Universidade Federal de Pernambuco - UFPE
Centro Acadêmico do Agreste - CAA
Departamento de Economia
Programa de Pós-Graduação em Economia - PPGECON

Edilberto Tiago de Almeida

Labor Pooling como fator de aglomeração: evidências para o Brasil no período 2002-2014

Dissertação de Mestrado

Caruaru, Pernambuco
2017

Universidade Federal de Pernambuco - UFPE
Centro Acadêmico do Agreste - CAA
Departamento de Economia
Programa de Pós-Graduação em Economia - PPGECON

Edilberto Tiago de Almeida

Labor Pooling como fator de aglomeração: evidências para o Brasil no período 2002-2014

Dissertação apresentada ao **Programa de Pós-Graduação em Economia - PPGECON** da **Universidade Federal de Pernambuco - UFPE/Centro Acadêmico do Agreste - CAA** como requisito parcial para obtenção do título de Mestre em Economia.

Área de Concentração: Economia Regional

Orientadora: Dra. Roberta de Moraes Rocha

Caruaru, Pernambuco
2017

Catalogação na fonte:
Bibliotecária – Simone Xavier CRB/4 - 1242

A447l Almeida, Edilberto Tiago de.
Labor Pooling como fator de aglomeração: evidências para o Brasil no período 2002-2014. / Edilberto Tiago de Almeida. – 2017.
72f., il. ; 30 cm.

Orientadora: Roberta de Moraes Rocha
Dissertação (Mestrado) – Universidade Federal de Pernambuco, CAA, Programa de Pós-Graduação em Economia, 2017.
Inclui Referências.

1. Aglomerados industriais. 2. Indústrias. 3. Labor Pooling. 5. Modelos econômicos. I. Rocha, Roberta de Moraes (Orientadora). II. Título.

330 CDD (23. ed.) UFPE (CAA 2017-050)

Edilberto Tiago de Almeida

Labor Pooling como fator de aglomeração: evidências para o Brasil no período 2002-2014

Dissertação apresentada ao **Programa de Pós-Graduação em Economia - PPGECON** da **Universidade Federal de Pernambuco - UFPE/Centro Acadêmico do Agreste - CAA** como requisito parcial para obtenção do título de Mestre em Economia. A Comissão Examinadora composta pelos professores abaixo considera o Candidato Edilberto Tiago de Almeida **APROVADO**.

Defesa realizada no **Centro Acadêmico do Agreste - CAA** da UFPE em Caruaru, Pernambuco no dia 14 de Fevereiro de 2017.

Professora Dra. Roberta de Moraes Rocha
Orientadora
Universidade Federal de Pernambuco - UFPE

Professor Dr. André Matos Magalhães
Examinador Interno
Universidade Federal de Pernambuco - UFPE

Professor Dr. Raul da Mota Silveira Neto
Examinador Interno
Universidade Federal de Pernambuco - UFPE

Professora Dra. Carla Calixto da Silva
Examinador Externo
Universidade Federal da Paraíba - UFPB

Dedico esta Dissertação ao meu Pai ***Edmilson*** *e a minha Mãe* ***Mara****.*

Agradecimentos

Agradeço a Deus por tudo. Por todas as conquistas ao longo desta trajetória acadêmica, por todo o aprendizado e amigos que tenho feito.

Agradeço aos meus pais pelo apoio sempre.

Agradeço a minha orientadora, a professora Dra. Roberta Rocha pelos ensinamentos durante os créditos e na confecção desta Dissertação e pelo apoio e confiança no período em que precisei me afastar do Projeto de Dissertação para me dedicar a seleção do doutorado.

Agradeço à Fundação de Amparo à Ciência e Tecnologia de Pernambuco (FACEPE), pelo apoio financeiro para o desenvolvimento desta pesquisa.

Agradeço aos membros da banca de qualificação do meu Projeto de Dissertação, o professor Dr. André Matos Magalhães, professor Dr. Raul da Mota Silveira Neto e a professora Dra. Carla Calixto da Silva pela disponibilidade e contribuições valiosas para o desenvolvimento da pesquisa. Ao mesmo tempo, agradeço novamente aos professores e professora pela disponibilidade para participar da banca de defesa da versão final da Dissertação.

Agradeço a todos os professores com quem tive a honra de aprender durante os créditos no curso de mestrado em economia do PPGECON/UFPE. A professora Monaliza Ferreira, Alane Alves, Sónia Gomes e Roberta Rocha, o professor Carlos Amorim, Diogo e Wellington Ribeiro.

Agradeço à secretária do PPGECON, Débora Pereira e ao bolsita Rafael pelo auxílio nos procedimentos durante todo o curso.

Agradeço a professora Dra. Carla Calixto da Silva, minha orientadora na graduação, pelo apoio sempre me incentivando e ensinando, com ela aprendi a gostar da pesquisa. Ela foi uma pessoa essencial que não mediu esforços para me ajudar nos procedimentos de antecipação da graduação para o ingresso no mestrado. Por isso, por ser um exemplo de profissionalismo e humildade, sou imensamente grato.

A todos os colegas da minha turma, e aos amigos que certamente levariei comigo, Alexsandra, Nanísia, Alan Umburana, Priscila, Andson e Josué.

Aos meus amigos que estão presentes desde a graduação Iza, Manoel e Rosy.

Agradeço a todos que contribuíram de forma direta ou indiretamente para o desenvolvimento desta pesquisa.

Resumo

Esta pesquisa tem como objetivo obter evidências sobre a importância do *labor pooling* para explicar a aglomeração industrial no Brasil. Com este propósito, a investigação empírica foi dividida em duas etapas. A primeira tem como objetivo quantificar a aglomeração da indústria da transformação nas microrregiões que compõem o país no período 2002-2014, através do índice de Ellison e Glaeser (1997). Na segunda etapa, utiliza-se como base o modelo teórico proposto por Krugman (1991) (*Labor Market Pooling Model*) com adaptações de Overman e Puga (2010), e a investigação empírica segue a estratégia destes últimos que analisa como as firmas reagem a choques no mercado de trabalho que influenciam a sua produtividade. A ideia é, conforme o modelo teórico, se as firmas se beneficiam da concentração de mão de obra especializada, choques exógenos mais heterogêneos permitem que estas ajustem melhor sua produção e emprego sem incorrer em mudanças significativas nos salários. Com o objetivo de testar esta hipótese, modelos de regressão são estimados regredindo o índice de Ellison e Glaeser (1997) em função de uma variável *proxy* para o *labor pooling*, que tenta captar choques exógenos no mercado de trabalho, controlando tanto para características setoriais observadas que variam no tempo quanto pelo efeito fixo setorial. Os resultados são consistentes com uma redução no nível de concentração industrial no período de 2002 a 2014. Quando consideradas trajetórias de setores específicos, foi possível observar que os mais intensivos em tecnologia, em média, são mais concentrados que os menos intensivos em tecnologia, evidências que estão de acordo com outros estudos na literatura. No que concerne ao *labor pooling*, foi possível obter evidências que corroboram a hipótese teórica do modelo de que este fator local se relaciona positivamente com a aglomeração espacial da indústria, indicando que os setores industriais podem se beneficiar da concentração de trabalhadores. Através destes resultados, as evidências apontam que as firmas tendem a se concentrar espacialmente próximas de onde estão localizados os mercados de trabalho que concentram trabalhadores com as características exigidas em seus respectivos processos produtivos. Intuitivamente tanto as firmas podem considerar a localização da força de trabalho como fator importante em sua localização, quanto os trabalhadores podem se concentrar espacialmente em função da localização das firmas. Nesse aspecto, este estudo se concentrou no primeiro caso e fornece resultados consistentes em relação ao problema da causalidade reversa e demais fontes de endogeneidade. Adicionalmente, as evidências aqui apresentadas podem auxiliar na formulação de políticas norteadoras do direcionamento da atividade produtiva no espaço.

Palavras-chave: *Labor Pooling.* Aglomeração. Indústrias.

Abstract

This research aims to obtain evidence about the importance of labor pooling to explain industrial agglomeration in Brazil. For this purpose, the empirical investigation was divided into two stages. The first one aims at quantifying the agglomeration of the transformation industry in the micro-regions that make up the country in the period 2002-2014, through the index of Ellison and Glaeser (1997). In the second step, the theoretical model proposed by Krugman (1991) (Labor Market Pooling Model) with adaptations of Overman and Puga (2010) is used as base, and the empirical investigation follows the latter's strategy that analyses how firms react to shocks in the labor market that influence their productivity. The idea is, according to the theoretical model, if firms benefit from the concentration of specialized labor, more heterogeneous exogenous shocks allow them to adjust their production and employment better without incurring significant changes in wages. In order to test this hypothesis, the models are estimated by regressing the index of Ellison and Glaeser (1997) as a function of a proxy variable for labor pooling, which attempts to capture exogenous shocks in the labor market , controlling for both observed sectoral characteristics that vary in time as well as the sectoral fixed effect. The results are consistent with a reduction in the level of industrial concentration in the period from 2002 to 2014. When considering specific sector trajectories, it was possible to observe that the most technology intensive ones, on average, are more concentrated than the less technology intensive ones, evidences which are in agreement with other studies in the literature. With regard to labor pooling, it was possible to obtain evidence to support the theoretical hypothesis of the model that this local factor is positively related to the spatial agglomeration of the industry, indicating that the industrial sectors can benefit from the concentration of workers. Through these results, the evidence indicates that firms tend to concentring spatially close to where labor markets are located that concentrate workers with the characteristics required in their respective production processes. Intuitively the firms can consider the location of the workforce as an important factor in their location, as workers can spatially concentrate in function on the location of firms. In this aspect, this study focused on the first case and provides consistent results regarding the problem of reverse causality and other sources of endogeneity. In addition, the evidence presented here may help formulate policies that guide the direction of productive activity in space.

Keywords: Labor Pooling. Agglomeration. Industries.

Sumário

1 Introdução **1**

2 Revisão de Literatura **5**
2.1 Experiência Internacional 5
2.2 Evidências para o Brasil 8

3 Novas Evidências da Aglomeração industrial no Brasil **14**
3.1 Índice de Ellison e Glaeser (1997) 14
3.2 Resultados 15

4 *Labor Pooling*: Modelo Teórico **24**

5 Estratégia Empírica **28**
5.1 Modelo Empírico 28
5.2 Dados 30

6 Evidências do *labor pooling* para o Brasil **33**

7 Considerações Finais **43**

Apêndices **51**

Apêndice A Considerações adicionais sobre o Modelo Teórico **52**

Apêndice B Tabelas Adicionais **56**

Lista de Figuras

3.1 Distribuição espacial do emprego na indústria da transformação por microrregiões do Brasil em 2002 16
3.2 Distribuição espacial do emprego na indústria da transformação por microrregiões do Brasil em 2014 17
3.3 IEG para o Brasil no período 2002-2014 (setores intensivos em tecnologia) 21
3.4 Heterogeneidade do índice de Ellison e Glaeser entre os setores 22

6.1 PEA nas áreas metropolitanas de São Paulo, Rio de Janeiro e Belo Horizonte no período 2002-2014 34
6.2 PEA nas áreas metropolitanas de Recife, Porto Alegre e Salvador no período 2002-2014 35
6.3 Índice de Ellison e Glaeser *versus Labor Pooling* 36
6.4 Heterogeneidade do *labor pooling* entre os setores 37

A.1 Diagrama de fases equilíbrio *labor market pooling* 52

Lista de Tabelas

3.1 Estatísticas descritivas do IEG para o Brasil, dados para o período de 2002 a 2014 18
3.2 Estatísticas descritivas do IEG para 2002 e 2014 19
3.3 IEG para os setores da indústria da transformação no Brasil e *ranking* de concentração 20

5.1 Resumo das variáveis do Modelo Empírico 29

6.1 Matriz de Correlação das variáveis do modelo empírico 36
6.2 Estatísticas descritivas das variáveis do modelo empírico 37
6.3 Resultados do modelo empírico 40

B.1 Intensidade tecnológica das divisões (CNAE dois dígitos) que compõem a indústria de transformação (CNAE 1.0) 56
B.2 Índice de concentração bruta por setores da indústria da transformação no período 2002-2014 57
B.3 Índice Herfindahl-Hirshman por setores da indústria da transformação no período 2002-2014 58
B.4 Índice de Ellison e Glaeser por setores da indústria da transformação no período 2002-2014 59
B.5 Resultados das Estimações *Fixed Effect* (FE) 60
B.6 Resultados Estimações *Random Effects* GLS (RE) 60
B.7 Teste de Hausman *Fixed Effect vs Random Effects* 60

Capítulo 1

Introdução

A aglomeração da atividade produtiva industrial não é uma característica única do Brasil, pode ser observada em diversos países (SOBRINHO; AZZONI, 2016). Dado que a distribuição das indústrias pode ser um fator determinante do crescimento e do desenvolvimento das regiões e, consequentemente, um dos fatores responsáveis pela heterogeneidade entre elas, os estudos sobre a concentração industrial têm apresentado avanços na literatura. Destaca-se, por exemplo, o esforço na construção de indicadores de concentração (ver Ellison e Glaeser (1997); Maurel e Sédillot (1999); e Duranton e Overman (2005)) e também estudos que exploram empiricamente o que está por trás da concentração dos agentes produtivos com ênfase nas economias de aglomeração (por exemplo, Rosenthal e Strange (2001); Resende e Wyllie (2004); Ellison, Glaeser e Kerr (2010); Overman e Puga (2010); e Burki e Khan (2013)).

Por outro lado, apesar dos avanços recentes, como destacado por Overman e Puga (2010), a literatura empírica tem sido menos eficaz em distinguir as fontes dos ganhos de produtividade derivados da aglomeração. Embora existam modelos teóricos microfundamentados que apresentam a formalização (ver Krugman (1991), Fujita e Thisse (2013) e Duranton e Puga (2004)) para a análise dos fatores locais como determinantes da aglomeração, do ponto de vista empírico, a distinção entre esses diferentes mecanismos é complexa. Essa complexidade, por sua vez, decorre da chamada "equivalência marshalliana", no sentido em que todos esses mecanismos sugerem "ganhos" de produtividade com a aglomeração, mas são difíceis de rastrear (Duranton e Puga (2004), Overman e Puga (2010)).

Dentre esses mecanismos, o *labor pooling* pioneiramente destacado por Marshall (1890), pode atuar como uma força de atração levando as indústrias à concentração espacial. Essa força centrípeta decorre das vantagens que podem ser obtidas pelas indústrias concentradas localmente onde existe um mercado de trabalho abundante em função da oferta de trabalho (MARSHALL, 1890). Uma formalização para esse argumento de Marshall foi proposta por Krugman (1991) (*Labor Market Pooling Model*), na qual um estabelecimento ajusta seu nível de produção e emprego em função de

choques idiossincráticos de produtividade. O modelo de Krugman (1991) sugere que os estabelecimentos tendem a se aglomerar espacialmente nos locais com um grande número de trabalhadores com as habilidades necessárias para sua produção. As firmas podem obter vantagens dessa oferta de mão de obra e se adaptar melhor a choques exógenos de produtividade sem reduções significativas no lucro esperado em função de ajustes nos salários locais. Como mostra o modelo, o efeito *labor pooling* será tanto maior quanto maior a heterogeneidade do choque no setor específico. De acordo com Overman e Puga (2010), a intuição desse modelo é que, beneficiando-se da concentração de trabalhadores, o estabelecimento pode ajustar melhor a quantidade de trabalho empregada em seu processo produtivo em função de um choque exógeno de produtividade sem incorrer em ajustes significativos de salários.

Para testar empiricamente a hipótese do modelo teórico para o Reino Unido, Overman e Puga (2010) regrediram uma medida de concentração espacial, o Índice de Ellison e Glaeser (1997), em função de uma variável que tenta captar o potencial *labor pooling* por setores, com controles para as características setoriais. O índice utilizado pelos autores apresenta a vantagem de ser ponderado pelo tamanho dos estabelecimentos e com isto o problema da dimensionalidade da indústria não influencia na magnitude do mesmo. Outra vantagem da utilização desse índice é o fato de ele ser formalmente desenvolvido a partir de um modelo microfundamentado.

Já a variável *labor pooling* é proposta por Overman e Puga (2010) para captar o efeito sugerido pelo modelo teórico de que os setores nos quais os choques idiossincráticos são mais heterogêneos tendem a ser espacialmente mais concentrados. Para exemplificar isso, os autores argumentaram que, se os salários são maiores quando o estabelecimento pretende aumentar a quantidade de trabalhadores em resposta a um choque positivo, por exemplo, isto reduz a capacidade de ajuste do estabelecimento em épocas de crise. Por outro lado, quando as mudanças no emprego do estabelecimento não são acompanhadas de aumentos significativos nos salários locais, maior a vantagem de localizar-se em regiões com uma grande concentração de trabalhadores com as habilidades necessárias para a sua produção. Partindo desta ideia, de acordo com Overman e Puga (2010), para isolar de forma direta[1] o efeito heterogeneidade, toma-se a diferença entre a variação percentual do emprego no estabelecimento e a variação percentual na indústria (setor) em valor absoluto. Essa variável assumirá valores maiores quando um estabelecimento estiver expandindo

[1]Os autores argumentam que estudos como o de Rosenthal e Strange (2001) incorporaram o efeito *pooling* indiretamente através das habilidades da mão de obra e que esse tipo de *proxy* pode não ser adequado para isolar a força atrativa derivada da concentração de trabalhadores, captando outros efeitos além do *labor pooling*.

seu emprego enquanto o resto da indústria está reduzindo ou vice-versa, captando a heterogeneidade dos choques.

No Brasil, como será detalhado em capítulos posteriores, de acordo com os dados da Relação Anual de Informações Sociais do Ministério do Trabalho (RAIS-MTE), das 557 microrregiões que compõem o país, apenas 15 foram responsáveis por concentrar 42,80% do emprego total na indústria da transformação no ano de 2002. Em 2014, a mesma quantidade de microrregiões concentrou 37,88% do emprego total na indústria da transformação. Intuitivamente, espera-se que a localização das firmas possa, de alguma forma, influenciar a decisão de escolha locacional dos trabalhadores. Principalmente nas grandes cidades esse fator pode ser importante dentre o conjunto de informações que os trabalhadores possuem e a partir dele definem seu local de moradia. Por outro lado, também é intuitivo pensar que as firmas levam em consideração o mercado de trabalho na decisão de onde instalar-se. Existe, desta forma, um possível problema de causalidade reversa neste contexto e essa é uma das principais dificuldades enfrentadas pelos estudos empíricos que abordam este tema, ou seja, é preciso cautela ao definir uma *proxy* para o *labor pooling*. Lidar com uma situação desse tipo é uma das vantagens de utilizar a estratégia de Overman e Puga (2010).

Estudos como este têm a finalidade de explorar os fatores intrínsecos ao mercado que funcionam como forças atrativas levando as indústrias à concentração espacial. Apesar de reconhecer que o setor público certamente atua de alguma forma determinando a distribuição geográfica das firmas, esta pesquisa aborda a questão de que a aglomeração industrial também pode ser explicada pelo próprio mercado. Entender estes mecanismos podem auxiliar na formulação de políticas regionais desenvolvimentistas com o objetivo de tornar as firmas mais produtivas. A partir da compreenção de como os mercado atua sobre a determinação da localização das firmas, as políticas podem ser traçadas no sentido de intensificar os "ganhos" gerados pelas economias externas. É neste sentido que pesquisas como esta, a partir de um modelo teórico, geram suas contribuições ao fornecer evidências sobre a aglomeração.

Para o Brasil, mesmo com a maior disponibilidade de dados desagregados geográfica e setorialmente, ainda são poucas as pesquisas que analisam a concentração industrial no país de forma mais criteriosa, com respaldo teórico e empírico. Além disto, na maioria dos estudos que utilizam o Índice de Ellison e Glaeser para o Brasil, o objetivo é apenas quantificar a aglomeração industrial, como será abordado nos próximos capítulos. Neste sentido, a presente pesquisa avança na literatura ao procurar evidências empíricas da hipótese teórica de Krugman (1991) com dados para o Brasil. Os estudos para o Brasil,

além de não abordarem diretamente o *labor pooling* de uma perspectiva exógena, os que o exploram o fazem sem um embasamento teórico para a ideia de Marshall. A pesquisa também difere, do ponto de vista empírico, em relação ao estudo de Overman e Puga (2010) ao utilizar uma estratégia que controla para as características setoriais não observadas fixas no tempo.

