

Por que as Indústrias se Coaglomeram? Evidências para o Brasil

Economia Regional e Urbana

Jullio Victor Pedrosa de Almeida

Mestrando em Economia pela Universidade Federal de Pernambuco (UFPE) – Campus Agreste. Rodovia BR-104, Km 59, s/n - Nova Caruaru – 55002-970 - Caruaru – Pernambuco - Brasil. Email: jullio.almeida@ufpe.br.

Roberta de Moraes Rocha

Doutora em Economia pelo Pimes/UFPE. Professora Associada do PPGECON/UFPE/Campus Agreste. Rodovia BR-104, Km 59, s/n - Nova Caruaru – 55002-970 - Caruaru – Pernambuco - Brasil. Email: roberta_rocha_pe@yahoo.com.br

Wellington Ribeiro Justo

Doutor em Economia pelo Pimes/UFPE. Professor associado da Universidade Regional do Cariri (URCA) e Professor do PPGECON/UFPE/Campus Agreste. Rua Nelson Alencar, 490, Centro – 63100-110 – Crato – Ceará – Brasil. Email: justowr@yahoo.com.br

Por que as Indústrias se Coaglomeram? Evidências para o Brasil

Economia Regional e Urbana

Resumo

O artigo buscou investigar os padrões de coaglomeração das indústrias brasileiras do setor de transformação e seus possíveis fatores de aglomeração a nível estadual, microrregional e municipal para setores divisão CNAE a 2 dígitos. Para tanto, utilizou-se do índice de aglomeração explorado por Ellison e Glaeser (1997) e refinado para pares de setores por Ellison, Glaeser e Kerr (2010). Através dos cálculos dos índices, testou-se inicialmente as hipóteses (1) de que há no Brasil coaglomeração de pares de setores e (2) que esse padrão de coaglomeração é regido pelas forças de aglomeração marshallianas. Os resultados apontam que diversos setores apresentam padrão de coaglomeração, e que as forças marshallianas tem papel fundamental nesse padrão. Os resultados das regressões múltiplas utilizando dados em painel mostram que as variáveis *labor pooling* e *input linkage* têm papel de destaque como fatores de aglomeração.

Palavras chave: Coaglomeração Industrial; Indústria da Transformação; Economias de Aglomeração; Forças Marshallianas; Índice de coaglomeração EG.

Abstract

The article sought to investigate the coagglomeration patterns of the Brazilian industries in the transformation sector and their available agglomeration data at the state, microrregional and municipal level for the 2 digit CNAE division sectors. Thus, the agglomeration index explored by Ellison and Glaeser (1997) and refined for sectors pairs by Ellison, Glaeser and Kerr (2010) was used. Through the calculations of the indices, was initially tested as hypotheses (1) that Brazil co-agglomerates pairs of sectors and (2) that this coagglomeration patterns is governed by marshall market forces. The results point that sets of sectors have a coagglomeration pattern, and that Marshallian forces play a fundamental role in this pattern. The results of the multiple regressions in cross-section data show that the variables *labor pooling* and *input linkage* play a prominent role as the agglomeration factors.

Keywords: Industry Coagglomeration; Manufacturing Industry; Economies of Agglomeration; Marshallian Forces; EG Coagglomeration Index.

Classificação JEL: R32, R23, L60.

Introdução.

Aglomerações industriais são estudadas a um longo tempo e há evidências que as firmas são geograficamente concentradas.¹ Através do aprimoramento de medidas de mensuração da aglomeração do emprego, foi possível investigar e testar empiricamente quais as forças que levam a ocorrência de clusters industriais. As principais ideias são apresentadas por Marshall (1920) que em seu trabalho aponta a existência de três forças de aglomeração, chamadas também de externalidades positivas ou *spillovers*, argumentando que as indústrias escolhem sua localização conforme as vantagens em diminuir os custos de transportes – de insumo-produto, trabalhadores e ideias. Marshall então sintetiza suas ideias acerca desses custos, como forças centrípetas que incentivam a aglomeração de indústrias numa mesma região, sendo conhecidas na literatura como *input-output linkages*, *labor pooling* e *intellectual spillovers*.

Para reforçar a importância das aglomerações na configuração econômica, diversos autores² investigaram a relação entre aglomeração e produtividade. As evidências empíricas encontradas mostram que setores mais concentrados tendem a ser mais produtivos e inovadores que os demais. Steingraber e Gonçalves (2015) analisam esta relação para o Brasil, identificando que as diferenças de produtividades são explicadas pela aglomeração e concentração de indústrias juntamente com as competências internas das empresas. No mesmo trabalho os autores evidenciam que alguns setores são mais sensíveis à aglomeração e a concentração setoriais sobre a produtividade.

A ideia de que vantagens naturais desempenham um papel fundamental para a configuração dos clusters industriais também é estudada. Setores que dependem largamente de recursos naturais tendem a se localizar em regiões próximas as fontes de recursos, levando a aglomeração. Esta aglomeração, porém, tende a ser mais forte na agricultura e indústria de extrativismo, não sendo necessariamente, as vantagens naturais o maior fator determinante nas indústrias manufatureiras. Ellison e Glaeser (1999) encontram evidências apontando que as vantagens naturais podem explicar cerca de 20% das concentrações industriais no setor de manufatura dos Estados Unidos da América (EUA).

Este artigo, porém, não investiga o papel das vantagens naturais nas aglomerações, dada a falta de informações referentes as vantagens naturais no Brasil, diferentemente dos Estados Unidos, onde Ellison e Glaeser (1999) construíram um índice elaborado baseado em diversas variáveis e informações. Entretanto, esta exclusão não significa que a análise será afetada. Ellison, Glaeser e Kerr (2010) mostram que a remoção da variável de vantagens naturais não afeta significativamente o modelo, sendo possível a investigação apenas das forças marshallianas como fatores determinantes a aglomeração. Logo, neste artigo o foco é a análise das forças de aglomeração propostas por Marshall.

Para investigação destes fatores, este trabalho se baseia no artigo de Ellison, Glaeser e Kerr – EGK (2010) onde a identificação destas vantagens foi melhor adequada utilizando-se a análise de coaglomeração de pares de setores na indústria manufatureira dos EUA. EGK propuseram um índice baseado no artigo de Ellison e Glaeser (1997) em que foi possível reduzir a análise

¹ Ver Krugman (1991a); Ellison e Glaeser (1997, 1999); Duranton e Overman (2005, 2008); Maurel e Sédillot (2009) para referências internacionais, Resende e Wyllie (2005); e Lautert e Araújo (2007); Silva e Silveira Neto (2007); e Rocha et. Al. (2013) para o caso brasileiro.

² Ver Henderson (2013); Rosenthal e Strange (2004); Barrios, Bertinelli e Strobl (2006); Andersson e Loof (2011); Puga et.al (2012);

das aglomerações para pares de setores diferentes, facilitando assim a identificação das externalidades marshallianas e sua influência na aglomeração entre setores.

O estudo direcionado ao conjunto de aglomerações, também chamado de coaglomerações, se intensificou principalmente nos anos recentes buscando entender por que diferentes setores tendem a se aglomerar. Avançar na compreensão dessas coaglomerações na indústria de transformação do Brasil é de relevante importância para a geografia econômica brasileira, pois desta forma é possível identificar os fatores de aglomeração de maior influência, e também quais setores tendem a se coaglomerar com maior intensidade. No Brasil esse tema tem sido pouco estudado e ainda há uma grande lacuna na literatura. Resende (2012) faz uma pesquisa exploratória para o caso do estado do Rio de Janeiro, obtendo resultados interessantes apontando que o *labor pooling* é importante para a coaglomeração entre pares de setores. Este trabalho visa contribuir para diminuir essa lacuna, investigando os padrões de coaglomeração para todo Brasil através do cálculo dos índices de EGK para dois setores, seguindo a linha de raciocínio de Ellison, Glaeser e Kerr (2010).

