0,02898***
	(0,000038)
Observações (NxT)	168.408.851
R^2 Ajustado	0,6276
Prob>F	0,0000
Efeito Fixo do trabalhador	
$ln(L_h)$	0,01447 ***
	(0,00006)
Observações (NxT)	168.408.851
N. de grupos	46.782.167
R^2 (overall)	0,4678
R^2 (within)	0,8587
R^2 (between)	0,3636
Prob>F	0,0000
Hausman χ^2	16.300.000
Prob>χ^2	0,0000
Efeito Fixo da firma	
$ln(L_h)$	-0,002941***
	(0,0008179)
Observações (NxT)	168.408.851
N. de grupos	3.897.242
R^2 (overall)	0,5496
R^2 (within)	0,6384
R^2 (between)	0,5744
Prob>F	0,0000
Hausman χ^2	274.589,46
Prob>χ^2	0,0000
Efeito Fixo de trabalhador e firma	
$ln(L_h)$	0,028945***
	(0,00005)
Observações (NxT)	168.408.851
N. de grupos	34.338.250
R^2 (overall)	0,5791
R^2 (within)	0,7383
R^2 (between)	0,4606
Prob>F	0,0000
Hausman χ^2	1.665.509,25
Prob>χ^2	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tabela B.11 – Decomposição do efeito da especialização cognitiva sobre os salários: *Pooled* de dados

Variável Dependente: Efeito da especialização cognitiva sobre os salários locais				
		***Pooled* OLS**		
Variável	**EEC**	**EEC(*i*)**	**EEC(*j*)**	**EEC(*ij*)**
Log. da densidade	-0,0273***	-0,0141***	-0,00024	0,00471
	(0,00537)	(0,00302)	(0,00357)	(0,00293)
Log. da área	-0,0401***	-0,0142***	0,0065	0,00082
	(0,00948)	(0,00423)	(0,0049)	(0,00461)
Fronteirio	0,05932**	-0,03669	-0,05138	-0,03874
	(0,02751)	(0,02429)	(0,0315)	(0,02529)
Turístico	-0,01484	-0,0368***	0,00952	-0,01717
	(0,01522)	(0,01348)	(0,01302)	(0,01162)
Norte	0,03695*	-0,00976	-0,04119*	0,00439
	(0,01922)	(0,02122)	(0,02258)	(0,01328)
Nordeste	0,103***	0,0357***	0,01029	0,01369
	(0,01349)	(0,00662)	(0,01617)	(0,01130)
Sul	-0,03161**	0,00171	-0,0213	-0,0122
	(0,01526)	(0,00753)	(0,01323)	(0,00945)
Centro-Oeste	-0,06269	-0,03359*	0,00017	0,00276
	(0,05853)	(0,01899)	(0,0117)	(0,01662)
Constante	0,5106***	0,3921***	0,020481	0,04948
	(0,06441)	(0,02509)	(0,03334)	(0,03336)
Dummies de ano	SIM	SIM	SIM	SIM
Observações	3.492	3.457	3.296	3.285
R^2 Ajustado	0,6844	0,5514	0,1749	0,2791
Prob>F	0,0000	0,0000	0,0000	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Notas: Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01; Resultados assumem peso analítico proporcional ao número de trabalhadores em cada Arranjo Populacional.

Tabela B.12 – Decomposição do efeito da especialização social sobre os salários: *Pooled* de dados

Variável Dependente: Efeito da especialização social sobre os salários locais				
		***Pooled* OLS**		
Variável	**EES**	**EES(*i*)**	**EES(*j*)**	**EES(*ij*)**
Log. da densidade	-0,0529***	-0,0264***	0,00475	-0,00158
	(0,01308)	(0,00403)	(0,0088)	(0,00527)
Log. da área	-0,03344**	-0,0189***	-0,00373	0,00815
	(0,0155)	(0,00547)	(0,0118)	(0,00868)
Fronteiriço	0,09401**	0,03025	0,01608	0,05018
	(0,04012)	(0,05752)	(0,04362)	(0,05783)
Turístico	-0,01995	-0,0519**	-0,0148	0,00246
	(0,04254)	(0,02533)	(0,0418)	(0,0289)
Norte	0,12213**	0,00841	0,1861***	-0,02098
	(0,05182)	(0,02395)	(0,0477)	(0,035)
Nordeste	0,3642***	0,14761***	0,0693***	-0,0391***
	(0,03748)	(0,01491)	(0,02438)	(0,01338)
Sul	-0,1187***	-0,03358**	-0,0266	-0,05568**
	(0,02589)	(0,0132)	(0,0319)	(0,02373)
Centro-Oeste	-0,0089	-0,03544**	-0,02176	-0,03967
	(0,03824)	(0,014)	(0,05448)	(0,02934)
Constante	0,587***	0,18204***	0,03177	0,03615
	(0,14078)	(0,03863)	(0,0788)	(0,05385)
Dummies de ano	SIM	SIM	SIM	SIM
Observações	3.686	3.672	3.604	3.598
R^2 Ajustado	0,7522	0,6579	0,2883	0,1340
Prob>F	0,0000	0,0000	0,0000	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018).
Notas: Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01; Resultados assumem peso analítico proporcional ao número de trabalhadores em cada Arranjo Populacional.