Na literatura nacional, estudos como os de Resende e Wyllie (2004), Silveira Neto (2005) e Resende (2015) tiveram como objetivo de pesquisa investigar os determinantes da concentração a partir de modelos econométricos usando como variável dependente índices de aglomeração. A presente pesquisa, no entanto, se diferencia dos estudos supracitados ao enfatizar o efeito *labor pooling* a partir de uma variável construída a nível da firma, que de acordo com Overman e Puga (2010), é capaz de captar a volatilidade idiossincrática enfrentada pelos estabelecimentos em cada setor, estando, deste modo, de acordo com a hipótese do modelo teórico. Com este escopo, a dissertação tem como objetivo de pesquisa obter evidências sobre a importância do *labor pooling* para explicar a concentração industrial no Brasil.

Especificamente:

i) Quantificar a aglomeração da indústria da transformação[2] no Brasil através do Índice de Ellison e Glaeser utilizando os microdados da Relação Anual de Informações Sociais do Ministério do Trabalho (RAIS-MTE) em nível da firma para o período 2002-2014;

ii) Estimar modelos de regressão considerando o Índice de Ellison e Glaeser (calculado para as microrregiões do país) como variável explicada e o *labor pooling* e demais fatores que podem influenciar a concentração espacial como variáveis explicativas.

A dissertação está dividida em sete capítulos incluindo esta introdução. No seguinte, são revisadas as evidências empíricas obtidas para o Brasil e demais países com relação aos padrões de concentração industrial, principalmente as que utilizaram o índice de Ellison e Glaeser (1997). No capítulo 3, são apresentadas novas evidências sobre os padrões de aglomeração da indústria da transformação a partir do índice de Ellison e Glaeser. No quarto capítulo, a ênfase é no modelo de Krugman (1991) e as considerações de Overman e Puga (2010) sobre o *labor pooling*. No quinto capítulo, a estratégia empírica é descrita. Nos dois últimos, apresentam-se os resultados do modelo empírico e os comentários finais, respectivamente.

[2]Divisões da Classificação Nacional de Atividades Econômicas - CNAE 1.0. Os setores podem ser verificados na Tabela B.1 no Apêndice.

Capítulo 2

Revisão de Literatura

Este capítulo tem como objetivo apresentar uma revisão dos estudos empíricos sobre os fatores importantes para explicar a concentração das indústrias no espaço utilizando um índice de concentração como variável dependente. Com este propósito, duas seções compõem o capítulo. A primeira com uma revisão dos estudos empíricos no cenário internacional e a segunda com os estudos empíricos feitos para o Brasil.

2.1 Experiência Internacional

Apesar da vasta literatura que busca quantificar a aglomeração industrial utilizando medidas de concentração microfundamentadas, principalmente após o estudo de Ellison e Glaeser (1997), os estudos com o escopo de pesquisa distinto, que além de quantificar buscam explicar o indicador, são mais escassos. Um dos principais problemas enfrentados nesta linha de pesquisa é a dificuldade de isolar os efeitos dos diferentes mecanismos que podem gerar benefícios através da concentração espacial às economias de aglomeração.

Buscando os efeitos das economias de aglomeração sobre a concentração da indústria nos EUA, Rosenthal e Strange (2001) utilizaram o índice de Ellison e Glaeser como variável dependente e uma série de variáveis explicativas representando as economias externas. Foram utilizadas *proxies* para *spillovers* de conhecimento, *labor pooling* e controles para custos de transporte e vantagens naturais. Os autores calcularam o índice utilizando três unidades geográficas distintas, *zip code*, municípios e estados no período 1982-1995. Para cada um destes níveis foram estimadas regressões distintas. No geral, as evidências encontradas pelos autores apontam para a importância das economias externas para explicar a aglomeração da indústria, principalmente o *labor pooling*, que mostrou efeitos positivos sobre o adensamento em todos os níveis geográficos estudados. Como *proxies* para esta variável, foram utilizadas as parcelas de trabalhadores por níveis de qualificação (participação de trabalhadores associados à gerência; participação dos trabalhadores com doutorado; participação dos trabalhadores com mestrado; e participação de trabalhadores bacharéis).

Um problema na forma como Rosenthal e Strange (2001) incluíram o *labor pooling* no modelo, através de parcelas de mão de obra por níveis de qualificação, é que estas proxies podem captar efeitos que vão além do labor pooling, no sentido, por exemplo, que a concentração de trabalhadores qualificados pode atrair mais firmas para uma determinada região; mas o inverso também pode ser verdadeiro, ou ainda, a concentração de trabalhadores qualificados pode aumentar a difusão de ideias gerando transbordamento de conhecimento. Por estes motivos, a estratégia de captar o efeito *labor pooling* de forma indireta pode ser ineficiente. Este é o argumento usado por Overman e Puga (2010) para dar embasamento e justificar a sua estratégia de isolar diretamente o *labor pooling* seguindo as premissas teóricas do modelo de Krugman (1991).

Com espoco de pesquisa similar, o estudo de Khan (2010) utiliza o índice de Ellison e Glaeser como variável explicada e dados para o período 1988-2003 nos EUA. Como variáveis explicativas, utilizou uma *proxy* para o *labor pooling* e demais controles para as economias de aglomeração. Os principais resultados sugerem que aumentos na produtividade dos trabalhadores estão positivamente relacionados com aumentos na concentração da indústria manufatureira nos EUA. Essa relação positiva é devido às externalidades geradas pela concentração de trabalhadores que possibilitam que as firmas localizadas nessas áreas se tornem mais produtivas.

A construção da variável *labor pooling* no estudo de Khan (2010) incorre no mesmo problema apontado por Overman e Puga (2010) para o trabalho de Rosenthal e Strange (2001). A pesquisa supracitada também utiliza a participação de trabalhadores por níveis de qualificação na produção como *proxy* (participação dos funcionários com grau de bacharel; percentual de empregados com qualificação inferior ao grau de bacharel; proporção de trabalhadores gerenciais em relação ao total de trabalhadores; valor acrescentado por trabalhador, líquido de custo dos materiais; e salários) para esse fator local, sujeitando-se, deste modo, a problemas com causalidade reversa e demais efeitos captados pela variável além do *labor pooling*. Os efeitos distintos que podem ser captados por *proxies* que mensuram indiretamente cada uma das economias de aglomeração pioneiramente destacadas por Marshall (1890), é um problema bastante conhecido na literatura e representa um importante desafio a ser superado. Além disto, para alguns países como o Brasil, a ausência de dados desagregados e informações suficientes aumenta a dificuldade para estudos dessa natureza.

A pesquisa de Alkay e Hewings (2012) quantifica a aglomeração industrial e obtém os padrões de concentração espacial através do índice de Ellison e Glaeser para a região metropolitana de Istambul, 22 divisões da indústria e três níveis geográficos no período 2000-

2005. No segundo estágio da pesquisa, os autores tiveram como objetivo obter os efeitos das economias de localização[3] e urbanização[4] sobre a concentração regredindo o índice supracitado sobre diversos controles para os fatores locais. Os possíveis determinantes da aglomeração foram obtidos a partir de modelos estimados por (2SLS) *Two-Stage Least Squares*. Os principais resultados obtidos apontam para a importância das economias de localização e urbanização para explicar a concentração espacial da indústria.

Voltando-se mais uma vez para a construção da variável que capta o efeito da concentração de trabalhadores, esta pesquisa também utiliza categorias para diferenciar a qualificação dos trabalhadores empregados na produção, três especificamente, alta, média e baixa qualificação. Ou seja, também utilizam controles que captam de forma indireta o efeito da concentração de trabalhadores sobre a aglomeração industrial.

Neste sentido, um dos principais avanços apresentados no estudo de Overman e Puga (2010) é a estratégia empírica utilizada por estes autores para isolar diretamente o efeito *labor pooling* a partir de dados longitudinais ao nível da firma. Os autores regrediram o índice de Ellison e Glaeser em função de uma *proxy* exógena para o *labor pooling* e uma série de características setoriais da indústria do Reino Unido no período 1992-2003 com objetivo de testar a hipótese teórica do modelo de Krugman (1991). O modelo teórico prediz a existência de uma relação positiva entre a concentração de trabalhadores com características necessárias para um determinado processo produtivo e a concentração de firmas pertencentes aos setores para os quais essas habilidades são importantes. Através da estratégia pioneira dos autores, foram obtidos resultados que comprovaram a hipótese teórica do modelo, mostrando que há de fato uma relação positiva entre o *labor pooling* e o índice de Ellison e Glaeser naquele país para o período analisado. Indicando que a concentração de trabalhadores pode explicar em parte a aglomeração industrial.

O diferencial do presente estudo em relação aos trabalhos até aqui referenciados com exceção de Overman e Puga (2010), é o uso da estratégia empírica destes últimos. Desta forma, valendo-se dos microdados da Relação Anual de Informações Sociais do Ministério do Trabalho (RAIS-MTE) foi possível isolar o efeito sugerido pelo modelo teórico para o

[3]Custos com aluguel, aquecimento, energia e água; custos totais com matéria-prima; área potencial de mercado de matérias-primas; percentual de trabalhadores altamente qualificados; percentual de trabalhadores com qualificação média; percentual de trabalhadores com baixa qualificação; área total do estabelecimento (m^2); existência de departamento de P&D (se sim = 1 caso contrário = 0); e produção de bens intermediários (se sim = 1 caso contrário = 0).

[4]População; densidade populacional; PIB; proporção da área industrial; número de empresas não industriais; número de universidades; taxa de emprego industrial; taxa de emprego no setor dos serviços; potencial de mercado com base na população; potencial do mercado de trabalho com base no emprego no sector industrial.

Brasil. Esta pesquisa também avança em relação ao estudo de Overman e Puga (2010) ao incorporar os estimadores de efeitos fixos. Com este método de estimação espera-se aumentar a robustez da análise controlando para as características não observadas fixas no tempo de cada setor da indústria da transformação. Neste estudo tentamos resolver o problema da causalidade reversa através da estratégia de Overman e Puga (2010) para construção da principal variável explicativa. Já para mitigar problemas com possíveis variáveis omitidas, são utilizados os demais controles para as características variantes no tempo de cada setor.

Adicionalmente, é importante ressaltar que o *labor pooling* tem sido explorado através de estratégias empíricas distintas. Como destacam Melo e Graham (2014), estas abordagens podem ser elencadas em dois grupos classificados como as "fontes de aglomeração" e as "fontes de produtividade". Atrelados ao primeiro conjunto, têm-se os estudos já citados no decorrer desta seção (Rosenthal e Strange (2001), Khan (2010), Overman e Puga (2010), Alkay e Hewings (2012)); pertencentes à segunda vertente têm-se, por exemplo, os estudos de Rigby e Essletzbichler (2002), Wheeler (2006), Andersson, Burgess e Lane (2007), Amiti e Cameron (2007), Mion e Naticchioni (2009) e Melo e Graham (2014). Esta pesquisa se enquadra no primeiro grupo de estudos.

2.2 Evidências para o Brasil

Os estudos empíricos para o Brasil podem ser divididos em dois grupos de acordo com o objetivo de pesquisa. Um grupo composto pelas pesquisas com o objetivo de mensurar a aglomeração da indústria (Biderman (2004), Resende e Wyllie (2005), Lautert e Araújo (2007), Maciente (2008), Araujo (2016) e outros) e outro grupo cujo o objetivo, além de mensurar, tenta explicar a concentração a partir dos fatores locais (Resende e Wyllie (2004), Silveira Neto (2005), Resende (2015)).

O estudo de Biderman (2004) apresenta evidências sobre a concentração para todas as divisões da CNAE nas microrregiões do Brasil para os anos de 1991 e 2000 com base no índice de Ellison e Glaeser (1997) (daqui em diante, IEG). Na pesquisa, o autor salienta a ausência de trabalhos para o Brasil baseados numa medida mais formal de aglomeração, em sua maioria de caráter descritivo[5], e fornece indicadores que corroboram os resultados de estudos anteriores, que evidenciaram a existência de

[5]Podem-se destacar, por exemplo, os estudos de Diniz, Crocco et al. (1996), Azzoni (1997), Pacheco (1999), Saboia (2000) e Suzigan et al. (2001).

concentração industrial no país. Como resultados, o grau de concentração, em média, não apresentou variação significativa no período estudado, sugerindo, desta forma, que o nível de concentração permaneceu relativamente constante. Como no estudo foram considerados setores bastante heterogêneos, o autor também apresentou uma análise baseada em amostra restrita, eliminando os setores para os quais foram obtidos valores díspares para o IEG. Ainda assim, os resultados se mantiveram, reforçando a ideia de concentração.

Resende e Wyllie (2005) empregaram três medidas de concentração, as construídas por Ellison e Glaeser (1997), Maurel e Sédillot (1999) e Devereux, Griffith e Simpson (2004), no caso brasileiro, para os dados da indústria da transformação em 1995 e 2001 fornecidos pela RAIS[6]. Para o cálculo dessas medidas, foram consideradas, em termos de desagregação setorial, a CNAE quatro dígitos, e desagregação geográfica, as microrregiões do país. Como principais resultados, destaca-se a elevada heterogeneidade entre os setores estudados; o IEG foi a medida de concentração que apresentou a menor variação média entre os anos 1995 e 2001, 0,0563 e 0,0527, respectivamente. Além disto, das 20 classificações ranqueadas de acordo com o grau de aglomeração, 14 pertencem às divisões da CNAE intensivas em tecnologia (alta e média-alta) de acordo com classificação proposta por Cavalcante (2014), sugerindo que esse tipo de atividade produtiva apresenta níveis mais elevados de concentração.

No mesmo contexto de Resende e Wyllie (2005), o estudo de Lautert e Araújo (2007) teve como objetivo calcular o IEG para a indústria da transformação no Brasil no período 1996-2001 com base na CNAE (divisões/2 dígitos e grupos/3 dígitos) com dados da RAIS. Nesse estudo, o nível de desagregação geográfica considerada foram os estados. Assim como no estudo anterior, as evidências encontradas sugerem que os setores intensivos em tecnologia apresentam maior grau de concentração relativamente aos demais setores. No *ranking* construído pelos autores, nas 10 primeiras posições em 1996 estão 6 dos 8 setores intensivos em tecnologia[7], com 5 mantendo-se no *ranking* para 2001, e a Divisão 31 entrando para as 10 primeiras posições nesse mesmo ano, ou seja, no

[6]Justificam o uso desta base argumentando sobre as limitações impostas aos dados obtidos junto ao Instituto Brasileiro de Geografia e Estatística (IBGE) como os da Pesquisa Industrial Mensal - Produção Física (PIM-PF), Pesquisa Industrial Mensal - Versão Dados Gerais (PIM-DG) e a Pesquisa Industrial Anual (PIA), e o fato de a RAIS possuir a vantagem de fornecer dados desagregados ao nível da firma, informações que são necessárias para a construção dos índices utilizados.

[7]Divisão 32, 3ª colocação em 1996 (0,119) e 4ª em 2001 (0,086); Divisão 34, 6ª colocação em 1996 (0,085) e 5ª em 2001 (0,074); Divisão 35, 7ª colocação em 1991 (0,048) e 23ª em 2001 (-0,032); Divisão 24, 8ª colocação em 1996 (0,044) e 8ª em 2001 (0,036); Divisão 29, 9ª colocação em 1996 (0,040) e 7ª em 2001 (0,039); e Divisão 33, 10ª colocação em 1991 (0,039) e 6ª em 2001 (0,043). IEG entre parênteses.

ranking para os dois anos, 6 dos 10 setores mais concentrados nos estados brasileiros são classificados como sendo intensivos em tecnologia.

Com ênfase na dinâmica da concentração do emprego por região e setores no Brasil, Maciente (2008) calculou o IEG para o período de 1994 a 2005. Através desse indicador, o autor encontrou evidências de que os setores com menores participações no emprego tendem a ser mais concentrados. Esses resultados estão de acordo com os dos outros estudos no sentido em que os setores intensivos em tecnologia, por exemplo, no Nordeste, não possuem parcelas expressivas do emprego. Assim, os resultados de Maciente (2008) sugerem, mesmo que indiretamente, a tendência à concentração desses setores. Com base no IEG por macrorregiões, os resultados indicaram uma tendência de redução da concentração do emprego nacional. Adicionalmente, outra discussão apresentada neste estudo refere-se à concentração da mão de obra por qualificação. Constatou-se que trabalhadores menos qualificados (até oito anos de estudo) encontram-se mais concentrados relativamente à mão de obra qualificada (educação superior).

O estudo de Vignandi, Parré e Guimarães (2016) contribuiu para o debate sobre a concentração da indústria no Brasil fornecendo índices de aglomeração espacialmente ponderados para os setores da indústria da transformação e mesorregiões brasileiras no período 1995-2010. A estratégia empregada na análise foi constituída a partir da incorporação dos efeitos de vizinhança nos índices de concentração. Os resultados encontrados pelos autores apontam para um processo de desconcentração da indústria da transformação no Brasil com exceção de alguns setores específicos, como o de fabricação de coque, refino de petróleo, elaboração de combustíveis nucleares e produção de álcool (Divisão 23). Como justificativa para explicar a discrepância da Divisão 23 relativamente aos demais setores, os autores argumentaram sobre a influência das atividades relacionadas à exploração do pré-sal. Sobre a inclusão do efeito vizinhança, de acordo com os autores, esse processo aumenta a robustez da análise aproximando-a mais da realidade.

Com escopo regional, Araujo (2016) fornece evidências sobre a concentração da indústria da transformação nas microrregiões e estados do Nordeste com base no IEG. O autor calculou o índice para os anos de 2005 e 2010 com dados da RAIS. Adicionalmente, o autor se propôs a fazer uma investigação considerando os níveis tecnológicos, através da qual obteve resultados que estão de acordo com indicadores apresentados nos estudos de Resende e Wyllie (2005), Lautert e Araújo (2007) e Maciente (2008), no sentido em que estes últimos sugeriram que os setores intensivos em tecnologia são mais concentrados que os demais no Brasil, e Araujo (2016) aponta que essa característica também pode ser observada para o Nordeste. Estudos para o Brasil baseados em análises mais descritivas,

como o de Pacheco (1999), já haviam indicado essa tendência. Assim, a indústria intensiva em tecnologia no Nordeste, de acordo com o autor, segue a tendência nacional.

No contexto do segundo grupo de estudos, a pesquisa de Resende e Wyllie (2004) avança na literatura e utiliza uma medida de aglomeração formalmente desenvolvida por Maurel e Sédillot (1999), que é similar ao IEG para o estudo da concentração industrial no estado do Rio de Janeiro. Os autores utilizaram diversas bases de dados[8] em seu estudo e informações para anos distintos com o objetivo de explicar a concentração industrial no estado. Para tanto, regrediram a medida de aglomeração sobre uma série de características setoriais e locais. Como a medida de concentração varia por setores, os autores adotaram a estratégia de Ellison e Glaeser (1997) para obter o peso de cada município no indicador setorial e assim compatibilizar as variáveis usadas no modelo empírico. Como principais resultados do estudo, destacam-se as evidências obtidas para as variáveis de infraestrutura local tendo um papel relativamente importante para explicar a aglomeração, enquanto as políticas de incentivo não foram expressivamente relevantes.

Silveira Neto (2005) fornece evidências sobre a especialização e concentração da atividade industrial no Brasil para o período 1950-1985 e 1985-2000. O exercício empírico do autor foi dividido em duas etapas. Na primeira, a partir de dados do Instituto Brasileiro de Geografia e Estatística (IBGE - Censos Industriais para o primeiro período destacado e da PIA para o segundo) para as macrorregiões do Brasil. Os resultados mostraram uma trajetória de desconcentração industrial a partir da década de 1970, ocorrendo de forma mais intensa até 1985 e, no período posterior (1985-2000), de forma mais atenuada. Nessa etapa também foi apresentada pelo autor uma análise desagregada por setores da indústria onde os intensivos em capital apresentaram uma tendência à concentração relativamente aos demais, evidências que estão de acordo com os resultados de Resende e Wyllie (2005). A análise de concentração desagregada setorialmente foi desenvolvida com base no Coeficiente de Localização de Hoover.