Deste modo, este trabalho procura usar o índice de EGK para testar duas hipóteses centrais. A primeira hipótese assumida no trabalho, é que existe padrões de coaglomeração nas indústrias de transformação no Brasil. A segunda hipótese, é que caso exista esses padrões, eles são regidos por forças de aglomeração marshallianas. Desta maneira, espera-se que este trabalho contribua de certa forma para atenuar a lacuna existente na literatura brasileira no que diz respeito ao estudo das coaglomerações industriais e seus determinantes, proporcionando um arcabouço empírico suficiente para gerar novas pesquisas futuras na área.

Tendo como objetivos principais o aprofundamento na análise dos padrões de coaglomerações dos pares de setores no Brasil a nível Regional, Estadual e Municipal, e a investigação de quais fatores de aglomeração estão presentes, assim o trabalho avança na literatura no sentido de investigar preliminarmente se há aglomeração entre pares de setores industriais no Brasil, e posteriormente, discutir quais as forças de mercado que atuam em maior presença nessas aglomerações. Para tanto, o trabalho utiliza o índice de EG adaptado para pares de setores utilizado no artigo de Ellison, Glaeser e Kerr (2010), fazendo uso de análise exploratória dos dados, para o primeiro passo, e a estimação de um modelo de regressão múltipla com dados em painel para o segundo passo.

Assim, divide-se este artigo em seis seções incluindo esta introdução. A próxima seção faz um levantamento das evidências empíricas encontradas sobre o tema. A seção 3 apresenta a base de dados adotada para investigação e decomposição dos fatores de aglomeração. A quarta e a quinta seção testam as hipóteses apresentadas. Por fim, na seção 6 apresentam-se os resultados e conclusões.

1. Evidências Empíricas

O grande passo dado para disseminação dos estudos empíricos ocorreu após os trabalhos de Krugman (1991) e Ellison e Glaeser (1997). O primeiro formalizou matematicamente a teoria das aglomerações urbanas, baseado na teoria internacional das vantagens comparativas. O segundo trabalho apresentou para a literatura um índice conciso com a teoria e formalização matemática, sendo prático para mensurar as aglomerações e coaglomerações de emprego. O índice de Ellison e Glaeser (1997), chamado de EG, trouxe avanços significativos e impulsionou os estudos, sendo uma ferramenta poderosa, principalmente no que tange a identificação dos fatores de aglomeração. Vale ressaltar os índices propostos por Overman e Duranton (2005) e Maurel e Sédillot (2009), que contribuíram para novas perspectivas de pesquisa.

Barrios, Bertinelli e Strobl (2002) utilizaram o índice de coaglomeração desenvolvido por Ellison e Glaeser (1997) com o intuito de testar a hipótese da existência de possíveis vantagens na proximidade da localização entre fábricas domésticas e multinacionais estrangeiras. A segunda hipótese a ser testada diz respeito a um provável impacto da presença estrangeira no crescimento regional, as duas hipóteses foram testadas para a Irlanda no período de 1983 a 1998. Por fim, chegaram à conclusão de que para um elevado número de setores, existe persistência na coaglomeração entre plantas domésticas e estrangeiras. Além disso, confirmase a presença de spillovers locais das multinacionais impactando positivamente no crescimento do emprego indígena, no entanto, estes spillovers somente acontecem em indústrias que apresentaram uma coaglomeração expressiva de plantas domésticas e estrangeiras.

Dohse e Steude (2003) mediram a distribuição das atividades de Neur Markt³ – segmento da bolsa alemã -, através do índice de coaglomeração de Ellison e Glaeser (1997), onde analisaram a concentração espacial do emprego na “nova economia”⁴, a coaglomeração de firmas de sub-setores variados de Neur Markt e os spillovers interregionais. Sendo assim, a Nova Economia – em referência ao Neur Markt -, o emprego e a capitalização de mercado estão sobre-representados no sul e no oeste da Alemanha, enquanto no leste está sub-representada. Nas regiões ricas e com elevadas produtividades de trabalho e densidade de atividades econômicas, tendem a se agrupar as empresas Neur Markt, bem como as regiões tecnologicamente fortes. As indústrias de Mídia e Entretenimento, Software, Indústria e Serviços Industriais e Biotecnologia apresentam um elevado grau de coaglomeração e são setores valiosos para o desenvolvimento regional, uma vez que geram ou recebem spillovers interindustriais em grande escala.

O índice de EG também foi empregado em Barrios, Bertinelli e Strobl (2006) aliado ao cálculo das estimativas de produtividade totais de fatores de plantas usando o procedimento de Levinsohn e Petrin (2003). Os resultados desse segundo trabalho corroboram os resultados do trabalho de 2002, onde a coaglomeração entre plantas domésticas e estrangeiras são persistentes na Irlanda e também a existência de spillovers locais das multinacionais aumentando a produtividade e o emprego de indígenas. Além disso, foi calculado o aumento no desvio padrão diante da presença externa, uma variação entre 3 e 6% na produtividade e 1,7 e 4,3% no emprego, essa variação dependeu da inclusão de Dublin e da medida da presença estrangeira escolhida. Esse resultado indica que uma parcela da coaglomeração entre plantas domésticas e estrangeiras é associada com a presença de spillovers positivos.

Ellison, Glaeser e Kerr (2010) empregaram o índice de EG para quantificar a coaglomeração de pares de indústrias para os Estados Unidos, além desse índice, para o mesmo fim, também foi utilizado uma aproximação “lumpy” ao índice contínuo criado por Duranton e Overman (2005). Os autores utilizam o índice simplificado para analisar separadamente pares de setores, e assim estudar as relações intersetoriais existentes, comparando as economias de aglomeração e vantagens naturais existentes nesses padrões de aglomeração. Os resultados do artigo apresentam um suporte às teorias marshallianas de aglomeração, uma vez que foram encontradas evidências consistentes de que os fatores marshallianos apresentam efeitos mais significativos nos padrões de coaglomeração, do que as vantagens naturais compartilhadas no setor manufatureiro dos EUA.

³ É um segmento da bolsa alemã voltado para o setor de tecnologia futura criado em 1997.

⁴ É a mudança de visão da economia, deixando de ser economia voltada a indústria e passando a ser uma economia baseada nos serviços, principalmente tecnológicos.

Em Gabe e Abel (2013), os índices de coaglomeração de Ellison e Glaeser (1997) são utilizados para as ocupações dos EUA em pares com o intuito de analisar os fatores que ajudam a determinar a concentração geográfica dessas ocupações. Tal análise foi feita em dois níveis: na área metropolitana e estadual. Em conjunto com o índice de EG as semelhanças de ocupações foram utilizadas para a análise dos chamados padrões geográficos da localização dos empregos. Descobriu-se que, mais nas áreas metropolitanas do que nos estados, o requisito de conhecimento semelhante é determinante para a coaglomeração das ocupações. Se aprofundando na área metropolitana, a coaglomeração sofre maior influência do conhecimento compartilhado nas áreas de engenharia e tecnologia, artes e humanidades, fabricação e produto, e matemática e ciência.

Mukim (2014) modificou os índices de EG e EGK para incluir uma medida de coaglomeração informal e formal dentro de cada indústria da manufatura da Índia. Concluiu-se que, dentro da mesma indústria, os vínculos comprador-vendedor entre as empresas formais e informais junto com os spillovers tecnológicos explicam em boa parte a coaglomeração. No entanto, para a agregação intra-industrial e a inter-industrial o efeito de coaglomeração não é muito grande. Para o setor formal, a coaglomeração influencia o nascimento de pequenas e médias empresas e as empresas formais tendem a se concentrarem em distritos com baixo nível de diversidade industrial. Enquanto as empresas informais se fincam próximas de empresas em indústrias similares e tendem a começar em distritos mais populosos.

