

DETERMINANTES DA LOCALIZAÇÃO INDUSTRIAL: UMA ANÁLISE PARA AS MESORREGIÕES BRASILEIRAS

Resumo: Conforme a literatura há inúmeros fatores capazes de influenciar a escolha do local de instalação de uma firma industrial, os quais na maioria das vezes diferem conforme o segmento da indústria. Existem evidências para o Brasil, bem como para outros países, de que a localização industrial tende a ocorrer de forma concentrada, implicando na heterogeneidade da distribuição das indústrias no espaço. Nesse sentido, o presente estudo buscou averiguar quais os fatores mais importantes para a localização industrial no Brasil, dentro os muitos apontados pela literatura. A análise foi feita por meio do modelo *logit* multinomial e multinível, por nível tecnológico das indústrias para as mesorregiões brasileiras. Como principais resultados, confirmou-se a importância dos níveis de tecnologia sobre a localização industrial, bem como das economias de aglomeração.

Palavras-Chave: Determinantes locais da indústria. Modelo *Logit Multinomial e Multinível*. Níveis tecnológicos. Mesorregiões brasileiras.

Abstract: According to the literature, there are many factors that influence the choice of the location of an industrial firm, which in most cases differ according to the industry segment. There is evidence for Brazil, as well as for other countries, that industrial location tends to occur in a concentrated way, implying the heterogeneity of the distribution of industries in space. In this sense, the present study sought to find out which are the most important factors for the industrial location in Brazil, within the many pointed out in the literature. The analysis was made using the multinomial and multilevel *logit* model, by technological level of the industries for the Brazilian mesoregions. As main results, the importance of the technology levels on the industrial location as well as the agglomeration economies was confirmed.

Keywords: Determinants of industrial location. *Multinomial and Multilevel Logit Model* Technological levels. Brazilian Mesoregions.

Introdução: A localização das atividades produtivas e mais especificamente das indústrias, tem sido uma questão de pesquisa importante em economia desde os trabalhos seminais sobre os “círculos concêntricos” de Johann Heinrich Von Thünen e o “Estado Isolado” de Alfred Weber. Contudo, dada a complexidade no entorno das questões locais, a literatura empírica, tanto nacional como internacional, costuma variar amplamente, de modo que não exista dois estudos com resultados exatamente iguais. A razão para a falta de consistência entre os estudos locais é explicada por um conjunto de fatores (BLAIR, PREMUS, 1987; BADRI, 2007).

A decisão local é um problema complexo, porque envolve escolha comportamental. Apesar de um dos interesses fundamentais da teoria econômica ser o comportamento das escolhas humanas, as análises econômicas não são capazes de observar e controlar todos os fatores que influenciam o comportamento. Essa tarefa se torna ainda mais difícil quando a tomada de decisão envolve aspectos qualitativos e até mesmo não econômicos, como questões políticas e motivações pessoais dos empresários. Além do aspecto comportamental, o entendimento da escolha local apresenta outro agravante: a disponibilidade dos fatores não é estática no tempo e no espaço, dado que variam conforme

ocorrem modificações geográficas na disponibilidade de recursos, mudanças tecnológicas, alteração nas preferências dos consumidores, dentre outros (MCFADDEN, 1974).

Os estudos sobre o tema também podem abordar diferentes níveis e unidades de análise, se estendendo de comunidades até países; dado que a firma industrial se depara com fatores como salários, impostos, alugueis, economias de aglomeração, e analisa como esses fatores se comportam a nível interurbano, inter-regional e internacional. Os fatores locais exercem influência distinta sobre decisões de investimentos regionais e internacionais (LEME, 1982; FERREIRA, 1989; GUIMARÃES *et al.*, 2004). Além disso, os fatores que impactam a decisão locacional podem mudar conforme o tipo da indústria. Segundo Hatzichronoglou (1997); Lima (2003), Passos e Lima, (1992) Staduto *et al.*, (2008) e Schettini (2010), as indústrias apresentam diferentes interesses e características de produção, conforme os diferentes segmentos; os diferentes bens produzidos por cada tipo de indústria tornam suas necessidades também diferentes.

Diante dessa variabilidade, algumas questões básicas continuam sendo reformuladas em diferentes modelos, inclusive para o caso brasileiro: A aglomeração é de fato a força dominante na decisão locacional? Custos de trabalho e da terra importam? Grandes mercados são mais eficientes para a localização de indústrias? Diferentes setores, com diferentes intensidades tecnológicas, de fato são tão diferentes quanto as suas escolhas locais? Os custos de transporte ainda são relevantes?

Apesar dos resultados da literatura empírica da localização industrial variarem, e dos estudos demonstrarem que os principais determinantes da localização diferem muito dependendo do país, região ou indústria em questão, alguma generalização pode ser feita. Um aspecto comum presente na literatura é o pressuposto de que as indústrias definem um local com base na maximização do lucro. A localização é considerada uma decisão de investimento de longo prazo, em que a empresa decide localizar-se no local em que o valor presente líquido de sua produção, seja o maior possível (BLAIR, PREMUS; 1987; GUIMARÃES *et al.*, 2000, 2004; FIGUEIREDO, 2002, 2003; CIEŚLIK, RYAN, 2005; HANSON, 2005).

Analisar fatores que influenciam a localização das indústrias no Brasil é, portanto, buscar indicações sobre o potencial crescimento futuro das regiões brasileiras. Isso porque, para que determinada região receba investimentos do setor industrial, deve apresentar níveis de lucratividade elevados, comparado a outras áreas. Em última instância, as indústrias escolhem entre as várias alternativas de localização, aquela que lhe proporcione maior lucro, tornando as desigualdades regionais um problema de diferentes oportunidades e níveis de progresso entre as regiões do país (HANSEN, 1987; HANSON, 2005; SCHETTINI, 2010).

Considerando o complexo debate teórico e empírico sobre tema da localização das indústrias no território brasileiro, temos a seguinte questão de pesquisa. Quais são os fatores locais que mais influenciam a localização industrial de acordo da densidade tecnológica? Nesse sentido, propõe-se investigar se determinado fator considerado pela literatura como um determinante da localização das indústrias é válido para o caso brasileiro ao nível mesorregional.

A metodologia empregada neste estudo para avaliar os determinantes locais das indústrias no Brasil parte de modelos de escolha discreta, os quais permitem que sejam estimadas as probabilidades de escolha com base na abordagem comportamental. Neste caso, a decisão de localização industrial é lançada como um problema de escolha discreta em que empresas selecionam os locais onde irão se localizar a partir de um conjunto distinto de regiões com características específicas, com o objetivo de maximizar seu lucro (GREENE, 2003).

O modelo de escolha discreta escolhido para a realização desta análise é o modelo *logit* multinomial e multinível. A escolha por um modelo multinível é um diferencial deste estudo, pois, o interesse é averiguar através desse método se os segmentos industriais, agrupados por nível tecnológico, influenciam nos interesses das indústrias quanto as características desejadas para o local. Estatisticamente, um modelo de regressão é entendido como multinível quando as variáveis explicativas são compostas de uma estrutura hierárquica e existe um efeito aleatório que interfere nas variáveis do primeiro nível. Neste estudo, o primeiro nível de análise corresponde ao segmento da indústria agrupados por nível tecnológico, e o segundo nível as características das mesorregiões.

Quadro analítico: A ideia principal por trás da teoria da localização é que o desempenho das empresas está altamente relacionado às características do território em está localizada. Sabe-se que os territórios influenciam as atividades econômicas que ocorrem dentro dele. Essas atividades têm características particulares (intensidade de uso do solo, uso de insumos energéticos, os requisitos de capital humano etc.) e necessitam de um ambiente específico que é fornecido de uma maneira diferente por cada local. Cada vez mais a literatura teórica e empírica vem se esforçando para gerar modelos que identifiquem os fatores locais dentre as mais numerosas alternativas espaciais, buscando melhor refletir as reais decisões de locais realizadas pelas empresas (BLAIR; PREMUS, 1987).

Além do reconhecido pioneirismo de obras como as de Weber [1909], Christaller [1933], Hoover [1948], Lösch [1940], Isard [1962] e Marshall [1985] para a formação da teoria da localização, o quadro teórico a Nova Geografia Econômica (NGE) proporcionou avanços relevantes ao tratar a localização através de modelos de equilíbrio geral. Influenciados por Marshall (1985), Masahisa Fujita, Paul Krugman, Andrew Mellinger, Jeffrey Sachs e Anthony Venables, atribuem às economias externas papel central na decisão locacional. Nessa abordagem admite-se que equilíbrios possam emergir de dinâmicas locais quando os agentes realizam o processo de maximização de seu bem-estar baseando-se nas decisões de outros agentes. As especificações orçamentárias, os preços, os recursos, a distribuição geográfica da população, a oferta e a demanda, e outros determinantes são especificados e incluídos nos modelos, dessa forma a NGE pode tratar das (des) economias de aglomeração e identificar a tensão entre forças centrípetas e centrífugas, que tendem a atrair e repelir atividades produtivas de determinado local, respectivamente (KALDOR, 1957; HIRSCHMAN 1958; HENDERSON, 1986; NAKAMURA, 1985, 2005; KRUGMAN, 1990, 1991, 1997, 1998).