Diferentemente dos trabalhos já citados nesta seção, exceto Resende e Wyllie (2004), Silveira Neto (2005) apresenta uma contribuição adicional para a literatura ao apresentar evidências no sentido de tentar explicar a concentração industrial. A estratégia empírica do autor, agora se referindo à segunda etapa do exercício, foi baseada em modelos de regressão com dados em painel para o período 1950-1985 e modelos de regressão *cross section* para 1996 e 2000, tendo como variável dependente o Índice de Hoover e como

[8]RAIS para a construção da medida de aglomeração e demais informações sobre a indústria; FIRJAN, CIRJ, SESI, SENAI e IEL para informações sobre políticas de incentivos ficais; Fundação CIDE para obter informações da matriz insumo-produto para o Rio de Janeiro e informações sociais e econômicas para as cidades do estado; e BNDES para dados sobre informações creditícias por setores da indústria.

regressores *proxies* para economias de escala e *linkages* setoriais. Os principais resultados sugeriram a importância desses fatores, embora, como ressaltado no estudo, neste tipo de modelo não se pode desconsiderar os efeitos das políticas públicas sobre a determinação da distribuição da indústria no território nacional.

O estudo de Maciente (2013) quantifica a aglomeração e coaglomeraçao no Brasil com dados da RAIS no período 1994-2010. Para isto, o autor utiliza o índice de aglomeração e coaglomeração de Ellison e Glaeser. No segundo estágio da pesquisa, o objetivo do autor foi verificar a relação aglomeração e coaglomeração industrial x economias de aglomeração. Os resultados obtidos apontam para uma redução da concentração da atividade econômica no país no período analisado, estando, deste modo, em consonância com estudos anteriores que chegaram a resultados similares. Essa característica, inclusive, já tinha sido exposta por trabalhos com escopo de pesquisa mais descritivo, não pautado em um índice microfundamentado. Segundo o autor, atividades ligadas à mineração apresentaram trajetória de concentração, diferentemente das atividades ligadas à agricultura, manufatura e imóveis, que apresentam trajetórias de desconcentração espacial. Embora tenham utilizado o IEG para quantificar a aglomeração, no modelo empírico usado para explicar a aglomeração, esta medida não foi usada. Para este propósito, nas regressões do modelo estimado através do método *Seemingly Unrelated Regression* (SUR), o emprego foi utilizado como variável explicada.

Já para explicar a concentração associada de pares de indústrias, o autor utilizou como variável dependente o índice de coaglomeração de Ellison e Glaeser. De modo geral, nas duas vertentes da análise, os fatores externos se mostraram importantes para explicar a aglomeração e a coaglomeração no Brasil, destacando ainda que as vantagens naturais e políticas fiscais regionais também podem influenciar a concentração das indústrias, principalmente a mineração e de transformação, e que os setores de serviços são mais influenciados por externalidades de urbanização.

A pesquisa de Costa (2014) também utilizou o índice proposto por Ellison e Glaeser para analisar a concentração da indústria no Brasil. Valendo-se dos dados RAIS, o autor calculou a referida medida para os anos de 1991, 1996, 2000, 2005 e 2011. Assim como os estudos anteriores, também foram encontradas evidências da desconcentração industrial no período compreendido na amostra. No entanto, ressaltou uma possível suavização desse processo após o ano 2000. Com relação aos setores de alta tecnologia, também em linha com resultados de estudos anteriores, as evidências apontam que são mais concentrados quando comparados a todos os setores conjuntamente. Além disto, apesar de também apresentarem uma tendência à desconcentração, esta parece ser mais lenta que a tendência

observada para os demais setores.

Adicionalmente, o autor também decompõe a concentração geográfica valendo-se da estratégia de Dumais, Ellison e Glaeser (2002), que utiliza um processo dinâmico originado de dois fatores - crescimento e contração de aglomerados industriais e aleatoriedade das taxas de crescimento - para analisar a aglomeração. No segundo estágio da pesquisa, os principais resultados indicaram que o surgimento de novas firmas contribui para a desconcentração da indústria já que esses novos empreendimentos ocorrem em maior quantidade fora dos aglomerados. Já com relação ao fechamento de empresas, os efeitos são positivos sobre aglomeração.

Ainda no contexto do segundo grupo de estudos, Resende (2015) avança na literatura ao utilizar um índice de coaglomeração industrial e relacioná-lo com o *labor pooling* no sentido de obter evidências da importância desse fator local para explicar a coaglomeração. Com base no índice que relaciona pares de setores industriais de Ellison, Glaeser e Kerr (2010), e dados sobre a indústria da transformação para o Rio de Janeiro no ano 2010, o autor quantifica a coaglomeração industrial e busca relacioná-la com uma *proxy* para o *labor pooling* e demais controles. O indicador foi calculado considerando a CNAE 5 dígitos e os municípios do referido estado com dados da RAIS. A *proxy* para o *labor pooling* foi obtida através de coeficientes de correlação entre o nível de emprego dos pares de setores. Como resultados do modelo econométrico, o autor sugere a importância do *labor pooling* e de características setoriais para explicar a coaglomeração.

Na presente pesquisa, busca-se fornecer evidências sobre a importância do *labor pooling* para explicar a concentração industrial com base no modelo de Krugman (1991). Além disto, a estratégia empírica segue a utilizada por Overman e Puga (2010), que permite testar a hipótese teórica do referido modelo ao mensurar de forma direta o efeito heterogeneidade dos choques de produtividade setoriais. Essa é a principal contribuição do presente estudo para o debate sobre o tema. É importante destacar o reconhecimento das políticas públicas como possível fator determinante da concentração industrial. Entretanto, a abordagem empírica reporta-se aos argumentos teóricos sobre os fatores externos relativos ao mercado.

Capítulo 3

Novas Evidências da Aglomeração industrial no Brasil

Este capítulo tem como objetivo apresentar novas evidências sobre os padrões de concentração da indústria da transformação no período 2002-2014 a partir do índice de Ellison e Glaeser (1997). A próxima seção apresenta de forma objetiva a medida de concentração supracitada e a segunda seção apresenta uma análise com base no cálculo do índice para o Brasil no período 2002-2014.

3.1 Índice de Ellison e Glaeser (1997)

O Índice de Concentração de Ellison e Glaeser (1997) pode ser definido de acordo com a seguinte equação:

$$\gamma_s = \frac{\frac{\sum_z (\delta_{zs} - x_z)^2}{1 - \sum_z x_z^2} - \sum_{k=1}^{N_s} \tau_{sk}^2}{1 - \sum_{k=1}^{N_s} \tau_{sk}^2} = \frac{G_s - H_s}{1 - H_s} \tag{3.1}$$

O IEG é composto por dois índices. Um é o índice de concentração bruta, G_s, que é definido de acordo com a equação abaixo:

$$G_s = \frac{\sum_z (\delta_{zs} - x_z)^2}{1 - \sum_z x_z^2} \tag{3.2}$$

Onde z indexa as subunidades geográficas; δ_{zs} é a participação de z no nível de emprego da indústria s nacional; e x_z é a participação de z no emprego industrial total.

Uma limitação apontada na literatura (Ellison e Glaeser (1997)) para indicador G_s é o viés oriundo da concentração via fatores característicos do processo produtivo.

Por exemplo, em função de um número reduzido de firmas com produção em grande escala. Neste caso, o fator escala de produção atuaria superestimando o indicador, fazendo com que setores com escalas de produção naturalmente maiores se mostrassem mais concentrados.

Outra medida base para a construção de índices de concentração mais elaborados é o índice Herfindahl-Hirshman. Trata-se de uma medida de escala que atribui pesos de acordo com as parcelas de mercado de cada firma e pertence necessariamente ao intervalo com limite inferior dado por $\frac{1}{n}$ e superior 1 (RESENDE; WYLLIE, 2005). Suponha que o setor industrial s possui N_s firmas, cada uma dessas com parcelas do emprego no setor dadas por $\tau_{s1}, \ldots, \tau_{sN_s}$. O índice de Herfindahl-Hirshman é definido de acordo com a equação abaixo:

$$H_s = \sum_{k=1}^{N_s} \tau_{sk}^2 \tag{3.3}$$

Deste modo, o IEG é uma ponderação de (3.2) por (3.3) no sentido de superar a limitação de G_s. Outra vantagem do IEG é que este índice é derivado de um modelo de escolha locacional microfundamentado desenvolvido por Ellison e Glaeser (1997), no qual os autores demostram que γ_s suporta três situações distintas no sentido de incorporar os efeitos dos fatores locais sobre a decisão das firmas. Uma que incorpora o efeito distorcivo das vantagens naturais, outra que considera os efeitos *spillovers* gerados pela concentração, e, por fim, uma situação na qual as duas forças podem atuar conjuntamente.

Uma limitação do IEG é o fato de este indicador não considerar a localização relativa das firmas e as distâncias que as separam. Por esse motivo, o índice é invariante até a permutação das firmas. Por exemplo, o efeito sobre o índice é o mesmo se se deslocarem firmas responsáveis por grande parte da atividade no setor que estão localizadas no centro da economia - subunidade geográfica - ou se estão localizadas na periferia (COMBES; GOBILLON, 2015).

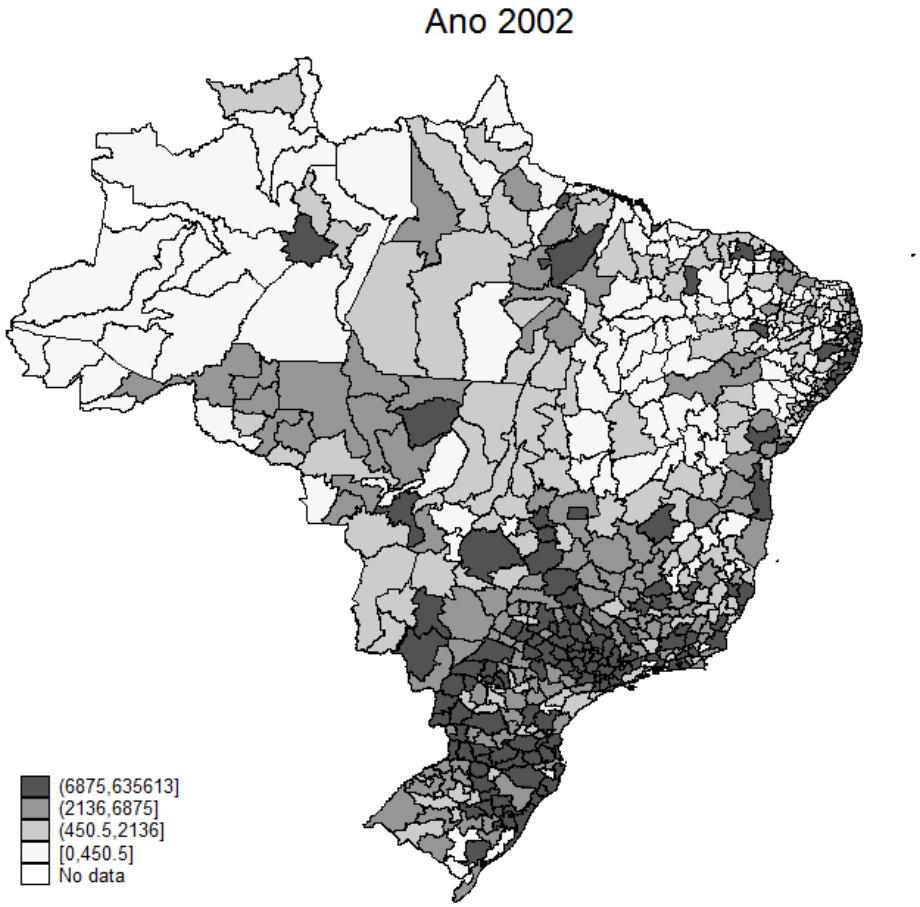
3.2 Resultados

No Brasil, como mostram os dados da Relação Anual de Informações Sociais do Ministério do Trabalho (RAIS-MTE), das 557 microrregiões que compõem o país, apenas 15 foram responsáveis por concentrar 42,80% do emprego total na indústria da transformação no ano de 2002. São elas: São Paulo, Porto Alegre, Rio de Janeiro,

Campinas, Belo Horizonte, Curitiba, Fortaleza, Caxias do Sul, Blumenau, Guarulhos, Joinville, Sorocaba, São José dos Campos, Osasco e Manaus. Já para o ano de 2014, a mesma quantidade de microrregiões, mas agora obedecendo à seguinte ordem: São Paulo, Rio de Janeiro, Campinas, Belo Horizonte, Porto Alegre, Curitiba, Fortaleza, Caxias do Sul, Joinville, Blumenau, Sorocaba, Manaus, Guarulhos, São José dos Campos e Osasco, concentraram 37,88% do emprego total na indústria da transformação.

Nesse mesmo período os empregos nesse tipo de atividade passaram de 5.130.893, no ano de 2002, para 7.900.289, no ano de 2014, apresentando uma taxa de crescimento de aproximadante 53%. Através das Figuras 3.1 e 3.2 é possível observar a distribuição do emprego por microrregiões do país para o ano de 2002 e 2014, respectivamente.

Figura 3.1: Distribuição espacial do emprego na indústria da transformação por microrregiões do Brasil em 2002



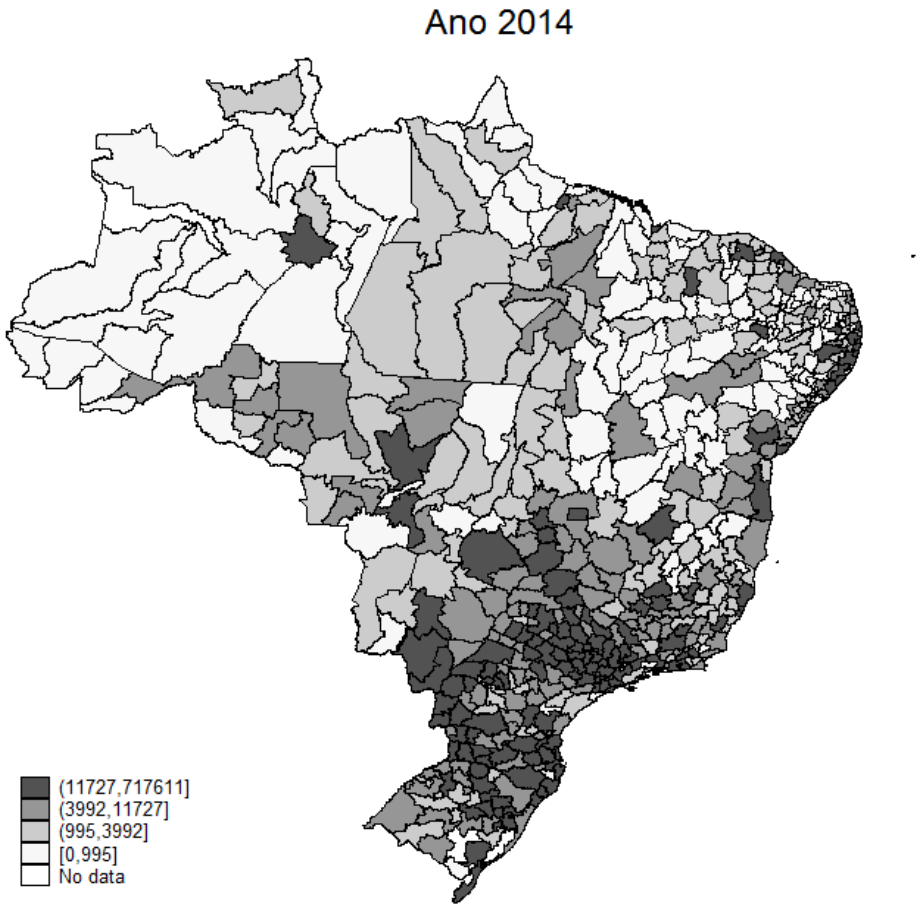
Nota: No data representa as microrregiões inexistentes no banco de dados da RAIS.
Fonte: Elaboração própria com base nos dados da RAIS para 2002 via *software* Stata 12.

Com auxílio dos mapas é possível verificar a dinâmica espacial da indústria da

transformação entre os doze anos que compõem a amostra. As microrregiões destacadas em cores mais escuras são as que concentraram mais empregos em cada ano. A maioria delas localizadas no Sudeste brasileiro.

No mapa para 2002, as 139 microrregiões que estão incluídas no primeiro intervalo (6875; 635613] (áreas mais escuras na Figura 3.1), concentraram 4.405.684 empregos, ou aproximadamente 85,86% do total de empregos. No mapa para 2014, a mesma quantidade de microrregiões pertencentes ao intervalo (11727; 717611] (áreas mais escuras na Figura 3.2), concentraram 6.567.408 empregos, o equivalente a aproximadamente 83% do total de empregos.

Figura 3.2: Distribuição espacial do emprego na indústria da transformação por microrregiões do Brasil em 2014



Nota: No data representa as microrregiões inexistentes no banco de dados da RAIS.
Fonte: Elaboração própria com base nos dados da RAIS para 2014 via *software* Stata 12.

Essa análise exploratória inicial sugere que mesmo considerando o crescimento do número de empregos entre 2002 e 2014, as microrregiões que mais concentraram empregos

perderam participação no emprego total no período. Mas, ainda não é possível indicar uma trajetória de desconcentração industrial no país somente com esses dados brutos. Certamente, existem diversos fatores que podem viesar essa análise preliminar, como, por exemplo, o tamanho das firmas. Por estes motivos, optou-se por utilizar uma medida microfundamentada de concentração espacial que não é afetada pelos efeitos do fator escala, como descrito na seção anterior.

A medida de concentração de Ellison e Glaeser descrita anteriormente foi calculada para o Brasil, por ano, no período 2002-2014 e por setores da indústria da transformação (divisões da CNAE 1.0). As subunidades geográficas consideradas na análise foram o conjunto de 557 microrregiões que compõem o país. O objetivo é verificar se as indústrias se encontram concentradas nessas microrregiões. Quando o IEG assume um valor positivo, isto pode indicar a existência de fatores distorcivos atuando sobre a escolha locacional das plantas produtivas levando-as à concentração. O índice, no entanto, não distingue esses efeitos, que podem estar associados a fatores naturais da microrregião e/ou *spillovers* gerados por firmas que escolheram a localidade em questão anteriormente (economias de aglomeração).

A Tabela 3.1 apresenta as estatísticas descritivas do IEG considerando toda a indústria da transformação, os setores não intensivos em tecnologia e os setores intensivos em tecnologia. A análise de concentração espacial por nível tecnológico, justifica-se dadas as evidências existentes na literatura, de que os setores mais intensivos em capital apresentam uma trajetória distinda do restante da indústria da transformação. A divisão dos setores por intensidade tecnológica foi realizada com base na taxonomia proposta por Cavalcante (2014).

Tabela 3.1: Estatísticas descritivas do IEG para o Brasil, dados para o período de 2002 a 2014

Média	Mediana	Mínimo	Máximo	Desvio Padrão	Coeficiente de Variação	Obs
Todos os setores da indústria da transformação						
0,02444	0,01108	0,00251	0,1617	0,03238	1,32463	299
Setores não intensivos em tecnologia						
0,02179	0,00869	0,002512	0,1617	0,03227	1,48075	195
Setores intensivos em tecnologia						
0,02942	0,01762	0,003294	0,13576	0,03215	1,09302	104

Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

Dadas as evidências empíricas existentes para o Brasil, sobre a concentração dos setores intensivos em tecnologia (Resende e Wyllie (2005); Lautert e Araújo (2007); Maciente (2008)), procura-se verificar se esse tipo de indústria ainda mantém esse padrão

mesmo considerando dados mais recentes. Considerando todos os anos e todos os setores, o IEG médio foi de 0,02444, indicando possivelmente, que a escolha locacional das firmas não é aleatória. Em média, os setores intensivos em tecnologia apresentaram um IEG maior que os demais setores. Isto pode indicar a existência de fatores que atuam mais fortemente sobre esses setores no sentido da concentração geográfica.

Também foi possível observar a diferença entre as médias do primeiro e do último ano considerados na amostra; com isso, obtêm-se indicativos de possíveis mudanças nos padrões de concentração no tempo. Considerando todos os setores da indústria de transformação, como mostra a Tabela 3.2, é possível identificar um aumento na média do IEG entre 2002 e 2014. Para os setores intensivos em tecnologia, observa-se uma pequena redução média dessa variável no mesmo período, embora os setores de tecnologia ainda se mostrem mais concentrados tanto em 2002 quanto em 2014. Essas evidências indicam que, assim como mostrado por Resende e Wyllie (2005) e Lautert e Araújo (2007), no Brasil os setores de tecnologia são mais concentrados que os demais setores da indústria de transformação.

