**Distribuição Espacial das Indústrias no Brasil: Uma Análise a partir de Modelos de Escolha Discreta**

**Roberta de Moraes Rocha – PPGECON/UFPE-CAA**

**Klebson Humberto de Lucena Moura – PIMES-UFPE**

**Resumo**

A pesquisa obteve evidências da influência de fatores locais, potenciais geradores de economias de localização e urbanização, na escolha locacional das indústrias da transformação das principais cidades brasileiras. A análise baseou-se na estimação de modelos de escolha discreta, a partir dos microdados da RAIS-MTE, que compreendeu um total de 60.005 estabelecimentos formais. Os principais resultados indicam que as indústrias tendem a se localizar nas cidades com uma maior concentração de indústrias da mesma divisão de atividades e naquelas que tem uma maior diversidade produtiva, este último suportado pelo modelo *mixed logit.* Também há indicações de que o custo de vida e o preço do aluguel atuam como fatores de repulsão da atividade produtiva. Finalmente, estabelecimentos maiores tendem a se aglomerar nas grandes cidades, especialmente em São Paulo.

**Palavras-chave:** Escolha Locacional; Economias de Aglomeração; Logit Condicional

**Spatial Distribution of Industries in Brazil: An Analysis from Discrete Choice Models**

**Abstract**

This research obtained evidence of the influence of local factors in the locational choice of firms. The analysis used discrete choice models, estimated from microdata RAIS-mte, totaling 60,005 formal establishments. The results show the existence of productive externalities generated by the concentration of industries in the same division, and productive diversity of cities; the cost of rent acts as a repulsive factor, but only the Mixed Logit model suggested an attraction for towns that have lower cost of labor and cost of living. Finally, larger establishments tend to cluster in large cities, especially in São Paulo.

.**Keywords:** Location Choice;Agglomeration Economics;Conditional Logit

**JEL: L6 ; R1**

(Área 10 - Economia Regional e Urbana)

**Distribuição Espacial das Indústrias no Brasil: Uma Análise a partir de Modelos de Escolha Discreta**

1. **Introdução**

As evidências empíricas para o Brasil (Silveira Neto, 2005; Feijo et. al., 2003; Diniz e Crocco, 1996; Benolli, 1980) e para outros países (Ellison e Glaeser, 1999; Ellison, 2007) apontam que a atividade produtiva tende a se concentrar no espaço. No Brasil, apenas 1,8% dos municípios (100 municípios) concentrou 50% dos empregos formais gerados pela indústria da transformação no ano de 2010[[1]](#footnote-1), e 40 municípios reponderam por 46% do Produto Nacional Bruto[[2]](#footnote-2) e por 30% da população. Nos Estados Unidos, em torno de 75% da população vive nas cidades que ocupam apenas 2% do território de todo o país (Rosenthal e Strange, 2006). Nesses países, também há indicações de que tanto firmas especializadas em uma mesma atividade produtiva, como aquelas que fazem parte de uma mesma cadeia produtiva, tendem a se aglomerar espacialmente (Lautert e Araujo, 1994; Resende e Wyllie, 2005; Resende, 2012; Ellison e Glaeser, 1999; Ellison e Glaeser, 1994).

Estas evidências sugerem, portanto, que mesmo com a existência de forças operando contra a aglomeração da atividade produtiva, como as associadas a uma maior concorrência ou a custos de produção mais elevados, parece que as vantagens produtivas obtidas pela maior parte das firmas espacialmente concentradas superam as primeiras, caso contrário, não haveria explicação para a concentração geográfica da atividade produtiva. Como consequência, políticas voltadas para a consolidação de complexos indústrias em países em desenvolvimento, como o Brasil por exemplo, aproveitando essas economias de aglomeração têm sido utilizadas como estratégias para o crescimento econômico regional no país[[3]](#footnote-3).

As economias de aglomeração decorrentes das externalidades produtivas geradas pelas firmas espacialmente concentradas são bem definidas e exploradas pela literatura aqui citada. Marshall (1920), pioneiramente, as definiu como sendo aqueles ganhos produtivos decorrente de fatores externos as firmas, os quais são gerados pelas firmas espacialmente concentradas, possibilitando um melhor acesso aos insumos, à mão de obra e serviços especializados, e aos *spillovers* tecnológicos. Porém, no campo empírico, esforços têm sido realizados no sentido de entender como esses fatores locais influenciam a distribuição da atividade produtiva, conforme estudos citados a seguir, o que constitui um passo importante para formulação de políticas públicas que visam à consolidação de polos industriais.

Para o Brasil, os estudos empíricos com esse escopo, ou são realizados para uma indústria especifica[[4]](#footnote-4), ou utilizam dados agregados por indústrias, cidades ou estados. Explorando esses últimos, destacam-se o trabalho de Silveira Neto (2005), o qual obtém indicações – a partir da estimação de modelos de regressão considerando como variável dependente um índice de concentração e como variáveis explicativas indicadores das economias de aglomeração – para as indústrias da transformação das grandes regiões brasileiras, de que os ganhos de escala na produção favorecem a concentração industrial, mas a integração vertical entre as indústrias atua como fator de repulsão da indústria. Com o objetivo de pesquisa semelhante, Silva et. al. (2005) tentam identificar a importância de fatores locais, especialmente os que geram economias de localização e urbanização, para explicar a distribuição da atividade produtiva a partir da análise da influência dessas variáveis no emprego industrial dos estados do país. Os autores obtêm evidências de que tanto as conexões de mercado para frente e para trás, quanto à diversidade industrial da região, influenciam positivamente o crescimento do emprego industrial nos estados.

Em relação a esses dois trabalhos, Rocha et. al. (2011) avançam ao expandir o estudo para cada uma das principais indústrias de transformação do Brasil e investigam qual dos fatores de aglomeração (externalidades marshalianas) – o acesso ao mercado, o acesso a matérias-primas, e a existências dos spillovers tecnológicos – é mais importante para explicar a concentração geográfica dessas indústrias. Os principais resultados da pesquisa sugerem que o modelo de acesso ao mercado é o mais indicado para explicar a localização espacial da maioria das indústrias considerada na análise. Essa evidência aponta que as firmas integradas a essas indústrias são atraídas para se estabelecerem em regiões onde concentram um maior mercado consumidor, as quais devem se beneficiar dos *linkages* de mercado (para trás e para frente) pelas firmas espacialmente concentradas.

Porém, destaca-se que esses estudos realizados para o Brasil utilizam, essencialmente, a variável “emprego” ou “preço do aluguel” como meio para identificar os fatores de aglomeração das indústrias brasileiras, além de utilizarem dados agregados geograficamente.

Por outro lado, avanços têm sido observados na literatura empírica internacional acerca desse tema, especialmente devido à existência e a uma maior facilidade ao acesso aos dados ao nível das firmas. Boa parte desses estudos utiliza como base modelos de escolha discreta (Arauzo e Manjon, 2004; Basile et. al., 2003; Baudewyns et. al., 2000; Cheng e Stough, 2006), mas se diferenciam quanto ao nível de desagregação setorial adotado e aos fatores locais considerados na análise. Cabe salientar que quando é possível identificar a localização das firmas para um ano ou mais, como também as suas características, torna-se viável a estimação de modelos probabilísticos – entre eles o *logit condicional* (Coughlin et al., 1991; Figueiredo et al., 2002; Friedman et al., 1992; Head et al., 1999), o *logit multinomial* (Araujo e Manjon, 2004; Baudewyns, 1999) ou o *nested logit* (Basele et. al., 2003; Hansen 1987) – a partir do qual pode-se investigar quais fatores locacionais influenciam positivamente as escolhas locacionais das firmas.

Com esse propósito de pesquisa e utilizando dados ao nível da firma, destaca-se um único trabalho já realiado para o Brasil, o de Hansen (1987). A partir de uma amostra para 360 indústrias do estado de São Paulo, o autor obtém evidências de que as economias de aglomeração influenciam as decisões locacionais das firmas, e a proximidade da grande São Paulo é um fator de atração. Porém, contrariando as expectativas, em alguns dos modelos estimados, firmas de setores tradicionais e intermediários são atraídas a se instalarem nas cidades que possuem maiores salários médios para essas indústrias.