Sem o uso dos índices de coaglomeração, Helsley e Strange (2010) encontraram resultados divergentes da literatura. Os autores analisaram a coaglomeração com base em modelos que incluem os casos intermediários, ou seja, cidades que possuem coaglomeração de algumas indústrias, mas não de todas. O modelo leva em consideração algumas hipóteses, tais como o fato de as firmas e as indústrias serem perfeitamente competitivas e a coaglomeração eficiente de Pareto. Os resultados indicam que a coaglomeração é, provavelmente, ineficiente. Acontecendo, em muitos casos, a coaglomeração excessiva, e em outros coaglomeração não realizada.

Em Helsley e Strange (2012), é assumido que os trabalhadores são perfeitamente móveis e as firmas são perfeitamente competitivas a fim de criar um modelo de formação de clusters de negócios, em cidades com várias indústrias, a migração dependerá da complementariedade da produção entre trabalhadores dessas diversas indústrias. As conclusões chegadas em Helsley e Strange (2010) são confirmadas, ressaltando a possível ineficiência das coaglomerações e atribuindo à fraqueza da migração como mecanismo de disciplina, apesar das forças de aglomeração no trabalho, essa ineficiência.

No caso do índice de EG para pares de setores industriais há dois trabalhos feitos no Brasil. Resende (2012) faz uma investigação dos determinantes de aglomeração para o estado do Rio de Janeiro no período de 2010. Utilizando-se de dados da RAIS (MTE), o autor procura relacionar o grau de coaglomeração entre dois setores e uma das forças de aglomeração teóricas de Marshall, o *labor pooling*. Seguindo a linha do artigo de Ellison, Glaeser e Kerr (2010), Resende obtém evidências interessantes para o caso, indicando que o papel do *labor pooling* é mais forte do que as outras forças de aglomeração para as variáveis aproximando-o a intensidade no uso de insumos.

2. Base de Dados

Os dados utilizados neste artigo são provenientes de duas fontes, sendo a principal delas a base de dados da Relação Anual de Informações Sociais-RAIS (MTE), na qual fornece uma gama

de informações ricas e diversificadas a respeito da situação do emprego no Brasil por regiões e setores de atividades a diferentes níveis de desagregação de classificação. A segunda base é advinda do Núcleo de Economia Regional e Urbana da USP – NEREUS, de onde será retirada informações referentes as matrizes de insumo-produto.

A RAIS é um relatório socioeconômico anual do ministério do trabalho e emprego, contendo informações extraídas junto aos empregadores e pessoas jurídicas, com o objetivo de controle das atividades trabalhistas e obtenção de dados para fins estatísticos. O relatório contém informações a respeito do número de trabalhadores, seus perfis, admissões e demissões, remunerações, etc. A RAIS pode ser disponibilizada em microdados ou dados agregados. Neste trabalho foi utilizado o número de trabalhadores por localidade e setores, baseados na Classificação Nacional de Atividade Econômicas – CNAE, para o cálculo do índice de coaglomeração; e informações referentes ao perfil (ocupação e escolaridade) dos trabalhadores para cálculo das *proxies* das variáveis *labor pooling* e *intellectual spillovers* usando dados agregados para o período 2006 a 2015.

A Matriz Insumo Produto foi calculada por Guilhoto e Sesso Filho (2005, 2010) a partir de dados das contas nacionais e disponibilizadas pelo NEREUS para o período 1995-2013. Neste trabalho foram utilizados os dados de 2010 a 2013 por se tratar de uma série com maiores detalhes dos dados desagregados. Essa base foi utilizada para o cálculo das *proxies* das variáveis *input e output linkages*.

3. Metodologia e Modelo Empírico

Neste tópico será apresentado a metodologia utilizada, explorando as variáveis contidas no modelo, como por exemplo, o índice de coaglomeração de EG. A fim de testar se a primeira hipótese é válida, será feita uma análise exploratória e descritiva dos dados obtidos a partir dos cálculos a nível estadual, microrregional e municipal para o período 2006 a 2015. Através dos resultados encontrados constatando-se a presença de coaglomeração em diversos setores, é possível dar um passo à frente e testar se o padrão de coaglomeração encontrado é regido pelas forças de aglomeração propostas por marshall. Para isso, faz-se uso das variáveis descritas na teoria marshalliana, inseridas numa regressão múltipla com dados em painel baseado no modelo empírico que será detalhado à frente.

3.1. Modelo Empírico - Equação

O modelo é basicamente a equação geral para dados em painéis:

$$Coagg = \beta_{i,t} + \beta_{1i,t}Labor_{1i,t} + \beta_{2i,t}Input_{2i,t} + \beta_{3i,t}Output_{3i,t} + \beta_{4i,t}Intellect_{4i,t} + e_{i,t}, \quad (1)$$

Onde o termo Coag representa o índice de coaglomeração; o termo α representa a constante; β_1 é o coeficiente representando o fator *labor pooling*; β_2 e β_3 são os coeficientes que representam o fator *input* e *output linkages*; β_4 é o coeficiente que representa o fator *intellectual Spillover*; onde o subscrito i denota os diferentes indivíduos e o subscrito t denota o período de tempo que está sendo analisado; e o termo de erro está representado pela notação $e_{i,t}$.

No modelo utilizado na literatura há o coeficiente de vantagens naturais, porém devido à falta de uma mensuração adequada para uma *proxy* no Brasil, foi retirado do modelo. EGK estimam regressões sem a variável de vantagens naturais e constata que, sem essa variável os fatores de Marshall apresentam valores maiores quando comparada com a estimação com o coeficiente de vantagens naturais presente. Porém, estes valores são considerados similares estatisticamente e que as vantagens naturais e os fatores de Marshall provavelmente são ortogonais uns com os outros. Deste modo, espera-se que a remoção da variável não afetará significativamente o

modelo. Para estimar o modelo utilizou-se o método de dados em painel, testando-se para efeitos fixos e aleatórios.

As regressões serão estimadas em dois períodos distintos, visto que algumas variáveis apresentam diferentes séries de tempo. Nas regressões univariadas, foi utilizado o período de 2006 a 2013 para estimações dos fatores *Labor e Intellectual*, enquanto que para as variáveis *Input* e *Output* foi utilizado o período de 2010 a 2013. Nas regressões multivariadas o período de 2010 a 2013 foi empregado, com um total de 1020 observações para 276 combinações de pares de setores.

3.2. Índice de Coaglomeração

A investigação empírica de aglomerações industriais conta hoje com diversas medidas refinadas, que partem de uma mesma base conceitual e utilizam-se de dados de emprego e pessoal ocupado. Nesse trabalho, optou-se pelo uso do índice de Ellison e Glaeser – EG adaptado para pares de setores dois a dois, demonstrado no artigo de Ellison, Glaeser e Kerr – EGK (2010). O índice de coaglomeração de EGK é a principal variável deste artigo e através do cálculo desta, é possível testar se a hipótese de que há pares de setores que tendem a se aglomerar é válida, e também testar a segunda hipótese de que caso exista tendência a se aglomerar, as forças de aglomeração de Marshall assumem papel relevante para explicar as variações nos padrões de coaglomeração. Sendo assim, esta será a variável dependente do modelo que se utilizará.