Tabela 3.2: Estatísticas descritivas do IEG para 2002 e 2014

	Todos os setores		Não intensivos em tecnologia		Intensivos em tecnologia	
	2002	2014	2002	2014	2002	2014
Média	0,02564	0,02686	0,02448	0,0266	0,02783	0,02733
Desvio padrão	0,03348	0,04123	0,03804	0,0447	0,02490	0,03668
Mínimo	0,0037	0,00257	0,0037	0,00257	0,0078	0,00329
Máximo	0,15163	0,1617	0,15163	0,1617	0,08466	0,11108

Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

Paralelamente a isto, essas estatísticas para o indicador de concentração sugerem que os possíveis efeitos distorcivos capazes de influenciar onde as plantas produtivas da indústria da transformação se instalam aumentaram no período analisado. É importante ressaltar ainda que essa análise considera informações médias para todos os setores, cada setor individualmente segue sua própria tendência que não necessariamente segue a tendência geral observada. A média para todos os setores na Tabela 3.2 sugere a existência de um processo de concentração da indústria no espaço no Brasil, ou seja, que houve um aumento da concentração da indústria da transformação no período 2002-2014. No entanto, como todos os setores são considerados, um único setor que por algum motivo específico apresentou aumento significativo da concentração no período pode superestimar a média.

De fato, por este motivo pode ser verificada ainda a medida de concentração para cada setor nos anos 2002 e 2014, como mostra a Tabela 3.3.

Tabela 3.3: IEG para os setores da indústria da transformação no Brasil e *ranking* de concentração

Divisão	Descrição	2002		2014	
		$\gamma_{s,2002}$	Posição	$\gamma_{s,2014}$	Posição
16	Fabricação de produtos do fumo	0,1516	1	0,0985	3
32	Fabricação de material eletrônico e de aparelhos e equipamentos de comunicações	0,0847	2	0,1111	2
19	Preparação de couros e fabricação de artefatos de couro, artigos de viagem e calçados	0,0613	3	0,0379	5
34	Fabricação e montagem de veículos automotores, reboques e carrocerias	0,0345	4	0,0192	8
35	Fabricação de outros equipamentos de transporte	0,0318	5	0,0464	4
20	Fabricação de produtos de madeira	0,0255	6	0,0153	10
22	Edição, impressão e reprodução de gravações	0,0247	7	0,0215	6
30	Fabricação de máquinas para escritório e equipamentos de informática	0,0220	8	0,0195	7
23	Fabricação de coque, refino de petróleo, elaboração de combustíveis nucleares e produção de álcool	0,0210	9	0,1617	1
33	Fabricação de equipamentos de instrumentação médico-hospitalares, instrumentos de precisão e ópticos, equipamentos para automação industrial, cronômetros e relógios	0,0183	10	0,0105	11
17	Fabricação de produtos têxteis	0,0167	11	0,0184	9
24	Fabricação de produtos químicos	0,0145	12	0,0048	17
15	Fabricação de produtos alimentícios e bebidas	0,0099	13	0,0059	14
25	Fabricação de artigos de borracha e de material plástico	0,0097	14	0,0051	16
31	Fabricação de maquinas, aparelhos e materiais elétricos	0,0092	15	0,0040	19
18	Confecção de artigos do vestuário e acessórios	0,0087	16	0,0091	12
37	Reciclagem	0,0087	17	0,0035	20
29	Fabricação de máquinas e equipamentos	0,0078	18	0,0033	21
26	Fabricação de produtos de minerais não metálicos	0,0078	19	0,0055	15
36	Fabricação de móveis e indústrias diversas	0,0068	20	0,0063	13
27	Metalurgia básica	0,0059	21	0,0047	18
28	Fabricação de produtos de metal - exclusive máquinas e equipamentos	0,0053	22	0,0026	23
21	Fabricação de celulose, papel e produtos de papel	0,0037	23	0,0031	22

Nota: Os setores (divisões) destacados podem ser classificados como sendo intensivos em tecnologia de acordo com a taxonomia proposta por Cavalcante (2014).
Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

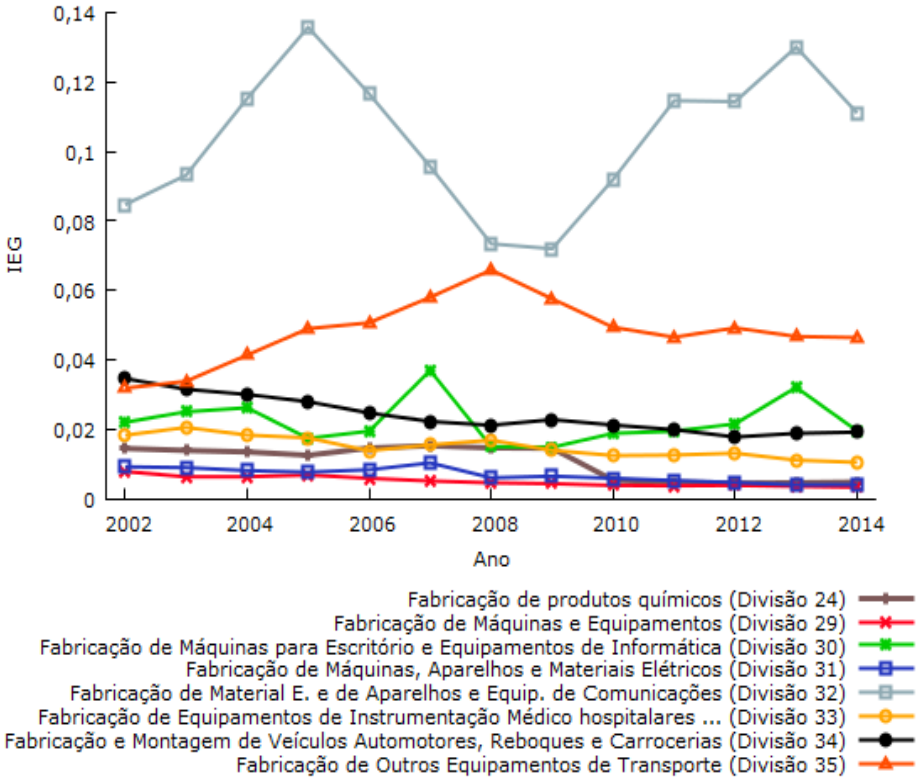
Para verificar se houve mudança no padrão de concentração em cada setor, os mesmos foram ranqueados de acordo com a magnitude do IEG nos dois anos.

Como pode ser visto na coluna com as posições para o ano de 2002 na Tabela 3.3, cinco dos oito setores classificados com de alta e média-alta tecnologia encontram-se entre

as 10 primeiras posições no *ranking*. Já em 2014, quatro setores de tecnologia estão entre as 10 primeiras posições no *ranking*. Parte da variação positiva na média do IEG para a indústria da transformação (como um todo e dos setores não intensivos em tecnologia) pode ser explicada pelo aumento significativo do IEG no setor de fabricação de coque, refino de petróleo, elaboração de combustíveis nucleares e produção de álcool (Divisão 23). Esse aumento no indicador de concentração associado a este setor especificamente pode ser consequência da exploração do pré-sal como haviam destacado Vignandi, Parré e Guimarães (2016).

A dinâmica do IEG por ano, no período 2002-2014, e por setores de alta e média-alta tecnologia pode ser observada a partir da Figura 3.3.

Figura 3.3: IEG para o Brasil no período 2002-2014 (setores intensivos em tecnologia)



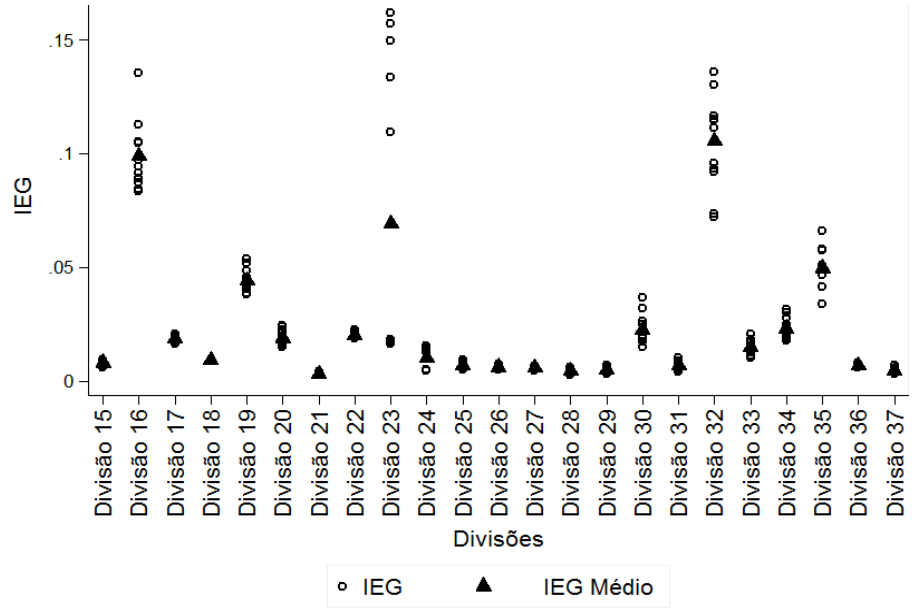
Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

É possível verificar, por exemplo, como o índice para o setor de fabricação de material eletrônico e de aparelhos e equipamentos de comunicações (Divisão 32), segundo no *ranking* em 2002 e 2014 se comportou ao longo dos anos e perceber seu comportamento ascendente com período de queda entre 2005 e 2009. Assim como também é possível notar a tendência decrescente do indicador para o setor de fabricação e montagem de veículos

automotores, reboques e carrocerias (Divisão 34).

Os dados indicam que existe uma heterogeneidade significativa entre os padrões de concentração setorial nas microrregiões do Brasil. De certo modo, isso já era esperado, dado que cada setor possui oferta e demanda específicas e a heterogeneidade do conjunto de microrregiões. Essas diferenças podem ser observadas a partir da Figura 3.4 que apresenta a variação média do índice de concentração entre os setores considerando o tempo fixo.

Figura 3.4: Heterogeneidade do índice de Ellison e Glaeser entre os setores



Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

A heterogeneidade verificada entre os setores foi significativa. Para justificar isto em sua pesquisa, Resende e Wyllie (2005) argumentaram que como seu estudo foi feito para o país, considerando deste modo uma elevada diversidade em termos econômicos, climáticos, geográficos e sociais as diferenças expressivas encontradas já eram esperadas.

As evidências empíricas encontradas nesse primeiro estágio da pesquisa sugerem de modo geral que, considerando toda a indústria da transformação, foi possível observar uma tendência de desconcentração industrial entre as microrregiões do Brasil. Essa tendência pode ser observada quando desconsiderados os possíveis fatores exógenos ao mercado - que podem ser importantes na determinação da localização das firmas, como as políticas de incentivos fiscais, mas não se enquadram no contexto das economias de aglomeração - que atuam sobre o aumento da média do índice de concentração. Além disto, como no caso da Divisão 23, existe ainda a exploração dos recursos naturais. No caso deste

setor especificamente, tanto a exploração dos recursos naturais quanto a atuação do setor público contribuíram para o aumento significativo do indicador de concentração. Por este motivo, apresenta uma trajetória diferente dos setores industriais de um modo geral.

Neste sentido, os resultados aqui apresentados estão de acordo com Silveira Neto (2005), Resende e Wyllie (2005) e Lautert e Araújo (2007), para o Brasil. Indicando que o padrão de concentração observado nacionalmente, desconsiderando setores possivelmente afetados por fatores exógenos ao mercado, fugindo desta forma do escopo das economias de aglomeração, segue a tendência observada pelos autores de que os setores intensivos em tecnologia são mais concentrados.

Neste estudo, essa heterogeneidade também pode ser consequência, em parte, das diferenças entre as microrregiões do país e das características de cada setor. Esta é a motivação para se verificar os possíveis efeitos dos fatores intrínsecos ao mercado sobre a concentração. O capítulo 6 traz uma discussão e evidências a partir do modelo empírico descrito no capítulo 5 sobre este tema.

Capítulo 4

Labor Pooling: Modelo Teórico

De acordo com Duranton e Puga (2004), os microfundamentos das economias de aglomeração podem ser distinguidos de três formas com base nos *linkages* de mercado (*sharing*), na eficiência no mercado de trabalho (*matching*) e mecanismos de aprendizagem (*learning*). Nesta pesquisa é explorada a força de aglomeração exercida a partir da concentração de trabalhadores, através da ideia introduzida por Marshall (1890) de *labor pooling*: a concentração de mão de obra especializada gera ganhos de produtividade para as firmas e trabalhadores reduzindo, por exemplo, o tempo de contratação da mão de obra e aumentado a capacidade das firmas de se ajustar aos choques de produtividade.

Krugman (1991) propõe um modelo teórico para explicar como essa força atua na distribuição da atividade produtiva. No modelo, a indústria é composta por um número discreto de firmas, que produzem utilizando como fator de produção o trabalho que é considerado homogêneo para a atividade na qual a indústria é especializada. A função de produção de cada empresa i pode ser descrita de acordo com:

$$y_i = (\beta + \varepsilon_i) l_i - \frac{1}{2}\psi l_i^2 \tag{4.1}$$

Onde β capta a mudança adicional na produção devido a um incremento marginal na força de trabalho (l_i); ε_i capta o efeito de choques de produtividade específicos da firma com $\varepsilon_i \sim (0, \sigma^2)$, e ψ mensura os retornos da produção. Cada estabelecimento tem o salário local como dado por:

$$w = \beta + \varepsilon_i - \psi l_i \tag{4.2}$$

Com isso, após o choque de produtividade, cada estabelecimento contrata mais ou menos trabalho até que o equilíbrio seja estabelecido novamente, o produto marginal do trabalho se iguale ao salário. A demanda por trabalho do estabelecimento i pode ser representada por:

$$l_i = \frac{\beta + \varepsilon_i - w}{\psi} \tag{4.3}$$

Em uma localidade, a força de trabalho consiste no somatório do total de trabalhadores empregados em cada firma. Considerando o equilíbrio no mercado de trabalho e n empresas podemos escrever o seguinte:

$$\sum_{i=1}^{n} l_i = \sum_{i=1}^{n} \frac{\beta + \varepsilon_i - w}{\psi} = \frac{1}{\psi}\left(n\beta + \sum_{i=1}^{n} \varepsilon_i - nw\right) = L \tag{4.4}$$

Utilizando (4.4) pode-se definir que o salário médio na localidade depende do número de trabalhadores, do número de firmas e do choque de produtividade em cada firma:

$$\frac{n\beta + \sum_{i=1}^{n} \varepsilon_i - nw}{n} = \psi\frac{L}{n} \equiv w = \beta - \psi\frac{L}{n} + \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i \tag{4.5}$$

O salário esperado é dado por:

$$E(w) = \beta - \psi\frac{L}{n} \tag{4.6}$$

considerando que $\varepsilon_i \sim (0, \sigma^2)$.

De (4.6), observa-se que o salário esperado se relaciona positivamente com o número de firmas, na medida em que o número de firmas aumenta na localidade, a relação força de trabalho/firma diminui, aumentando, portanto, a produtividade marginal desse fator e consequentemente os salários. Por outro lado, relaciona-se inversamente com o número de trabalhadores, por conta da redução da produtividade, e com a intensidade dos retornos decrescentes (ψ). A variância do salário depende do número de firmas:

$$var(w) = \frac{\sigma^2}{n} \tag{4.7}$$

A covariância do salário com o choque de produtividade específico da firma também depende do número de firmas:

$$cov(w, \varepsilon_i) = \frac{\sigma^2}{n} \tag{4.8}$$

Considerando os custos com mão de obra, o lucro da firma pode ser denotado por:

$$\pi_i = y_i - wl_i = (\beta + \varepsilon_i)l_i - \frac{1}{2}\psi l_i^2 - (\beta + \varepsilon_i - \psi l_i)l_i = \frac{(\beta - w - \varepsilon_i)^2}{2\psi} \tag{4.9}$$

Da equação anterior, o lucro é uma função convexa dos choques idiossincráticos e

dos salários. O lucro esperado é dado por:

$$E(\pi_i) = \frac{[\beta - E(w)]^2 + var(\varepsilon_i - w)}{2\psi} \tag{4.10}$$

Substituindo (4.6) e desenvolvendo o termo $var(\varepsilon_i - w)$ em (4.10) obtém-se:

$$E(\pi_i) = \frac{\psi}{2}\left(\frac{L}{n}\right)^2 + \frac{var(\varepsilon_i) + var(w) - 2cov(w, \varepsilon_i)}{2\psi} \tag{4.11}$$

Da equação anterior, na ausência de choques, o lucro seria dado pelo primeiro termo do lado direito da equação. Existiria, neste caso, uma correlação positiva entre a razão trabalho/estabelecimento e o lucro esperado em função da redução do salário esperado. O segundo termo do lado direito em (4.11) capta o efeito *labor pooling*. "Cada empresa se beneficia da partilha do mercado de trabalho com outras empresas em face aos choques idiossincráticos" (DURANTON; PUGA, 2004, p.17).

O lucro esperado é positivamente correlacionado com $var(\varepsilon_i)$ e $var(w)$ por ser convexo em relação a ε_i e w. Por outro lado, é negativamente correlacionado com o termo $cov(w, \varepsilon_i)$. Neste último caso, como destacado por Overman e Puga (2010), se os salários são maiores quando o estabelecimento pretende expandir sua produção em resposta a um choque positivo e menores quando o estabelecimento pretende diminuir sua produção em função de um choque negativo, a capacidade de ajuste por parte dos estabelecimentos fica limitada, no sentido em que as expectativas de lucro caem em função dos ajustes nos salários. Esta é a ideia fundamental do modelo que ressalta os micro fundamentos do *labor pooling* como fonte de aglomeração (DURANTON; PUGA, 2004; OVERMAN; PUGA, 2010). Ou seja, os estabelecimentos preferem os locais onde os choques não são refletidos fortemente nos salários.

Usando a $var(w) = \frac{\sigma^2}{n}$ e $cov(w, \varepsilon_i) = \frac{n}{\sigma^2}$ e considerando, além destes termos, o fato que $var(\varepsilon_i) = \sigma^2$ na equação (4.11) obtém-se:

$$E(\pi) = \frac{\psi}{2}\left(\frac{L}{n}\right)^2 + \left(1 - \frac{1}{n}\right)\frac{\sigma^2}{2\psi} \tag{4.12}$$

Os lucros esperados são iguais para estabelecimentos de um mesmo local e setor. Além disto, o efeito *labor pooling* mensurado pelo segundo termo do lado direito em (4.12) é mais forte quando σ^2 é maior. Isto implica que os benefícios da concentração de trabalhadores são maiores quando a heterogeneidade do choque no setor é maior. Assim, setores com choques mais heterogêneos tendem a ser mais propensos à aglomeração.

O modelo de Krugman (1991) sugere que os estabelecimentos tendem a se aglomerar

espacialmente nos locais com um grande número de trabalhadores com as habilidades necessárias para sua produção. As firmas podem obter vantagens dessa oferta de mão de obra e se adaptar melhor a choques exógenos de produtividade sem reduções significativas no lucro esperado em função de ajustes nos salários locais. Como mostra o modelo, o efeito *labor pooling* será tanto maior quanto maior a heterogeneidade do choque no setor específico. Partindo desta hipótese teórica, de acordo com Overman e Puga (2010), para isolar de forma direta o efeito *labor pooling*, toma-se a diferença entre a variação percentual do emprego no estabelecimento e a variação percentual na indústria (setor) em valor absoluto. Essa variável assumirá valores maiores quando um estabelecimento está expandindo seu emprego enquanto o resto da indústria está reduzindo ou vice-versa, captando a heterogeneidade dos choques. Esta será estratégia empírica adotada nesta pesquisa para a construção do modelo empírico.

Cabe ressaltar ainda que o potencial *labor pooling* é calculado a partir dos microdados da RAIS; através dessa base é possível identificar cada estabelecimento através do CNPJ.

Capítulo 5

Estratégia Empírica

5.1 Modelo Empírico

Como destacado anteriormente, o modelo de Krugman (1991) sugere uma relação positiva entre a concentração de trabalhadores e firmas no espaço. Mais precisamente, que as firmas podem se ajustar melhor ao mercado aumentando ou reduzindo a quantidade de mão de obra na produção frente a choques exógenos, beneficiando-se da concentração de trabalhadores. E, essa capacidade de ajuste das firmas será tanto maior, quanto mais elevada a heterogeneidade com que as demais firmas do mesmo setor ajustam a parcela de trabalho na produção. Diante desta hipótese e da elevada heterogeneidade com que as indústrias se distribuem geograficamente no Brasil, esta pesquisa busca verificar empiricamente a hipótese do modelo, com dados para o país, no sentido de obter evidências da importância do *labor pooling* para explicar a aglomeração industrial no mesmo.