Buscando contribuir para a literatura nacional, a presente pesquisa, a partir de uma base de dados ao nível das firmas para o ano de 2010 da Relação Anual de Informações Sociais do Ministério do Trabalho (RAIS-Mte) compreendendo 60.005 estabelecimentos industriais formais, objetiva identificar quais fatores locais, em especial, aqueles que podem gerar economias de localização e urbanização identificados na literatura, atuam como fatores de atração das indústrias. Além dessas variáveis, também são considerados indicadores de custos de produção, salário e custo do aluguel, o índice de custo de vida e atributos individuais das firmas que variam no tempo. Para tanto, será estimado modelos de escolha discreta, o *logit* condicional e o *mixed* *logit*, com dados ao nível da firma, para as principais cidades brasileiras considerando as principais indústrias da transformação (2-dígitos) segundo a CNAE do país[[5]](#footnote-5). Ressalta-se que esta estimação apenas é possível porque a base de dados utilizada, os microdados da RAIS do Ministério do Trabalho, disponibiliza os dados por firma e é possível identificar cada uma delas, como também a sua localização, através de um código único e imutável, o CNPJ.

1. **Evidências da Distribuição Espacial das Indústrias da Transformação no Brasil**

Antes de averiguar a importância dos fatores de aglomeração e localização para a escolha locacional das firmas, é importante avaliar a distribuição dos estabelecimentos industriais no Brasil. Na literatura nacional há vários estudos empíricos sobre a concentração da atividade industrial no país, os quais utilizando como variável de análise o emprego, número de estabelecimentos ou o valor transformado, entre outras variáveis (Silveira Neto, 2005; Domingues 2005; Lautert e Araújo, 2007). Resende e Wyllie (2005) utilizando os dados da RAIS entre 1995 e 2001, por exemplo, calculam índices de concentração através de várias medidas, entre elas, o índice de Ellison e Glaser, e observam que, durante este período de análise, houve mudanças na ordenação das indústrias com mais elevados índices de concentração, mais que há uma tendência geral do país para concentração da atividade produtiva[[6]](#footnote-6).

Com o intuito de apresentar as indústrias de transformação com maior representatividade no país, o Gráfico 1 apresenta o número de estabelecimentos para cada divisão de atividades (2 dígitos da CNAE). De acordo com o Gráfico 1 pode-se perceber que as divisões com maior participação no total de estabelecimentos industriais são: as divisões D15 (Fabricação de Produtos Alimentícios e Bebidas), D18 (Confecção de Artigos do Vestuário e Acessórios), D22 (Edição, Impressão e Reprodução de Gravações), D28 (Fabricação de Produtos de Metal Exceto Máquinas e Equipamentos) e finalmente D36 (Fabricação de Móveis e Indústrias Diversas).

**Gráfico 1. Número de Estabelecimentos da Indústria no País por divisão CNAE (2010).**

Fonte: Elaboração própria a partir dos dados da Rais-Mte (2010)

O referido gráfico indica que as três indústrias no Brasil com maior número de estabelecimentos, divisão 15, 18 e 28, são responsáveis por 41,25% do total de estabelecimentos do país. Porém, convém salientar que essa representatividade em termos de números de estabelecimentos não implica de forma idêntica na importância em termos de geração de emprego, já que pode haver uma grande participação de micro e pequenas empresas nessas atividades.

Como o foco da análise é explicar a decisão de localização dos investidores/indústrias entre as principais capitais brasileiras para quais foram possíveis obter um conjunto de informações dos atributos locais utilizados nos modelos empíricos, as informações a seguir são apresentadas para essas cidades, a saber: Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba, Porto Alegre, Goiânia e Brasília. Com respeito à distribuição das indústrias entre capitais do país, os dados da RAIS-Mte para o ano de 2010 apontam que 22%[[7]](#footnote-7) dos estabelecimentos industriais se localizam nessas capitais, percentual elevado do ponto de vista das opções de escolha existentes no país, dado que essas capitais representam 0,5% das opções de localização[[8]](#footnote-8).

O Gráfico 2 apresenta a mesma informação do gráfico 1, o número de estabelecimentos formais por divisão de atividades segunda CNAE-95, mas apenas para as cidades citadas consideradas na análise. Conforme pode ser visualizado no gráfico, as indústrias de Confecção (Divisão 18), de Edição, Impressão e Reprodução de Gravações (Divisão 22), de Alimentos e Bebidas (Divisão 15) e de Fabricação de Produtos de Metal (Divisão 28) são as mais representativas nessas cidades, as quais respondem por mais de 50% dos estabelecimentos da indústria da transformação nelas instaladas.

**Gráfico 2. Número de Estabelecimentos nas 11 Capitais por Divisão CNAE-95**

Fonte: Elaboração própria a partir dos dados da Rais-Mte (2010)

Outra informação, a partir da qual se pode avaliar a distribuição da atividade produtiva, pode ser obtida com base no número de empregos gerados pela indústria, já que para algumas indústrias não intensivas em mão de obra, uma elevada representatividade em termos de estabelecimentos não significa que sejam as mais importantes para a geração de emprego.

Desse modo, analisando a distribuição dos empregos em cada uma das divisões da indústria da transformação, verifica-se que, em relação a geração de empregos, a indústria de confecção (CNAE 18) também é a que se encontra mais concentrada nas 11 capitais, seguida das indústrias de Alimentos e Bebidas (CNAE 15) e de Edição, Impressão e Reprodução de Gravações (Divisão 22).

**Gráfico 3. Número de Vínculos da Indústria de Transformação nas 11 Capitais por divisão CNAE-95.**

Fonte: Elaboração própria a partir dos dados da Rais-Mte (2010)

Quanto ao tamanho dos estabelecimentos[[9]](#footnote-9), característica influente de forma diferenciada nas escolhas locacionais dos investidores, a maioria das firmas localizadas nas capitais em análise é micro (82%), e pequenas empresas (15%), sendo apenas 3% médias empresas e menos de 1% Grandes empresas.

A respeito das características das cidades consideradas na análise, observa-se com base na Tabela 1, que as cidades com maior população, São Paulo e Rio de Janeiro, também são as que possuem os maiores custos de alugueis.

**Tabela 1. Características Econômicas, Sociais e Geográficas das 11 Capitais – 2010**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Município | População  Residente (h) | Densidade  Demog. (h/km2) | Alugueis  (R$) | Rendimento  Domiciliar(R$) | Desemprego\* | Custo de Vida\*\* | Índice de Herfindahl | |
| Belém | 1.393.399 | 1315,26 | 437,10 | 2951,16 | 0,104 | 3262,04 | 0,188 |
| Fortaleza | 2.452.185 | 7786,44 | 284,00 | 2771,98 | 0,078 | 2838,14 | 0,248 |
| Recife | 1.537.704 | 7039,64 | 382,31 | 3487,90 | 0,127 | 3104,61 | 0,133 |
| Salvador | 2.675.656 | 3859,44 | 384,99 | 2843,45 | 0,131 | 3075,01 | 0,099 |
| Belo Horizonte | 2.375.151 | 7167,00 | 514,21 | 4480,00 | 0,066 | 3344,53 | 0,079 |
| Rio de Janeiro | 6.320.446 | 5265,82 | 599,00 | 4131,57 | 0,074 | 3334,96 | 0,080 |
| São Paulo | 11.253.503 | 7398,26 | 594,48 | 4388,05 | 0,042 | 3294,47 | 0,082 |
| Curitiba | 1.751.907 | 4027,04 | 551,05 | 4594,80 | 0,049 | 3065,59 | 0,085 |
| Porto Alegre | 1.409.351 | 2837,53 | 571,51 | 4742,65 | 0,055 | 3039,92 | 0,083 |
| Goiânia | 1.302.001 | 1776,74 | 406,80 | 4035,64 | 0,054 | 3297,72 | 0,129 |
| Brasília | 2.570.160 | 444,66 | 583,26 | 5397,10 | 0,082 | 3193,19 | 0,154 |

Fonte: Rais-Mte -2010; (\*) Percentual de pessoas desocupadas na capital, calculado através dos microdados do Censo 2010; (\*\*) O custo de vida foi calculado como uma média do custo de vida mensal disponível no IPEADATA.

Também cabe destacar que as cidades com mais elevada renda domiciliar estão localizadas no sul do país, a exceção de Brasília. E, as cidades com mais elevadas taxas de desemprego, Belém, Recife e Salvador, também estão entre as cidades com mais baixa renda domiciliar. Além disso, destaca-se que as cidades no norte, nordeste e centro-oeste são as que apresentaram menor diversidade produtiva, medida pelo índice de Herfindahl.