No quadro empírico, o aumento nas últimas décadas de estudos que investigam as forças que existem por trás das decisões locacionais das atividades produtivas, sobretudo as indústrias, proporcionaram avanços metodológicos importantes, principalmente quanto a função de maximização de utilidade. A maioria das contribuições recentes para a teoria da localização consiste em novas evidências sobre alguns determinantes, como impostos, salários, custos de transporte, economias de aglomeração etc. Muitas vezes tais evidências são resultado de novos conjuntos de dados, como áreas geográficas menores, e as “novas” investigações se diferenciam uma das outras em termos de especificações econométricas, variáveis e características de amostragem. Alguns estudos focalizam uma determinada região e, portanto, chegam a conclusões que diferem dos estudos com uma perspectiva nacional. Estudos realizados para novas indústrias e para novas plantas também apresentam resultados distintos. Os determinantes também não são estáticos no tempo, e mudam à medida que as condições de produção mudam (BLAIR, PREMUS, 1987; FIGUEIREDO, 2002; CIEŚLIK,

RYAN, 2005; BADRI, 2007; CAROD *et al.*, 2010; BRUYNE, 2003, ELLISON, GLAESER, 1997; FIGUEIREDO *et al.*, 2002; 2003; GUIMARÃES *et al.*, 2000; 2004).

Metodologia: O presente estudo aplicou métodos de escolha discreta para testar quais fatores são determinantes na localização das indústrias. Os modelos de escolha discreta permitem que sejam estimadas as probabilidades de escolha a partir da abordagem comportamental.

Neste caso, a decisão de localização industrial é lançada como um problema de escolha discreta em que empresas selecionam os locais onde irão se localizar a partir de um conjunto distinto de regiões e localidades. Uma grande vantagem da abordagem da escolha discreta em pesquisa de localização industrial é que ela pode ser testada contra uma ampla variedade de dados de aspectos espaciais, mantidos por governos nacionais e regionais.

Os dados empregados nos modelos de escolha discreta podem ser obtidos por meio de dois procedimentos, o de escolha revelada e o de escolha declarada. As preferências reveladas baseiam-se nas escolhas efetivamente realizadas pelos indivíduos diante de uma situação existente e contêm informação sobre a importância relativa das diversas variáveis que influenciam a sua decisão de consumo, por exemplo. As preferências declaradas referem-se às escolhas em relação a um conjunto de opções, por exemplo, são apresentados cenários hipotéticos ao consumidor para que ele indique a sua escolha. Neste trabalho a fonte de informação empregada é a escolha revelada, uma vez que são obtidas em banco de dados e não através de entrevistas (SILVA *et al.*, 2014).

Um dos objetivos econômicos fundamentais é entender o comportamento que leva à tomada de decisão dos agentes. Essa, contudo, não é uma tarefa simples, pois, os modelos econométricos não podem observar e controlar todos os fatores que influenciam o comportamento humano. As análises de comportamento se tornam ainda mais complexas quando as alternativas para a tomada de decisão são de natureza qualitativa (MCFADDEN, 1974).

Apesar das dificuldades, muitos modelos de escolha comportamental individual foram desenvolvidos utilizando dados de escolhas de uma população (MDFADDEN, 1974; MCFADDEN, 1997; FIGUEIREDO *et al.*, 2002; GUIMARÃES *et al.*, 2004; HANSON, 1998; 2005). Nas análises de escolha comportamental com alternativas de natureza qualitativa, como é o caso deste trabalho, assume-se como regra que todos os indivíduos em uma população tenham um comportamento comum, exceto por algumas observações puramente aleatórias (erros). Também se assume que variações sistemáticas nas escolhas agregadas refletem variações na escolha individual (MCFADDEN, 1974).

Neste trabalho foi aplicado o modelo *logit* de escolha comportamental para a decisão locacional das indústrias nas mesorregiões brasileiras, a exemplo de Guimarães *et al.* (2000; 2004), Hanson (2005), dentre outros. O modelo *logit* é um método econométrico de seleção qualitativa, uma vez que gera respostas de procedimentos qualitativos do tipo presença ou ausência de um determinado atributo. Nesses modelos em que a variável dependente (Y) é de natureza qualitativa o objetivo é encontrar a probabilidade de que algo aconteça, como a probabilidade de uma empresa se localizar em um local devido as suas características (GREENE, 2003).

Assim, conforme Greene (2003) o modelo *logit* utiliza a função distribuição acumulada logística, dada pela equação (1):

$$P_i = E(Y = 1 \vee X_i) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}} \quad (1)$$

em que P_i é a probabilidade, X_i são as variáveis explicativas do modelo, os β são os parâmetros das variáveis explicativas e e base do logaritmo natural.

A partir de algumas transformações algébricas obtém-se:

$$L_i = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_1 + \beta_2 X_i \quad (2)$$

em que o logaritmo da razão de chance é linear em X e nos parâmetros. Lé denominado *logit* e, em consequência, o modelo é denominado modelo *logit*.

Tal modelo não pode ser estimado pelo procedimento normal de Mínimos Quadrados Ordinários (MQO) visto que os dados não estão em nível individual ou micro. Nessa situação pode-se utilizar o método de máxima verossimilhança (MV) para estimar os parâmetros (GREENE, 2003). Por meio da obtenção dos coeficientes das variáveis explicativas, pode-se perceber a influência dessas na variação do logaritmo da probabilidade da variável dependente.

No caso deste estudo, o interesse é saber, por meio do modelo *logit*, a probabilidade de uma indústria escolher determinado local, a partir de uma série de alternativas e características. Muitos estudos empíricos com objetivos semelhantes a este utilizam como variável resposta, o número de indústrias que surgem (nascem) em uma região em dado intervalo temporal. Esse é o caso dos trabalhos de Guimarães *et al.* (2000; 2004). Para o presente estudo, no entanto, não foi possível prospectar informações sistematizadas sobre o nascimento de indústrias nas regiões brasileiras. Diante disso, a estratégia utilizada foi de calcular o saldo líquido de estabelecimentos industriais, isto é, a diferença entre o número de indústrias que abriram e fecharam, em determinado intervalo de tempo. Dessa forma, a variável resposta é categorizada, podendo ser de saldo negativo, zero, positivo e menor que 50, ou positivo e maior ou igual que 50. A escolha desses intervalos para a variável resposta ocorreu a partir de vários testes, em que os valores citados foram os que apresentaram melhores resultados e ajustamento do modelo¹.

A variável resposta é categorizada, com quatro categorias, mas não ordenada. Nos modelos *logit* ordenados a variável de resposta tem mais de duas categorias ordenadas ou hierarquizadas. Mas, há situações em que o regressando não é ordenado, que é o caso deste estudo. Pode-se citar o exemplo da escolha do meio de transporte para ir ao trabalho, que pode ser bicicleta, motocicleta, automóvel, etc. Embora sejam respostas categoriais, não há classificação ou ordenamento (GREENE, 2003).

Assim, no caso deste estudo a variável dependente pode assumir os seguintes valores: Saldo de estabelecimentos industriais: Variável categórica = 0 se saldo < 0; Variável categórica = 1 se saldo = 0; Variável categórica = 2 se 0 < saldo < 50; Variável categórica = 3 se saldo >= 50.

O modelo para escolha locacional será dado por:

$$Prob(Y_i = j) = \frac{e^{\beta_j' X_i}}{\sum_{k=0}^3 e^{\beta_k' X_i}}, j = 0, 1, 2, 3. \quad (3)$$

O modelo (3) é um modelo *logit* multinomial. A estimação da equação resulta em um conjunto de probabilidades para J+1 escolhas. Esse modelo implica a possibilidade de computar J log-odds ratios (razão de chances).

$$\ln\left[\frac{P_{ij}}{P_{ik}}\right] = X_i'(\beta_j - \beta_k) = X_j' \beta_{jsek} = 0 \quad (4)$$

Do ponto de vista da estimação, é desejável que a *odds ratio*, P_j/P_k , não dependa de outras alternativas, o que significa que o pressuposto de irrelevância de alternativas independentes (IIA) é respeitado.

¹Apesar da descrição das variáveis e dos bancos de dados ser apresentada na sequência do trabalho, cabe mencionar aqui algumas informações a esse respeito, como forma de explicar e justificar o modelo adotado.