A estratégia empírica utilizada na presente pesquisa segue o estudo de Overman e Puga (2010), que usa o IEG como variável dependente numa regressão em função da variável que busca captar o efeito da concentração de trabalhadores e demais controles para as características setoriais. Com base nos argumentos teóricos de que o *labor pooling* é importante para explicar a concentração da atividade produtiva industrial, o modelo estimado é estruturado na forma de painel estático e pode ser representado de acordo com a seguinte equação:

$$\gamma_{st} = x'_{st}\theta_{st} + \epsilon_{st} \tag{5.1}$$

Onde θ_{st} é um vetor $(kx1)$ de parâmetros que captam os efeitos das variáveis exógenas sobre a variável explicada, γ_{st}; x_{st} é uma matriz $(kx1)$ de variáveis explicativas; e ϵ_{zt} o termo de erro estocástico i.i.d.$(0, \sigma^2)$.

Especificando as variáveis exógenas:

$$\gamma_{st} = \alpha + \psi P_{st} + \lambda L_{st} + \varphi \Lambda_{st} + \epsilon_{st} \tag{5.2}$$

A variável dependente, γ_{st}, é o Índice de Concentração de Ellison e Glaeser; s indexa os setores da indústria da transformação (15, ..., 37); t indexa os anos (t = 2003, ..., 2014); P_{st} é uma *proxy* representado o *labor pooling*, construída da forma proposta por Overman e Puga (2010), que consiste na diferença entre a variação percentual do emprego no estabelecimento e na indústria em valor absoluto; L_{st} é uma variável controle e representa dependências intersetoriais, incorporadas no modelo seguindo o estudo de Silveira Neto (2005) que aplicou a metodologia usada por Kim (1995) ao caso brasileiro; Λ_{st} é um vetor de controles com características setoriais que variam no tempo; e ϵ_{st} é o termo de erro estocástico $i.i.d.(0, \sigma^2)$.

A Tabela 5.1 apresenta um resumo das variáveis incluídas nas regressões.

Tabela 5.1: Resumo das variáveis do Modelo Empírico

Variável	Descrição	Fonte	Período
γ_{st}	Índice de Concentração de Ellison e Glaeser	RAIS (dados agregados e microdados)	2003 - 2014
P_{st}	*Labor pooling*	RAIS (microdados)	2003 - 2014
L_{st}	Dependências Intersetoriais	IBGE - Pesquisa Industrial Anual - Empresa (PIA)	2003 - 2014*
gp_{st}	Gastos de pessoal	IBGE - Pesquisa Industrial Anual - Empresa (PIA)	2003 - 2014*
cec_{st}	Gastos com energia elétrica e consumo de combustíveis	IBGE - Pesquisa Industrial Anual - Empresa (PIA)	2003 - 2014*
$alug_{st}$	Gastos com aluguéis e arrendamentos	IBGE - Pesquisa Industrial Anual - Empresa (PIA)	2003 - 2014*
imp_{st}	Gastos com impostos e taxas	IBGE - Pesquisa Industrial Anual - Empresa (PIA)	2003 - 2014*

Nota: * Para os setores de Fabricação de material eletrônico e de aparelhos e equipamentos de comunicações; Fabricação de equipamentos de instrumentação médico-hospitalares, instrumentos de precisão e ópticos, equipamentos para automação industrial, cronômetros e relógios; e Reciclagem, as informações não estão disponíveis para o período após 2007.
Fonte: Elaboração própria.

No vetor Λ_{st} estão incluídas as seguintes variáveis controles: gp_{st} = Gastos de pessoal (Mil Reais), incluída seguindo a estratégia de Khan (2010), que utiliza salários como um possível fator determinante da localização das indústrias; cec_{st} = Gastos da indústria com compras de energia elétrica e consumo de combustíveis (Mil Reais), controle para captar a parcela dos gastos com enegia elétrica na produção, incluída com base em Alkay e Hewings (2012) e Overman e Puga (2010); $alug_{st}$ = Gastos da indústria com aluguéis e arrendamentos (Mil Reais), que intuitivamente pode afetar de alguma forma a concentração da indústria, que foi incluída no modelo com base em Alkay e Hewings (2012); imp_{st} = Dispêndios da indústria com impostos e taxas (Mil Reais); intuitivamente

espera-se que a arrecadação de impostos possa influenciar o modo como as firmas se organizam no espaço. Esta última variável, no entanto, não é ideal para captar os efeitos dos incentivos fiscais sobre a concentração porque varia por indústria e no tempo e não por microrregião; por esse motivo os modelos serão estimados com e sem essa variável, tratando-se, deste modo, de um teste de robustez. Todas as variáveis estão em valores reais, deflacionadas pelo deflator do PIB (2000 = 100).

A variável *proxy* para o *labor pooling* foi obtida de acordo com a seguinte expressão:

$$P_s = |\Delta\%emp_{estab} - \Delta\%emp_s| \tag{5.3}$$

Onde $\Delta\%emp_{estab}$ = variação percentual do emprego no estabelecimento; $\Delta\%emp_s$ = variação percentual do emprego no setor (indústria). A heterogeneidade do choque de produtividade é mensurada pelo valor absoluto da diferença entre essas duas medidas de variação. Da forma descrita, este indicador varia por estabelecimento, para agregar a variável por setores, seguindo Overman e Puga (2010), tomou-se a média por indústria.

Por fim, o controle para dependências intersetoriais foi construído conforme o estudo de Silveira Neto (2005) para o Brasil e pode ser representado pela equação abaixo:

$$L_{st} = \frac{C_{st}}{VTI_{st}} \tag{5.4}$$

Em que C_{st} = Custos com consumo de matérias-primas, materiais auxiliares e componentes no setor s no período t; VTI_{st} = Valor da transformação industrial no setor s no período t.

5.2 Dados

A principal base de dados utilizada é a Relação Anual de Informações Sociais do Ministério do Trabalho (RAIS-MTE)[10], os dados agregados setorialmente por divisões da CNAE e por microrregiões, e também os microdados, onde é possível identificar cada firma através do código único e imutável, o número do Cadastro Nacional de Pessoa Jurídica (CNPJ), bem como acompanhar sua trajetória anual. Por meio deste banco de dados, tornou-se possível a implementação da estratégia de Overman e Puga (2010) ao caso brasileiro, através dos dados longitudinais disponíveis para todas as firmas da indústria da

[10]BRASIL. Ministério do Trabalho e Emprego. RAIS - Relação Anual de Informações Sociais. Disponível em: http://www. rais. gov. br. Acesso em: 20 de junho de 2016.

transformação no país no período 2002-2014. O período analisado foi definido mediante a disponibilidade dos dados. Os microdados da RAIS são provenientes de um registro anual que cobre todos os estabelecimentos formais do Brasil. O objetivo desta base é monitorar a trajetória das firmas. Os dados estão organizados por estabelecimento, fornecendo suas principais características ao final de cada ano.

Adicionalmente, também são considerados dados da Pesquisa Industrial Anual - Empresa (PIA) fornecidos pelo IBGE[11] sobre as características dos setores da indústria da transformação.

O modelo geral apresentado na equação (5.1) pode ser especificado de várias formas (GREENE, 2003; BALTAGI, 2008; WOOLDRIDGE, 2010). Dentre elas podemos destacar o modelo dos efeitos fixos (EF), também conhecido como transformação *within* e o dos efeitos aleatórios (EA). No primeiro caso, admite-se que os parâmetros, $\theta's$, são os mesmos para todos os setores (CNAE dois dígitos), exceto θ_{1s}, ou seja : $\theta_{kst} = \theta_k$, $\forall$ s, t, exceto para $k = 1$ onde $\theta_{1st} = \theta_{1s}$. Na segunda estratégia, a heterogeneidade individual é tratada de forma aleatória, ou seja, incorporada no termo de perturbações que é particionado, uma parte comum e outra individual. A definição de qual método é mais adequado aos dados é feita com base no Teste de especificação proposto por Hausman (1978).

Para testar a homocedasticidade dos erros utilizou-se o teste de Wald modificado, que também pode ser usado para painéis não balanceados. Foram implementados testes para dependência *cross-section*, apesar de esse tipo de problema ser mais comum em macropainéis (BALTAGI, 2008). Com este objetivo, em geral, utiliza-se o teste do Multiplicador de Lagrange (LM) desenvolvido por Breusch e Pagan (1980). A eficiência desse teste está condicionada à estrutura dos dados, ou seja, $T > N$, onde T representa as informações de série temporal e N as unidades *cross-section*. Como neste estudo estima-se um modelo no qual o número de observações de série temporal é menor que as de *cross-section* $T < N$, utilizou-se o teste proposto por Pesaran (2004) que também pode ser implementado em painéis não balanceados.

Realizaram-se testes de hipóteses para efeitos fixos no tempo. E, também, foram implementados testes de raiz unitária para dados em painel. Para esta finalidade existem alguns tipos de testes. Por exemplo, o teste de Levin, Lin e Chu (2002), o de Harris e Tzavalis (1999), o teste Fisher-Type proposto por Choi (2001), o de Breitung (1999) e Breitung e Das (2005), e o teste IPS de Im, Pesaran e Shin (2003). A definição do teste a ser aplicado depende das características assintóticas do painel. Isto é, as proporções

[11] BRASIL. Sistema IBGE de Recuperação Automática. Disponível em: https://sidra.ibge.gov.br. Acesso em: 15 de agosto de 2016.

de N e T e das restrições impostas aos dados. Por fim, para testar a correlação serial, utilizou-se o teste proposto por Wooldridge (2010).

Capítulo 6

Evidências do *labor pooling* para o Brasil

Com objetivo de explicar a concentração da indústria a partir de efeitos distorcivos gerados por fatores locais, especialmente o derivado da concentração de trabalhadores, esta pesquisa testa a hipótese teórica do modelo de Krugman (1991) utilizando dados para o Brasil. A estratégia empírica segue a sugerida por Overman e Puga (2010) para a construção da variável *labor pooling*. A sinalização de que o acesso ao mercado de trabalho seria um importante fator determinante da concentração da atividade industrial existe desde Marshall (1890). Apesar disto, somente com a contribuição de Krugman (1991) esses argumentos ganharam uma formalização (*Labor Pooling Model*). A hipótese teórica é que setores com choques de produtividade mais heterogêneos tendem a aglomerar-se espacialmente.

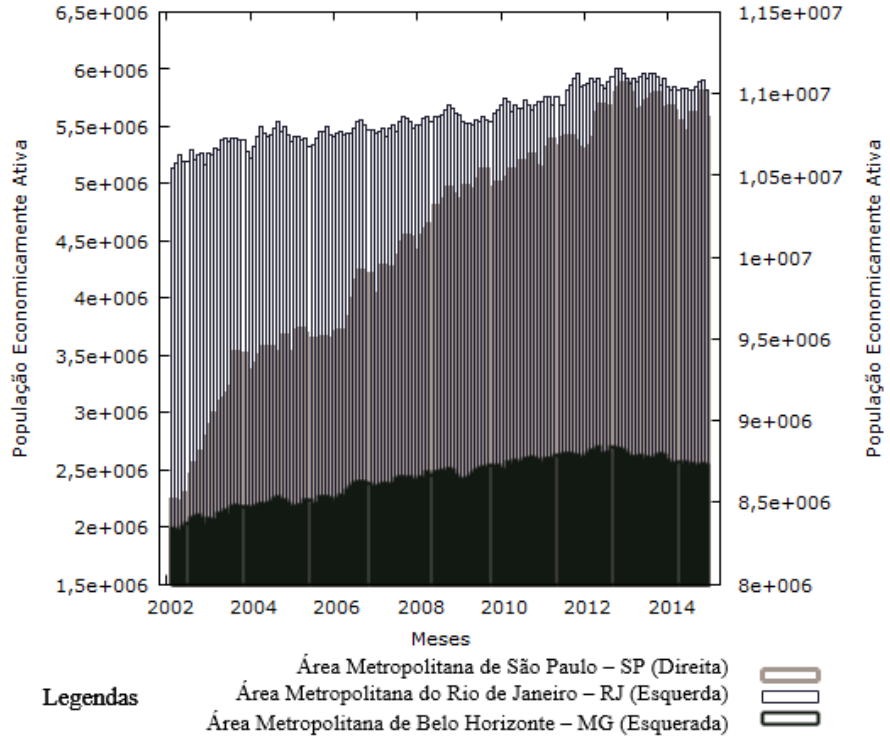
É intuitivo pensar que as firmas consideram o tamanho do mercado de trabalho onde irão instalar suas plantas produtivas como um dos fatores importantes na tomada de decisão e que esse mecanismo pode atuar como uma força centrípeta levando-as à concentração espacial. Sendo um fator de produção, a oferta de mão de obra em abundância ou, por outro lado, a escassez dela pode afetar de forma significativa o planejamento das indústrias, principalmente naquelas intensivas em trabalho ou nas que precisam de algum tipo específico de mão de obra qualificada. Como discutido anteriormente, além de intuitiva, a relação positiva entre a concentração espacial de trabalhadores e firmas pode ser verificada em diferentes países do ponto de vista empírico. No entanto, explicitar essa relação não parece uma tarefa trivial, principalmente em função da falta de dados e, na maioria das vezes, de variáveis que forneçam informações adequadas para esse propósito.

No Brasil, de acordo com os dados do Instituto Brasileiro de Geografia e Estatística (IBGE) - Pesquisa Mensal de Emprego População Economicamente Ativa (PMEPEA), nas principais regiões metropolitanas[12] a força de trabalho aumentou no período 2002-2014. Como pode ser visto nas Figuras 6.1 e 6.2 a seguir para as áreas metropolitanas de São Paulo - SP, Rio de Janeiro - RJ e Belo Horizonte - MG e áreas metropolitanas de Recife - PE, Porto Alegre - RS e Salvador - BA, respectivamente. Em todos os casos é

[12] As regiões apresentadas foram definidas de acordo com a disponibilidade dos dados.

possível observar uma trajetória ascendente da População Economicamente Ativa (PEA), destacando-se o expressivo crescimento em São Paulo - SP (8.515.647 em 02/2002 e 10.851.561 em 12/2014).

Figura 6.1: PEA nas áreas metropolitanas de São Paulo, Rio de Janeiro e Belo Horizonte no período 2002-2014

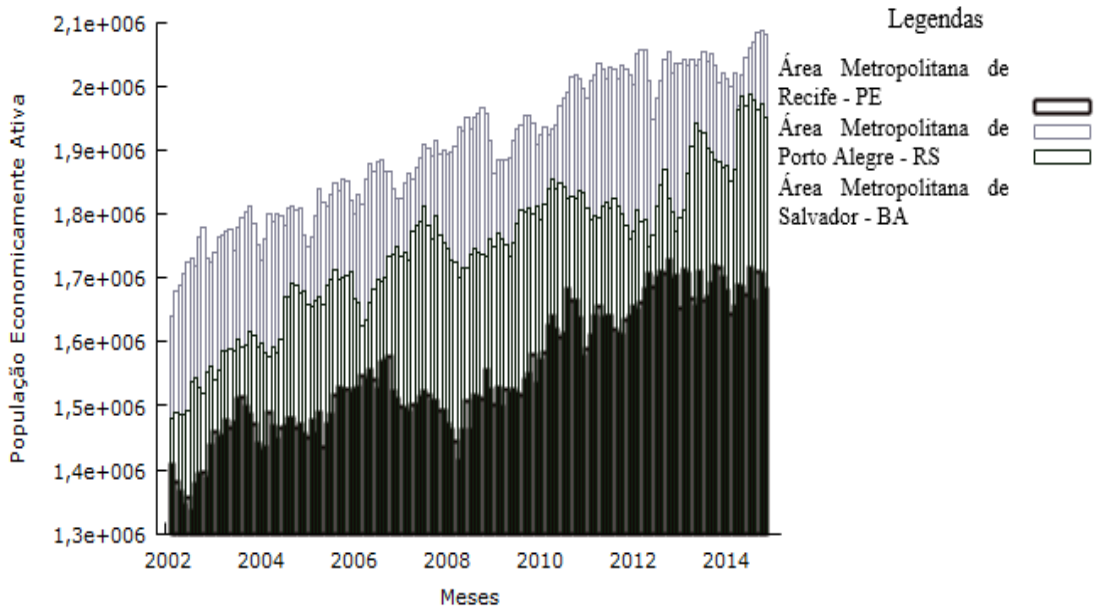


Fonte: IBGE - Pesquisa Mensal de Emprego População Economicamente Ativa (PMEPEA).

Diversos fatores podem ser responsáveis pelo crescimento da PEA nessas cidades, como o crescimento demográfico e a imigração. Neste último caso, é intuitivo pensar que a oferta de emprego nas grandes cidades pode ser um dos fatores determinantes da imigração regional de trabalhadores. De fato, uma série de estudos trata essa questão na literatura nacional; abordá-los, no entando, foge do escopo da presente pesquisa. Mas, chegamos à principal dificuldade enfrentada em estudos como este: a causalidade reversa. Se por um lado a concentração de trabalhadores pode ser uma fonte de aglomeração industrial, por outro, a aglomeração de indústrias pode ser um importante fator determinante da imigração regional de mão de obra. Por este motivo, considerar parcelas específicas do mercado de trabalho, como é feito pela maioria dos estudos empíricos nessa área, pode não ser uma estratégia adequada. É neste ponto que Overman e Puga (2010) avançam ao proporem uma maneira de estudar os efeitos do *labor pooling* a partir da diferença entre

a variação do emprego ao nível do estabelecimento e a variação do emprego ao nível da indústria (nesta pesquisa, CNAE dois dígitos).

Figura 6.2: PEA nas áreas metropolitanas de Recife, Porto Alegre e Salvador no período 2002-2014



Fonte: IBGE - Pesquisa Mensal de Emprego População Economicamente Ativa (PMEPEA).

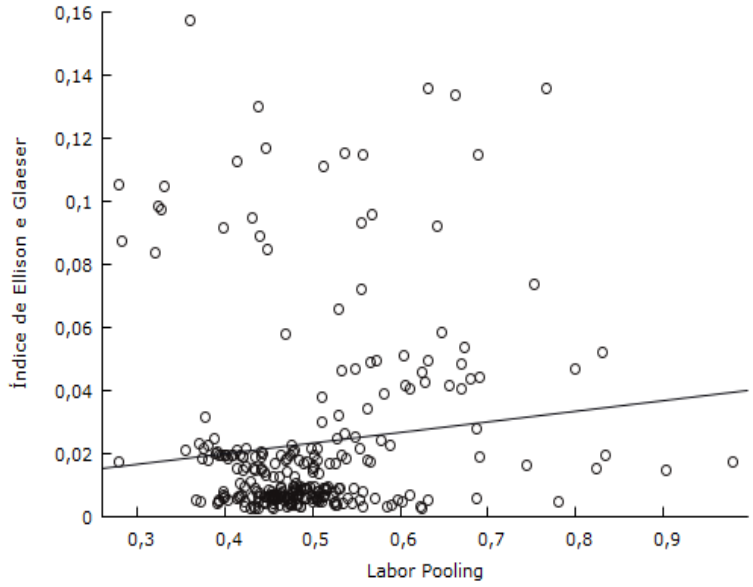
Com base na medida de concentração de Ellison e Glaeser e a estratégia de Overman e Puga (2010), tornou-se possível o estudo da relação de concentração firmas x trabalhadores através de uma associação unilateral, ou seja, minimizando o problema da causalidade reversa. A dispersão do IEG x *labor pooling* na Figura 6.3 indica uma associação positiva entre a concentração da indústria e o *labor pooling*.

A variável *labor pooling* mostrou-se bastante heterogênea entre os setores como pode ser observado na Figura 6.4. Considera-se o tempo fixo e se obtém a média da variável para cada setor; com isso é possível obter indicativos de como essa variável se comporta entre os setores.

Da forma que foi construída esta variável é capaz de isolar o efeito causado pela aglomeração de trabalhadores e será maior quando um estabelecimento específico estiver expandindo o emprego (contratando) enquanto o resto da indústria está reduzindo o número de trabalhadores, ou vice-versa. Ou seja, beneficiando-se da concentração de trabalhadores com as características necessárias para seu processo produtivo, as firmas podem se ajustar melhor aos choques de produtividade sem incorrer em grandes ajustes nos salários locais. E este "ganho" será tanto maior quanto maior for a heterogeneidade

com que as demais firmas do mesmo setor ajustam a parcela de trabalho empregada em seus respectivos processos produtivos.