A variável será denotada como $Coagg_{i,j}$, sendo definida como a correlação da participação do emprego entre pares de setores i e j , em determinada área geográfica. O índice proposto é definido entre dois setores i e j e é dado por:

$$Coagg_{i,j} = \frac{\sum_{m=1}^M (S_{mi} - X_m)(S_{mj} - X_m)}{1 - \sum_{m=1}^M X_m^2}, \quad (2)$$

Onde m indexa as localidades, S_{mi} e S_{mj} denotam a participação do emprego da indústria i e j , respectivamente, na localidade m , e X_m representa o tamanho agregado das localidades m mensurados em termo da participação média do emprego da região entre o total das indústrias. O indicador proposto pelos autores tem como base teórica a “abordagem de jogo de dardos”, que mostra o fato de que as empresas tomam suas decisões de localização baseadas num jogo de dardos, se o valor de γ for zero. Isto é, as empresas tenderiam a se localizar aleatoriamente, como num jogo de dardos.

A base de dados utilizada para o cálculo do índice é advinda da RAIS, onde se utilizam os setores da indústria da transformação classificados a dois dígitos (CNAE2), compatibilizando um total de 24 setores, gerando 276 combinações dois a dois, para os anos de 2006 até 2015. A desagregação a dois dígitos foi escolhida para se ajustar às variáveis explicativas que, por limitações de informações, estão classificadas a dois dígitos em sua maioria. O período foi escolhido dado a disponibilidade dos dados encontrados, visto que antes de 2006 não havia dados detalhados para fazer esta análise. Serão calculados os índices a nível estadual, microrregional e municipal.

3.3. Derivação do Índice de Coaglomeração

A grande maioria dos índices partem de duas medidas básicas, propostas na década de 1940, e que são usadas até hoje: o índice *Herfindahl-Hirshman* (HH) e o indicador g - que são a base dos refinamentos metodológicos trazidos por Ellison e Glaeser (1997) e Dumais, Ellison e

Glaeser (2002). O índice de *Herfindahl-Hirshman* é um indicador de competitividade do setor e de competição do mercado, tendo como objetivo a análise da participação da planta no setor produtivo. Trata-se de uma medida do tamanho das plantas de dada região em relação ao setor. O índice HH é definido como:

$$HH_{i,t} = \sum_{(k=1)}^{(n)} Z^2_{i,k,t} , \quad (3)$$

Onde $Z_{i,k,t}$ é a parcela de mercado da k -ésima planta do setor i , que é composto por n plantas no período t . Será medida a participação de mercado da planta pela proporção do emprego formal desta planta no total do setor em dada região de análise. A medida situa-se entre $1/n$ e 1 , sendo que 1 refere-se a um caso no qual apenas uma planta abastece o mercado todo, e quando está próximo de zero refere-se a um mercado muito competitivo.

Costa e Biderman (2014) definem o índice g como uma medida de dispersão dada pela variância da participação de cada região nos diversos setores de atividade em relação ao valor esperado caso não houvesse sobre ou subconcentração. O índice $g_{i,t}$ pode ser escrito como:

$$g_{i,t} = \sum_{m=1}^M (S_{m,i,t} - X_{m,t})^2, \quad (4)$$

Onde, $X_{m,t}$ é a participação da microrregião m no emprego industrial nacional, obtido pelo quociente entre o total do emprego industrial da microrregião e o total do emprego industrial nacional no momento t ; e $S_{m,i,t}$ é a participação do emprego da indústria i na microrregião m no total do emprego desta indústria no país, calculada por meio do quociente entre o total do emprego indústria i da microrregião m e o total do emprego desta indústria em nível nacional.

Segundo Costa e Biderman (2014), o indicador informa o quanto a participação da região m no emprego do setor i está próxima da participação desta região no emprego industrial como um todo, ou seja, ele avalia se a participação da região na composição de um determinado setor segue o padrão de participação geral da região na composição dos empregos industriais. Ou seja, se todas as regiões tiverem exatamente a mesma composição que o país, o índice será zero. Quanto mais alto o valor de $G_{i,t}$, maior é o grau de concentração geográfica das atividades do setor i . Ellison e Glaeser (1997) propõem uma normalização para $g_{i,t}$ que equivale a:

$$G_{i,t} = \frac{\sum (S_{m,i,t} - X_{m,t})^2}{1 - \sum X^2_{m,t}}, \quad (5)$$

O índice $G_{i,t}$ então assume a forma geral do setor i , no momento t . Porém, para evitar problemas de mensuração em relação ao tamanho das firmas, Ellison e Glaeser propuseram a normalização do índice g pelo índice de HH, para que o caso de poucas plantas industriais não influencie na mensuração na concentração, como por exemplo, o caso dos monopólios naturais. Dado que o índice G não incorpora o índice de HH, sendo comumente chamado de indicador de concentração bruta, Ellison e Glaeser propuseram a partir do índice G e do HH um novo índice:

$$\gamma_{i,t} = \frac{(G_{i,t} - HH_{i,t})}{(1 - HH_{i,t})}, \quad (6)$$

Substituindo (5) em (6), tem-se:

$$\gamma_{i,t} = \frac{g_{i,t}/(1 - \sum_m x_m^2) - HH_{i,t}}{(1 - HH_{i,t})}, \quad (7)$$

Assim, simplificando o índice de EG encontrado anteriormente para dois pares de setores, quando $i = 2$, pode-se encontrar $g_{i,t} = \sum_{m=1}^M (S_{m1} - X_m)(S_{m2} - X_m)$, e assim com a remoção do índice de HH, simplifica-se a fórmula:

$$\gamma^c = \frac{\sum_{m=1}^M (S_{m1} - X_m)(S_{m2} - X_m)}{1 - \sum_{m=1}^M X_m^2}, \quad (8)$$

A fórmula deixa claro que o índice de coaglomeração de EG é estreitamente relacionado com a covariância do emprego industrial nas duas indústrias, enquanto que o índice de aglomeração é uma medida de variância. O denominador reflete a simples covariância para eliminar o caráter sensível da desagregação geográfica. Desse modo, o índice de HH não entra na fórmula de coaglomeração, pois a irregularidade nas plantas afeta apenas a variância da participação do emprego que poderia ser confundida com aglomeração dentro da indústria, mas não conduz, por si só, a um aumento espúrio na covariância inter-indústria.

3.4. Labor Pooling

A movimentação de trabalhadores é um dos fatores de aglomeração abordados por Marshall (1920), que enfatiza a importância das economias de escala referentes ao mercado de trabalho. Diversas teorias e modelos foram propostas para sintetizar os benefícios do *labor pooling*. Diamond e Simon (1990) e Krugman (1991a) argumentam que ao decorrer do tempo, as empresas passam a ser mais ou menos produtivas e a mobilidade dos trabalhadores entre os empregos proporcionam ganhos de produtividade e diminuição do diferencial salarial. Helsley e Strange (1990) propõe que as aglomerações facilitam o *match* entre emprego e trabalhador e suas aptidões. Duranton e Combes (2006) e Dahl e Klepper (2007) sugerem que aglomerações com maior facilidade de mobilidade de trabalhadores incentiva a localização de novos empreendedores decorrente da força de trabalho adequada levando ao aumento da concentração do emprego.

Todas essas teorias partem do pressuposto que as aglomerações ocorrem por que os trabalhadores são capazes de mover-se entre as firmas e indústrias. Porém, isso pode acontecer apenas, quando os trabalhadores apresentarem o mesmo perfil necessário para essas indústrias. Desta forma, este trabalho baseia-se o *labor pooling* como a correlação dos perfis semelhantes de trabalhadores entre as atividades dos pares de setores.