Modelo Logit Multinomial e Multinível: Além de ser caracterizado como um modelo de variável resposta categórica não ordenada, este trabalho apresenta outra característica que deve ser considerada em sua modelagem econométrica. De acordo com Carod *et al.* (2010), os determinantes locacionais da indústria podem ser diferentes conforme o nível tecnológico do segmento industrial. Dessa forma, uma modelagem cujo objetivo é identificar os determinantes locacionais, deve considerar essa diferenciação setorial.

Uma forma de tratar essa questão é por meio de análises multiníveis. Nesses casos, a base de dados possui uma estrutura hierárquica e quando esses dados são tratados de forma convencional podem acarretar um viés sistemático de subestimação da variável agregada. Isto porque, haveria a quebra de um dos pressupostos convencionais de uma análise de regressão – o de que a variância do erro é constante. A utilização de modelos multiníveis tem sido uma das alternativas para resolver este problema (PONTILI, 2015).

Estatisticamente, um modelo de regressão é entendido como multinível quando as variáveis explicativas são compostas de uma estrutura hierárquica e existe um efeito aleatório que interfere nas variáveis do primeiro nível. Com isso, as interações entre as covariâncias medidas nos diferentes níveis afetam o resultado da variável dependente. Nesses modelos o intercepto e/ou a inclinação não são os mesmos para todas as unidades de nível 2, e a variação pode se dar pelo efeito de alguma variável explicativa de nível 2 e/ou por componente aleatório. Os modelos multiníveis possibilitam a análise de como as variáveis explicativas de um nível modificam o efeito das variáveis explicativas em outro nível (GUO; ZHAO, 2000; RIANI, 2005). Abaixo, um modelo linear multinível, com uma variável explicativa e a suposição de que a variabilidade do nível 2 afeta apenas o intercepto:

$$y_{zj} = \beta_{0j} + \beta_{1j} \cdot x_{1zj} + e_{zj} \quad (5)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (6)$$

$$\beta_{1j} = \gamma_{10} \quad (7)$$

Em que: y_{zj} é uma variável dependente contínua da z -ésima unidade do nível 1 e j -ésima unidade do nível 2. β_{0j} é o intercepto que sofre efeito aleatório da j -ésima unidade do nível 2, cuja representação está na equação (6). β_{1j} é o coeficiente da inclinação, o qual não se altera em cada unidade do nível 2.

Substituindo as equações (6) e (7), na equação (5), tem-se:

$$y_{zj} = \gamma_{00} + \gamma_{10} \cdot x_{1zj} + u_{0j} + e_{zj} \quad (8)$$

Na equação (6) o efeito aleatório do nível 2 é expresso pelo termo u_{0j} , enquanto o erro aleatório é dado por e_{zj} . Parte-se do pressuposto de que ambos são independentes e seguem distribuição normal com média zero e variâncias constantes: $\sigma_{u_0}^2$ e σ_e^2 . Sendo que $\sigma_{u_0}^2$ é a variabilidade entre grupos e σ_e^2 é a variabilidade dentro do grupo. A variância de y_{zj} é resultado dessas duas variâncias:

$$VAR(y_{zj}) = \sigma_{u_0}^2 + \sigma_e^2 \quad (9)$$

A partir de $\sigma_{u_0}^2$ e σ_e^2 é possível calcular o coeficiente de correlação intra classes, cuja medida indica em quanto a variabilidade da variável dependente é dada pela variação entre grupos:

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2} \quad (10)$$

A estatística ρ varia entre 0 e 1; quanto mais próximo de 1, maior a variabilidade de y_{zj} , que resulta das diferenças entre as unidades do nível 2; e quanto mais próximo de 0, maior é a homogeneidade entre as unidades desse nível e pouca influência intergrupos sobre y_{zj} .

Mesmo que na equação (6), os parâmetros da parte fixa do modelo, comum a todas as observações, são representados por γ_{00} e γ_{10} ; o parâmetro da inclinação é o mesmo para todas as unidades do nível 2; e o intercepto é afetado pelo efeito aleatório, variando de uma unidade j para outra e representado por $\gamma_{00} + u_{0j}$.

Pode-se ainda considerar que a inclinação também varia de uma unidade para outra do nível 2. Assim, o sistema de equação passa a ser:

$$y_{zj} = \beta_{0j} + \beta_{1j} \cdot x_{1zj} + e_{zj} \quad (11)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (12)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (13)$$

Substituindo as equações (12) e (13) na (11), tem-se:

$$y_{zj} = \gamma_{00} + \gamma_{10} \cdot x_{1zj} + u_{1j} \cdot x_{1zj} + u_{0j} + e_{zj} \quad (14)$$

No caso do modelo representado pela equação (14), o efeito aleatório do nível 2 afeta tanto o intercepto u_{0j} , quanto a inclinação u_{1j} ; o terceiro componente do efeito aleatório e o termo de erro do nível 1 é e_{zj} . Igualmente supõe-se que esses termos de erro possuem distribuição normal, com média zero e variâncias constantes: $\sigma_{u_0}^2$, $\sigma_{u_1}^2$ e σ_e^2 . Os resíduos do nível 2 são considerados independentes dos resíduos do nível 1, mas, correlacionados entre si, com covariância dada por $\sigma_{u_{01}}$ (RIANI, 2005; PONTILI, 2015).

Ampliando o modelo e considerando variáveis explicativas nos diversos níveis, tem-se o modelo abaixo:

$$y_{zj} = \beta_{0j} + \beta_{1j} \cdot X_{1zj} + e_{zj} \quad (15)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \cdot Z_j + u_{0j} \quad (16)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \cdot Z_j + u_{1j} \quad (17)$$

No nível individual tem-se a variável dependente y_{zj} e a variável independente X_{1zj} ; no nível de grupo tem-se o vetor das variáveis explicativas Z_j . Nesse sistema de equações: β_{0j} e β_{1j} são, respectivamente, parâmetros estimados do intercepto e da variável explicativa do nível 1, assumidos como aleatórios a partir das equações 16 e 17; u_{0j} e u_{1j} são os efeitos aleatórios do nível 2; γ_{01} e γ_{11} são os coeficientes da regressão associados com os efeitos das variáveis explicativas do nível do grupo na relação estrutural do nível da indústria; γ_{00} e γ_{10} são, respectivamente, os valores do intercepto estimado e da inclinação estimada quando Z_j é igual a zero.

Substituindo (16) e (17) em (15):

$$y_{zj} = \gamma_{00} + \gamma_{10} \cdot x_{1zj} + \gamma_{01} \cdot Z_j + \gamma_{11} \cdot Z_j \cdot X_{1zj} + u_{1j} \cdot x_{1zj} + u_{0j} + e_{zj} \quad (18)$$

O primeiro lado da equação, na parte direita, indica a parte fixa do modelo – $\gamma_{00} + \gamma_{10} \cdot x_{1zj} + \gamma_{01} \cdot Z_j + \gamma_{11} \cdot Z_j \cdot X_{1zj}$ – e a segunda parte do lado direito é conhecida como parte aleatória – $u_{1j} \cdot x_{1zj} + u_{0j} + e_{zj}$. O termo é a interação entre os efeitos aleatórios das inclinações e o vetor de variáveis independentes (PONTILI, 2015).

No modelo *logit* multinível a equação geral com diversas variáveis explicativas é expressa pela equação (19):

$$\eta_{zj} = \ln \left(\frac{\pi_{zj}}{1 - \pi_{zj}} \right) \gamma_{00} + \gamma_{10} \cdot x_{1zj} + \gamma_{01} \cdot Z_j + \gamma_{11} \cdot Z_j \cdot X_{1zj} + u_{1j} \cdot x_{1zj} + u_{0j} + e_{zj} \quad (19)$$

Em que η_{zj} é o log da chance de sucesso e π_{zj} a probabilidade de sucesso para a indústria z .

Para a estimação do modelo utiliza-se o *software* estatístico *HLM (Hierarchical Linear and Nonlinear Modeling)*, desenvolvido por Stephen Raudenbush e Anthony S. Bryk.

Base de dados e variáveis: Considerando o modelo conceitual RUM, a estimação multinível (dois níveis) permite verificar a probabilidade de uma indústria de transformação z , de um dado segmento industrial dentre 24 setores, localizar-se em uma das 137 mesorregiões brasileiras de acordo com as suas características e do segmento industrial. O primeiro nível corresponde ao segmento da indústria e o segundo nível indica as características das mesorregiões. A variável depende e as variáveis independentes estão descritas a seguir, bem como, as bases de dados, considerando os dois níveis.