Figura 6.3: Índice de Ellison e Glaeser *versus Labor Pooling*



Nota: *Outliers* removidos da amostra (valores maiores que 1 para o *labor pooling*).
Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

O coeficiente de correlação entre o IEG e o *labor pooling* apresentado na Tabela 6.1 também sugere uma relação positiva entre essas variáveis. Isto pode indicar que a hipótese teórica do modelo de Krugman pode ser verificada empiricamente para a indústria da transformação no Brasil.

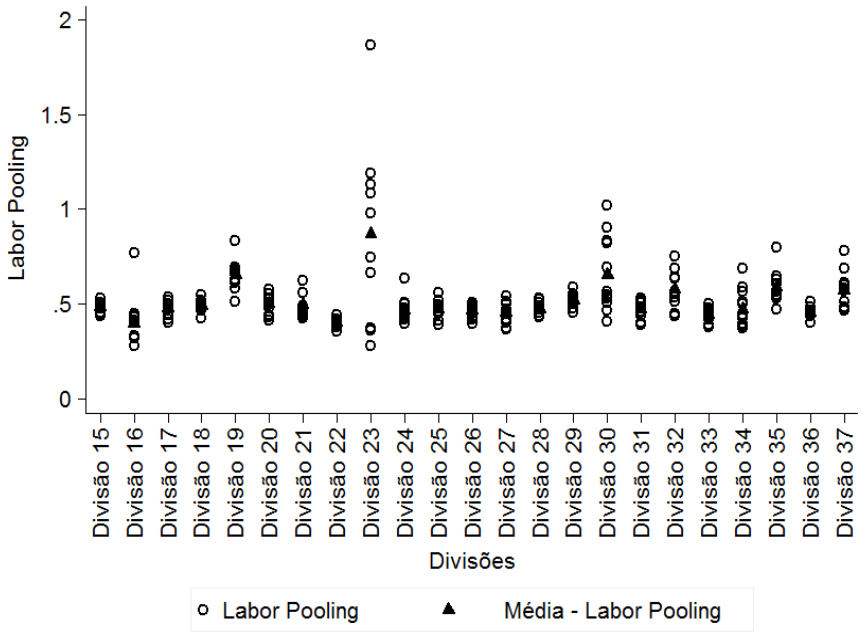
Tabela 6.1: Matriz de Correlação das variáveis do modelo empírico

	γ_{st}	P_{st}	L_{st}	gp_{st}	cec_{st}	$alug_{st}$	imp_{st}
γ_{st}	1						
P_{st}	$0,2621$	1					
L_{st}	0.0176	$-0,0211$	1				
gp_{st}	$-0,0921$	$-0,0114$	$0,1349$	1			
cec_{st}	$-0,0674$	$-0,0915$	$0,1751$	$0,0769$	1		
$alug_{st}$	$0,046$	$0,1602$	$0,0326$	$0,5576$	$0,6069$	1	
imp_{st}	$-0,0479$	$0,1025$	$0,233$	$0,5719$	$0,627$	$0,8422$	1

Fonte: Elaboração própria com base nos dados da pesquisa.

A Tabela 6.2 apresenta as estatísticas descritivas das variáveis incluídas no modelo empírico. Adicionalmente, o modelo empírico apresentado no capítulo 5 foi estimado considerando dados que variam por setores e no tempo, sendo possível acompanhar a trajetória de uma variável para um determinado setor no período 2003-2014. O IEG foi

Figura 6.4: Heterogeneidade do *labor pooling* entre os setores



Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

calculado para o período 2002-2014. No entanto, como a variável *labor pooling* foi obtida a partir da variação do nível de emprego por estabelecimento e setor, a observação inicial foi perdida.

Tabela 6.2: Estatísticas descritivas das variáveis do modelo empírico

Variável	Média	Desvio Padrão	Mínimo	Máximo	Obs. ausentes
γ_{st}	0,0243	0,0323	0,0025	0,1617	0
P_{st}	0,5114	0,1486	0,2779	1,865	2*
L_{st}	1,1635	0,4362	0,37022	2,4029	21**
gp_{st}	10,3443	24,4922	0,0506	267,4135	21**
cec_{st}	8,4227	27,7634	0,0051	251,8339	21**
$alug_{st}$	0,4085	0,5544	0,0025	3,5891	21**
imp_{st}	0,3454	0,4063	0,0015	2,5366	21**

Nota: * observações díspares eliminadas da amostra, ** informações não disponíveis para o período após 2007. As variáveis gp_{st}, cec_{st}, $alug_{st}$ e imp_{st} foram ponderadas por $\frac{1}{1.000.000}$.
Fonte: Elaboração própria com base nos dados da pesquisa.

Os resultados das regressões estimadas são apresentados na Tabela 6.3. Inicialmente, desconsiderando-se a estrutura de painel, os resultados dos estimadores *Pooled* OLS podem ser observados nas colunas 2, 3 e 4 que apresentam os resultados dos modelos 1 ao 3, respectivamente. Nas três especificações do modelo, os parâmetros estimados foram significativos para a variável *labor pooling*.

Nos modelos 2 e 3, após incluir os controles, o parâmetro associado à variável *labor pooling* permanece positivo e significativo a 5%. Isto indica que, mesmo controlando para as características setoriais que variam no tempo e diferentes especificações do modelo, a hipótese do modelo teórico permanece válida para os dados do Brasil. No modelo 1, que pode ser interpretado como uma associação linear entre o *labor pooling* e o IEG, o coeficiente estimado para a variável de interesse foi 0,0565, indicando uma correlação positiva. Isto é, na medida em que a heterogeneidade captada pela *proxy* para o *labor pooling* varia em 1%, a concentração medida pelo IEG varia (aumenta) em 0,0565. Nos demais modelos estimados por OLS (2 e 3), os parâmetros foram 0,052 e 0,5439, respectivamente. A magnitude da estimativa não varia muito em relação ao modelo 1, principalmente no modelo 3, quando incluídos os demais controles e o efeito fixo temporal.

Nos modelos 4 ao 8, os parâmetros foram obtidos a partir dos estimadores de efeitos fixos. Assim, controla-se para as características não observadas e fixas no tempo. O parâmetro associado ao *labor pooling* também foi significativo estatisticamente em todos os modelos estimados por este método. Ou seja, mesmo controlando-se para as características observadas variáveis no tempo e não observadas fixas no tempo, o parâmetro se manteve positivo e significativo. Isto pode sugerir que esse fator local gerador de economias de aglomeração, pioneiramente destacado por Marshall, pode ser importante para explicar a aglomeração da indústria da transformação nas microrregiões do Brasil.

A estimação dos modelos de efeitos fixos é suportada pelo teste Hausman. Em todos os modelos a hipótese de homocedasticidade foi rejeitada a 5%. Deste modo, para lidar com este problema e obter estimativas robustas, todas as regressões foram estimadas com correção de heterocedasticidade.

No modelo 4, o parâmetro estimado para o *labor pooling* foi igual a 0,02243. Isto indica que a correlação positiva entre a heterogeneidade dos choques de produtividade do trabalho e o IEG, prevista no modelo 1, é mantida mesmo com a inclusão do efeito fixo setorial.

Nos modelos de 5 a 8, quando se controla para outros possíveis fatores determinantes da concentração, a importância do *labor pooling* aumenta do ponto de vista estatístico, o parâmetro associado a essa fonte de aglomeração se torna significativo a 0,1%. Diante destas evidências, os resultados obtidos a partir de diferentes estimadores indicam a importância do *labor pooling* para explicar a aglomeração, de acordo com a hipótese teórica do modelo de Krugman. Ou seja, as firmas consideram o tamanho do mercado de trabalho disponível na região onde se instalam. Este fator local pode ser importante, pois permite que as firmas se adaptem aos choques de produtividade reduzindo ou

aumentando a quantidade de trabalho empregado em sua produção, sem incorrer em variações significativas nos gastos com salários e consequentemente nos lucros.

Nos modelos 5 e 6, as estimativas para a variável em questão foram iguais 0,02718 e 0,02387, respectivamente. As demais variáveis controles não foram significativas em ambas as especificações. Assim, as evidências obtidas indicam que mesmo considerando o efeito fixo setorial e demais características setoriais variantes no tempo, um aumento absoluto de 1% na variável *labor pooling* implica em um aumento de magnitude 0,02718 e 0,02387 no IEG, respectivamente.

Nos dois últimos modelos, além de o parâmetro associado ao *labor pooling* ser significativo, as estimativas associadas às variáveis *proxies* para os *linkages* intersetoriais (L_{st}), gastos com pessoal (gp_{st}), gastos com consumo de energia elétrica e consumo de combustíveis (cec_{st}) e gastos com aluguéis e arrendamentos ($alug_{st}$) foram significativas estatísticamente.

Para os *linkages* intersetoriais, como destaca Silveira Neto (2005), o sinal esperado é negativo. O autor argumenta que uma maior dependência intersetorial pode beneficiar a menor concentração. No entanto, considerando a aglomeração da indústria em nível microrregional, é intuitivo pensar, por exemplo, que os setores nos quais as etapas de produção estejam mais ligadas, as firmas tendem a se localizar próximas umas das outras, reduzindo, por exemplo, os custos de transporte. Apesar disto, vale destacar que embora esta variável seja incluída em 6 das 8 especificações estimadas, somente em uma delas as evidências apontam para correlação positiva entre o IEG e L_{st}. Neste caso, um aumento de uma unidade em L_{st}, implicaria em um aumento de 0,01127 no IEG.

Para inclusão ou exclusão de variáveis nos diferentes modelos estimados, utilizou-se como critério a matriz de correlação apresentada na Tabela 6.1. Com relação à variável gastos com pessoal, nos dois modelos, 7 e 8, os parâmetros foram negativos e significativos a 1% e 5%, respectivamente. Isto pode indicar que quanto maiores os gatos com mão de obra, por exemplo, menor será a tendência à concentração das firmas no espaço. Este resultado encontra respaldo na ideia que está por trás do modelo de Krugman. O modelo sugere que as firmas mais concentradas estão menos sujeitas aos ajustes nos salários locais relativamente às firmas que estão mais dispersas geograficamente.

Tabela 6.3: Resultados do modelo empírico

Variável	*Pooled* OLS					Efeito Fixo		
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7	Modelo 8
P_{st}	$0,0565^{**}$	$0,052^{*}$	$0,05439^{*}$	$0,02243^{**}$	$0,02718^{***}$	$0,02387^{***}$	$0,02012^{***}$	$0,02202^{***}$
	$(0,01024)$	$(0,02224)$	$(0,02259)$	$(0,00784)$	$(0,00628)$	$(0,00435)$	$(0,00436)$	$(0,00473)$
L_{st}		$0,00308$	$0,00438$		$0,00540$	$0,00663$	$0,01127^{*}$	$0,01132$
		$(0,00677)$	$(0,00588)$		$(0,00414)$	$(0,00362)$	$(0,00492)$	$(0,00589)$
gp_{st}		-0.00010	$-0,00011^{*}$		$0,00005$	$-0,00002$	$-0,00017^{**}$	$-0,00019^{*}$
		$(0,00009)$	$(0,00005)$		$(0,00007)$	$(0,00003)$	$(0,00006)$	$(0,00007)$
cec_{st}		$-0,00004$	$-0,00015^{**}$		$0,00004$	$-0,00005$	$-0,00022^{*}$	$-0,00031^{*}$
		$(0,00007)$	$(0,00005)$		$(0,00006)$	$(0,00004)$	$(0,00008)$	$(0,00012)$
$alug_{st}$							$0,02508^{*}$	$0,02723^{**}$
							$(0,0114)$	$(0,00825)$
imp_{st}		$-0,00106$				$0,01231$		$0,00566$
		$(0,0124)$				$(0,01458)$		$(0,0072)$
Const.	$-0,00484$	$-0,00558$	$-0,00696$	$0,01259^{**}$	$0,00180$	$-0,00072$	$-0,00709$	$-0,00443$
	$(0,01024)$	$(0,01291)$	$(0,01222)$	$(0,004)$	$(0,00468)$	$(0,00581)$	$(0,00657)$	$(0,00845)$
Dum.Tempo	*não*	*não*	*sim*	*não*	*não*	*não*	*não*	*sim*
R^2	$0,0687$	$0,0798$	$0,0994$	$0,0336$	$0,0569$	$0,0742$	$0,2183$	$0,2633$
F	$7,9050$	$4,5911$	$2,0403$	$8,1859$	$13,4390$	$10,8196$	$28,2231$	$111,0393$
Num Obs	274	253	253	274	253	253	253	253

Nota: *, ** e *** indicam que os coeficientes são significativos a 5%, 1% e 0,1% respectivamente. Desvio padrão das estimativas entre parênteses. As variáveis gp_{st}, cec_{st}, $alug_{st}$ e imp_{st} foram ponderadas por $\frac{1}{1.000.000}$.

Fonte: Elaboração própria com base nos dados da pesquisa.

A estimativa associada à variável gastos com consumo de energia elétrica e consumo de combustíveis foi negativa. Isto pode indicar que esse componente dos custos variáveis atua favorecendo a desconcentração. Quanto menores os gastos com enégia elétrica, maior a concentração. Não é possível determinar claramente se o sinal esperado para essa estimativa seria positivo ou negativo. Dependendo do setor industrial, os efeitos podem ser distintos. Considerando os setores intensivos em trabalho, se os custos com energia elétrica não representam uma parcela significativa dos custos totais, este componente terá pouca importância relativa na decisão de escolha locacional da firma. No estudo de Overman e Puga (2010), a estimativa associada à participação dos gastos com energia elétrica nos insumos foi negativa, mas não significativa.

Por fim, a variável gastos com aluguéis e arrendamentos. Nos modelos 7 e 8, as estimativas associadas a esta variável foram significativas a 5% e 1%, respectivamente. O sinal do parâmetro foi positivo, indicando possivelmente que quanto maior o custo com aluguéis maior a tendência à concentração. Assim como no caso da variável controle anterior, não é clara a relação esperada entre a concentração industrial e os gastos com aluguéis. A depender da localização do polo de desenvolvimento, por exemplo, o custo com aluguéis e arrendamentos poderia representar um fator de repulsão da atividade. Por outro lado, se se epera, de modo geral, que os aluguéis sejam mais elevados em locais com grande concentração de indústrias, mas esse custo seja pequeno relativamente aos demais; mesmo com o preço alto do aluguel a indústria tenderia à concentração porque poderia se beneficiar de outras forças atrativas.

É importante ressaltar ainda que os fatores locais destacados por Marshall são características que podem explicar apenas de forma parcial a aglomeração, de modo que políticas desenvolvimentistas certamente possuem alguma influência nesse processo. Apesar da tentativa de controlar para este mecanismo exógeno através da variável imp_{st} (dispêndios da indústria com impostos e taxas), o parâmetro associado a esse controle não foi significativo. Como dito anteriormente, a inclusão dessa variável no modelo pode ser interpretada como um teste de robustez, já que por limitações impostas pelos dados não foi possível obter uma variável que fornecesse as características das políticas de incentivos fiscais por microrregiões. Além disto, o foco da pesquisa são os fatores intrínsecos ao mercado, especificamente o *labor pooling*; demais controles devem ser tratados em segundo plano.

Com relação à magnitude do parâmetro para o *labor pooling*, no geral, as estimativas obtidas na presente pesquisa são menores que as apresentadas por Overman e Puga (2010) para o Reino Unido. Isto pode sugerir que no Brasil, apesar de a concentração de

trabalhadores influenciar a distribuição espacial das firmas, a importância desse fator local é aparentimente menor no país. Para determinar em qual dos dois casos este tipo de externalidade poderia ser mais forte, seria necessário uma análise mais detalhada que considerasse uma grande parte das diferentes características entre as duas economias.

Também foi possível observar que, nos modelos estimados considerando o efeito fixo setorial, as estimativas para o *labor pooling* foram relativamente menores quando comparadas com as obtidas por OLS. Essa redução pode estar associada às características setoriais não observadas fixas no tempo, no sentido em que cada setor, em função dessas características, pode ser mais ou menos influenciado pela concentração de trabalhadores. Ademais, de acordo com a estatística F, os modelos com efeitos fixos são melhores na perspectiva da significância conjunta dos parâmetros.

O índice de Ellison e Glaeser já icorpora as economias de escala, que seria um possível fator determinante da aglomeração industrial, já que o índice é formado a partir das informações sobre o nível de emprego nas firmas. Por este motivo, não foram incluídas variáveis com a finalidade de captar este efeito nas regressões estimadas, assim como em Rosenthal e Strange (2001).

Como destacado na revisão de estudos empíricos, esta pesquisa se enquadra na abordagem do *labor pooling* como fonte de aglomeração. Apesar de adotar uma estratégia empírica distinta dos trabalhos referenciados, exceto o de Overman e Puga (2010), que permitiu isolar diretamente o efeito do fator em questão, os resultados convergem no sentido em que as evidências empíricas apresentadas por Rosenthal e Strange (2001), Khan (2010) e Alkay e Hewings (2012) também sugerem uma relação positiva entre o *labor pooling* e a aglomeração das indústrias. Os resultados para o Brasil também estão em consonância com os de Overman e Puga (2010) com relação ao sentido do efeito. Em geral, apesar das diferenças entres os países analisados nesses estudos, parece haver uma relação positiva que prevalece em todos eles fortalecendo a ideia de Marshall de que a concentração de trabalhadores é importante para explicar a distribuição espacial das firmas.

Capítulo 7

Considerações Finais

A presente pesquisa teve como objetivo analisar os padrões de concentração industrial no Brasil e investigar a importância do *labor pooling* para explicá-los. Para tanto, o nível de concentração da indústria da transformação foi quantificado através do IEG. Por meio desse indicador, foi possível obter evidências sobre os níveis de concentração industrial, bem como sobre a sua dinâmica. Os resultados obtidos indicam que houve uma redução média do índice entre 2002 e 2014, quando desconsiderados possíveis efeitos exógenos sobre o indicador, o que sugere que houve uma tendência à desconcentração industrial no país. No entanto, quando se considera as trajetórias por divisão de atividades da indústria da transformação, alguns destes apresentaram uma tendência à aglomeração, como é o caso do setor de Fabricação de coque, refino de petróleo, elaboração de combustíveis nucleares e produção de álcool (Divisão 23).

Paralelamente, também foi possível observar que as indústrias intensivas em tecnologia, em média, são mais concentradas que os demais setores da indústria da transformação. Esses resultados estão de acordo com os obtidos por Resende e Wyllie (2005), Silveira Neto (2005), Lautert e Araújo (2007) e Maciente (2008). Apesar de apresentarem índices de concentração médios maiores que os demais setores, as indústrias classificadas com sendo de alta e média-alta tecnologia de acordo com o método de Cavalcante (2014) também apresentaram tendência à desconcentração espacial, embora essa trajetória seja mais suave relativamente à apresentada por todos os setores conjuntamente e os demais não intensivos em tecnologia.

A análise a partir do índice de Ellison e Glaeser fornece evidências de que a distribuição espacial das firmas nas microrregiões do Brasil pode ser mais concentrada do que deveria se ocorresse de forma aleatória. Isto é, sugere a existência de fatores locais e/ou exógenos ao mercado atuando sobre a decisão locacional da indústria. A partir desta constatação inicial, surge a necessidade de explorar quais fatores são responsáveis por gerar benefícios às indústrias levando-as à concentração espacial.

Com base nos argumentos de Marshall (1890) sobre o *labor pooling* e baseando-se na formalização microfundamentada proposta por Krugman (1991) para esse fator de

aglomeração (*Labor Market Pooling Model*), foi realizado um teste empírico para obter evidências da importância desse fator local para explicar a concentração industrial no Brasil. Seguindo a estratégia empírica de Overman e Puga (2010), foi possível obter evidências que corroboram a hipótese teórica do modelo de que o *labor pooling* está relacionado positivamente com a concentração, indicando que os setores industriais podem se beneficiar da concentração de trabalhadores. A hipótese foi verificada a partir de dois estimadores distintos, e para todos os casos os resultados foram estatisticamente significantes e de acordo com o esperado, mesmo com a inclusão de controles para características não observadas fixas no tempo e controles para economias de escala, *linkages* intersetorias e características setoriais que variam no tempo.