Para esta variável, foi escolhida uma *proxy* baseada na literatura existente, que fosse capaz de captar os efeitos na variação dos índices de coaglomeração. Dentre as opções disponíveis na literatura, optou-se por utilizar a mesma ideia dos artigos de Ellison, Glaeser e Kerr (2010) e Resende (2012), onde se calcula a correlação entre as atividades de ocupações entre cada par de setores, da mesma forma feita para o cálculo do índice de coaglomeração, sendo possível captar quais pares de setores apresenta maior similaridade no uso de trabalhadores com perfis semelhantes.

A variável será denotada como *Labor* e é definida como: a fração do emprego do setor i na ocupação o . Mensura-se a similaridade dos trabalhadores nos setores i e j através da correlação entre $Labor_{io}$ e $Labor_{jo}$ dentre as ocupações de maneira análoga ao índice de coaglomeração.

Os dados utilizados neste cálculo são provenientes da RAIS segmentados por ocupações obtidos da Classificação Brasileira de Ocupações – CBO ao nível dos principais subgrupos, contabilizando 48 ocupações. O período observado compreende os anos de 2006 a 2013, tendo a mesma quantidade de observações que o índice de coaglomeração.

3.5. Input-Output Linkages

A proximidade entre fornecedores e clientes favorece a redução de custos de transportes, levando a verticalização entre os setores e incentivando a aglomeração. Marshall (1920) argumenta que as firmas decidem aonde se localizar baseado nos custos de transporte entre insumos e produtos. Desta forma, a teoria aponta para uma aproximação entre as indústrias que tem forte relação insumo-produto. Fujita, Krugman e Venables (1999) apontam a redução dos custos de transportes dos bens (insumo ou produto) como o principal motivo a aglomerar.

Para melhor identificação a variável foi dividida em duas composições: *forward linkage* e *backward linkage*. Deste modo, é possível captar os fatores *input* e *output* de maneira individuais, analisando se o fator de interrelação entre os pares de setores depende mais de um que do outro. Utilizou-se como base a Matriz de Insumo Produto - MIP calculada pelo Núcleo de Economia Regional e Urbana da USP para os anos de 2006 até 2013, haja vista a disponibilidade dos dados.

Define-se $Input_{j \rightarrow i}$ como a participação dos insumos do setor i advindos do setor j . E também se define $Output_{i \rightarrow j}$ como a participação dos produtos do setor i que são vendidos para o setor j . Essas participações são calculadas entre todos os setores, podendo não ser indústrias de transformação, desse modo o valor do índice será entre 0 e 1. Dado que há relações de ambas as direções, define-se as variáveis como a média das participações dos pares de setores $i \rightarrow j$ e $j \rightarrow i$.

A matriz está composta em sua maioria pela classificação a dois dígitos. Porém, no período 2006-2009 a agregação é maior que o período 2010-2013. Assim, se faz necessária a manipulação dos dados para agregar todos os setores a dois dígitos, mantendo a padronização necessária. Um problema foi encontrado, visto que dois setores estão agregados a um dígito, isto é, o nível de agregação é maior que o utilizado, sendo extremamente complicado de desagregar, diferentemente da agregação feita. Dessa forma, na estimação do modelo estes dois setores foram excluídos da análise, para evitar problemas de mensuração.

3.6. Intellectual Spillovers

A última razão que leva as firmas a se aglomerarem diz respeito a flutuação de conhecimento entre as firmas. Marshall (1920) enfatiza que trabalhadores podem aprender novas técnicas uns com os outros em clusters industriais, levando ao aumento da produtividade. Trabalhos empíricos reforçam essa ideia mostrando que setores tecnológicos e com alto grau de capital humano influenciam na aglomeração. [Ver Saxenian (1996); Glaeser e Kahn (2001)].

Infelizmente, esta variável detém um grau de dificuldade de mensuração extremo, principalmente para países subdesenvolvidos como é o caso do Brasil, visto que não há informações consistentes o suficiente para se ter uma boa *proxy*, como há nos países desenvolvidos onde existe dados de matrizes de flutuações de conhecimento, como a matriz de M. Scherer's (1984) utilizada no artigo de ELLISON, GLAESER e KERR (2010). Uma primeira alternativa seria o uso de patentes por indústrias para se utilizar como uma *proxy*, porém, não há dados disponíveis para todos os setores e o período analisado.

A alternativa escolhida foi basear-se na literatura teórica de Marshall (1920) onde o conhecimento tende a flutuar entre os setores com alto grau tecnológico, que, por conseguinte detém maior proporção de trabalhadores com alto grau de estudo. Desta forma utilizou-se como *proxy* a proporção de trabalhadores com grau de estudo igual ou maior ao Ensino Superior.

Define-se a variável como a correlação entre as proporções de trabalhadores com escolaridade acima ou igual ao Ensino Superior nos pares de setores. Tendo $Intellect_{i,j}$ sido calculado como a média da proporção ao longo dos pares de setores i e j .

4. Análise do Índice de Coaglomeração

Para análise exploratória dos índices de coaglomeração de pares de setores industriais, analisaram-se 24 setores correspondentes à divisão da CNAE 2.0 – conhecido na literatura como nível de dois dígitos – no qual resultou em 276 combinações dois a dois para o período de 2006 até 2015. A análise foi construída a nível estadual, contemplando os 26 estados mais o distrito federal; a nível microrregional, no qual em 2006 o número era de 556 microrregiões e passando a ser em 2015 o número de 558, assim agruparam-se as duas novas microrregiões e analisou-se todo o período pelas 556 microrregiões iniciais; para a análise geográfica considerando os municípios, deparou-se com um problema, pois em 2006 havia no Brasil 4496 municípios e em 2015 este número saltou para 4911 municípios. Assim, foi necessário testar se os valores para dados agrupados e não agrupados para o ano de 2015 apresentariam diferenças, no qual os cálculos mostraram que não há diferenças significantes nos valores dos índices, portanto calcularam-se os índices a nível municipal com os dados não agrupados para cada período.

Tabela 1 – Índice Médio dos Pares de Coaglomeração por Ano e Área

<i>Ano/Nível</i>	<i>Estadual</i>	<i>Microrregional</i>	<i>Municipal</i>
2006	-0.000840	-0.000231	-0.000001
2007	-0.000918	-0.000311	-0.000002
2008	-0.001071	-0.000306	-0.000005
2009	-0.001091	-0.000263	-0.000006
2010	-0.001076	-0.000060	-0.000007
2011	-0.001175	-0.000255	-0.000009
2012	-0.001252	-0.000286	-0.000010
2013	-0.001349	-0.000356	-0.000012
2014	-0.001229	-0.000317	-0.000015
2015	-0.000995	-0.000270	-0.000016

Fonte: Elaboração própria com base nos cálculos do índice de EGK.

Através dos resultados observados na tabela 1 pode-se inferir e colaborar para um tema que vem sendo discutido nas últimas décadas – a desconcentração industrial. Analisando a coaglomeração média ao longo dos anos de todos os pares de setores obtidos do índice especificado na equação (1), encontram-se valores negativos e próximos de zero em todos os níveis, e crescentes negativamente a nível estadual e municipal. Ainda na mesma tabela, fica mais fácil identificar e compreender a evolução dos índices médios ao longo do tempo.

Analisando a tabela 2, a primeira inferência mostra que na média os pares de setores tendem a ser aglomerados aleatoriamente nos três níveis geográficos considerados, isto é, os pares de setores na média não apresentam padrões de aglomeração. Este resultado já era esperado e é compatível com os resultados obtidos por Ellison et al. (2010) e Resende (2012), visto que na

média os valores tendem a ser próximo de zero pelo efeito compensatório, porém há de se observar que o valor do efeito médio é negativo dado pela maior força a não aglomeração.