A variável dependente foi calculada para ser uma *proxy* das indústrias que nascem em uma região em determinado período de tempo, a exemplo do estudo desenvolvido por Figueiredo *et al.* (2002, 2003) e Guimarães *et al.*, (2000, 2004), cuja análise foi sobre os determinantes locacionais da indústria para os municípios americanos. No caso brasileiro, no entanto, não foi possível sistematizar informações sobre nascimento industrial e optou-se por utilizar um saldo de estabelecimento das mesorregiões para os 24 setores da indústria de transformação da CNAE 2.0 (dois dígitos), calculado para o período 2006 e 2014.

A estratégia utilizada foi de calcular o saldo líquido de estabelecimentos industriais, isto é, o diferencial entre indústrias que abriram e fecharam no período de 2006 a 2014. Esse período justifica-se pois, proporciona variabilidade suficiente para realizar a regressão e ao mesmo tempo, não é tão longo a ponto que tenham ocorrido mudanças estruturais na economia, e nas características das mesorregiões.

Dessa forma, a variável resposta é categorizada, podendo ser: Saldo de estabelecimentos industriais: variável categórica = 0 se saldo < 0; variável categórica = 1 se saldo = 0; variável categórica = 2 se $0 < \text{saldo} < 50$; variável categórica = 3 se saldo ≥ 50 .

Os dados para o cálculo da variável dependente foram extraídos do Relatório de Informações Sociais – RAIS, entre os anos 2006 e 2014, para os 24 setores que correspondem ao nível de agregação dois dígitos da CNAE 2.0, para o setor da indústria de transformação.

a) Variáveis explicativas de primeiro Nível:

No primeiro nível analisa-se a probabilidade de localização em determinada mesorregião segundo a intensidade tecnológica. Foi considerado que a intensidade tecnológica pode impactar na decisão locacional em razão das indústrias apresentarem diferentes interesses e características de produção conforme seu padrão tecnológico. Os diferentes bens produzidos por cada tipo de indústria tornam as necessidades também diferentes. O nível de tecnologia determina as técnicas disponíveis e a combinação de insumos (HATZICHRONOGLU, 1997; FIGUEIREDO *et al.*, 2002, 2003; GUIMARÃES *et al.* 2000, 2004; LIMA 2003; PASSOS; LIMA, 1992; STADUTO *et al.*, 2008; SCHETTINI, 2010).

Para a consideração dos efeitos da intensidade tecnológica os 24 setores da indústria de transformação da Classificação Nacional de Atividades Econômicas 2.0 (dois dígitos) foram agrupados em quatro categorias: alta, média-alta, média-baixa e baixa tecnologia. Tal classificação baseou-se na classificação tecnológica da Organização para Cooperação e Desenvolvimento Econômico (OCDE) que define os níveis tecnológicos dos setores econômicos com base na relação entre despesas com P&D e valor agregado, bem como na tecnologia incorporada nas compras de bens intermediários e bens de capital. A classificação tecnológica da OCDE corresponde a uma agregação dos setores econômicos em quartis, de acordo com dados objetivos coletados sobre as empresas que os compõem (HATZICHRONOGLU, 1997; CAVALCANTE, 2014; OCDE, 2016).

Os 24 setores da indústria de transformação foram classificados da seguinte forma: i) Alta Tecnologia (ALTA): Impressão e Reprodução de Gravações, Fabricação de Produtos Farmoquímicos e Farmacêuticos, Fabricação de Equipamentos de Informática, Produtos

Eletrônicos e Ópticos; ii) Média Alta Tecnologia (MED_ALT): Fabricação de Produtos Químicos, Fabricação de Máquinas, Aparelhos e Materiais Elétricos, Fabricação de Máquinas e Equipamentos, Fabricação de Veículos Automotores, Reboques e Carrocerias, Fabricação de Outros Equipamentos de Transporte, Exceto Veículos Automotores; iii) Média baixa Tecnologia (MED_BAIX): Fabricação de Coque, de Produtos Derivados do Petróleo e de Biocombustíveis, Fabricação de Produtos de Borracha e de Material Plástico, Fabricação de Produtos de Minerais Não-Metálicos, Metalurgia, Fabricação de Produtos de Metal, Exceto Máquinas e Equipamentos, Manutenção, Reparação e Instalação de Máquinas e Equipamentos; iv) Baixa Tecnologia (BAIXA): Fabricação de Produtos Alimentícios, Fabricação de Bebidas, Fabricação de Produtos do Fumo, Fabricação de Produtos Têxteis, Confecção de Artigos do Vestuário e Acessórios, Preparação de Couros e Fabricação de Artefatos de Couro, Artigos para Viagem e Calçados, Fabricação de Produtos de Madeira, Fabricação de Celulose, Papel e Produtos de Papel, Fabricação de Móveis, Fabricação de Produtos Diversos.

b) Variáveis explicativas de segundo Nível

No segundo nível foram consideradas as variáveis regionais. Foram analisadas decisões de localização das firmas para muitas opções de unidade espaciais, mas optou-se pela unidade mesorregional, mais especificamente 137 mesorregiões brasileiras. A escolha pela unidade espacial de análise mesorregional é justificada porque ela possui a agregação adequada para este estudo, já que níveis menos agregados como o microrregional ou municipal não revelariam uma variabilidade dos saldos de estabelecimento suficientes para se estimar as regressões. Ao mesmo tempo, ela é pequena o suficiente para expressar o grande número de alternativas de localização no território brasileiro.

As variáveis independentes incluem os fatores escolhidos para serem testados neste trabalho como determinantes da localização das indústrias no Brasil. Ao todo foram testadas sete variáveis. Conforme Blair e Premus (1987), Badri (2007), os fatores considerados na decisão locacional podem ser divididos em dois grandes grupos: fatores que influenciam os custos da firma (custo de transporte, custo da mão de obra, aluguéis e outros) e fatores que influenciam a receita, nesse caso consideram-se todas as variáveis que podem ser úteis na mensuração do tamanho de mercado.

Do lado das variáveis de custo é testado o custo da mão de obra, através da renda média – em número de salários mínimos – para o ano de 2014. A renda média é calculada para os empregos da indústria de transformação, sendo uma variável que caracteriza o custo da mão de obra para cada mesorregião.

Vários autores (WEBER, 1929; HOOVER, 1948; GREENHURT; GREENHURT, 1975) chamam a atenção para a importância dos custos de transporte e sugerem que a indústria escolhe o local que minimiza esses custos. Para mensurar essa variável leva-se em conta as distâncias rodoviárias de cada mercado. Considerou-se como áreas de mercado as capitais estaduais, de modo que para mensurar a distância rodoviária de cada mesorregião com sua respectiva capital estadual, fez-se necessário, em primeiro lugar, estipular um ponto em cada mesorregião e a partir desse calcular a distância. Os pontos das mesorregiões foram os mesopolos, identificados como os municípios de maior Produto Interno Bruto (PIB) de cada mesorregião. Os mesopolos são áreas centrais para suas regiões, via de regra possuem maior população e densidade na oferta de serviços e, portanto, mostram-se adequados para os objetivos deste trabalho (GARCIA, 2007; CROCCO *et al.*, 2006)².

²Os polos mesorregionais identificados neste trabalho são apresentados no Apêndice X, Tabela X.

Após a definição das áreas centrais – centrais em sentido econômico e não necessariamente espacial – das mesorregiões e calculadas as distâncias rodoviárias, realizou-se uma ponderação dessas distâncias com o objetivo de corrigir vieses de mensuração. Uma mesorregião pode apresentar um custo de transporte – distância – pequeno com relação a área de mercado de sua respectiva capital, mas, a área de mercado em questão pode não ser tão importante em comparação às outras áreas. Por exemplo, comparando a distância da mesorregião Vale do Acre para Rio Branco, com a mesorregião Oeste paranaense para Curitiba, a mesorregião do Acre apresenta uma distância muito menor e, portanto, um “custo de transporte” menor, entretanto, a importância da área de mercado de Curitiba em relação à Rio Branco é expressivamente maior. A forma encontrada para lidar com esse problema é expressa pela equação abaixo:

$$\text{proxymdistância} = \frac{\text{PIBcapitalestadual}}{\text{PIBnacional}} \cdot \text{distância}_{ij} \quad (20)$$

Onde i são os mesopolos e j as capitais dos estados brasileiros. Portanto, distância_{ij} é a distância rodoviária mensurada em km do mesopolo i para a capital estadual j .

Para mensurar o tamanho de mercado optou-se pelo PIB mesorregional. As firmas tendem a se localizar nos centros das áreas de mercado e não a se dispersar. Dessa forma, quanto maior o PIB de uma região, espera-se que maior seja a probabilidade da localização industrial. Os PIB's mesorregionais foram prospectados para o ano de 2010 e com o objetivo de obter um melhor ajustamento da regressão, optou-se por utilizar o logaritmo do PIB, o que reduz a magnitude da série.