Com relação às diferenças entre a estrutura produtiva das cidades, fazendo a análise com base na variável emprego, em seis (Belém, Recife, Salvador, Belo Horizonte, Rio de Janeiro e Brasília) das 11 cidades consideradas na análise, a indústria de Alimentos e Bebidas (DIVISÃO 15) é a mais importante em termos relativos (Tabela A2). Nas cidades de Fortaleza, onde há um polo industrial de confecção, de São Paulo e Goiânia, tem-se a Indústria de Confecção (DIVISÃO 18) com a maior participação relativa no emprego industrial local. E, nas cidades de Porto Alegre e Curitiba, as que mais se distanciam da média para todas as cidades, é a indústria de Edição, Impressão e Reprodução de Gravações (DIVISÃO 22) e a Indústria de Maquinas e Equipamentos (DIVISÃO 29), respectivamente, com maiores participação no emprego local.

Cabe destacar ainda que, apenas para a cidade do Recife, a indústria com maior participação no emprego industrial local (Indústria de Alimentos e Bebidas), também é a mais importante, em termos relativos, quando comparado para o conjunto das cidades. (ver tabela A3)

Com base nessa análise descritiva, a partir da qual se pôde observar como as cidades consideradas na análise se diferenciam quanto ao potencial econômico e, também, na sua estrutura produtiva, cabe investigar a importância dos potenciais fatores locais geradores de economias de aglomeração, para a escolha locacional dos investidores/indústrias.

1. **Modelos de Escolha Discreta: Escolha Locacional das Firmas**

O modelo de escolha locacional dos tomadores de decisão/investidores/firmas baseia-se na hipótese que as firmas escolhem a cidade j para se instalar de forma a maximizar o lucro, dados os atributos locais das cidades e os preços dos fatores de produção, capital e trabalho. Assumindo a forma funcional Cobb-Douglas para a função de produção e que tal função exibe as propriedades de uma função neoclássica, a função lucro pode ser descrita como:

**(1)**



inclui os fatores locais da cidade j que influenciam exogenamente a produtividade da firma; é custo fator de produção, trabalho (L), em j; eé o custo de produção do capital (K) em j.

Apenas para simplificação, sem alterações nas principais conclusões que se quer obter, assume-se, assim como Hansen (1986), que o fator de produção “capital”, embora necessário para a produção, é disponível sem custos para as firmas, de modo que a equação 1 pode ser reescrita na forma:

 **(2)**

Resolvendo o problema de maximização do lucro da firma obtém-se a quantidade ótima do fator de produção “trabalho”:

 **(3)**

Substituindo na função lucro tem-se o lucro ótimo na forma:

 **(4)**

Da equação (4), observa-se que o lucro ótimo correlaciona-se positivamente com , que inclui os fatores locais que influenciam o lucro das firmas, e negativamente com o preço do fator de produção, trabalho.

O modelo acima pode ser transformado em um problema de escolha discreta (Hansen, 1986), adaptando-o ao modelo de maximização da utilidade aleatória de McFadden (1974) conhecido como *logit* condicional. Dados os k atributos locais, as firmas/investidores (i) escolhem a localidade j para se instalar entre as J opções de escolha de forma a maximizar os lucros  que assume a seguinte forma linear:

, e  **(5)**

 é um vetor 1 x M das características locais, é o vetor dos parâmetros, é o termo de erro aleatório.

A maximização da função lucro pela firma “i” sujeita a escolha locacional implica que: se *j’* é preferido à alternativa disponível *j’’*, significa que o lucro obtido pela firma em *j’* ()*,* é maior do que o que o investidor teria em *j’’* () de modo que a probabilidade do investidor *i* escolher a localidade *j’* para se instalar é dada por:

 onde  **(6)**

os resíduos têm distribuição do tipo “valor extremo” 

McFadden (1974) mostra que se os erros da função  são identicamente e independentemente distribuídos e tem uma função de distribuição do tipo “*valor-extremo*”, a probabilidade da região *j’* ser escolhida é:

 **(7)**

 representa a escolha da firma *“i”* e *J* é o número de escolhas possíveis.

A maximização da probabilidade do investidor *i* escolher a localidade *j’* é obtida maximizando a função log-verossimilhança:

 **(8)**

o indicador é igual a um se a firma *i* escolhe a localidade *j’*; caso contrário,  é igual a zero.

O modelo descrito acima, o *logit* condicional, impõe que a relação da probabilidade de escolha entre duas alternativas quaisquer não depende das demais alternativas disponíveis. Este resultado é conhecido na literatura econômica como propriedade da Independência das Alternativas Irrelevantes (IIA)[[10]](#footnote-10) (Wooldridge, 2002; Greene, 2003) a qual assume que a razão entre a probabilidade do indivíduo *“i”* escolher a localidade *j’* e a probabilidade de escolha da localidade *j’’* não depende do conjunto de escolha das demais localidades. Porém, se a hipótese da Independência das Alternativas Irrelevantes (IIA) for rejeitada, outros modelos que relaxam essa hipótese podem ser estimados com o mesmo propósito da pesquisa.

Neste caso, uma situação possível é quando os coeficientes do modelo (1) variam sobre a população, compatível com o fato dos fatores locais (variáveis explicativas) influenciarem de forma diferenciada as escolhas locacionais dos investidores. Neste caso, a condição derivada da hipótese de coeficiente fixos do modelo *logit* condicional da equação (5) não será válida e a hipótese da Independência das Alternativas Irrelevantes será rejeitada. Assim, o termo de erro estocástico será correlacionado sobre as alternativas. Ou seja, a parte estocástica da função lucro passa a ser , com o vetor “” contém as variáveis explicativas do modelo e o vetor dos coeficientes “” mensurando a parte não observada das preferências dos investidores, a qual varia sobre a população (Revelt e Train, 1998, Revelt e Train, 1999).

Considerando esse caso, o modelo a ser estimado pode ser o *mixed logit*, e a equação (1) deve ser reescrita:

, com  **(9)**

Ou,

 **(10)**

 é a média da distribuição das preferências da população;  é uma medida de variação das preferências individuais; e  é vetor das variáveis explicativas do modelo.

Para o caso da presente pesquisa, a probabilidade de cada uma das onze capitais ser escolhida, dado as demais alternativas de escolha disponíveis, é igual ao produto da probabilidade do *logit* condicional avaliada para cada alternativa de escolha, na forma:

 **(11)**

 é a probabilidade do investidor observar todas as onze alternativas disponíveis para alocar o seu empreendimento.

Contudo, como não se tem a informação das preferências dos investidores/estabelecimentos pelos fatores locais (variáveis explicativas), o vetor “” para cada estabelecimento é desconhecido, sendo necessário integrar a probabilidade do *logit* condicional sobre todos os possíveis valores de “” e então obtem-se a probabilidade de escolha do *mixed logit*.

Assim, a probabilidade condicional integrada sobre todos os possíveis valores de  fornece o valor esperado da probabilidade de escolha do *logit* condicional, que pode ser representada como:

 **(12)**

 é a função densidade. A função densidade pode tomar diferentes distribuições a depender das características das variáveis em estudos (Train, 2003).

Como pode ser observado pela expressão acima, agora é preciso integrar a probabilidade dos investidores observarem todas as alternativas de escolha e, portanto, os denominadores da probabilidade de escolha entre duas alternativas quaisquer não podem ser mais cancelados como no *logit* condicional; não impondo assim a propriedade das Alternativas Irrelevantes. (Train, 2003)

Logo, a função *log-verossimilhança* derivada da probabilidade não condicional é:

 **(13)**

Como a probabilidade de escolha do *mixed logit* () não pode ser resolvida analiticamente e, portanto, a função log-verossimilhança não pode ser maximizada, é utilizado o método de simulação descrito em Revelt e Train (1997) e em Train (2003) para obter a função log-verossimilhança simulada, .

Desta forma, a função de probabilidade simulada, , é aproximada em três etapas. Primeiro, para um dado valor dos parâmetros de *,* no caso da distribuição normal,a média () e a covariância (), o valor de  é extraído desta distribuição. Utilizando este , é calculado o produto da probabilidade do modelo *logit* padrão *(**)*. As etapas 1 e 2 são repetidas várias vezes e a média dos resultados é utilizada como uma aproximação de :

 **(14)**

N é o número de repetições;  é o enésimo valor de  construído com base na função de densidade ;  é o estimador de  e representa a probabilidade simulada da seqüência de escolhas dos indivíduos. [[11]](#footnote-11)

1. **Modelo Empírico e Dados**

A identificação dos fatores locais que podem influenciar as decisões locacionais dos investidores é investigada a partir da estimação dos modelos probabilísticos descritos na seção anterior utilizando dados ao nível da firma. A ideia é examinar como esses atributos locais, identificados com base na literatura, influenciam essas escolhas. Ressalta-se que esta estimação apenas é possível devido a base de dados utilizada, os microdados da RAIS do Ministério do Trabalho, disponibiliza os dados por firma e é possível identificar cada uma delas, como também a sua localização, através de um código único e imutável, o CNPJ.