Tabela 2 – Estatísticas Descritivas dos Índices de Coaglomeração

Estatísticas Descritivas	Observações	Média	Desvio Padrão	Mínimo	Máximo
Índices de EG					
Nível Estadual	276	-0.00110	0.02248	-0.07207	0.13704
Nível Microrregional	276	-0.00027	0.00577	-0.01869	0.03242
Nível Municipal	276	-0.00001	0.00007	-0.00036	0.00037

Fonte: Elaboração própria, baseado nas regressões.

A segunda inferência é mais interessante, pois se encontra uma tendência ao longo dos anos a não aglomeração dos pares de setores, isto é, os valores médios negativos estão ficando maiores ao longo do período de 2006 a 2015. Este resultado corrobora com a desconcentração industrial que se identificou nas últimas décadas, visto que os fatores exógenos, como exemplo, políticas fiscais, podem superar as forças de aglomeração, levando assim a não aglomeração dos pares de setores e consequentemente a desconcentração.

Investigando mais a fundo⁵, encontra-se que 46,9% dos valores encontrados para os índices a todos os níveis são positivos, indicando assim que pouco menos da metade dos pares de setores são aglomerados por algum fator, o que ajuda a explicar por que os valores médios são negativos. Quando se olham os valores que são relevantes – EGK consideram valores acima de 0,01 para análise de três dígitos – 11% dos pares de setores são coaglomerados de forma mais relevante, sendo o nível estadual responsável pela presença de 92,37% dos valores acima de 0,01. Vale ressaltar que ao nível de dois dígitos esperam-se valores não muito altos, e a título de comparação o valor máximo para o nível estadual a três dígitos no trabalho para os EUA foi no valor de 0,207, enquanto neste presente trabalho o valor máximo obtido foi de 0,137.

Analisando individualmente cada nível⁶, observa-se que a nível estadual os valores apresentam valores esperados em comparação aos outros trabalhos da área, apresentando 87 pares de setores com índices superiores a 0,01. A nível microrregional é possível perceber a inconstância ao longo dos anos, apresentando uma significativa variação no índice médio ao longo do tempo. A nível municipal esperava-se valores maiores quando comparado às evidências encontradas para os EUA, visto que os índices máximos e mínimos foram muito próximos a zero, indicando assim uma aleatoriedade na decisão de coaglomerar. Esse fato talvez possa ser explicado pela quantidade alta de municípios existentes no Brasil somado a desconcentração dos últimos anos, onde a literatura aponta a desconcentração das atividades econômicas.

Quando se observam as tabelas 3 e 4, interessantes resultados aparecem numa análise preliminar dos dados. Pode-se perceber que os fatores de aglomeração tratados neste trabalho, aparentemente, têm forte importância na coaglomeração dos 10 pares de setores com maior índice de coaglomeração. A repetição da aparição dos setores Fabricação de Produtos do Fumo e Fabricação de Produtos Químicos colabora para reforçar o aspecto de forte evidência para presença das forças de aglomeração de Marshall.

Analisando atentamente é possível inferir sobre alguns pares específicos que demonstram fortemente os fatores de aglomeração presentes. A fabricação de Produtos do Fumo (12) parece

⁵ Por falta de espaço, serão omitidas as tabelas, podendo ser solicitadas posteriormente.

⁶ Por falta de espaço, serão omitidas as tabelas, podendo ser solicitadas posteriormente.

estar relacionada a setores intensivos em trabalho, principalmente de baixo nível de qualificação, como por exemplo, aos setores Fabricação de Móveis e a Preparação de Couros e Fabricação de Artefatos de Couro, Artigos para Viagem e Calçados (15), indicando assim que o fator *labor pooling* pode estar presente nesses setores. O *labor pooling* aparenta estar presente também nos setores relacionados à Fabricação de Produtos Químicos (20), visto que estes setores demandam mão de obra qualificada e especializada.

Tabela 3 – Maiores Valores dos Índices Médios de Coaglomeração dos Pares de Setores ao Longo dos Anos a Nível Estadual – Setores a 2 dígitos

	Setor 1 (CNAE Divisão)	Setor 2 (CNAE Divisão)	Índice Médio de Coaglomeração
1	Fabricação de Produtos do Fumo (12)	Preparação de Couros e Fabricação de Artefatos de Couro, Artigos para Viagem e Calçados (15)	0.137045
2	Fabricação de Produtos Químicos (20)	Fabricação de Veículos Automotores, Reboques e Carrocerias (29)	0.053908
3	Fabricação de Produtos Químicos (20)	Fabricação de Produtos Farmoquímicos e Farmacêuticos (21)	0.045921
4	Fabricação de Produtos do Fumo (12)	Fabricação de Móveis (31)	0.043799
5	Fabricação de Máquinas e Equipamentos (28)	Fabricação de Veículos Automotores, Reboques e Carrocerias (29)	0.043012
6	Fabricação de Equipamentos de Informática, Produtos Eletrônicos e Ópticos (26)	Fabricação de Outros Equipamentos de Transporte, Exceto Veículos Automotores (30)	0.041058
7	Fabricação de Produtos do Fumo (12)	Fabricação de Produtos de Madeira (16)	0.040086
8	Fabricação de Produtos de Madeira (16)	Fabricação de Móveis (31)	0.038702
9	Fabricação de Produtos Químicos (20)	Fabricação de Veículos Automotores, Reboques e Carrocerias (29)	0.038611
10	Fabricação de Produtos Farmoquímicos e Farmacêuticos (21)	Fabricação de Produtos Diversos (32)	0.037570

Fonte: Elaboração própria a partir dos cálculos com dados da RAIS baseado na CNAE.

Em relação ao fator de aglomeração *input-output linkages* é fácil ver que muito provavelmente está presente na aglomeração entre Fabricação de Produtos de Madeira (16) e Fabricação de Móveis (31), assim como Fabricação de Produtos Químicos (20) e Fabricação de Produtos Farmoquímicos e Farmacêuticos (21). Nesta análise preliminar há algumas evidências intuitivas, porém não se deve restringir a apenas a um fator, pois é muito provável que todas as forças de Marshall estejam presentes nas coaglomerações identificadas nas tabelas 3 e 4. Apenas através da análise econométrica, na qual será feita na próxima seção, é possível identificar se esses fatores realmente estão influenciando na coaglomeração desses setores investigados.

Tabela 4 – Maiores Valores dos Índices Médios de Coaglomeração dos Pares de Setores ao Longo dos Anos a Nível Microrregional – Setores a 2 dígitos

	Setor 1 (CNAE Divisão)	Setor 2 (CNAE Divisão)	Índice Médio de Coaglomeração
1	Fabricação de Equipamentos de Informática, Produtos Eletrônicos e Ópticos (26)	Fabricação de Outros Equipamentos de Transporte, Exceto Veículos Automotores (30)	0.032416
2	Impressão e Reprodução de Gravações (18)	Fabricação de Produtos Farmoquímicos e Farmacêuticos (21)	0.018211
3	Fabricação de Produtos Químicos (20)	Fabricação de Produtos Farmoquímicos e Farmacêuticos (21)	0.016029
4	Fabricação de Produtos Farmoquímicos e Farmacêuticos (21)	Fabricação de Veículos Automotores, Reboques e Carrocerias (29)	0.015989
5	Fabricação de Produtos Farmoquímicos e Farmacêuticos (21)	Fabricação de Produtos de Borracha e de Material Plástico (22)	0.011340
6	Fabricação de Produtos Farmoquímicos e Farmacêuticos (21)	Fabricação de Produtos Diversos (32)	0.011245
7	Impressão e Reprodução de Gravações (18)	Fabricação de Produtos Químicos (20)	0.010085

Fonte: Elaboração própria a partir dos cálculos com dados da RAIS baseado na CNAE.