Tabela B.13 – Decomposição do efeito da especialização manual sobre os salários: *Pooled* de dados

Variável Dependente: Efeito da especialização manual sobre os salários locais				
		***Pooled* OLS**		
Variável	**EEM**	**EEM(*i*)**	**EEM(*j*)**	**EEM(*ij*)**
Log. da densidade	-0,0224**	-0,0101*	0,0055	-0,00019
	(0,01131)	(0,0055)	(0,00813)	(0,00543)
Log. da área	0,0222	0,0231***	0,0267**	0,02795***
	(0,01584)	(0,00758)	(0,0116)	(0,00851)
Fronteiriço	0,1006**	0,07111*	0,1242**	0,05834
	(0,04918)	(0,03772)	(0,0558)	(0,0471)
Turístico	0,06958	0,06099**	0,0558*	0,02936
	(0,04981)	(0,02648)	(0,0323)	(0,03406)
Norte	0,1947***	0,05581	0,04178	-0,0427*
	(0,0696)	(0,04619)	(0,05638)	(0,02323)
Nordeste	0,3711***	0,19523***	-0,0296	0,0541***
	(0,03547)	(0,01975)	(0,0434)	(0,01383)
Sul	-0,0956***	-0,02847	0,00105	-0,00689
	(0,03096)	(0,01762)	(0,02799)	(0,02529)
Centro-Oeste	0,06121	0,02046	0,0541**	0,024
	(0,03921)	(0,01872)	(0,02645)	(0,03051)
Constante	-0,2322*	-0,2828***	-0,511***	-0,36097***
	(0,12007)	(0,04687)	(0,07214)	(0,05319)
Dummies de ano	SIM	SIM	SIM	SIM
Observações	3.669	3.653	3.588	3.582
R^2 Ajustado	0,4861	0,4358	0,1002	0,1167
Prob>F	0,0000	0,0000	0,0000	0,0000

Fonte: Elaboração própria (2021), com base nos dados da RAIS – ME (2009-2018)
Notas: Erro-padrão robusto em parênteses. * p < 0.10, ** p < 0.05, *** p < 0.01; Resultados assumem peso analítico proporcional ao número de trabalhadores em cada Arranjo Populacional.

ANEXOS

ANEXO A – Diferencial salarial no Modelo Ricardiano do Mercado de Trabalho de Acemoglu e Autor (2011)

Para a compreensão do diferencial salarial entre trabalhadores se faz necessário avaliar a hipótese de compensação do mercado, isto é, como a oferta de trabalhadores está organizada no equilíbrio, além de se compreender a determinação do índice de preços das tarefas desempenhadas. No que se refere à oferta de trabalhadores, avaliando a condição de compensação do mercado, define-se:

$\int_0^1 l(i)di \leq L$, $\int_0^1 m(i)di \leq M$ *e* $\int_0^1 h(i)di \leq H$,

Em que, L, M e H correspondem ao total de trabalhadores com habilidades baixa, média e alta, respectivamente. Considerando $\int_0^{I_L} l(i)di \leq L$ e definindo I como um conjunto de ocupações/tarefas, obtém-se:

$$l(i)I_L - l(i)0 = L$$

$$l(i) = \frac{L}{I_L} \; para\; todo\; i < I_L$$

Considerando $\int_{I_L}^{I_H} m(i)di \leq M$, tem-se:

$$m(i)I_H - m(i)I_L = M$$

$$m(i)(I_H - I_L) = M$$

$$m(i) = \frac{M}{I_H - I_L} \; para\; todo\; I_H > \; i > I_L$$

Considerando $\int_{I_H}^{1} h(i)di \leq H$, tem-se:

$$h(i)1 - h(i)I_H = H$$

$$h(i)(1 - I_H) = H$$

$$h(i) = \frac{H}{1 - I_H} \; para\; todo\; i > I_H$$

Assim, determina-se a oferta de equilíbrio para os três tipos de trabalhadores de habilidade baixa, média e alta. A condição de oferta categorizada por habilidades dos trabalhadores é de fundamental importância para a determinação dos diferenciais salariais entre os mesmos. No que se refere ao índice de preços das tarefas desempenhadas, denomina-se como aquele que

reflete a compensação entre a diferença no salário de indivíduos de mesma habilidade que desempenha tarefas diferentes (*i* e *i'*), isto é:

$$w_L = p(i')A_L\alpha_L(i')$$
$$w_L = p(i)A_L\alpha_L(i)$$

Logo:

$$P_L \equiv p(i')A_L\alpha_L(i') = p(i)A_L\alpha_L(i)$$

Em que, P_L representa o índice de preço das tarefas desempenhadas pelos trabalhadores de baixa habilidade. O mesmo pode ser aplicado para os demais tipos de habilidades:

$$P_M \equiv p(i')A_M\alpha_M(i') = p(i)A_M\alpha_M(i)$$
$$P_H \equiv p(i')A_H\alpha_H(i') = p(i)A_H\alpha_H(i)$$