O modelo a ser estimado segue em sua especificação geral:

  **(15)**

A variável dependente, , assume valor igual a “1” se a firma “i” está localizada na cidade “j”; o vetor  inclui as características controles das firmas, tais como, tamanho e nível médio da escolaridade dos trabalhadores. No vetor  são inclusos as características das cidades que podem explicar a localização das firmas, como: i. o potencial de mercado da cidade/acesso ao mercado que tem como *proxies* a renda média e o salário médio da cidade; ii. indicadores de custos de produção e preço, a média do aluguel residência da cidade, o índice de custo de vida, e o salário médio da indústria; iii. economias de urbanização/localização, participação da indústria no emprego total, o quociente locacional, e Hirshman-Herfindal. E, , é o termo de erro do modelo.

A base de dados a partir da qual foram obtidas as informações das indústrias é proveniente do Relatório Anual de Informações Sociais (RAIS)[[12]](#footnote-12) do Ministério do Trabalho e do Emprego, mais especificamente, os microdados da base que contém informações de todos os estabelecimentos e trabalhadores formais do Brasil.

Este banco de dados deriva da RAIS, um registro administrativo anual que permite identificar os estabelecimentos e trabalhadores formalmente inseridos no mercado de trabalho que recebem benefícios sociais, a fim de monitorá-los. O objetivo desta base de dados é o acompanhamento geográfico, setorial e ocupacional da trajetória dos trabalhadores ao longo do tempo e, também, registrar informações dos estabelecimentos formais. A cobertura da base abrange todos os trabalhadores formais do país, correspondendo a uma espécie de censo do mercado de trabalho formal brasileiro. A base está organizada por trabalhador e cada registro armazena a situação do indivíduo ao final de cada ano da trajetória, através das características do empregador, do trabalhador e do vínculo. Além disso, como na base é possível identificar o estabelecimento que os trabalhadores estão empregados, através do CNPJ, também é possível obter uma base com características de todos os estabelecimentos formais do país.

A análise é realizada para o ano de 2010, considerando as principais capitais brasileiras – Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba, Porto Alegre, Goiana e Brasília –, para as quais foi possível obter informações sobre o índice de custo de vida, variável de extrema importância para a análise. O recorte setorial considerado foi para a indústria da transformação, para as divisões de atividade a 2-dígitos de acordo com a Classificação Nacional de Atividades Econômicas (CNAE).

Quanto as varáveis explicativas incorporadas na análise, tiveram como base o Censo Demográfico do ano de 2010, além da base de dados da própria RAIS-Mte. O Quadro 1 descreve cada variável com sua respectiva fonte e sinal esperado com base na literatura.

**Quadro 1. Descrição das Variáveis Explicativas Consideradas no Modelo Empírico**

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis Explicativas** | **Descrição** | **Sinal Esperado** | **Fonte** |
| Salário médio da Manufatura | Somatório do salário de todos os trabalhadores formais dividido pelo numero de trabalhadores formais de cada indústria que o estabelecimento faz parte | (-) | RAIS |
| Média do Aluguel Residencial | Somatório dos alugueis residências dividido pelo total de imóveis alugados nas principais capitais | (-) | Censo 2010 |
| Salário médio da Cidade | Somatório do salário de todos os trabalhadores formais dividido pelo numero de trabalhadores formais | (+) | RAIS |
| Renda Média das Famílias | Somatório da renda de todas as famílias dividido pelo numero de famílias nas principais capitais | (+) | Censo 2010 |
| Índice de Custo de Vida (INPC-IBGE) | Índice Nacional de Preço ao Consumidor nas principais capitais | (+) | IPEADATA 2010 |
| Desemprego | Percentual de pessoas desocupadas | (+) ou (-) | Censo 2010 |
| Quociente Locacional | Participação do emprego formal do município j na indústria i ( dividido pela participação em todas as indústrias ( | (+) | RAIS 2010 |
| Índice de *Herfindahl* | Somatório da participação no emprego de cada divisão de atividades da indústria da transformação elevado ao quadrado | (-) | RAIS 2010 |
| Participação no Emprego (PART) | É a parcela do emprego de cada indústria em cada Capital | (+) | RAIS 2010 |
| TamanhoXEscolha Locacional | Número de Empregos por Firma | (-) | RAIS 2010 |
| Escolaridade(Anos)XEscolha  Locacional | Nível de Instrução Médio dos trabalhadores da firma | (-) | RAIS 2010 |

Fonte: Elaboração Própria

A escolha das variáveis explicativas apresentadas no Quadro 1 baseou-se na literatura citada ao longo do artigo e os sinais dos coeficientes especificados segue a teoria e os resultados dos trabalhos empíricos comentados a seguir.

As variáveis *proxies* de custo de produção, o salário e custo do aluguel, considerados nos modelos neoclássicos de maximização de lucros, devem atuar como fator de repulsão da atividade produtiva, de modo que os investidores prefiram as cidades que apresentam os menores custos de produção, considerando demais fatores constantes. Nesse sentido, boa parte das evidências empíricas para outros países, de fato, obtém indicações que corroboram com a teoria de que a variável salário é um fator de dispersão da atividade produtiva (Coughlin et al., 1991; Crozet et al., 2004; Friedman et al., 1992; Head et al., 1999). Porém, para o Brasil, Hansen (1987) obtém indicações para as indústrias tradicionais e de bens intermediários de que tal fator atua atraindo as firmas do Estado de São Paulo para as cidades com os maiores salários. Já a variável custo aluguel, apesar de menos explorada nos estudos empíricos, indicam que essa variável é um fator de repulsão das indústrias (Figueiredo *et al.* 2002 e Hansen, 1987).

Espera-se que os indicadores do potencial do mercado, o salário médio e a renda média das famílias da cidade, apresentem sinais positivos indicando que as firmas são atraídas a se instalar nas cidades onde a população tem um maior poder aquisitivo. Coughlin et al. (1991) encontram evidências de que as indústrias são atraídas para as cidades Norte-Americanas que apresentam uma maior renda per capita.

Se o índice de custo de vida estiver correlacionado com o preço dos produtos indústrias, já que tal informação pode ser utilizada para o cálculo do índice, espera-se que tal variável apresente o sinal positivo, indicando que as cidades com maior custo de vida também são aquelas que as firmas podem cobrar preços mais elevados pelos seus produtos. Porém, vale salientar que esta variável por estar relacionada com outros fatores de aglomeração pode captar a influência dos mesmos nas escolhas locacionais das firmas.

A variável desemprego, de acordo com a teoria, seria um fator de atração da atividade produtiva já que cidades com taxas de desemprego mais elevadas também seriam aquelas onde os empresários poderiam pagar menores salários. Os estudos de Basile et al. (2003) para os países da União Europeia e o de Coughlin et al. (1991) e Friedman et al. (1992) para os Estados Unidos corroboram esta hipótese. Porém, especialmente nos países em desenvolvimento, como o Brasil, é possível que tal variável também esteja relacionada com uma mão de obra com pouca qualificação, de baixa produtividade. Neste último caso, a variável deve apresentar o sinal negativo.

As variáveis locais, indicadoras das economias de localização e urbanização, bem exploradas pela literatura aqui citada, devem atuar como fatores de aglomeração da atividade produtiva, são eles: o índice que busca medir a diversidade do emprego na capital *j*, índice de *Hirshman-Herfindal*, o quociente locacional e a participação da indústria no total do emprego industrial da cidade.

Os controles para as características do estabelecimento, o tamanho (pessoal ocupado) e o nível de escolaridade médio dos trabalhadores, foram incorporados no modelo interagindo-os com o município onde o estabelecimento está instalado, de modo a ser possível a estimação dos modelos de escolha discreta que exigem variações nas informações nas j cidades. Porém, tais variáveis têm importantes interpretações. Araujo e Manjon (2004) obtêm indicações de que o tamanho do estabelecimento influenciam na escolha locacional das firmas das cidades da Espanha, e os estabelecimentos maiores tendem a localizar-se em Barcelona, entre as opções de escolha consideradas no estudo. Se as interações com o tamanho do estabelecimento, tendo como categoria omitida a cidade de São Paulo, apresentarem o sinal negativo, haverá indícios de que as firmas maiores são atraídas a se instalarem na cidade de maior importância financeira para o país, São Paulo. Da mesma forma, se os coeficientes das interações com a escolaridade média dos trabalhadores das firmas forem positivos pode-se concluir que os estabelecimentos mais intensivos em capital humano são os mais propensos a se instalarem na cidade de São Paulo.