Além dos custos, outras características da mão de obra podem ser importantes na decisão locacional, como qualificação da mão de obra, nível de escolaridade, estabilidade do trabalhador, dentre outros. Muitos trabalhos empíricos, inclusive os que utilizam fonte de dados primário, apontam para a importância da qualificação da mão de obra para as indústrias. Uma pesquisa realizada nos Estados Unidos, no ano de 1984, fez um levantamento com mil executivos de indústrias de vários ramos e identificou que a produtividade do trabalho, extremamente ligada à qualificação da mão de obra, era o segundo fator considerado de “vital importância” para a escolha do local da empresa. (BLAIR; PREMUS, 1987; BADRI, 2007).

Dada a relevância desse fator, adiciona-se ao modelo uma variável que expressa a qualificação da mão de obra, com o objetivo de averiguar se mesorregiões com mão de obra qualificada são mais atrativas no que diz respeito à decisão locacional das indústrias do que as mesorregiões com mão de obra menos qualificada. Para formulação dessa variável empregou-se a proporção de empregados com determinada escolaridade para cada mesorregião. A equação (24) apresenta a fórmula para o cálculo da proporção de empregados por nível de escolaridade.

$$P_{ki} = \frac{\text{Número de empregados}_{ki}}{\text{Número de empregados}_{total_i}} \quad (21)$$

Em que k conota o nível de escolaridade e i a mesorregião; P_{ki} é a proporção de empregados da mesorregião para um nível de escolaridade k . Os níveis de escolaridade foram divididos em analfabeto, ensino fundamental, ensino médio e ensino superior.

Ao longo dos anos os modelos teóricos e empíricos sobre localização industrial têm incorporado aglomeração junto com os fatores de custo e mercado. A aglomeração inclui tanto as economias de localização como as de urbanização. O arcabouço teórico da economia de aglomeração admite que este tipo de economia exerce efeitos sobre os salários urbano-industriais e afeta o crescimento das regiões. No entanto, não há um consenso entre as teorias de aglomeração de que estes efeitos são provocados pelas economias externas de

especialização produtiva ou se os mesmos remetem à aglomeração urbana – diversificação produtiva (GUIMARÃES, 2002; CAROD *et al.*, 2008).

Como método para investigar se as economias de aglomeração, sejam elas de localização ou de urbanização, influenciam a localização das indústrias de transformação no Brasil, utilizou-se o indicador conhecido como Índice de *Hirschman-Herfindahl* (*IHH*) ajustado, o qual demonstra a variabilidade da especialização produtiva entre as mesorregiões. É calculado através da equação (22):

$$IHH_i = \sum_{j=i}^n \left[\left(\frac{E_{ij}}{E_i} \right) - \left(\frac{E_j}{E_p} \right) \right] \quad (22)$$

Em que:

E_{ij} é o emprego na mesorregião i no setor j .

E_i é o total de empregos industriais na mesorregião i .

E_j é o emprego nacional no setor j .

E_p é o total de empregos industriais no Brasil.

$n = 1, 2, 3, \dots, 24$, são os setores industriais.

O *IHH* ajustado varia entre 0 e 2. Quando o *IHH* for igual à zero, uma mesorregião i qualquer será considerada perfeitamente diversificada. Quando apresentar seu valor máximo, *IHH* igual a 2, ocorrerá o oposto, a mesorregião em questão será totalmente especializada (GALINARI *et al.*, 2007). Dessa forma, valores menores para o *IHH* podem indicar economias de aglomeração do tipo urbanização; valores mais elevados podem ser interpretados como indicadores da presença de economias de aglomeração do tipo localização. Com uma única variável é possível analisar se as economias de aglomeração, e qual delas, influenciam a probabilidade de localização industrial nas mesorregiões brasileiras.

A Tabela 1 apresenta as variáveis explicativas levantadas neste estudo, com a nomeação que cada uma delas recebeu, bem como uma breve descrição.

Tabela 1 – Variáveis explicativas

Nome da variável	Descrição
ANALF	Proporção de empregados analfabetos
ENS_FUND	Proporção de empregados com ensino fundamental
ENS_MED	Proporção de empregados com ensino médio
ENS_SUP	Proporção de empregados com ensino superior
<i>IHH</i>	Índice de <i>Hirschman-Herfindahl</i> ajustado
DIST	Índice de distância (Proxy de custos de transporte)
RENDM_SM	Renda média dos empregos da indústria de transformação por número de salários mínimos
<i>LNPIB</i>	Logaritmo do PIB mesorregional

Fonte: Elaborado pelos autores.

Com relação aos dados utilizados para a formulação das variáveis, a maior parte deles, referente aos estabelecimentos industriais, à renda, às informações sobre emprego industrial e escolaridade (anos de estudos), foram coletados na base de dados da Relação Anual de Informações Sociais – RAIS, publicada pelo Ministério do Trabalho e Emprego – MTE. As informações sobre o Produto Interno Bruto (PIB), população e km^2 foram extraídas da base de dados do Instituto de Pesquisa Econômica Aplicada – IPEADATA. As informações referentes às distâncias foram prospectadas no *Google Maps* (2016). O *Google Maps* é um serviço que fornece as distâncias de deslocamento para uma matriz de origens e destinos. As informações são baseadas no trajeto rodoviário entre pontos de partida e de

chegada, conforme calculado pelo *Google Maps*, consistindo em linhas que contem valores de distâncias para cada par, mais especificamente, quilômetros de distância para cada par. A coleta dessas informações, portanto, é feita dois a dois, isto é, coleta-se a distância entre duas cidades de cada vez. Os dados que formaram as variáveis foram coletados para as mesorregiões brasileiras.

As equações abaixo expressam o modelo a ser estimado, com as variáveis formuladas para este estudo:

Nível 1 – Características dos setores industriais – intensidade tecnológica

$$\ln \left[\frac{P(0)}{P(3)} \right] = \beta_0(0) + \beta_1(0).MED_{ALT} + \beta_2(0).MED_{BAIX} + \beta_3(0).BAIXA \quad (23)$$

$$\ln \left[\frac{P(1)}{P(3)} \right] = \beta_0(1) + \beta_1(1).MED_{ALT} + \beta_2(1).MED_{BAIX} + \beta_3(1).BAIXA \quad (24)$$

$$\ln \left[\frac{P(2)}{P(3)} \right] = \beta_0(2) + \beta_1(2).MED_{ALT} + \beta_2(2).MED_{BAIX} + \beta_3(2).BAIXA \quad (25)$$

Nível 2 – Características das mesorregiões

$$\beta_0(0) = \gamma_{00}(0) + \gamma_{01}(0).ANALF + \gamma_{02}(0).ENS_{FUND} + \gamma_{03}(0).ENS_{MED} + \gamma_{04}(0).IHH + \gamma_{05}(0).DIST + \gamma_{06}(0).RENDM_{SM} + \gamma_{07}(0).lnPIB + u_0(0) \quad (26)$$

$$\beta_0(1) = \gamma_{00}(1) + \gamma_{01}(1).ANALF + \gamma_{02}(1).ENS_{FUND} + \gamma_{03}(1).ENS_{MED} + \gamma_{04}(1).IHH + \gamma_{05}(1).DIST + \gamma_{06}(1).RENDM_{SM} + \gamma_{07}(1).lnPIB + u_0(1) \quad (27)$$

$$\beta_0(2) = \gamma_{00}(2) + \gamma_{01}(2).ANALF + \gamma_{02}(2).ENS_{FUND} + \gamma_{03}(2).ENS_{MED} + \gamma_{04}(2).IHH + \gamma_{05}(2).DIST + \gamma_{06}(2).RENDM_{SM} + \gamma_{07}(2).lnPIB + u_0(2) \quad (28)$$

Pelo fato de o modelo ser multinomial, a regressão é feita tomando sempre uma categoria como referência. No caso da variável dependente, tanto no nível 1 como no nível 2, a categoria de referência foi o saldo de estabelecimentos industriais maior ou igual a 50 (SALDO3). Com relação aos níveis de intensidade tecnológica (nível 1), a variável de referência é a alta tecnologia (ALTA). No modelo de nível 2, a única variável que possui mais de uma categoria é o nível de escolaridade e a categoria de referência neste caso é o ensino superior (ENS_SUP).

Resultados e discussões: O presente estudo buscou inovar ao considerar que a análise para a decisão locacional das indústrias depende não somente das características dos locais considerados, mas, também, do tipo de indústria a partir da sua classificação por intensidade tecnológica. Essa diferenciação de níveis busca captar os efeitos da tecnologia e inovação sobre os resultados esperados das empresas, e, conseqüentemente, no que diz respeito a seus determinantes locacionais. O processo de inovação tem importância estratégica para as empresas na busca de vantagens que melhorem seu desempenho e diferenciais competitivos em suas atividades (PIA, 2004). Alguns resultados deste estudo são apresentados pelas Tabelas 2 e 3.