Fazendo uma comparação entre os níveis estadual e municipal, percebe-se que os índices microrregionais maiores que 0,01, dos quais são apenas 7 contra 85 do estadual, tem a mesma intuição em respeito as forças de aglomeração. A repetição dos setores de Fabricação de Produtos Químicos (20) e Fabricação de Produtos Farmoquímicos e Farmacêuticos (21) evidencia a ideia apresentada anteriormente. Outro aspecto interessante a notar é a presença de 3 pares de setores nos maiores índices em ambos os níveis, o que pode representar que esses pares têm forte tendência a se aglomerar. Uma ressalva importante é que essa repetição também pode ser advinda das vantagens naturais, na qual não irá se abordar neste trabalho, porém deve ser objeto de estudo futuramente em trabalhos posteriores.

5. Fatores associados a Coaglomeração

Tabela 5 – Estatísticas Descritivas das Variáveis Utilizadas nas Regressões

Variável	Observações	Média	Desvio Padrão	Mínimo	Máximo
<i>Coagg</i>	2208	-0.00111	0.021995	-0.08014	0.14296
<i>Labor</i>	2208	-0.08174	0.055839	-0.19004	0.26892
<i>Intellect</i>	2208	0.010929	0.011405	0.000309	0.09253
<i>Input</i>	1020	0.011198	0.021906	0.000011	0.22440
<i>Output</i>	1020	0.012983	0.026106	0.000020	0.23945

Fonte: Elaboração própria baseada nos resultados da regressão.

Numa primeira análise, as tabelas 5 e 6 apresentam estatísticas referentes as variáveis utilizadas no modelo. Através das informações, é possível perceber que a relação das variáveis explicativas entre pares de setores acompanha o esperado pela literatura, apresentando índices médios próximos de zero. Por outro lado, identifica-se valores máximos interessantes que, comprovam a importância das relações intersetoriais, através dos fatores marshallianos.

Percebe-se também, uma correlação moderada entre o fator explicativo *Labor* e as demais variáveis explicativas, e uma forte correlação entre as variáveis *Input* e *Output*, que intuitivamente era esperado pela relação das variáveis abordada ao longo do trabalho.

Tabela 6 – Correlações dos Fatores Explicativos e Coaglomeração

	<i>Coagg</i>	<i>Labor</i>	<i>Intellect</i>	<i>Input</i>	<i>Output</i>
<i>Coagg</i>	1.0000				
<i>Labor</i>	0.0774	1.0000			
<i>Intellect</i>	0.1109	0.1768	1.0000		
<i>Input</i>	0.0793	0.3832	-0.0993	1.0000	
<i>Output</i>	0.0647	0.3646	-0.0566	0.8537	1.0000

Fonte: Elaboração própria baseada nos resultados da regressão.

A tabela 7 traz os resultados das estimações da equação (6), regredidas primeiramente de forma univariada, e posteriormente multivariada comparando três diferentes regressões. Foi colocado a regressão OLS apenas para título de comparação. Analisando primeiramente as regressões univariadas, percebe-se rapidamente a relação inversa ao esperado na literatura para o sinal dos coeficientes das variáveis *Intellect* e *Output*. No artigo de Resende (2012), encontra-se o mesmo padrão negativo para a variável *Output*, porém, neste presente trabalho a variável não é estatisticamente significativa. De todo modo, esta evidência reforça um possível efeito contrário do fator de produção entre os setores industriais no incentivo a aglomerar. Uma possível explicação para esse fato é que, dado que o setor de infraestrutura brasileiro não é completamente adequado, as indústrias não levam em consideração a diminuição dos custos de transporte nos transportes dos bens finais.

Em relação a variável *Intellect*, duas explicações surgem como prováveis para o sinal negativo. A primeira é que a *proxy* não esteja adequada para captar os efeitos da variável, visto que esta já é conhecida na literatura como de difícil mensuração. A segunda explicação pode ser analisada pela regressão multivariada, visto que na regressão OLS a variável assume coeficiente positivo, e após descontar o efeito fixo ao longo do tempo o coeficiente torna-se negativo. É possível que o fato do nível de escolaridade ter crescido bastante – principalmente médio e superior – no período estudado em todas as regiões do Brasil, possa ter incentivado as empresas a não decidir se aglomerar baseado nos *spillovers tecnológicos*, imaginando-se que todas as regiões detêm mão de obra qualificada.

Entretanto, as regressões univariadas confirmam a importância dos fatores *Labor* e *Input*, no padrão de configuração das coaglomerações, seguindo a linha dos resultados encontrados em Ellison, Glaeser e Kerr (2010). Sendo o fator *Labor* mais destacado, dado que apresentou maior coeficiente a um nível de significância maior em comparação aos demais fatores. Além de ter apresentado maior poder explicativo dentre todas as outras variáveis explicativas, com o R^2 apontando que cerca de 4,7% das variações no índice de coaglomeração são explicadas pela força de aglomeração *labor pooling*.

Observando as regressões multivariadas, pode-se inferir primeiramente que a regressão em OLS não captura os efeitos corretamente, sendo apenas para análise da variável *Intellect*. Comparando-se com as regressões anteriores é fácil ver que todos os coeficientes das variáveis são significantes na regressão anterior permaneceram significativas ao modelo, porém com variações nos valores dos coeficientes e respectivos erros padrões. O poder explicativo do modelo também aumentou, sendo agora 7%. Ainda comparando com as regressões univariadas, percebe-se que os fatores *Labor* e *Input* apresentam coeficientes mais próximos e maiores que

os demais, ressaltando novamente os resultados encontrados em Ellison, Glaeser e Kerr e Resende (2012).

Tabela 7 – Comparativo das Regressões dos Determinantes da Coaglomeração nas Indústrias de Transformação

Regressões Univariadas em Paineis para Efeitos Fixos (E. F)				
	<i>Labor</i>	<i>Intellect</i>	<i>Input</i>	<i>Output</i>
Coeficiente (Desvio Padrão)	0.21525*** (0.022)	-0.09471*** (0.026)	0.11299** (0.048)	-0.00850 (0.1095)
R ²	0.0472	0.0066	0.0061	0.0013
Observações	2208	2208	1020	1020
Regressões Multivariadas				
	Estimação Base OLS	Estimação Base Painel E. F.	Estimação Painel sem a variável <i>Intellect</i> E. F	
<i>Labor</i>	0.0108 (0.012)	0.1574*** (0.026)	0.1611*** (0.027)	
<i>Intellect</i>	0.2213** (0.100)	-0.1144*** (0.030)		
<i>Input</i>	0.1036* (0.061)	0.1750** (0.071)	0.1768** (0.072)	
<i>Output</i>	-0.0226 (0.042)	-0.0074 (0.011)	-0.0093 (0.011)	
R ²	0.0213	0.0709	0.0704	
<i>Observações</i>	1020	1020	1020	

Fonte: Elaboração própria baseada nos resultados da regressão. *** a 1%, ** a 5%, * a 10%.

6. Conclusões

O trabalho testa os objetivos propostos inicialmente, e conclui que há coaglomeração entre pares de setores e que parte dessa coaglomeração é explicada pelas forças de mercado apresentadas por Marshall. Os fatores *labor pooling* e *input linkage* se mostraram mais relevantes dentre os estudados, sendo condizente com a literatura, onde encontra-se resultados similares. Os resultados da regressão múltipla em dados de painel suportam alguns resultados encontrados em Resende (2012) e evidencia a importância desses dois fatores na economia brasileira.