A Tabela 2 informa sobre as estatísticas descritivas das variáveis definidas no Quadro 1.

**Tabela 2. Estatísticas Descritiva dos Dados**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variáveis Explicativas** | **Média** | **Desvio Padrão** | **Maximo** | **Mínimo** |
| Salário médio da Cidade | 1976.663 | 439.7652 | 2787.74  (Rio de Janeiro) | 1340.58  (Fortaleza) |
| Salário médio da Manufatura | 1959.406 | 605.4858 | 3042.95  Rio de Janeiro | 998.02  (Fortaleza) |
| Desemprego | .0782733 | .0291115 | 0.1312  (Salvador) | 0.04175  (São Paulo) |
| Média do Aluguel Residencial | 482.6101 | 102.969 | 598.999  (Rio de Janeiro) | 284  (Fortaleza) |
| Índice de Custo de Vida | 3168.198 | 150.6663 | 3344.53  Belo Horizonte | 2838.14  (Fortaleza) |
| Renda Média das Famílias | 3984.028 | 823.6241 | 5397.099  (Brasília) | 2771.984  (Fortaleza) |
| Quociente Locacional | .9753627 | .8294018 | 17.55016 | 0(\*) |
| Índice de *Herfindahl* | .1240459 | .0522664 | .248477  (Fortaleza) | .079033  (Belo Horizonte) |
| Participação no Emprego | .0843581 | .0869485 | .454543 | 0(\*) |
| Tamanho | 1.997317 | 1.413511 | 0 | 9 |
| Escolaridade | 6.107593 | 1.137715 | 11 | 1 |

Nota: Elaboração própria com base nos dados; (\*) existem divisões que não apresentam firmas em determinadas capitais de forma que o quociente locacional e a participação no emprego são nulas.

Os valores máximo e mínimo se referem diretamente a cidade com o maior e menor valor da variável considerada respectivamente. As exceções são o quociente locacional e a participação no emprego, que são observados por setor e cidade.

1. **Análise dos Resultados**

Os resultados do Modelo *Logit* *Condicional* encontram-se na Tabela 3, na primeira coluna estão os resultados das estimações considerando as variáveis de custo de produção (Salário Manufatura e Aluguel), do potencial do mercado (Salário Cidade e Renda das Famílias), desemprego e custo de vida. Um segundo modelo foi estimado com a inclusão ao modelo 1 dos potenciais fatores de aglomeração (Participação no Emprego, Índice de Herfindahl, Quociente Locacional). No terceiro modelo adicionaram-se variáveis de interações, entre a escolha locacional e o tamanho da firma e entre a média de escolaridade dos trabalhadores da firma.

A tabela informa os coeficientes dos modelos estimados a partir da equação 17, a qual tem como objetivo identificar quais fatores locais influenciam nas escolhas locacionais das firmas consideradas na amostra, assim como se atuam como fator de atração ou repulsão. Deste modo, um coeficiente positivo indica que o atributo local atua como fator de aglomeração das firmas, caso contrário, seria uma força de dispersão da atividade industrial. Quanto ao ajuste do modelo, a maioria das variáveis apresentam coeficientes significantes a menos de 5%, e os coeficientes mostram-se conjuntamente significantes. Além disso, a maioria dos coeficientes estimados apresentou o sinal conforme esperado, independentemente da especificação do modelo.

As variáveis que buscam mensurar os custos de produção das firmas, salário médio da indústria na qual a firma pertence e o preço do aluguel apresentaram, em todos os 3 modelos estimados, coeficientes com o mesmo sinal e estatisticamente significantes a menos de 5%. Contrariando a teoria, os resultados indicam que as firmas da amostra considerada na análise são atraídas a se estabelecerem naquelas cidades que apresentam uma maior média salarial para a indústria a qual a firma faz parte (mesma divisão de atividades). Resultados semelhantes foram encontrados por Head et al. (1999) e Smith e Florida (1994), e são consistentes com a hipótese de que tal variável esteja captando o nível de produtividade dos trabalhadores. Ou seja, as firmas só estarão dispostas a pagar maiores salários, gerando um maior custo de produção, se os ganhos de produtividade oriundos em localidades com trabalhadores mais produtivos superarem o aumento no custo de produção. O sinal negativo do custo do aluguel, conforme esperado, indica que este é um fator de repulsão da atividade produtivo.

Já as variáveis “média salarial da cidade” e a “renda média” apresentaram os sinais de acordo com o esperado, a exceção do sinal negativo da média salarial do terceiro modelo, sugerindo que as cidades que possuem um maior potencial de mercado, maior capacidade de gastos financeiros das famílias, atraem investimentos industriais.

Pelo menos para as cidades consideradas no estudo, os resultados encontrados parecem corroborar com o fato de que cidades com uma maior taxa de desemprego são preteridas pelas firmas. Resultados semelhantes foram encontrados por Cieslik (2005) para as firmas das províncias de Polônia, por Egelnet al. (2004) para as firmas intensivas em conhecimento das regiões da Germânia, e por Woodwar (1992) para as indústrias Japonesas instaladas nos Estados Unidos. Neste caso, se a taxa de desemprego está relacionada com um menor aquecimento da economia local, assim como, com uma mão de obra menos qualificada, assim esta variável deve atuar como um fator de repulsão das firmas.

**Tabela 3. Resultados do Modelo *Logit*, todas as Indústrias**

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis Explicativas** | Modelo 1 | Modelo 2 | Modelo 3 |
| Salário Manufatura | 0,0004\* | 0,0018\* | 0,0033\* |
|  | (0,0000) | (0,0001) | (0,0005) |
| Aluguel | -0,0165\* | -0,0145\* | -0,0151\* |
|  | (0,0002) | (0,0002) | (0,0013) |
| Salário Cidade | 0,0024\* | 0,0006\* | -0,0015\* |
|  | (0,0001) | (0,0001) | (0,0006) |
| Renda das Famílias | 0,0008\* | 0,0010\* | 0,0016\* |
|  | (0,0000) | (0,0000) | (0,0001) |
| Desemprego | -31,069\* | -30,796\* | -33,246\* |
|  | (0,2402) | (0,2578) | (1,3861) |
| Custo de Vida | 0,0000 | 0,0004\* | 0,0047\* |
|  | (0,0000) | (0,0000) | (0,0003) |
| Participação no Emprego |  | 3,6760\* | 3,7975\* |
|  |  | (0,0842) | (0,0844) |
| Índice de Herfindahl |  | 4,8125\* | 2,1951\* |
|  |  | (0,3757) | (1,9891) |
| Quociente Locacional |  | 0,1184\* | 0,1073\* |
|  |  | (0,0062) | (0,0061) |
| Tamanho\*Belém |  |  | -0,0538\* |
| Tamanho\*Fortaleza |  |  | -0,0488\* |
| Tamanho\*Recife |  |  | -0,0843\* |
| Tamanho\*Salvador |  |  | -0,1081\* |
| Tamanho\*Belo Horizonte |  |  | -0,1152\* |
| Tamanho\*Rio de Janeiro |  |  | 0,0214\* |
| Tamanho\*Curitiba |  |  | -0,1391\* |
| Tamanho\*Porto Alegre |  |  | -0,1826\* |
| Tamanho\*Goiânia |  |  | -0,2211\* |
| Tamanho\*Distrito Federal |  |  | -0,3009\* |
| Escolaridade\*Belém |  |  | -0,1337\* |
| Escolaridade\*Fortaleza |  |  | -0,0144 |
| Escolaridade\*Recife |  |  | 0,0116 |
| Escolaridade\*Salvador |  |  | 0,2993\* |
| Escolaridade\*Belo Horizonte |  |  | -0,0611\* |
| Escolaridade\*Rio de Janeiro |  |  | -0,0592\* |
| Escolaridade\*Curitiba |  |  | 0,1433\* |
| Escolaridade\*Porto Alegre |  |  | 0,1683\* |
| Escolaridade\*Goiânia |  |  | -0,0637\* |
| Escolaridade\*Distrito Federal |  |  | -0,0048 |
| Log-Verossimilhança | -116571.15 | -114013.15 | -112644.43 |
| N. Observações | 660055 | 660055 | 660055 |

Nota: \* significante à menos de 5%. Valores entre parênteses se referem ao erro padrão dos coeficientes.