Este estudo vem contribuir no campo empírico para o Brasil ao apresentar evidências de que as indústrias espacialmente concentradas apresentam vantagens produtivas ao se beneficiarem com a concentração de mão de obra. Beneficiando-se do *labor pooling*, as indústrias internalizam as externalidades positivas geradas por esse fator local tornando-se mais produtivas, quando comparadas com aquelas que estão isoladas espacialmente. Este é um importante resultado norteador de políticas regionais desenvolvimentistas que buscam tornar as indústrias nacionais mais produtivas, e estão em consonância com aquelas que estão direcionadas ao fortalecimento de Arranjos Produtivos Locais. Dadas estas evidências, recomenda-se que as políticas industriais que objetivem o aumento da produtividade no país passem a levar em consideração as potencialidades das regiões, a vocação produtiva dos estabelecimentos e trabalhadores locais, tendo-se em mente que a concentração industrial não deve ser considerada como "prejudicial" à economia; pelo contrário, em geral são resultados de forças de mercado que geram externalidades produtivas podendo alavancar a produtividade do país.

Espera-se contribuir para o debate sobre este tema no Brasil, fornecendo evidências que diferem das existentes previamente na literatura em dois pontos principais. A questão da contração de trabalhadores é tratada com respaldo teórico. Para o país, apesar de não abordar a aglomeração, mas a coaglomeração industrial, o estudo de Resende (2015) aborda o *labor pooling* de uma perspectiva empírica. O segundo ponto em que este estudo avança na literatura reporta-se à estratégia empírica adotada com dados ao nível da firma. Como discutido anteriormente, a causalidade reversa é um dos principais problemas enfrentados nesse tipo de pesquisa. Pela definição da *proxy* para o *labor pooling* utilizada, este problema foi minimizado.

Como sugestões para futuras pesquisas, pretende-se expandir o estudo no sentido de considerar a heterogeneidade da qualificação da mão de obra, incorporar no modelo

variáveis que captem a "qualidade" das habilidades dos trabalhadores, além do fator "quantidade". Com este avanço espera-se trabalhar com as três forças de aglomeração *sharing*, *matching* e *learning*. Na presente pesquisa a primeira força de atração foi explorada através do *labor pooling* e dos *linkages* intersetoriais em segundo plano. Ressaltam-se, ainda, as limitações impostas pelo conjunto de dados. Além do período compreendido na amostra, que foi definido de acordo com a disponibilidade dos dados, não foi possível encontrar, por exemplo, uma variável por microrregiões para controlar o efeito das políticas de incentivos fiscais. Esse tipo de política desenvolvimentista pode atuar como uma força atrativa de investimentos exógena ao mercado.

Referências

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. 2.1, 5.1, 6

AMITI, M.; CAMERON, L. Economic Geography and Wages. *The Review of Economics and Statistics*, MIT Press, v. 89, n. 1, p. 15–29, 2007. 2.1

ANDERSSON, F.; BURGESS, S.; LANE, J. I. Cities, matching and the productivity gains of agglomeration. *Journal of Urban Economics*, Elsevier, v. 61, n. 1, p. 112–128, 2007. 2.1

ARAUJO, J. E. S. *Novas Evidências da Distribuição espacial das indústrias do Nordeste brasileiro: 2005-2010*. Dissertação (Mestrado) — Universidade Federal de Pernambuco, Programa de Pós-Graduação em Economia - PPGECON, Pernambuco, 2016. 2.2

AZZONI, C. R. Concentração regional e dispersão das rendas per capita estaduais: análise a partir de séries históricas estaduais de pib, 1939-1995. *Estudos Econômicos*, v. 27, n. 3, p. 341–393, 1997. 5

BALTAGI, B. *Econometric Analysis of Panel Data*. [S.l.]: John Wiley & Sons, 2008. 5.2

BIDERMAN, C. Concentração de atividades produtivas no Brasil: diferenças intersetoriais, regionais e sua dinâmica. 2004. 2.2

BRASIL. Ministério do trabalho e emprego. rais relação anual de informações sociais, disponível em: http://www. rais. gov. br. *Acesso em 2016*, 2016. 10

BRASIL. *Sistema IBGE de Recuperação Automática, Disponível em: https://sidra.ibge.gov.br*. [S.l.]: Brasília, 2016. 11

BREITUNG, J. The local power of some unit root tests for panel data. Humboldt-Universität zu Berlin, Wirtschaftswissenschaftliche Fakultät, 1999. 5.2

BREITUNG, J.; DAS, S. Panel unit root tests under cross-sectional dependence. *Statistica Neerlandica*, Wiley Online Library, v. 59, n. 4, p. 414–433, 2005. 5.2

BREUSCH, T. S.; PAGAN, A. R. The Lagrange Multiplier Test and its applications to model specification in econometrics. *The Review of Economic Studies*, JSTOR, v. 47, n. 1, p. 239–253, 1980. 5.2

BURKI, A. A.; KHAN, M. A. Agglomeration economies and their effects on technical inefficiency of manufacturing firms: Evidence from Pakistan. *Sciences (contact: mushtaq@ lums. edu. pk)*, 2013. 1

CAVALCANTE, L. R. Classificações tecnológicas: uma sistematização. Instituto de Pesquisa Econômica Aplicada (IPEA), 2014. 2.2, 3.2, 3.3, 7, B

CHOI, I. Unit Root Tests for Panel Data. *Journal of international money and Finance*, Elsevier, v. 20, n. 2, p. 249–272, 2001. 5.2

COMBES, P.-P.; GOBILLON, L. The empirics of agglomeration economies. *Henderson, J.V., Duranton, G., Strange, W. (Eds.), Handbook of Regional and Urban Economics, vol.5.*, North Holland, Amsterdam, 2015. 3.1

COSTA, A. B. *Mobilidade da indústria e as forças urbanas de aglomeração: uma análise do processo dinâmico da concentração industrial no Brasil*. Dissertação (Mestrado) — Fundação Getulio Vargas - FGV/Escola de Administração de Empresas de São Paulo - EAESP, São Paulo, 2014. 2.2

DEVEREUX, M. P.; GRIFFITH, R.; SIMPSON, H. The geographic distribution of production activity in the UK. *Regional Science and Urban Economics*, Elsevier, v. 34, n. 5, p. 533–564, 2004. 2.2

DINIZ, C. C.; CROCCO, M. A. et al. Reestruturação econômica e impacto regional: o novo mapa da indústria brasileira. *Nova Economia*, Economics Department, Universidade Federal de Minas Gerais (Brazil), v. 6, n. 1, p. 77–103, 1996. 5

DUMAIS, G.; ELLISON, G.; GLAESER, E. L. Geographic concentration as a dynamic process. *Review of Economics and Statistics*, MIT Press, v. 84, n. 2, p. 193–204, 2002. 2.2

DURANTON, G.; OVERMAN, H. G. Testing for localization using micro-geographic data. *The Review of Economic Studies*, Oxford University Press, v. 72, n. 4, p. 1077–1106, 2005. 1

DURANTON, G.; PUGA, D. Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics*, Elsevier, v. 4, p. 2063–2117, 2004. 1, 4, 4, A, A.1, A

ELLISON, G.; FUDENBERG, D. *Knife Edge of Plateau: When Do Market Models Tip?* [S.l.], 2003. A

ELLISON, G.; GLAESER, E. L. Geographic Concentration in US Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, v. 105, n. 5, 1997. (document), 1, 2.1, 2.2, 3, 3.1, 3.1, 3.1

ELLISON, G.; GLAESER, E. L.; KERR, W. R. What causes industry agglomeration? evidence from coagglomeration patterns. *The American Economic Review*, American Economic Association, v. 100, n. 3, p. 1195–1213, 2010. 1, 2.2

FUJITA, M.; THISSE, J.-F. *Economics of Agglomeration: cities, industrial location, and globalization*. [S.l.]: Cambridge university press, 2013. 1

GREENE, W. H. *Econometric Analysis*. [S.l.]: Pearson Education India, 2003. 5.2

HARRIS, R. D.; TZAVALIS, E. Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of econometrics*, Elsevier, v. 91, n. 2, p. 201–226, 1999. 5.2

HAUSMAN, J. A. Specification Tests in Econometrics. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 1251–1271, 1978. 5.2

IM, K. S.; PESARAN, M. H.; SHIN, Y. Testing for unit roots in heterogeneous panels. *Journal of econometrics*, Elsevier, v. 115, n. 1, p. 53–74, 2003. 5.2

KHAN, A. M. *Essays on Agglomeration Trends in the US Manufacturing Industries, 1988-2003*. Dissertação (Mestrado), 2010. 2.1, 5.1, 6

KIM, S. Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860-1987. *The Quarterly Journal of Economics*, JSTOR, p. 881–908, 1995. 5.1

KRUGMAN, P. R. *Geography and Trade*. [S.l.]: MIT press, 1991. (document), 1, 2.1, 2.2, 4, 4, 5.1, 6, 7, A, A, A.1, A

LAUTERT, V.; ARAÚJO, N. C. M. d. Concentração industrial no Brasil no período 1996-2001: uma análise por meio do índice de Ellison e Glaeser (1994). *Economia Aplicada*, SciELO Brasil, v. 11, n. 3, p. 347–368, 2007. 2.2, 3.2, 3.2, 7

LEVIN, A.; LIN, C.-F.; CHU, C.-S. J. Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, Elsevier, v. 108, n. 1, p. 1–24, 2002. 5.2

MACIENTE, A. *The determinants of agglomeration in Brazil: input-output, labor and knowledge externalities*. Tese (Doutorado) — University of Illinois at Urbana-Champaign, 2013. 2.2

MACIENTE, A. N. A Concentração do emprego no Brasil entre 1994 e 2005: aspectos regionais e setoriais. Instituto de Pesquisa Econômica Aplicada (IPEA), 2008. 2.2, 3.2, 7

MARSHALL, A. *Principles of Economics*. [S.l.]: Macmillan London, 1890. 1, 2.1, 4, 6, 7

MAUREL, F.; SÉDILLOT, B. A measure of the geographic concentration in french manufacturing industries. *Regional Science and Urban Economics*, Elsevier, v. 29, n. 5, p. 575–604, 1999. 1, 2.2

MELO, P. C.; GRAHAM, D. J. Testing for labour pooling as a source of agglomeration economies: Evidence for labour markets in england and wales. *Papers in Regional Science*, Wiley Online Library, v. 93, n. 1, p. 31–52, 2014. 2.1

MION, G.; NATICCHIONI, P. The spatial sorting and matching of skills and firms. *Canadian Journal of Economics/Revue canadienne d'économique*, Wiley Online Library, v. 42, n. 1, p. 28–55, 2009. 2.1

OVERMAN, H. G.; PUGA, D. Labor pooling as a source of agglomeration: An empirical investigation. In: *Agglomeration Economics*. [S.l.]: University of Chicago Press, 2010. p. 133–150. (document), 1, 2.1, 2.2, 4, 5.1, 5.1, 5.1, 5.1, 5.2, 6, 6, 6, 6, 7, A

PACHECO, C. A. Novos Padrões de Localização Industrial? Tendências Recentes dos Indicadores da Produção e do Investimento Industrial. Instituto de Pesquisa Econômica Aplicada (IPEA), 1999. 5, 2.2

PESARAN, M. H. General Diagnostic Tests for Cross Section Dependence in Panels. CESifo working paper series, 2004. 5.2

RESENDE, M. Industrial coagglomeration: some state-level evidence for Brazil. *Nova Economia*, SciELO Brasil, v. 25, n. 1, p. 181–194, 2015. 1, 2.2, 7

RESENDE, M.; WYLLIE, R. Ubicación Industrial y Políticas de Incentivos Locales en Brasil: una investigación empírica. *Perspectivas*, v. 2, p. 9–24, 2004. 1, 2.2

RESENDE, M.; WYLLIE, R. Aglomeração industrial no Brasil: um estudo empírico. *Estudos Econômicos (São Paulo)*, SciELO Brasil, v. 35, n. 3, p. 433–460, 2005. 2.2, 3.1, 3.2, 3.2, 7

RIGBY, D. L.; ESSLETZBICHLER, J. Agglomeration economies and productivity differences in US cities. *Journal of Economic Geography*, Oxford Univ Press, v. 2, n. 4, p. 407–432, 2002. 2.1

ROSENTHAL, S. S.; STRANGE, W. C. The determinants of agglomeration. *Journal of Urban Economics*, Elsevier, v. 50, n. 2, p. 191–229, 2001. 1, 2.1, 6

SABOIA, J. Desconcentração industrial no brasil nos anos 90: um enfoque regional. Instituto de Pesquisa Econômica Aplicada (IPEA), 2000. 5

SILVEIRA NETO, R. d. M. Concentração industrial regional, especialização geográfica e geografia econômica: evidências para o brasil no período 1950-2000. *Revista Econômica do Nordeste, Fortaleza*, v. 36, n. 2, p. 189–208, 2005. 1, 2.2, 3.2, 5.1, 5.1, 6, 7

SOBRINHO, E. M. G.; AZZONI, C. R. Aglomerações industriais relevantes do Brasil em 2010. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 9, n. 1, p. 1–18, 2016. 1

SUZIGAN, W. et al. Aglomerações industriais no estado de São Paulo. *Economia Aplicada*, v. 5, n. 4, p. 695–717, 2001. 5

VIGNANDI, R. S.; PARRÉ, J. L.; GUIMARÃES, P. de F. Measures of Industry Agglomeration in Brazil: a Study Addressing Neighboring Effects. In: . [S.l.: s.n.], 2016. v. 34, n. 65. 2.2, 3.2

WHEELER, C. H. Cities and the growth of wages among young workers: Evidence from the NLSY. *Journal of Urban Economics*, Elsevier, v. 60, n. 2, p. 162–184, 2006. 2.1

WOOLDRIDGE, J. M. *Econometric Analysis of Cross Section and Panel Data*. [S.l.]: MIT press, 2010. 5.2

Apêndices

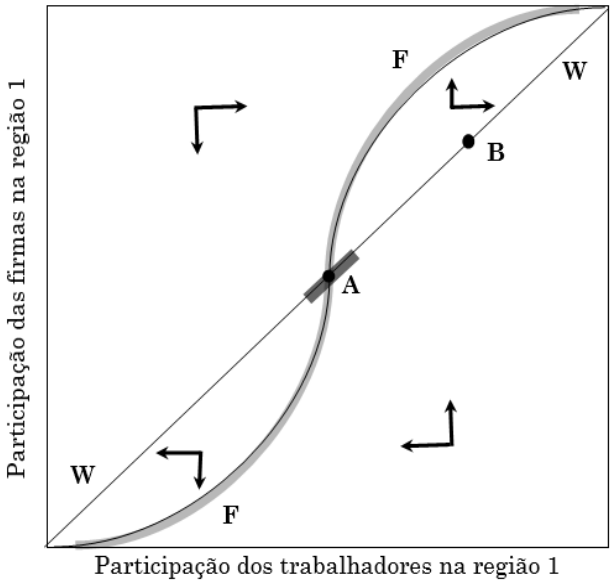
Capítulo A

Considerações adicionais sobre o Modelo Teórico

Das equações para o lucro esperado, é possível observar ainda que o efeito pooling decresce em função da intensidade dos retornos decrescentes (ψ). De acordo com Duranton e Puga (2004), isso ocorre porque a demanda por trabalho torna-se menos elástica aos choques idiossincráticos. Em termos de equilíbrio, Krugman (1991) propõem uma análise considerando dois mercados de trabalho locais (1,2), com empresas e trabalhadores neutros ao risco e que escolhem sua localização antes dos choques de produtividade. O equilíbrio interior definido pelo autor é tal que os lucros e salários esperados sejam equalizados nos dois locais: $E(\pi_1) - E(\pi_2) = 0$ e $E(w_1) - E(w_2) = 0$

A Figura A.1 apresenta o diagrama de fases equilíbrio *labor market pooling* proposto por Krugman (1991).

Figura A.1: Diagrama de fases equilíbrio *labor market pooling*



Fonte: Krugman (1991) com adaptações de Duranton e Puga (2004).

O *locus* **WW** representa a indiferença dos trabalhadores entre os dois locais ($E(w_1) - E(w_2) = 0$). Como os salários esperados dependem apenas da razão trabalhadores/empresas, o *locus* de indiferença entre os dois locais será uma reta com inclinação $\frac{L}{n}$. Para as empresas, o *locus* de indiferença é representado por **FF** ($E(\pi_1) - E(\pi_2) = 0$). O equilíbrio simétrico ($n_1 = n_2 = \frac{n}{2}$ e $L_1 = L_2 = \frac{L}{2}$) satisfaz as duas condições e está representado pelo ponto **A**. Além disto, (4.12), a esquerda do ponto **A**, **FF** é convexa e fica abaixo de **WW**, e a direita de **A**, **FF** é côncava e situa-se acima de **WW**. A intuição do modelo pode ser melhor entendida considerando o ponto adicional **B** no *locus* **WW**. Em **B**, a proporção de trabalhadores/empresas é a mesma em ambas as localizações, no entanto, na localização 1 (ponto **A**) existem mais de ambos. Assim, pela equação (4.12), os lucros são superiores na localização 1 (DURANTON; PUGA, 2004).

Como representado pelas setas no diagrama, o equilíbrio simétrico é *saddle-path* (trajetória instável) com respeito a pequenas perturbações simultâneas na quantidade de trabalhadores e empresas. Em função disto, como os agentes sabem disso ex-ante, empresas e trabalhadores, tendem necessariamente a aglomeração em um único mercado de trabalho local (DURANTON; PUGA, 2004).

O diagrama de fases na Figura A.1 é uma adaptação apresentada em Duranton e Puga (2004) que incorpora a ressalva feita por Ellison e Fudenberg (2003) com relação ao conceito de equilíbrio em Krugman (1991). De acordo com Ellison e Fudenberg (2003), a condição de equilíbrio para um número discreto de empresas é a condição de que as essas empresas não aumentarão seus lucros esperados ao mudar de localidade, e não a condição de que os lucros esperados sejam iguais entre os locais. Assim, esta condição está representada no diagrama pela área sombreada em torno de **FF**, que como destacado em Duranton e Puga (2004), implica que se a empresa desviar do equilíbrio simétrico, os salários esperados aumentam, e todas as empresas, incluindo a própria, terrão de pagar por isso, por este motivo, não poderão obter lucros mais elevados ao desviar do equilíbrio.

Ellison e Fudenberg (2003) também demostraram que o modelo de Krugman faz parte de uma classe de modelos na qual existem patamares de equilíbrio. Assim, qualquer alocação pertencente a área representada em (A.1) é um equilíbrio. No diagrama, este patamar de equilíbrio está representado pelo seguimento diagonal centrado pelo equilíbrio simétrico.

$$\frac{n_1}{n} = \frac{L_1}{L} \in \left\{ \frac{\sigma^2}{2\psi^2\left(\frac{L}{n}\right) + \sigma^2}, 1 - \frac{\sigma^2}{2\psi^2\left(\frac{L}{n}\right) + \sigma^2} \right\} \tag{A.1}$$

Overman e Puga (2010) propõem ainda uma discussão baseada em duas definições

alternativas de equilíbrio urbano para o modelo de Krugman em uma versão multisetorial e multilocacional. Um equilíbrio com realocação simultânea e outro com livre entrada e prêmio salarial, nos quais a intuição chave do modelo é preservada.

No primeiro caso, a análise é desenvolvida considerando dois estágios. Na etapa inicial, todos os estabelecimentos e trabalhadores escolhem simultaneamente sua localização. O número de estabelecimentos em cada local é exógeno. Na etapa dois, os estabelecimentos recebem os choques de produtividade. Como existe um grande número de trabalhadores, mudanças na localização de um trabalhador individual não afeta os salários e nem os lucros, os salários são, portanto, equalizados em todos os locais e nenhum trabalhador tem incentivos para mudar. Pela a equação (4.6), no equilíbrio, a proporção de trabalhadores por estabelecimentos, $(\frac{L}{n})$, tem de ser a mesma em todos os locais. Diferentemente dos trabalhadores, o número estabelecimentos é discreto, de modo que uma mudança na sua localização altera salários e lucros. Os autores enfatizam que esta é uma suposição essencial do modelo, pois é ela que suporta a ideia do efeito *pooling*.