Tabela 2 – Resultados da aplicação do modelo *logit* multinomial e multinível para o saldo da indústria de transformação no Brasil - coeficiente

Variáveis Explicativas	Variável dependente (Saldo = ou > 50 foi omitido)		
	Categoria 1: Saldo < 0	Categoria 2: Saldo = 0	Categoria 3: Saldo > 0 e < 50
Média alta tecnologia	-0.549382***	-0.109439	0.428284
Média baixa tecnologia	-2.033755*	-2.750192*	-1.865642*
Baixa tecnologia	-1.289752*	-2.022730*	-1.189162*

Nível Mesorregional
(Ensino superior foi omitido)

Proporção de empregados analfabetos	4.475403	14.698628	15.992974**
Proporção de empregados com ensino fundamental	3.015817	2.816304	2.607807
Proporção de empregados com ensino médio	6.332704**	6.249213**	5.824539*
Índice de Hirschman-Herfindahl ajustado	2.327040***	3.595489*	2.129610***
Índice de distância (Proxy de custos de transporte)	0.006713	0.001633	0.003389
Renda média dos empregos da indústria de transformação por número de salários mínimos	0.871473*	1.128173*	0.821594*
Logaritmo do PIB mesorregional	-0.873278*	-2.093934*	-1.147165*

Fonte: Resultados da pesquisa.

Nota: * Denota significância ao nível de 1%, ** Denota significância ao nível de 5%, *** Denota significância ao nível de 10%.

Tabela 3 – Coeficiente dos interceptos calculados

	Coeficiente nível 1	Coeficiente nível 2
Intercepto1(0)	0.935548*	0.983837*
Intercepto1(1)	0.388123*	0.353178***
Intercepto1(2)	2.043214*	2.097870*

Fonte: Resultados da pesquisa.

Nota: * Denota significância ao nível de 1%, ** Denota significância ao nível de 5%, *** Denota significância ao nível de 10%.

O estudo realizado em dois níveis mostrou-se eficiente, visto que, confirma a existência de efeito aleatório na média de saldo de estabelecimento industrial, indicado pela significância do intercepto da regressão, como apresentado na Tabela 3. Dessa forma, pode-se controlar a dispersão dos dados considerando que a dispersão das informações sobre características mesorregionais pode apresentar diferença conforme o nível de intensidade tecnológica analisado. Por exemplo, podemos analisar duas mesorregiões com características muito parecidas, mas, que se mostrarão totalmente diferentes com relação à decisão locacional de indústrias com níveis de intensidade tecnológica distintos.

Com relação ao primeiro nível analisado, a saber, se a intensidade tecnológica influencia a decisão locacional das empresas, as variáveis apresentaram ao menos algum resultado estatisticamente significativo para todas as categorias. Esse resultado confirma a hipótese de que a decisão locacional das indústrias e seus determinantes mudam conforme sua intensidade tecnológica.

Na Tabela 2 é possível realizar uma interpretação mais desagregada, primeiramente com relação à média alta intensidade tecnológica para a categoria 1, ou seja, saldo negativo, que foi a única estatisticamente significativa para esse nível tecnológico. Observa-se que a probabilidade de empresas de média alta tecnologia apresentarem saldo negativo é menor do que empresas com esse mesmo nível tecnológico apresentarem saldo positivo maior ou igual a 50; comparativamente a empresas de alta tecnologia. A contrapartida disso é que empresas de alta tecnologia tem uma probabilidade maior de apresentarem saldo negativo.

Esse resultado se repete para os outros dois níveis tecnológicos que aparecem na regressão – média baixa e baixa tecnologia – cujos coeficientes apresentaram significância estatística para todas as categorias. Dessa forma, pode-se interpretar que a probabilidade de saldos mais baixos de estabelecimentos industriais é sempre maior para as indústrias de alta tecnologia. A probabilidade de saldo maior ou igual a 50 é maior para empresas com média alta, média baixa e baixa tecnologia, comparativamente a alta tecnologia.

Isso leva a interpretação de que os segmentos industriais nasceram com mais intensidade nas mesorregiões brasileiras de 2006 a 2014 foram setores menos intensivos em tecnologia, como era esperado *a priori*. Esperava-se que segmentos menos intensivos em tecnologia apresentassem uma probabilidade maior quanto ao saldo de estabelecimentos positivos. Isso porque, estudos como os de Schettini (2010) mostram que há maior distribuição, bem como maior número de indústrias, pertencentes a setores mais tradicionais nas mesorregiões brasileiras. Os níveis de média baixa e baixa intensidade tecnológica, também apresentaram variação positiva para essas variáveis, e em alguns casos a variação foi maior do que a do primeiro nível citado, como apresenta as figuras a seguir.

Na análise do segundo nível, a escolaridade dos empregados do setor industrial, o qual indica qualificação da mão de obra, também apresentou resultados consistentes. Analisando cada nível individualmente, nota-se que o analfabeto apresentou resultado estatisticamente significativo para a categoria 3. Uma maior proporção de empregados analfabetos, em relação à proporção de empregados de nível superior, aumenta as chances de que o saldo de empresas fique entre 0 e 50 e diminui a probabilidade de que haja empresas com saldo maior ou igual a 50.

Para o ensino médio, tem-se uma maior probabilidade de que as categorias observáveis (saldo negativo, saldo zero, saldo entre 0 e 50) ocorram, quando aumenta a proporção de empregados com esse nível de escolaridade, em comparação à proporção de empregados com ensino superior. Tanto os resultados para analfabeto como para ensino médio indicam que o aumento do nível de escolaridade para ensino superior, torna a probabilidade de obter-se saldo de empresas positivo e maior ou igual a 50, maior.

A forma de medir os efeitos das aglomerações e analisar qual tipo de economia de aglomeração é mais relevante na decisão locacional, foi através de uma *proxy*, o índice *Hirschman-Herfindahl* ajustado, que indica se uma mesorregião é diversificada ou especializada para o total das atividades analisadas. As mesorregiões com *IHH* mais próximo de zero são consideradas diversificadas e as com *IHH* próximo a 2 especializadas. Com essa análise é possível revelar se é a diversidade ou a especialização produtiva das mesorregiões que influenciam na localização das indústrias.

O *IHH* ajustado apresentou resultado estatisticamente significativo, fornecendo evidências de que as economias de aglomeração são um fator determinante na localização das indústrias de transformação no Brasil. Esse resultado é bastante importante, e costuma ser confirmado frequentemente nos estudos empíricos. Em muitos trabalhos, as aglomerações aparecem como principal fator na definição de localização industrial, a exemplo de Smith e Florida (1994) e Guimarães *et al.* (2000).

Os efeitos da aglomeração podem ser tanto positivos como negativos para as firmas. Do ponto de vista teórico, há certa generalização de que a concentração de atividade econômica em um local específico e o grau de atração que tal concentração tem para novas atividades, revela um perfil em formato de U invertido: é inicialmente positivo, mas, uma vez atingido um determinado valor, torna-se negativo. Isso seria uma função do tamanho das aglomerações, sendo que as deseconomias de aglomeração podem ser tão intensas perante as economias que se abre a possibilidade da ocorrência de economias de aglomeração líquidas negativas (GALINARI *et al.*, 2007; CAROD *et al.*, 2010).

Os impactos positivos da aglomeração ocorrem quando a comunicação, os *spillovers* informacionais são importantes, bem como mão de obra especializada e proximidade com fornecedores. Por outro lado, a aglomeração pode desestimular os investidores quando geram externalidades negativas, como congestionamentos, poluição, elevação nos preços da terra, dentre outros (SMITH, FLORIDA, 1994; CIEŚLIK, RYAN, 2005).

Essas questões remetem as forças centrípetas e centrífugas analisadas pela Nova Geografia Econômica. Essas forças resumem-se nos efeitos do tamanho de mercado na geração de vínculos que promovam concentração geográfica, de um lado, e a força de oposição de elementos imóveis trabalhando contra essa concentração, de outro (Krugman (1990, 1991, 1997, 1998).

Ainda com relação ao *IHH*, analisando os resultados para cada categoria, é possível isolar qual tipo de economia de aglomeração é mais importante na decisão locacional. Isso porque, a variável *IHH* apresentou coeficiente com sinal positivo para todas as categorias de saldo analisadas (Categoria 1, 2 e 3), tendo como referência a categoria 4 (saldo maior ou igual a 50). Em termos técnicos, isso significa que a probabilidade de saldo negativo, zero, e positivo menor que 50, aumenta para valores mais elevados do *IHH* em comparação a saldo positivo maior ou igual a 50.