O sinal positivo e estatisticamente significante do índice de custo de vida para os modelos 2 e 3 sugere que as firmas são atraídas para as cidades com os maiores índices. Neste caso, a variável seria uma força de atração das firmas. Este resultado é consistente na situação em que um maior índice de custo de vida é associado com maiores preços para os produtos finais da firma. Além disso, o custo de vida das cidades pode refletir outras vantagens de aglomeração que são correlacionadas positivamente com a variável, como por exemplo, o tamanho do mercado.

A participação da indústria no emprego local – variável que busca captar as economias de localização pela concentração das firmas especializadas na mesma atividade industrial a qual o estabelecimento faz parte – apresentou coeficiente com o sinal positivo sinalizando a importância desse atributo local como fator de aglomeração das indústrias. Corroborando essa relação, os resultados também indicam que uma maior participação relativa da indústria no emprego da cidade, medido pelo quociente locacional, atraem os estabelecimentos para a cidade. Porém, os resultados não são consistentes com a presença de economias de urbanização como fator de atração das firmas consideradas na amostra, já que o índice de diversidade apresentou o sinal positivo, sugerindo que os investidores/firmas preferem as cidades com uma menor diversidade produtiva industrial.

Além desses resultados, as variáveis de interação incluídas no modelo, assinalam que as firmas maiores buscam se instalarem nas cidades maiores, São Paulo em especial, a única exceção foi a cidade do Rio de Janeiro. Além disso, as firmas com trabalhadores mais qualificados preferem se instalar na cidade de São Paulo, a exceção das cidades de Curitiba e Porto Alegre.

Os resultados expostos na Tabela 3 foram obtidos considerando todas as divisões da indústria de transformam instaladas nas cidades consideradas na análise. Porém, como as divisões de atividades da indústria de transformação se diferem quanto à intensidade dos fatores de produção, capital intensivo, trabalho ou recursos naturais, os fatores locais também podem atuar de modo diferenciado sobre as escolhas locacionais agentes produtivos. Deste modo, o segundo modelo, o qual apresentou os melhores resultados, foi estimado para cada uma das principais divisões da indústria de transformação: as indústrias de Fabricação de Produtos Alimentícios e Bebidas (CODIGO 15), Confecção de Artigos do Vestuário e Acessórios (CODIGO 18), Edição, Impressão e Reprodução de Gravações (CODIGO 22), Fabricação de Artigos de Borracha e Plástico (CODIGO 25) e Fabricação de Produtos de Metal Exceto Máquinas e Equipamentos (CODIGO CNAE 28).

Recorrendo a classificação de Moreira e Najberg (1998) das indústrias de transformação com respeito à intensidade dos fatores de produção, a indústria de Fabricação de Produtos Alimentícios e Bebidas seriam intensivas em Recursos Naturais, a indústria de Confecção de Artigos do Vestuário e Acessórios em trabalho, e as indústrias de Edição, Impressão e Reprodução de Gravações, Fabricação de Artigos de Borracha e Plástico e Fabricação de Produtos de Metal Exceto Máquinas e Equipamentos são intensivas em capital.

Os resultados encontram-se na Tabela 4. Observa-se inicialmente que os modelos estimados por indústrias também apresentaram um bom grau de ajuste, a maioria das variáveis com coeficiente estatisticamente significante a menos de 5%, e são conjuntamente significante. Quanto ao sinal dos coeficientes estimados, verifica-se que há algumas diferenças a depender da divisão de atividades para o qual o modelo foi estimado.

Os sinais dos coeficientes estimados das variáveis de custo de produção, salário da indústria e custo do aluguel, para todas as cinco indústrias, são consistentes com os resultados do modelo 2, a primeira variável atuando com um fator de repulsão e a segunda como um fator de atração dos investimentos industriais.

A variável média salarial da cidade apresentou sinal positivo para a indústria de Confecção de Artigos do Vestuário e Acessórios (CODIGO 18) e para a indústria de Edição, Impressão e Reprodução de Gravações (CODIGO 22), indústrias estas intensivas em mão de obra e capital, respectivamente. Mas, tal atributo local atua como fator de dispersão das indústrias de Fabricação de Produtos Alimentícios e Bebidas (CODIGO 15) e Fabricação de Produtos de Metal Exceto Máquinas e Equipamentos (CODIGO 28), resultado este não esperado.

O custo de vida, corroborando os resultados obtidos para o modelo 2, atua como um fator de atração para as indústrias de Fabricação de Produtos Alimentícios e Bebidas (CODIGO 15) e para as indústrias Confecção de Artigos do Vestuário e Acessórios (CODIGO 18). Mas, seria um fator de repulsão para as indústrias de Edição, Impressão e Reprodução de Gravações (CODIGO 22) e para as indústrias de Fabricação de Produtos de Metal Exceto Máquinas e Equipamentos (CODIGO 28).

Quanto ao índice de diversidade industrial, tal atributo local parece atuar como atributo de dispersão das indústrias de Fabricação de Produtos Alimentícios e Bebidas (CODIGO 15), de Edição, Impressão e Reprodução de Gravações (CODIGO 22), e de Fabricação de Produtos de Metal Exceto Máquinas e Equipamentos (CODIGO 28). E, há indicações de que apenas as indústrias de Confecção de Artigos do Vestuário e Acessórios (CODIGO 18) e as indústrias de Edição, Impressão e Reprodução de Gravações (CODIGO 22) são beneficiadas pelas economias produtivas geradas pela concentração relativa de indústrias da mesma divisão de atividades.

**Tabela 4. Resultados do Modelo *Logit* por Indústria**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variáveis Explicativas** | Divisão 15 | Divisão 18 | Divisão 22 | Divisão 25 | Divisão 28 |
| Salário da Indústria | 0,0020\* | 0,0008\* | 0,0015\* | 0,0021\* | 0,0030\* |
|  | (0,0002) | (0,0002) | (0,0003) | (0,0008) | (0,0003) |
| Aluguel | -0,0094\* | -0,0218\* | -0,0101\* | -0,0194\* | -0,0084\* |
|  | (0,0006) | (0,0007) | (0,0007) | (0,0032) | (0,0010) |
| Salário Cidade | -0,0010\* | 0,0026\* | 0,0007\* | 0,0011 | -0,0017\* |
|  | (0,0003) | (0,0003) | (0,0003) | (0,0014) | (0,0003) |
| Renda das Famílias | 0,0008\* | 0,0009\* | 0,0009\* | 0,0011\* | 0,0010\* |
|  | (0,0001) | (0,0001) | (0,0001) | (0,0002) | (0,0001) |
| Desemprego | -11,9324\* | -40,0918\* | -22,6726\* | -41,3829\* | -30,2916\* |
|  | (0,6976) | (1,1536) | (0,9665) | (1,4440) | (0,8777) |
| Custo de Vida | 0,0007\* | 0,0008\* | -0,0001 | -0,0010\* | -0,0005\* |
|  | (0,0001) | (0,0001) | (0,0001) | (0,0003) | (0,0001) |
| Índice de Herfindahl | 8,1555\* | -11,1696 | 7,4938\* | 0,7679 | 11,4873\* |
|  | (0,9682) | (1,0875) | (1,1247) | (5,5295) | (1,2205) |
| Quociente Locacional | -0,1676\* | 0,4714\* | 0,0329 | 0,6042 | 0,9673\* |
|  | (0,0364) | (0,0637) | (0,1023) | (0,3612) | (0,0854) |
| Log-Verrosimilhança | -11.964 | -24974.778 | -14.537 | -4.513 | -11.027 |
| N. Observações | 57156 | 161095 | 80322 | 33044 | 65912 |

Nota: \* significante à menos de 5%. Valores entre parênteses se referem ao erro padrão dos coeficientes.

1. **Teste de Robustez**

O teste de Hausman foi aplicado aos modelos estimados na seção anterior para verificar a validade da hipótese da Independência das Alternativas Irrelevantes, excluindo a cidade de São Paulo da amostra, sendo que todos os modelos rejeitaram a hipótese. Desse modo, a tabela 5 apresenta os resultados do melhor modelo obtido a partir da estimação do *Mixed Logit* (coluna 2) com o intuito de confrontar os resultados com os estimados supondo coeficientes estimados fixos para os estabelecimentos (*Logit* condicional, coluna 1). Destaca-se que para o modelo *Mixed Logit*, tem-se a média e o desvio padrão dos coeficientes estimados já que agora os coeficientes variam sobre a população. Para a estimação do modelo considerou-se que os coeficientes das variáveis, Salário Manufatura, Quociente Locacional, Participação do Emprego, Índice de Herfindahl, tem distribuição normal, e que as variáveis, custo de vida, aluguel e renda das famílias, têm coeficientes fixos.