Os estabelecimentos mudam buscando aumentar seus lucros, como os salários são equalizados em todos locais, essa mudança fará com que $\frac{L}{n}$ se reduza no local de destino, se não existir uma oferta de trabalho de trabalho suficientemente grande neste local, o salário esperado aumenta e os lucros esperados decrescem, e isto é captado pelo primeiro termo do lado direito da equação (4.12). Alternativamente, se após a mudança, a oferta de trabalho é suficiente para que não exista mudanças significativas no salário esperado, o estabelecimento poderá se adaptar melhor aos choques de produtividade e obter lucros esperados maiores, esse é o efeito pooling do trabalho, captado pelo segundo termo do lado direito da equação (4.12). Considere que os setores diferem apenas em termos de variância nos choques, σ^2, então uma oferta de trabalho local reduzida atua como um fator de repulsão igualmente forte em todos os setores, favorecendo, deste modo, a dispersão. Por outro lado, no caso da concentração da força de trabalho, o efeito *labor pooling* é tanto maior quanto maior for σ^2.

No segundo caso, o número de estabelecimentos em cada local é endógeno tendo o tamanho do mercado de trabalho local como dado. O equilíbrio do número de estabelecimentos em cada local é determinado pela entrada livre de estabelecimentos até que a entrada da n-ésima firma faça com que as firmas nesta localidade não consigam cobrir seus custos fixos de entrar no mercado. Denotado estes custos por ϕ, usando $E(\pi) = \phi$ na equação (4.12) e resolvendo para n, chega-se ao número de estabelecimentos de equilíbrio em cada cidade (ignorando restrições de integridade):

$$n = \frac{\sqrt{(\sigma^2)^2 + 4(2\psi\phi - \sigma^2)\psi^2 L^2} - \sigma^2}{2(2\psi\phi - \sigma^2)} \tag{A.2}$$

Desta equação, resulta que nos locais onde a oferta de trabalho é maior, existirá um número maior de estabelecimentos e consequentemente menor será $\frac{L}{n}$:

$$\frac{\partial\left(\frac{L}{n}\right)}{\partial L} = \frac{-2\sigma^2(2\psi\phi - \sigma^2)}{\sqrt{(\sigma^2)^2 + 4(2\psi\phi - \sigma^2)\psi^2 L^2}\left(\sqrt{(\sigma^2)^2 + 4(2\psi\phi - \sigma^2)\psi^2 L^2} - \sigma^2\right)} < 0 \tag{A.3}$$

Ou seja:

$$\frac{-\sigma^2}{n\sqrt{(\sigma^2)^2 + 4(2\psi\phi - \sigma^2)\psi^2 L^2}} < 0 \tag{A.4}$$

Pela equação (4.6), isto cria um prêmio salarial que compensa o efeito *labor pooling* garantindo que os lucros sejam exauridos em todos os lugares:

$$\frac{\partial E(w)}{\partial L} = -\psi\frac{\partial\left(\frac{L}{n}\right)}{\partial L} = -\psi\frac{-\sigma^2}{N\sqrt{(\sigma^2)^2 + 4(2\psi\phi - \sigma^2)\psi^2 L^2}} > 0 \tag{A.5}$$

O equilíbrio com livre entrada e prêmio salarial é, portanto, bastante diferente do equilíbrio com realocação simultânea. Mas, a intuição chave do modelo é mantida. Para observar isso, toma-se a segunda derivada de $\frac{L}{n}$ com relação a σ^2 para obter:

$$\frac{\partial^2\left(\frac{L}{n}\right)}{\partial L \partial \sigma_s} = -\frac{1}{2\psi^2 L^2}\left\{1 + \frac{\sigma^2[(\sigma^2)^2 + 2\psi^2 L^2(8\psi\phi - 2\sigma^2)]}{[(\sigma^2)^2 + 4(2\psi\phi - \sigma^2)\psi^2 L^2]^{\frac{3}{2}}}\right\} < 0 \tag{A.6}$$

Isto implica que cidades com mercados de trabalho maiores atraem um número desproporcional de estabelecimentos e esse efeito é tanto maior quanto maior for σ^2 no setor. Assim, chega-se a mesma conclusão de antes, os setores que possuem choques de produtividade mais heterogêneos tendem a aglomerar-se.

Capítulo B
Tabelas Adicionais

Tabela B.1: Intensidade tecnológica das divisões (CNAE dois dígitos) que compõem a indústria de transformação (CNAE 1.0)

Divisão	Descrição	Class. OCDE
15	Fabricação de produtos alimentícios e bebidas	Baixa
16	Fabricação de produtos do fumo	Baixa
17	Fabricação de produtos têxteis	Baixa
18	Confecção de artigos do vestuário e acessórios	Baixa
19	Preparação de couros e fabricação de artefatos de couro, artigos de viagem e cal.	Baixa
20	Fabricação de produtos de madeira	Baixa
21	Fabricação de celulose, papel e produtos de papel	Baixa
22	Edição, impressão e reprodução de gravações	Baixa
23	Fabricação de coque, refino de petróleo, elaboração de combustíveis nucleares e produção de álcool	Média-baixa
24	Fabricação de produtos químicos	Média-alta
25	Fabricação de artigos de borracha e de material plástico	Média-baixa
26	Fabricação de produtos de minerais não metálicos	Média-baixa
27	Metalurgia básica	Média-baixa
28	Fabricação de produtos de metal - exclusive maquinas e equipamentos	Média-baixa
29	Fabricação de maquinas e equipamentos	Média-alta
30	Fabricação de maquinas para escritório e equipamentos de informática	Alta
31	Fabricação de maquinas, aparelhos e materiais elétricos	Média-alta
32	Fabricação de material eletrônico e de aparelhos e equipamentos de comunicações	Alta
33	Fabricação de equipamentos de instrumentação médico-hospitalares, instrumentos de precisão e ópticos, equipamentos para automação industrial, cronômetros e relo.	Alta
34	Fabricação e montagem de veículos automotores, reboques e carrocerias	Média-alta
35	Fabricação de outros equipamentos de transporte	Média-alta
36	Fabricação de moveis e industrias diversas	Baixa
37	Reciclagem	Baixa

Fonte: Cavalcante (2014).

Tabela B.2: Índice de concentração bruta por setores da indústria da transformação no período 2002-2014

							G_{st}						
Div. CNAE	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
15	0,01072	0,01003	0,00991	0,00960	0,00972	0,00954	0,00938	0,00877	0,00875	0,00830	0,00744	0,00718	0,00654
16	0,17706	0,15833	0,13507	0,11720	0,10791	0,12978	0,11857	0,11509	0,11455	0,11252	0,12570	0,13944	0,13290
17	0,01849	0,01829	0,01850	0,01963	0,01982	0,02072	0,02166	0,02236	0,02189	0,02066	0,01929	0,02023	0,01989
18	0,00924	0,00913	0,00928	0,00910	0,00929	0,00978	0,01007	0,01020	0,00987	0,00966	0,01005	0,00997	0,00964
19	0,06419	0,05700	0,05491	0,05156	0,04912	0,04786	0,04754	0,04837	0,04671	0,04493	0,04480	0,04361	0,04254
20	0,02610	0,02524	0,02454	0,02292	0,02204	0,02005	0,01832	0,01801	0,01728	0,01639	0,01593	0,01566	0,01597
21	0,00631	0,00672	0,00554	0,00537	0,00474	0,00456	0,00454	0,00495	0,00510	0,00512	0,00511	0,00544	0,00578
22	0,02600	0,02378	0,02219	0,01965	0,02043	0,02064	0,01977	0,02022	0,02174	0,02030	0,02060	0,02055	0,02253
23	0,04110	0,03943	0,03662	0,03690	0,02978	0,02876	0,02639	0,02653	0,15211	0,16902	0,18499	0,18966	0,19507
24	0,01546	0,01484	0,01437	0,01343	0,01583	0,01630	0,01569	0,01567	0,00659	0,00611	0,00585	0,00593	0,00616
25	0,01059	0,01011	0,00961	0,00908	0,00822	0,00859	0,00779	0,00710	0,00650	0,00640	0,00611	0,00651	0,00588
26	0,00826	0,00733	0,00746	0,00793	0,00704	0,00635	0,00679	0,00592	0,00595	0,00570	0,00562	0,00605	0,00586
27	0,01362	0,01384	0,01285	0,01329	0,01310	0,01213	0,01237	0,01286	0,01295	0,01248	0,01269	0,01248	0,01296
28	0,00592	0,00612	0,00651	0,00612	0,00619	0,00561	0,00475	0,00502	0,00449	0,00386	0,00355	0,00300	0,00294
29	0,00950	0,00808	0,00818	0,00834	0,00723	0,00642	0,00575	0,00557	0,00507	0,00488	0,00492	0,00461	0,00438
30	0,03400	0,03697	0,04520	0,04180	0,04650	0,05963	0,04897	0,04591	0,04863	0,04671	0,04766	0,05646	0,05386
31	0,01313	0,01297	0,01223	0,01127	0,01165	0,01384	0,00949	0,00977	0,00908	0,00858	0,00821	0,00705	0,00701
32	0,09363	0,10191	0,12353	0,14447	0,12442	0,10376	0,08117	0,07895	0,09967	0,12427	0,12512	0,14335	0,12426
33	0,02288	0,02489	0,02260	0,02167	0,01807	0,02014	0,02108	0,01787	0,01597	0,01607	0,01586	0,01352	0,01290
34	0,04246	0,03844	0,03654	0,03427	0,03037	0,02775	0,02647	0,02782	0,02590	0,02474	0,02252	0,02360	0,02398
35	0,08910	0,08386	0,08648	0,08938	0,08938	0,10021	0,10182	0,08537	0,07525	0,07129	0,07038	0,06778	0,06863
36	0,00734	0,00668	0,00800	0,00705	0,00831	0,00845	0,00771	0,00751	0,00711	0,00691	0,00689	0,00679	0,00674
37	0,01296	0,01002	0,00815	0,00734	0,00534	0,00678	0,00765	0,00853	0,01255	0,00892	0,01207	0,01167	0,01103

Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

Tabela B.3: Índice Herfindahl-Hirshman por setores da indústria da transformação no período 2002-2014

							H_{st}						
Div. CNAE	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
15	0,00087	0,00089	0,00091	0,00091	0,00088	0,00086	0,00085	0,00079	0,00081	0,00080	0,00070	0,00070	0,00067
16	0,02997	0,02643	0,02513	0,02505	0,02553	0,02732	0,02971	0,03059	0,02816	0,03171	0,03181	0,03874	0,03813
17	0,00182	0,00176	0,00174	0,00173	0,00158	0,00160	0,00155	0,00158	0,00155	0,00152	0,00149	0,00154	0,00154
18	0,00056	0,00043	0,00042	0,00043	0,00044	0,00067	0,00092	0,00087	0,00093	0,00067	0,00055	0,00052	0,00053
19	0,00313	0,00352	0,00339	0,00342	0,00354	0,00406	0,00408	0,00644	0,00532	0,00460	0,00443	0,00501	0,00478
20	0,00062	0,00067	0,00064	0,00066	0,00065	0,00060	0,00061	0,00061	0,00065	0,00062	0,00064	0,00069	0,00071
21	0,00262	0,00254	0,00233	0,00228	0,00216	0,00204	0,00203	0,00206	0,00206	0,00212	0,00216	0,00243	0,00267
22	0,00130	0,00126	0,00122	0,00114	0,00114	0,00115	0,00107	0,00106	0,00109	0,00107	0,00105	0,00107	0,00106
23	0,02056	0,02256	0,01971	0,01908	0,01292	0,01260	0,01042	0,00954	0,04811	0,04105	0,04180	0,03865	0,03979
24	0,00102	0,00098	0,00094	0,00095	0,00127	0,00109	0,00108	0,00119	0,00134	0,00137	0,00138	0,00137	0,00140
25	0,00091	0,00090	0,00085	0,00083	0,00076	0,00076	0,00078	0,00074	0,00076	0,00076	0,00075	0,00079	0,00081
26	0,00051	0,00055	0,00057	0,00052	0,00076	0,00047	0,00046	0,00041	0,00041	0,00040	0,00038	0,00038	0,00040
27	0,00777	0,00770	0,00699	0,00725	0,00708	0,00652	0,00682	0,00676	0,00672	0,00730	0,00776	0,00773	0,00833
28	0,00063	0,00060	0,00065	0,00063	0,00051	0,00058	0,00054	0,00076	0,00077	0,00051	0,00042	0,00037	0,00037
29	0,00171	0,00183	0,00188	0,00154	0,00136	0,00135	0,00120	0,00127	0,00126	0,00117	0,00111	0,00109	0,00109
30	0,01230	0,01228	0,01953	0,02489	0,02759	0,02362	0,03449	0,03161	0,03038	0,02786	0,02671	0,02524	0,03504
31	0,00397	0,00410	0,00416	0,00369	0,00337	0,00358	0,00350	0,00332	0,00322	0,00338	0,00358	0,00309	0,00301
32	0,00980	0,00962	0,00936	0,01008	0,00886	0,00894	0,00835	0,00752	0,00849	0,01087	0,01211	0,01525	0,01482
33	0,00471	0,00448	0,00436	0,00425	0,00436	0,00475	0,00431	0,00405	0,00359	0,00354	0,00281	0,00247	0,00247
34	0,00823	0,00715	0,00672	0,00655	0,00584	0,00563	0,00544	0,00519	0,00481	0,00496	0,00485	0,00488	0,00491
35	0,05918	0,05183	0,04709	0,04249	0,04078	0,04480	0,03850	0,02952	0,02724	0,02598	0,02231	0,02209	0,02333
36	0,00053	0,00049	0,00053	0,00052	0,00051	0,00055	0,00051	0,00051	0,00050	0,00048	0,00048	0,00047	0,00044
37	0,00432	0,00335	0,00278	0,00251	0,00243	0,00232	0,00319	0,00406	0,00712	0,00402	0,00790	0,00820	0,00752

Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

Tabela B.4: Índice de Ellison e Glaeser por setores da indústria da transformação no período 2002-2014

Div. CNAE	2002	2003	2004	2005	2006	2007	γ_{st} 2008	2009	2010	2011	2012	2013	2014
15	0,00986	0,00915	0,00901	0,00869	0,00885	0,00868	0,00854	0,00799	0,00795	0,00751	0,00675	0,00649	0,00588
16	0,15163	0,13547	0,11278	0,09452	0,08454	0,10534	0,09158	0,08717	0,08889	0,08346	0,09698	0,10476	0,09853
17	0,0167	0,01655	0,01679	0,01792	0,01827	0,01915	0,02014	0,02081	0,02037	0,01916	0,01784	0,01872	0,01838
18	0,00869	0,00871	0,00886	0,00867	0,00885	0,00912	0,00916	0,00934	0,00894	0,009	0,0095	0,00945	0,00911
19	0,06125	0,05366	0,05169	0,04831	0,04574	0,04398	0,04364	0,0422	0,04161	0,04051	0,04055	0,03879	0,03795
20	0,0255	0,02459	0,02391	0,02228	0,02141	0,01946	0,01772	0,01741	0,01664	0,01578	0,0153	0,01498	0,01528
21	0,0037	0,00418	0,00322	0,00309	0,00258	0,00253	0,00251	0,00289	0,00304	0,00301	0,00296	0,00301	0,00312
22	0,02473	0,02255	0,021	0,01854	0,01932	0,01952	0,01872	0,01917	0,02067	0,01925	0,01957	0,01949	0,02149
23	0,02097	0,01725	0,01725	0,01817	0,01708	0,01637	0,01615	0,01716	0,10925	0,13345	0,14944	0,15708	0,16171
24	0,01446	0,01387	0,01344	0,0125	0,01459	0,01522	0,01462	0,0145	0,00526	0,00474	0,00448	0,00457	0,00477
25	0,00969	0,00922	0,00877	0,00826	0,00746	0,00784	0,00701	0,00637	0,00574	0,00565	0,00536	0,00572	0,00507
26	0,00775	0,00678	0,00689	0,00742	0,00628	0,00588	0,00634	0,00551	0,00554	0,00531	0,00524	0,00567	0,00546
27	0,00589	0,00618	0,0059	0,00609	0,00606	0,00565	0,00559	0,00614	0,00628	0,00522	0,00496	0,00479	0,00467
28	0,00529	0,00553	0,00586	0,00549	0,00568	0,00503	0,00421	0,00427	0,00372	0,00336	0,00314	0,00264	0,00257
29	0,00781	0,00626	0,00631	0,00681	0,00588	0,00507	0,00455	0,00431	0,00382	0,00371	0,00382	0,00353	0,00329
30	0,02197	0,025	0,02618	0,01735	0,01944	0,03688	0,015	0,01476	0,01882	0,01939	0,02153	0,03203	0,01951
31	0,00919	0,0089	0,0081	0,00761	0,00831	0,0103	0,00601	0,00647	0,00588	0,00522	0,00464	0,00397	0,00402
32	0,08466	0,09318	0,11525	0,13576	0,1166	0,09568	0,07343	0,07197	0,09196	0,11465	0,1144	0,13008	0,11109
33	0,01825	0,0205	0,01832	0,01749	0,01378	0,01547	0,01684	0,01388	0,01243	0,01257	0,01308	0,01108	0,01046
34	0,03452	0,03152	0,03001	0,0279	0,02467	0,02224	0,02114	0,02276	0,02119	0,01988	0,01776	0,01881	0,01916
35	0,03179	0,03379	0,04134	0,04898	0,05067	0,05801	0,06586	0,05755	0,04936	0,04652	0,04916	0,04671	0,04638
36	0,00682	0,0062	0,00747	0,00653	0,0078	0,0079	0,00721	0,00701	0,00661	0,00643	0,00641	0,00633	0,00629
37	0,00867	0,0067	0,00539	0,00485	0,00291	0,00446	0,00448	0,00449	0,00546	0,00492	0,00421	0,0035	0,00353

Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

Tabela B.5: Resultados das Estimações *Fixed Effect* (FE)

γ_{st}	Coef	Std. Err.	t	$P > t$	[95% Conf.	Interval]
P_{st}	0,0215	0,0057	3,7600	0,0000	0,0102	0,0328
E_{st}	-0,0022	0,0001	-14,6400	0,0000	-0,0025	-0,0019
L_{st}	0,0099	0,0052	1,8900	0,0590	-0,0004	0,0202
gp_{st}	0,0000	0,0000	-0,5300	0,5990	0,0000	0,0000
cmp_{st}	0,0000	0,0000	2,9100	0,0040	0,0000	0,0000
cec_{st}	0,0000	0,0000	-1,0000	0,3200	0,0000	0,0000
Const	0,0439	0,0073	6,0300	0,0000	0,0295	0,0582

Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

Tabela B.6: Resultados Estimações *Random Effects* GLS (RE)

γ_{st}	Coef	Std. Err.	z	$P > z$	[95% Conf.	Interval]
P_{st}	0,0282	0,0068	4,1300	0,0000	0,0148	0,0416
E_{st}	-0,0013	0,0002	-8,8800	0,0000	-0,0016	-0,0010
L_{st}	0,0086	0,0056	1,5300	0,1250	-0,0024	0,0197
gp_{st}	0,0000	0,0000	0,5200	0,6030	0,0000	0,0000
cmp_{st}	0,0000	0,0000	0,7400	0,4600	0,0000	0,0000
cec_{st}	0,0000	0,0000	0,4100	0,6800	0,0000	0,0000
Const	0,0279	0,0096	2,8900	0,0040	0,0090	0,0468

Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

Tabela B.7: Teste de Hausman *Fixed Effect vs Random Effects*

	Coeficientes			
	(b)	(B)	(b-B)	sqrt(diag($V_b - V_B$))
	Fixed Effect	*Random Effects*	Difference	S.E.
P_{st}	0,0215	0,0282	-0,0067	0,0011
E_{st}	-0,0022	-0,0013	-0,0009	0,0001
L_{st}	0,0099	0,0086	0,0012	0,0028
gp_{st}	0,0000	0,0000	0,0000	0,0000
cmp_{st}	0,0000	0,0000	0,0000	0,0000
cec_{st}	0,0000	0,0000	0,0000	0,0000

Test: H_0: difference in coefficients not systematic

$\chi^2(3) = (b-B)'[(V_b - V_B)^{-1}](b-B) = 73{,}11$

$Prob > \chi^2 = 0{,}0000$ ($V_b - V_B$ is not positive definite)

Fonte: Elaboração própria com base nos dados da RAIS para o período 2002-2014.

Este volume foi tipografado em LaTeX na classe UFPEThesis (<www.cin.ufpe.br/~paguso/ufpethesis>). Para detalhes sobre este documento, clique aqui.