É possível perceber que a recente desconcentração industrial brasileira na última década pode ter influenciado o grau de coaglomeração entre os setores, sendo este um futuro estudo interessante para se aprofundar, relacionando este estudo com as políticas de incentivos fiscais desse período. É possível com os resultados deste trabalho auxiliar na tomada de decisão quando houver incentivos a determinados setores a desconcentrar, visto que alguns setores apresentaram recorrente alta relação com outros setores.

Um resultado já esperado, mas pouco evidente no Brasil, que foi obtido com os resultados, é que com o nível de desagregação geográfica os índices sofreram modificações, o que é uma crítica recorrente ao índice criado por Ellison e Glaeser. Para isto seria interessante ter uma segunda medida para testar como contraponto, como exemplo, o índice de Overman e Duranton, porém infelizmente para o Brasil até então, não há dados necessários para este índice.

Por fim, o trabalho contribui empiricamente na literatura nacional, analisando os padrões de coaglomeração existentes no Brasil a nível estadual, microrregional e municipal, e apresenta fortes evidências para as forças marshallianas presentes nas coaglomerações industriais. Para trabalhos futuros, é interessante a ampliação da análise a 3 dígitos em todas as variáveis e também testar um possível problema de causalidade reversa. Além de uma análise da possível força das vantagens naturais na coaglomeração dos setores.

Referências

- Andersson M.; e Loof H.** (2011). Agglomeration and productivity: evidence from firm-level data, *Regional Science and Urban Economics*.
- Alecke, B.; Asleben, C.; Scharr, F.; Untiedt, G.** (2006). Are there really high-tech clusters? The geographic concentration of German manufacturing industries and its determinants, *Annals of Regional Science*, 40, 19-42.
- Alecke, B.; Asleben, C.; Scharr, F.; Untiedt, G.** (2008). Geographic concentration of sectors on the German Economy: some unpleasant macroeconomic evidence for regional cluster policy. In: U. Blien; Mauer, U. (eds.), *The Economics of Regional Clusters. Networks, Technology and Policy*, Cheltenham: Edward Elgar.
- Alonso-Villar, O.; Chamorro-Rivas, J. M.; González-Cerdeira, X.** (2004). Agglomeration economies in manufacturing industries: the case of Spain, *Applied Economics*, 36, 2103–2116.
- Audretsch, D. B.** (1998). Agglomeration and the location of innovative activity. *Oxford Review of Economic Policy*, 14, 1998.
- Barrios, S.; Bertinelli, L.; Strobl, E.** (2006). Coagglomeration and spillovers, *Regional Science and Urban Economics*, 36, 467-481.
- Combes, Pierre-Philippe, and Gilles Duranton.** (2006). “Labour Pooling, Labour Poaching, and Spatial Clustering.” *Regional Science and Urban Economics*, 36(1): 1–28.
- Combes, P.; Gobillon, L.; Puga, D.; Roux, S.; and Gilles Duranton.** 2012. THE PRODUCTIVITY ADVANTAGES OF LARGE CITIES: DISTINGUISHING AGGLOMERATION FROM FIRM SELECTION. *Econometrica*, Vol. 80, No. 6, 2543–2594.
- Dohse, D.; Steude, S.** (2003). CONCENTRATION, COAGGLOMERATION AND SPILLOVERS: THE GEOGRAPHY OF NEW MARKET FIRMS IN GERMANY. *43rd European Congress of the Regional Science Association*
- Duranton, Gilles, and Henry G. Overman.** 2005. “Testing for Localization Using Micro-Geographic Data”. *Review of Economic Studies*, 72(4): 1077–106.
- Duranton, Gilles, and Henry G. Overman.** 2008. “Exploring the Detailed Location Patterns of U.K. Manufacturing Industries Using Microgeographic Data.” *Journal of Regional Science*, 48(1): 213–43.
- Ellison, G.; Glaeser, G.** (1997). Geographic concentration in U.S. manufacturing industries: a dartboard approach, *Journal of Political Economy*, 105, 889-927.
- Ellison, G.; Glaeser, G.** (1999). The Geographic concentration of industry: does natural advantage explain agglomeration?, *American Economic Review Papers and Proceedings* 89, 311-316.
- Ellison, G.; Glaeser, G.; Kerr, W.** (2010). What causes industry agglomeration? Evidence from coagglomeration patterns, *American Economic Review*, 100, 1195- 1213.
- Fujita, Masahisa, Paul Krugman, and Anthony J. Venables.** 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.
- Gabe, T.; Abel, J.** (2013). SHARED KNOWLEDGE AND THE COAGGLOMERATION OF OCCUPATIONS. *Federal Reserve Bank of New York Staff Reports*, no. 612

Helsley, Robert W., and William C. Strange. 1990. "Matching and Agglomeration Economies in a System of Cities". *Regional Science and Urban Economics*, 20(2): 189–212.

Helsley, Robert W., and William C. Strange. (2010).

Helsley, Robert W., and William C. Strange. (2012). Coagglomeration and the Scale and Composition of Clusters.

Henderson, J. Vernon. 2003. "Marshall's Scale Economies". *Journal of Urban Economics*, 53(1): 1–28.

Krugman, P. (1991). *Geography and Trade*, Cambridge-MA: MIT Press, 1991.

Lautert, V.; Araújo, N.C.M. (2007). Concentração industrial no Brasil no período 1996-2001: uma análise por meio do índice de Ellison e Glaeser (1994), *Economia Aplicada*, 11, 347-368.

Levinsohn, James and Amil Petrin. (2003). "Estimating Production Functions Using Inputs To Control For Unobservables," *Review of Economic Studies*, 2003, v70(2, Apr), 317-341.

Marshall, A. (1920). *Principles of Economics*, London: MacMillan.

Maurel, F.; Sédillot, B. (1999). A measure of the geographic concentration in French manufacturing industries, *Regional Science and Urban Economics*, 29, 575-604.

Mega Mukhim. (2014) Coagglomeration of formal and informal industry: evidence from India *Journal of Economic Geography* (2014) pp. 1–23.

Puga, D. (2010). The magnitude and causes of agglomeration economies, *Journal of Regional Science*, 50, 203-219.

Resende, M.; Boff, H. (2002). Concentração industrial. In: D. Kupfer & L. Hasenclever (eds.), *Economia Industrial: Teorias e Prática no Brasil*, Editora Campus: Rio de Janeiro, 73-90.

Resende, M.; Wyllie, R. (2005). Aglomeração industrial no Brasil, *Estudos Econômicos*, 35, 433-460.

RESENDE, Marcelo. (2012) Co-aglomeração industrial no estado do Rio de Janeiro: um estudo empírico. *Nova econ. [online]*. 2015, vol.25, n.1, pp.181-194

ROCHA, Roberta; BEZERRA, Fernanda Mendes; DE MESQUITA, Cristiane Soares. Uma Análise dos Fatores de Aglomeração da Indústria de Transformação Brasileira. *Revista EconomiA*, 2013

Rosenthal, S. S.; Strange, W. C. (2001). The determinants of agglomeration, *Journal of Urban Economics*, 50, 191-229.

Rosenthal, S. S.; Strange, W.C. (2004). *Evidence on the nature and sources of agglomeration economies*. In: J.V. Henderson & J. F. Thisse (eds.), *Handbook of Urban and Regional Economics*, Vol. 4, Amsterdam: North-Holland.

Silva, M. V. B.; SILVEIRA NETO, R. M. (2007) Crescimento do emprego industrial no Brasil e geografia econômica: evidências para o período pós-real. *Economia: Revista da Anpec, Brasília (DF)*, v. 8, n. 2, p. 269-288.

Steingraber, R.; Gonçalves, F. O. (2015). A influência da aglomeração e da concentração da indústria sobre a produtividade total dos fatores das empresas industriais brasileiras. *Nova econ. vol.25 no.2*.