Comparando os resultados obtidos pelos dois modelos, os sinais dos coeficientes, observa-se que apenas duas variáveis do modelo *Mixed Logit*, apresentaram sinais distintos dos obtidos pela estimação do *Logit* condicional. O custo de vida, segundo o modelo *Mixed Logit*, seria uma força de repulsão da atividade produtiva. Ou seja, quando considerado a heterogeneidade das preferências dos investidores, tal variável, em média, não é uma força de aglomeração das indústrias. Assim como, contrariando os resultados obtidos pelo modelo *Logit* Condicional, o Modelo *Mixed Logit* sugere que o índice de diversidade de Herfindahl atua atraindo os estabelecimentos para as cidades que tem uma maior diversidade produtiva. As demais variáveis consideradas nas estimações apresentaram os mesmos sinais em ambos os modelos estimados.

**Tabela 5. Resultados do Modelo *Logit* Condicional e Mixed Logit**

|  |  |  |
| --- | --- | --- |
| **Coeficientes** | Logit Condicional | Mixed Logit |
| Custo de Vida | 0,0004\* | -0,0016\* |
|  | (0,0000) | (0,0001) |
| Aluguel | -0,0067\* | -0,0212\* |
|  | (0,0002) | (0,0003) |
| Renda das Famílias | 0,0010\* | 0,0027\* |
|  | (0,0000) | (0,0000) |
| Salário da Manufatura | 0,0018\* | 0,0033\* |
|  | (0,0000) | (0,0000) |
| Quociente Locacional | 0,0859\* | 0,1519\* |
|  | (0,0052) | (0,0094) |
| Participação no Emprego | 6,3091\* | 7,8739\* |
|  | (0,0817) | (0,1491) |
| Índice de Herfindahl | 4,0602\* | -22,4668\* |
|  | (0,1818) | (0,4636) |
| **Desvio Padrão** |  |  |
| Salário Manufatura |  | -0,0032\* |
|  |  | (0,0000) |
| Quociente Locacional |  | 0,1609\* |
|  |  | (0,0183) |
| Participação do Emprego |  | 0,1261 |
|  |  | (0,6609) |
| Índice de Herfindahl |  | 2,9707\* |
|  |  | (0,4713) |
| Log-Verossimilhança | -124470.47 | -118527.7 |
| N. de Observações | 660055 | 660055 |

Nota: \* significante à menos de 5%. Valores entre parênteses se referem ao erro padrão dos coeficientes.

Cabe salientar que se optou em primeiro apresentar os resultados obtidos pela estimação do modelo *Logit* condicional devido a mais fácil convergência na estimação em relação ao modelo *Mixed Logit* que impõe limitações computacionais mais severas. Além disso, como o objetivo na análise foi estabelecer relações entre as variáveis, identificar quais fatores atuam como fator de atração e repulsão das indústrias, a maioria dos resultados do modelo *Logit* condicional foram consistentes com o esperado pela teoria e pelas evidências empíricas obtidas em outros estudos.

1. **Considerações Finais**

A presente pesquisa buscou obter evidências dos fatores locais, especialmente aqueles que podem gerar externalidades produtivas para as firmas, que influenciam a distribuição espacial das indústrias de transformação nas principais capitais brasileiras. A análise baseou-se nas estimações de modelos de escolha discreta, o *Logit* condicional e o *Mixed Logit*, utilizando os microdados da RAIS-Mte como fonte de informação dos estabelecimentos industriais brasileiros.

Os principais resultados obtidos para o setor industrial, tanto pelo modelo *logit* condicional quanto pelo *mixed logit*, corroboram com o esperado, isto é, a existência de externalidades produtivas geradas pela concentração de indústrias da mesma divisão de atividades. Esses resultados são consistentes com o fato de ganhos de produtividades obtidos para as firmas inseridas em aglomerações industriais. Isto possivelmente decorre de uma maior facilidade de interação entre firmas, propiciando trocas de informações e geração de *spillovers* tecnológicos por estarem sujeitas a um melhor *matching* no mercado de trabalho, entre outras vantagens produtivas. Já as evidências para a importância das economias de urbanização apenas foram suportadas pelo modelo *mixed logit*.

A respeito dos custos de produção, há indicações que o custo do aluguel atua como fator de repulsão da atividade produtiva, e que as indústrias são atraídas a se instalarem nas cidades que apresentam uma mão de obra com maior custo de contratação. Esse último resultado sugere que para as cidades consideradas na análise, a remuneração média da indústria estaria captando o potencial produtivo dos seus trabalhadores. Reforçando este resultado, a taxa de desemprego, contrariando parte das evidências empíricas obtidas para outros países (Coughlin et al., 1991; Crozet et al., 2004; Friedman et al., 1992; Head et al., 1999) apresentou sinal negativo, sugerindo que esta variável pode estar correlacionada com a existência de uma mão de obra de baixa qualificação e produtividade, o que deve desencorajar investimentos nas cidades com taxas elevadas de desemprego. Quanto a variável *proxy* para custo de vida, apenas o modelo *mixed logit* indica que tal variável atua como fator de repulsão das indústrias.

Os resultados também apontam que os estabelecimentos maiores (grandes empresas) tendem a se aglomerar nas maiores cidades, em especial, na cidade de São Paulo, em comparação as micros e pequenas empresas.

As evidências da importância dos fatores locacionais para explicar a distribuição de cada divisão das principais indústrias da transformação do país sugerem que, a depender do fator de produção que a indústria é intensiva e das próprias características da indústria, esses atributos podem influenciar de forma diferenciada as decisões locacionais dos investidores. Porém, parte relevante dos resultados corroborou os encontrados para toda a indústria da transformação como um todo.

Como sugestões para trabalhos futuros, pretende-se testar outros modelos de escolha discreta relaxando a hipótese da independência das alternativas irrelevantes. Por exemplo, é possível que um investidor primeiro decida qual será a grande região onde pretende instalar uma indústria, e em um segundo momento, a cidade específica. Outra extensão do trabalho será no sentido de explorar os resultados obtidos por indústria, tentar explicar porque algumas variáveis atuam como fator de atração para certas divisões industriais, enquanto para outras seria um fator de repulsão.

As evidências obtidas para a influência dos fatores locais que geram economias de localização e urbanização nas escolhas locacionais dos investidores/indústrias ao menos indicam para a importância de se fortalecer no país as políticas voltadas para a consolidação de arranjos produtivos industriais com vista a elevação da produtividade da indústria no país.

**Referencias Bibliográficas**

BASILE, Roberto; CASTELLANI, Davide; ZANFEI, Antonello. Location choices of multinational firms in Europe: The role of EU cohesion policy. **Journal of International Economics**, v. 74, n. 2, p. 328-340, 2008.

BONELLI, Regis. Concentração industrial no Brasil: indicadores da evolução recente. **Pesquisa e Planejamento Econômico**, v. 10, n. 3, p. 851-884, 1980.

CIEŚLIK, Andrzej. Location of foreign firms and national border effects: the case of Poland. **Tijdschrift voor Economische en Sociale Geografie**, v. 96, n. 3, p. 287-297, 2005.

COUGHLIN, Cletus C.; TERZA, Joseph V.; ARROMDEE, Vachira. State characteristics and the location of foreign direct investment within the United States. **The Review of Economics and Statistics**, p. 675-683, 1991.Crozet et al., 2004

DINIZ, Clélio Campolina; CROCCO, Marco Aurélio. Reestruturação econômica e impacto regional: o novo mapa da indústria brasileira. **Nova economia**, v. 6, n. 1, p. 77-103, 1996.

DOMINGUES, E. P. Aglomerações e Periferias Industriais no Brasil e no Nordeste. Revista Econômica do Nordeste, Fortaleza, v. 36, nº 4, out-dez. 2005.

EGELN, Jürgen; GOTTSCHALK, Sandra; RAMMER, Christian. Location decisions of spin-offs from public research institutions. **Industry and innovation**, v. 11, n. 3, p. 207-223, 2004.

ELLISON, Glenn; GLAESER, Edward L. The geographic concentration of industry: does natural advantage explain agglomeration?. **American Economic Review**, p. 311-316, 1999.

ELLISON, Glenn; GLAESER, Edward L.; KERR, William. **What causes industry agglomeration? Evidence from coagglomeration patterns**. National Bureau of Economic Research, 2007.

ELLISON, Glenn; GLAESER, Edward L. **Geographic concentration in US manufacturing industries: a dartboard approach**. National Bureau of economic research, 1994.