Comparando duas tarefas realizadas por trabalhadores de alta e média habilidade $I_M < i < I_H < i'$ e considerando a função de produção de escolha de bem final como numerário: $p(i)y(i) = Y, para\ todo\ i\ \in [0,1]$, tem-se:

$$p(i)A_M\alpha_M(i)m(i) = p(i')A_H\alpha_H(i')h(i')$$

Utilizando as expressões de índices de preço e compesação no mercado, obtém-se:

$$P_M A_M m(i) = P_H A_H h(i)$$

$$P_M A_M(\frac{M}{I_H - I_L}) = P_H A_H(\frac{H}{1 - I_H})$$

$$P_M \frac{A_M M}{I_H - I_L} = P_H(\frac{A_H H}{1 - I_H})$$

$\frac{P_H}{P_M} = \frac{A_M M}{I_H - I_L}(\frac{A_H H}{1 - I_H})^{-1}$Agora, comparando duas tarefas realizadas por trabalhadores de média e baixa habilidade $I_L < i < I_M < i'$ e considerando a função de produção de escolha de bem final como numerário: $p(i)y(i) = Y, para\ todo\ i\ \in [0,1]$, tem-se:

$$p(i)A_L\alpha_L(i)l(i) = p(i')A_M\alpha_M(i')m(i')$$

Utilizando as expressoões de índices de preço e compesação no mercado, obtém-se:

$$P_L A_L l(i) = P_M A_M m(i)$$

$$P_L A_L (\frac{L}{I_L}) = P_M A_M (\frac{M}{I_H - I_L})$$

$$P_L (\frac{A_L L}{I_L}) = P_M (\frac{A_M M}{I_H - I_L})$$

$$\frac{P_M}{P_L} = \frac{A_L L}{I_L} (\frac{A_M M}{I_H - I_L})^{-1}$$

O equílibrio salarial, mais especificamente, os diferenciais salariais podem ser obtidos considerando as seguintes definições de salários:

$$w_L = P_L A_L$$
$$w_M = P_M A_M$$
$$w_H = P_H A_H$$

A relação entre os níveis salarias expressam a desigualdade salarial existente entre os trabalhadores de diferentes perfis de habilidades. Considerando as equações de preços relativos pode-se expressar a relação entre os salários de alta e média habilidade (ou o diferencial salarial dos trabalhadores de alta e média habilidade como):

$$\frac{w_H}{w_M} = \frac{P_H A_H}{P_M A_M}$$

Substituindo a relação de preços $\frac{P_H}{P_M}$ na equação anterior tem-se:

$$\frac{w_H}{w_M} = \frac{A_H}{A_M} \frac{A_M M}{I_H - I_L} (\frac{A_H H}{1 - I_H})^{-1}$$

$$\frac{w_H}{w_M} = \frac{A_H}{A_M} \frac{A_M M}{I_H - I_L} \Big/ \frac{A_H H}{1 - I_H}$$

$$\frac{w_H}{w_M} = \frac{M}{I_H - I_L} \Big/ \frac{H}{1 - I_H}$$

$$\frac{w_H}{w_M} = \frac{M}{I_H - I_L} \frac{1 - I_H}{H}$$

$$\frac{w_H}{w_M} = \frac{M}{H} \frac{1 - I_H}{I_H - I_L}$$

$$\frac{w_H}{w_M} = (\frac{H}{M})^{-1} \frac{1 - I_H}{I_H - I_L}$$

Considerando as equações de preços relativos, pode-se expressar a relação entre os salários de média e baixa habilidade (ou o diferencial salarial dos trabalhadores de média e baixa habilidade como):

$$\frac{w_M}{w_L} = \frac{P_M A_M}{P_L A_L}$$

Substituindo a relação de preços $\frac{P_M}{P_L}$ na equação anterior tem-se:

$$\frac{w_M}{w_L} = \frac{A_M}{A_L} \frac{A_L L}{I_L} (\frac{A_M M}{I_H - I_L})^{-1}$$

$$\frac{w_M}{w_L} = \frac{A_M L}{I_L} \Big/ \frac{A_M M}{I_H - I_L}$$

$$\frac{w_M}{w_L} = \frac{L}{I_L} \Big/ \frac{M}{I_H - I_L}$$

$$\frac{w_M}{w_L} = \frac{L}{I_L} \frac{I_H - I_L}{M}$$

$$\frac{w_M}{w_L} = \frac{L}{M} \frac{I_H - I_L}{I_L}$$

$$\frac{w_M}{w_L} = (\frac{M}{L})^{-1} \frac{I_H - I_L}{I_L}$$