Como valores mais elevados do *IHH* expressam aglomerações do tipo urbanização/diversificação, enquanto valores mais baixos, aglomerações de concentração/especialização, nota-se que as economias de especialização parecem superar as economias de diversificação para o caso brasileiro. As aglomerações do tipo especialização apresentam maior importância, pois, aumentam a probabilidade de saldo de estabelecimentos iguais ou maiores que 50.

Os resultados para a *proxy* de custo de transporte chamam atenção neste estudo, pois, esperava-se que quanto maior as distâncias em relação as áreas de mercado, maiores os custos com transporte, e que isso seria uma implicação negativa sobre a decisão locacional. No entanto, os coeficientes dessa variável foram não significantes do ponto de vista estatístico, indo contra os princípios básicos da teoria de localização urbana e regional. Esse resultado, no entanto, não necessariamente expressa uma irrelevância dos custos de transporte para a decisão do local de instalação das indústrias, mas, pode ter ocorrido como uma falha dessa variável quanto a capacidade de representar os custos de transporte.

A variável Renda Média (número de salários mínimos) dos trabalhadores da indústria de transformação, utilizada neste estudo com objetivo de averiguar os efeitos dos custos da mão de obra sobre as decisões locacionais, mostrou-se estatisticamente significativa. De acordo com a teoria neoclássica de localização, as diferenças de custo de mão de obra, e outros custos, no espaço, podem influenciar a decisão de investimento industrial, além dos fatores de aglomeração. Para as mesorregiões brasileiras, este estudo demonstra que há evidências de que as diferenças de custo de mão de obra influenciam a localização.

A evidência empírica dos custos da mão de obra sobre a decisão de localização industrial varia amplamente. Alguns autores descobriram que salários mais elevados

desestimulam o investimento industrial; enquanto outros não encontraram uma relação específica (FIGUEIREDO *et al.*, 2002, 2003; GUIMARÃES *et al.*, 2000, 2004; CIEŚLIK, RYAN, 2005). Neste estudo, a variável renda média dos empregos industriais apresentou uma relação positiva para todas as categorias de saldo, o que significa que a probabilidade do saldo de estabelecimentos industriais nas mesorregiões brasileiras ser negativo, zero e positivo e menor que 50, aumenta conforme aumentam os salários médios, comparativamente à categoria de saldo positivo, maior ou igual a 50. Ou seja, os resultados mostram que em mesorregiões com salários médios menores há uma maior probabilidade de encontrar saldo de empresas positivo e maior ou igual a 50. No Brasil, a escolha locacional recai sobre as mesorregiões que pagam menores salários aos trabalhadores.

A literatura teórica e empírica, indica que maiores níveis salariais são negativamente associados com a instalação de novas indústrias ou novas plantas industriais, e isso se confirma também para o Brasil. Como sugerem os resultados desta pesquisa, as indústrias preferem mão de obra qualificada, mas, desejam pagar baixos salários. Em suas análises de maximização de lucro as indústrias realizam um *trade-off* entre baixo custo e qualificação da mão de obra (SMITH, FLORIDA, 1994).

O tamanho da área de mercado é mensurado pelo PIB, mais especificamente, pelo Log do PIB; expresso em Reais para o ano de 2010. Neste estudo a influência PIB sobre a localização das indústrias de transformação nas mesorregiões brasileiras mostrou-se estatisticamente significativa ao nível de 1%. O PIB geralmente aparece nos estudos empíricos como um fator positivo na decisão locacional. A interpretação dessa variável pode conter tanto fatores ligados a demanda por produtos industriais como fatores ligados a oferta de insumos. Do lado da demanda o PIB reflete o potencial de demanda de uma região e dá uma estimativa do tamanho de mercado para os tomadores de decisão; do lado da oferta, as indústrias criam demanda por insumos, cuja oferta é proporcional ao tamanho de uma base econômica regional.

Os resultados para o PIB também indicam que a probabilidade de localização industrial diminui para valores mais elevados do PIB, comparativamente a um saldo mais elevado de estabelecimentos industriais. Esse resultado pode ser interpretado como uma probabilidade maior de se ter um saldo mais elevado de estabelecimentos industriais nas mesorregiões que apresentam maiores valores para o PIB. Tal resultado é consistente com as bases teóricas, uma vez que, conforme a abordagem da Nova geografia Econômica, as indústrias são atraídas para regiões em que tenham a possibilidade de servir grandes mercados locais, a partir de algumas plantas com baixo custo de transporte (KRUGMAN, 1991; FUJITA *et al.*, 2002; HANSON, 2005).

Conclusões: Uma das principais conclusões desta análise foram as evidências de que os níveis de intensidade tecnológica das indústrias são relevantes quanto às decisões locais. Isso significa que as características locais desejadas podem mudar conforme o nível tecnológico ao qual o segmento da indústria pertence. Além disso, observou-se as características das mesorregiões brasileiras atraem mais indústrias com menos intensidade tecnológica, uma vez que esses segmentos foram os que mais se localizaram nas mesorregiões brasileiras de 2006 a 2014.

Os principais resultados do segundo nível são evidências de que as indústrias preferem locais com mão de obra mais qualificada, como mostrou a maior probabilidade de saldo de estabelecimento mais elevados em locais com maiores níveis de escolaridade dos empregos industriais. As indústrias preferem locais com mão de obra qualificada, mas, ao

mesmo tempo, que tenham baixo custo de trabalho, visto que a probabilidade de localização industrial aumenta conforme a renda média do emprego industrial diminui.

As economias de aglomeração também foram testadas neste estudo, muitas pesquisas empíricas, tanto para o Brasil como para outros países, sugerem que as indústrias são de fato atraídas por aglomerações produtivas. Em linha com as evidências relatadas na literatura, os resultados indicam que as economias de aglomeração são um fator importante na determinação da localização de indústrias de transformação. Neste trabalho, a utilização do *IHH* como *proxy* de aglomeração permitiu que se analisasse a importância das aglomerações na decisão locacional, bem como que tipo de aglomeração é mais importante na decisão locacional das indústrias de transformação no Brasil. Os resultados para essa variável permitem uma comparação dos diferentes tipos de economias de aglomeração como fatores nas decisões seleção locacional. Aparentemente, as economias de aglomeração de localização/especialização têm o efeito mais forte sobre a localização de indústrias no Brasil.

O PIB das mesorregiões, como uma medida de tamanho de mercado, expressa uma variável para o lado da receita. As evidências encontradas são que o tamanho de mercado é importante para a localização industrial. Isso foi confirmado pelo resultado estatisticamente significativo do PIB das mesorregiões. Quanto maior o PIB mesorregional maior é a probabilidade de que o saldo de estabelecimento seja igual ou maior que 50.

Os custos de transporte foram incluídos no modelo por uma *proxy*, um índice de distâncias entre mesorregiões e suas capitais estaduais. O resultado para essa *proxy* não foi estatisticamente significativo. Estatisticamente, isso indica que os custos de transporte não são relevantes quanto à localização industrial. No entanto, não se pode descartar completamente esse fator como um determinante locacional, uma vez que a *proxy* utilizada pode não ter expressado fielmente os custos de transporte para a decisão locacional das empresas. Também pode ser que o saldo dos estabelecimentos majoritariamente está associado à produção direcionada ao mercado local, portanto, o custo de transporte não estará refletido em custo para empresa. Desta forma, essa *proxy* não capta claramente o custo intrarregional ou mesmo local.

Exceto os transportes, as demais variáveis têm aderência com as abordagens teóricas e empíricas sobre localização industrial, principalmente com o modelo da Nova Geografia Econômica; mesmo para um país em desenvolvimento como o Brasil, com a particularidade da grande extensão territorial e com majoritária participação de indústria de baixa tecnologia em todas as mesorregiões estudadas. No mais, apesar do trabalho ter cumprido com os objetivos propostos, isso não significa que se encerram as possibilidades de análises sobre o tema, ao contrário, a realização deste estudo abriu horizonte para novas perspectivas de pesquisas, ou até mesmo a consideração de outras variáveis como taxas e impostos locais e, disponibilidade de matérias primas.

Por fim, cabe lembrar esse tipo de análise tem por objetivo final servir como base e suporte para o desenvolvimento de políticas públicas, pois, acredita-se que um cenário de concentração e desigualdade não se rompe de forma aleatória, uma vez que a tendência do capital é a concentração. Dessa forma, a intervenção nesse cenário através de políticas públicas que gerem vantagens comparativas para as mesorregiões, pode ter grande efeito para que estas atraiam investimentos industriais e ampliem as possibilidades de melhores condições econômicas e sociais.

Referências:

BADRI, M. A. (2007). Dimensions of industrial location factors: review and exploration. **Journal of business and public affairs**, v. 1, n. 2, p. 1-26 .