FEIJO, Carmem Aparecida; CARVALHO, Paulo Gonzaga M.; RODRIGUEZ, Maristella Schaefers. Concentração industrial e produtividade do trabalho na indústria de transformação nos anos 90: evidências empíricas. **Economia**, v. 4, n. 1, p. 19-52, 2003.

FIGUEIREDO, Octávio; GUIMARAES, Paulo; WOODWARD, Douglas. Home-field advantage: location decisions of Portuguese entrepreneurs. **Journal of Urban Economics**, v. 52, n. 2, p. 341-361, 2002.

FRIEDMAN, Joseph; GERLOWSKI, Daniel A.; SILBERMAN, Johnathan. What attracts foreign multinational corporations? Evidence from branch plant location in the United States. **Journal of Regional science**, v. 32, n. 4, p. 403-418, 1992.

THISSE, Jacques-François. Economics of Agglomeration, Cities, Industrial Location and Regional Growth. 2002.

FUJITA, M.; KRUGMAN, P.; VENABLES, A. Economia espacial: urbanização, prosperidade econômica e desenvolvimento humano no mundo. São Paulo: Editora Futura, 2002.

GREENE, William H. **Econometric analysis.** Pearson Education India, 2003.

HANSEN, Eric R. Industrial location choice in Sao Paulo, Brazil: a nested logit model. **Regional Science and Urban Economics**, v. 17, n. 1, p. 89-108, 1987.

HEAD, Keith; RIES, John; SWENSON, Deborah. Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. **Journal of international economics**, v. 38, n. 3, p. 223-247, 1995.

KRUGMAN, Paul R. **Geography and trade**. MIT press, 1991.

LAUTERT, Vladimir; ARAÚJO, Nilton Clóvis Machado de. Concentração industrial no Brasil no período 1996-2001: uma análise por meio do índice de Ellison e Glaeser (1994). **Economia Aplicada**, v. 11, n. 3, p. 347-368, 2007.

MARSHALL, Alfred. 1890. **Principles of Economics**. MacMillan.

MCFADDEN, D. Conditional logit analysis of qualitative choice behavior, Zarembka P., Frontiers in Econometrics, 1974, 105-142.

MOREIRA, M., NAJBERG, S. “Abertura comercial: criando ou exportando empregos?” **Pesquisa e Planejamento Econômico**, v. 28, n. 2, p. 371-398, 1998.

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

RESENDE, Marcelo; WYLLIE, Ricardo. Aglomeração industrial no Brasil: um estudo empírico. **Estudos Econômicos (São Paulo)**, v. 35, n. 3, p. 433-460, 2005.

RESENDE, Marcelo. Co-aglomeração industrial no estado do Rio de Janeiro: um estudo empírico. 2012.

SILVA, Magno Vamberto Batista ; [SILVEIRA NETO, Raul da Mota](http://lattes.cnpq.br/3921888911294306) . Determinantes da localização industrial no brasil e geografia econômica: evidências para o período pós-real. In: X Encontro Regional deEeconomia, 2005, Fortaleza.

SILVEIRA, Raul. Concentração industrial regional, especialização geográfica e geografia econômica: evidências para o Brasil no período 1950-2000. **Revista Econômica do Nordeste, Fortaleza**, v. 36, n. 2, 2005.

VENABLES, Anthony J. Equilibrium Locations of Vertically Linked Industries. **International Economic Review**, v.37, n.2, p.341-359, maio 1996.

WOOLDRIDGE, Jeffrey M. 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press.

**Apêndice**

**Quadro A.1. Divisões de Atividades das Indústrias da Transformação Segundo CNAE-95**

|  |  |
| --- | --- |
| [15](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=15&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE PRODUTOS ALIMENTICIOS E BEBIDAS |
| [16](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=16&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE PRODUTOS DO FUMO |
| [17](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=17&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE PRODUTOS TEXTEIS |
| [18](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=18&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | CONFECÇAO DE ARTIGOS DO VESTUARIO E ACESSORIOS |
| [19](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=19&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | PREPARAÇAO DE COUROS E FABRICAÇAO DE ARTEFATOS DE COURO, ARTIGOS DE VIAGEM E CAL. |
| [20](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=20&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE PRODUTOS DE MADEIRA |
| [21](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=21&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE CELULOSE, PAPEL E PRODUTOS DE PAPEL |
| [22](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=22&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | EDIÇAO, IMPRESSAO E REPRODUÇAO DE GRAVAÇOES |
| [23](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=23&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE COQUE, REFINO DE PETROLEO, ELABORAÇAO DE COMBUSTIVEIS NUCLEARES E PRODUÇAO DE ALCOOL |
| [24](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=24&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE PRODUTOS QUIMICOS |
| [25](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=25&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE ARTIGOS DE BORRACHA E DE MATERIAL PLASTICO |
| [26](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=26&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE PRODUTOS DE MINERAIS NAO-METALICOS |
| [27](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=27&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | METALURGIA BASICA |
| [28](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=28&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE PRODUTOS DE METAL - EXCLUSIVE MAQUINAS E EQUIPAMENTOS |
| [29](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=29&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE MAQUINAS E EQUIPAMENTOS |
| [30](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=30&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE MAQUINAS PARA ESCRITORIO E EQUIPAMENTOS DE INFORMATICA |
| [31](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=31&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE MAQUINAS, APARELHOS E MATERIAIS ELÉTRICOS |
| [32](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=32&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE MATERIAL ELETRONICO E DE APARELHOS E EQUIPAMENTOS DE COMUNICAÇOES |
| [33](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=33&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE EQUIPAMENTOS DE INSTRUMENTAÇAO MÉDICO-HOSPITALARES, INSTRUMENTOS DE PRECISAO E OPTICOS, EQUIPAMENTOS PARA AUTOMAÇAO INDUSTRIAL, CRONOMETROS E RELO. |
| [34](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=34&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO E MONTAGEM DE VEICULOS AUTOMOTORES, REBOQUES E CARROCERIAS |
| [35](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=35&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE OUTROS EQUIPAMENTOS DE TRANSPORTE |
| [36](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=36&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | FABRICAÇAO DE MOVEIS E INDUSTRIAS DIVERSAS |
| [37](http://www.cnae.ibge.gov.br/divisao.asp?coddivisao=37&CodSecao=D&TabelaBusca=CNAE_110@CNAE%201.0) | RECICLAGEM |

1. Fonte de dados RAIS-Mte 2010. [↑](#footnote-ref-1)
2. Fonte Contas Nacionais 2010 do IBGE obtido no IPEADATA. [↑](#footnote-ref-2)
3. Citam-se, por exemplo, as pesquisas financiadas pelo BNDES (disponível em http://www.bndes.gov.br) e pelo SEBRAE (disponível em http://www.redesist.ie.ufrj.br/) para a identificação de Arranjos Produtivos Locais no Brasil. [↑](#footnote-ref-3)
4. Nesse sentido, cita-se, por exemplo, os estudos desenvolvidos com base na “abordagem” de “Arranjos Produtivos Locais”, os quais objetivam apresentar explicações das forças de aglomeração identificadas especificamente para uma localidade, que abrange o APL, e para um setor produtivo, o setor de especialização do APL. [↑](#footnote-ref-4)
5. Na tabela A1 apresenta as divisões de atividade da Indústria da Transformação segunda a Classificação Nacional de Atividades (CNAE-95). [↑](#footnote-ref-5)
6. O artigo apresenta em seu apêndice tabelas de vários índices de concentração para classificação de até 5 dígitos da indústria de transformação. [↑](#footnote-ref-6)
7. O percentual deve ser tornar ainda mais elevado se consideramos regiões metropolitanas, já que geralmente muitos distritos industriais se localizam no entorno das capitais. [↑](#footnote-ref-7)
8. Partindo da hipótese de que todos os municípios do país são escolhas possíveis para a escolha locacional [↑](#footnote-ref-8)
9. Utilizou-se aqui a definição do SEBRAE: micro empresas são as que possuem até 19 funcionários; pequenas empresas são as que têm de 20 a 99 funcionários, as médias de 100 a 499, e as grandes acima de 500 funcionários. [↑](#footnote-ref-9)
10. A hipótese da IIA pode ser testada a partir do teste de Hausman. [↑](#footnote-ref-10)
11. Train (2003) discute as propriedades de eficiência dos estimadores que maximizam a função de log-verossimilhança simulada. [↑](#footnote-ref-11)
12. Maiores informações sobre esta base constam em http://www.mte.gov.br/pdet [↑](#footnote-ref-12)