- BLAIR, J. P.; PREMUS, R. (1987). Major factors in industrial location: A review. **Economic Development Quarterly**, v. 1, n. 1, p. 72-85.
- BRUYNE, K. de. (2003) The location of economic activity. Is there a spatial employment structure in Belgium? **Processed CES-KU, Leuven**.
- CAROD, J. M. A.; SOLIS, D. L.; ANTOLÍN, M. M. (2010). Empirical studies in industrial location: an assessment of their methods and results. **Journal of Regional Science**, v. 50, n. 3, p. 685-711.
- CIEŚLIK, A.; RYAN, M. (2005). Location Determinants of Japanese Multinationals in Poland: Do Special Economic Zones Really Matter for Investment Decisions? **Journal of Economic Integration**, p. 475-496.
- CROCCO, M. A. *et al.* (2006). Metodologia de identificação de aglomerações produtivas locais. **Nova Economia**, v. 16, n. 2, p. 211-241.
- ELLISON, G.; GLAESER, E. L. (1997). Geographic concentration in US manufacturing industries: a dartboard approach. **Journal of Political Economy**, vol. 105, no. 5.
- FERREIRA, C. M. de C. (1989). As teorias da localização e a organização espacial da economia. In: **Economia Regional: Teorias e Métodos de Análise**. Haddad P. R. (org.) Fortaleza: BNB/Etene.
- FIGUEIREDO, O.; GUIMARAES, P.; WOODWARD, D. (2002). Modeling industrial location decision in U.S. counties. **Working paper series – Universidade do Minho**, n. 18, April.
- FIGUEIREDO, O.; GUIMARAES, P.; WOODWARD, D. (2003). A tractable approach to the firm location decision problem. **Review of Economics and Statistics**, v. 85, n. 1, p. 201-204.
- GALINARI, R.; CROCCO, M. A.; LEMOS, M. B.; BASQUES, M. F. D. (2007). O efeito das economias de aglomeração sobre os salários industriais: uma aplicação ao caso brasileiro. **Revista de Economia Contemporânea**, Rio de Janeiro: Instituto de Economia - UFRJ, v. 11, n. 3, p. 391-420.
- GARCIA, R. A. A. (2007). Regionalização Econômico-Demográfica dos Pólos Econômicos Brasileiros. **Anais: Encontros Nacionais da ANPUR**, v. 12.
- GOOGLE MAPS. (2016). **Matriz de distâncias do Google**. Disponível em: <<https://developers.google.com/maps/documentation/distancematrix/#Introduction>>. Acesso em: 13 de Junho.
- GREENE, W. H. (2003). **Econometric analysis**. Pearson Education.
- GREENHUT, J. G; GREENHUT, M. L. (1975). Spatial Price Discrimination, Competition and Locational Effects. **New Series**, Vol. 42, No. 168. pp. 401-419.
- GUIMARÃES, Paulo; FIGUEIREDO, Octávio; WOODWARD, Douglas. (2000). Agglomeration and the location of foreign direct investment in Portugal. **Journal of Urban Economics**, v. 47, n. 1, p. 115-135.
- GUIMARAES, P.; FIGUEIREDO, O.; WOODWARD, D. (2004). Industrial Location Modeling: Extending the Random Utility Framework. **Journal of Regional Science**, v. 44, n. 1, p. 1-20.
- GUO, G.; ZHAO, H. (2000). Multilevel modeling for binary data. **Annual Review of Sociology**, v. 26, n. 1, p. 441-462.
- HANSEN, E. R. (1987). Industrial location choice in Sao Paulo, Brazil: a nested logit model. **Regional Science and Urban Economics**, v. 17, n. 1, p. 89-108.
- HANSON, G. H. (2005). Market potential, increasing returns and geographic concentration. **Journal of international economics**, v. 67, n. 1, p. 1-24.
- HATZICHRONOGLU, T. (1997). Revision of the High-technology Sector and Product Classification. **STI Working Paper - OECD**, n. 2.

- HENDERSON, J. V. (1986). Efficiency of resource usage and city size. **Journal of Urban economics**, v. 19, n. 1, p. 47-70.
- HIRSCHMAN, A. O. (1958). Interregional and International transmission of economic growth. In: **The strategy of economic development**. New Haven, Yale University Press, p.183-201.
- HOOVER, E. M. (1948). **The location of economic activity**. McGraw-Hill.
- KALDOR, N. (1957). A model of economic growth. **The economic journal**, v. 67, n. 268, p. 591-624.
- KRUGMAN, P. R. (1990). Increasing returns and Economic Geography. **National Bureau of Economic Research**.
- KRUGMAN, P. R. (1991). **Geography and trade**. MIT Press.
- KRUGMAN, P. (1997). **Development, geography, and economic theory**. Vol. 6. MIT press.
- KRUGMAN, P. (1998). What's new about the new economic geography? **Oxford review of economic policy**, v. 14, n. 2, p. 7-17.
- LEME, R. A. Da S. (1982). **Contribuições à teoria da localização industrial**. São Paulo: Instituto de Pesquisas Econômicas (IPE).
- LIMA, R. (2003). Notas à margem da estrutura industrial do Rio Grande do Sul. **Ensaios FEE**. Porto Alegre, v. 24, n. 1, p. 49-92.
- MARSHALL, A. (1985). **Princípios de economia: tratado introdutório**. Volume I. Abril Cultural.
- MCFADDEN, D. (1974). Conditional logit analysis of qualitative choice behavior. In: ZAREMBKA, P. (ed.). **Frontiers of econometrics**. Academic Press.
- MCFADDEN, D. (1978). Modeling the choice of residential location. **Transportation Research Record**, n. 673.
- NAKAMURA, R. (1985). Agglomeration economies in urban manufacturing industries: a case of Japanese cities. **Journal of Urban Economics**, v. 17, n. 1, p. 108-124.
- NAKAMURA, R *et al.* (2005). Agglomeration Economies and Linkage Externalities in Urban Manufacturing Industries: A Case of Japanese Cities. In: **ERSA conference papers**. European Regional Science Association.
- OECD Broadband Portal. (2005). **Directorate for Science, Technology and Industry**.
- STAN INDICATORS. Disponível em: < <http://www.oecd.org/sti/>>. Acesso em: 17 de jun. de 2016.
- PASSOS, M. C.; LIMA, R. S. (1992). Entre perdas e ganhos apontamentos sobre a indústria gaúcha. **Ensaios FEE**. Porto Alegre, v.13 n.2 p. 485-517.
- PESQUISA INDUSTRIAL ANUAL – PIA. (2004). **Série Relatórios Metodológicos**. IBGE. V. 22.
- PONTILI, R. M. (2015). **Determinantes do abandono e atraso escolar, de adolescentes no ensino médio: uma análise para a região Sul do Brasil**. Tese (Doutorado) – Programa de Pós-Graduação em Desenvolvimento Regional e Agronegócio, Universidade Estadual do Oeste do Paraná, Toledo – PR.
- RIANI, J. D. L. R. (2005). **Determinantes do resultado educacional no Brasil: Família, perfil escolar dos municípios e dividendo demográfico numa abordagem hierárquica e espacial**. 218 f. Tese (Doutorado) - Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte - MG.
- SILVA, T.; MENDES, F. B.; FARIA, C. A. (2010). **Aplicação de um modelo de escolha discreta para análise da divisão modal em cidades de porte médio**. Lisboa, 2010. Disponível em:<<http://redpgv.coppe.ufrj.br/index.php/pt-BR/producao-da-rede/artigos->



cientificos/2010-1/445-aplicacao-de-um-modelo-de-escolha-discreta-para-analise-da-divisao-modal-em-cidades-de-porte-medio/file>. Acesso em 25 de jan. 2014.

SCHETTINI, D. C. D. (2010). **Eficiência produtiva da indústria de transformação nas regiões brasileiras**: uma análise de fronteiras estocásticas e cadeias espaciais de Markov. 201 f. Tese (Doutorado) - Departamento de Economia da Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo - SP.

SMITH JR, D. F.; FLORIDA, R. (1994). Agglomeration and industrial location: An econometric analysis of Japanese-affiliated manufacturing establishments in automotive-related industries. **Journal of Urban Economics**, v. 36, n. 1, p. 23-41.

STADUTO, J. A. R.; FERRERA DE LIMA, J. MALDANER, I. STAMM, C. (2008). Análise locacional das ocupações nas regiões metropolitana e não metropolitana do estado do Paraná. **Revista de Economia**. Curitiba, v. 34, n. 2, p. 117-139.

WEBER, A. (1929). **Theory of the Location of Industries**. Translated by Carl J. Friedrich from Weber's 1909 book. Chicago: The University of Chicago Press.