

**Economias de Aglomeração e Escolha Locacional das Indústrias de Alimentos e Bebidas no Brasil:
Uma Aplicação do *Mixed Logit***

Roberta de Moraes Rocha (UFPE/PPGECON)
Klebson Humberto de Lucena Moura (PIMES-UFPE)

Resumo

A pesquisa apresenta evidências da importância das economias de aglomeração, urbanização e localização, para a distribuição espacial das indústrias de alimentos e bebidas entre as principais cidades brasileiras. A análise baseou-se na estimação de modelos de escolha discreta, o *logit* condicional e o *mixed Logit*, a partir dos microdados da RAIS-Mte, para uma amostra de 5.196 estabelecimentos formais, uma cobertura de 92% do total com registros no Ministério do Trabalho. Os resultados da pesquisa indicam que, para a indústria em específica, é preciso considerar a heterogeneidade das firmas: atributos locais influenciam de forma diferenciada a escolha locacional das firmas e, conseqüentemente, a sua produtividade e lucro. Os resultados são consistentes com a presença de economias de localização e urbanização nas cidades que possuem uma maior concentração relativa de empregos na indústria, assim como, naquelas que têm uma estrutura produtiva mais diversificada, os quais atuam como fatores de aglomeração. Além disso, há indicações de que as conexões para trás são importantes para explicar a distribuição espacial desses estabelecimentos, as quais são mais propensos a se instalarem próximos aos seus fornecedores de insumos. Os resultados também apontam que, mesmo após estes controles, as indústrias de alimentos e bebidas são atraídas a se instalarem na cidade de São Paulo.

Palavras-Chaves: Modelos Discretos, *Mixed Logit*, Escolha Locacional, Economias Aglomeração

Abstract

This research obtain evidences of the importance of agglomeration factors for the spatial distribution of food and drink industries of the main Brazilian cities. The analysis based on the estimation of discrete choice models, conditional logit and mixed logit, from microdata RAIS-Mte. It was possible because each establishment was identified through of the CNPJ them. The sample totalized 5,196 establishment. The main results indicate that cities with a greater relative concentration of jobs in the food and drink industries as well as with greater industrial diversity act as a force of attraction of the Food and drink industry. Moreover, there are indications that the back connections and market access are important to explain the spatial distribution of industries. The results indicated that food and drink industries prefer to locate the city of São Paulo.

Keys-Words: Discrete Model, Mixed Logit, Locational Choice, Agglomeration Economies

JEL: L6 ; R11.

1. Introdução

Três importantes evidências têm norteado estudos em economia regional e urbana acerca da distribuição da atividade produtiva no espaço. Primeiro, firmas e trabalhadores se distribuem irregularmente no espaço (Ellison, 2007; Ellison e Glaeser, 1999; Lautert e Araujo, 1994; Resende e Wyllie, 2005; Resende, 2012; Ellison e Glaeser, 1999; Ellison e Glaeser, 1994). No Brasil, apenas 1,8% dos municípios (100 municípios) concentraram 50% dos empregos formais gerados pela indústria da transformação no ano de 2010¹; e 40 municípios responderam por 46% do Produto Nacional Bruto² e por 30% da população³. 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).

Segundo, grandes cidades concentram trabalhadores mais qualificados (Glaeser e Mare, 2001; Rocha et. al., 2011) e firmas mais produtivas (Combes et. al., 2009) o que pode potencialmente explicar, ao menos em parte, essa tendência à concentração das indústrias. Combes et. al. (2009) observam que quando se duplica o tamanho das cidades dos Estados Unidos, a produtividade dessas cidades aumenta, em média, entre 2% e 7%. No Brasil, no ano de 2010, apenas os 50 maiores municípios, que concentraram 31% da população, apresentou uma produtividade média cerca de 64% maior que o conjunto dos demais municípios⁴. E, terceiro, esses ganhos de produtividade obtidos pelas indústrias espacialmente concentradas parece perder a importância com a distância (Rosenthal e Strange, 2003), o que reforça as duas primeiras evidências empíricas.

Com respeito à segunda evidência⁵, a correlação entre o tamanho das cidades e a produtividade dos agentes econômicos, é possível identificar na literatura ao menos duas possíveis explicações para tal fato. Remontando ao trabalho seminal de Marshall (1989), estudo ainda bem explorado pela literatura mais recente aqui citada, a concentração de firmas e trabalhadores geram economias de aglomeração que elevam a produtividade das firmas devido a possibilidade de um melhor acesso ao mercado de insumos, à uma mão de obra e serviços mais especializados, e ao beneficiamento dos *spillovers* tecnológicos, entre outros fatores. Do outro lado, cidades com uma melhor infraestrutura e acesso a uma maior variedade e qualidade de bens de consumo e serviços, assim como, com um melhor aporte de amenidades naturais tendem a atrair trabalhadores mais qualificados (Glaeser e Maré 2001; Rocha, et. al., 2011) e firmas mais produtivas (Combes et. al., 2009), possivelmente aquelas que possuem um maior nível de diferenciação dos produtos que é especializada (Melitz, 2003).

De acordo com as evidências empíricas para o Brasil (Silveira Neto, 2005; Feijo et. al., 2003; Diniz e Crocco, 1996; Bonelli, 1980) e para outros países (Ellison, 2007; Ellison e Glaeser, 1999) parece não haver dúvidas, de um modo geral, que há uma tendência para a concentração das industriais, o que deve variar na magnitude, conforme a intensidade dos fatores de produção para o qual a indústria é especializada. O fato é que mesmo existindo forças contrárias a aglomeração da atividade produtiva, como aquelas associadas a uma maior concorrência ou a custos de produção mais elevados, parece que as vantagens produtivas obtidas pela maior parte das firmas aglomeradas superam as primeiras, pois caso contrário não haveria explicação para a concentração geográfica das indústrias. Como consequência, políticas voltadas para a consolidação de complexos industriais em países como o Brasil têm sido utilizadas como estratégias para o aumento da produtividade da indústria nacional e para o crescimento regional⁶.

Uma vez constatada essa tendência à concentração industrial, uma segunda etapa de investigação é analisar a importância das economias de aglomeração para explicar a distribuição espacial das indústrias. Porém, esta análise não é tão simples e direta como a primeira e impõe algumas dificuldades empíricas dado que os agentes econômicos, trabalhadores e firmas, podem ser positivamente selecionados para as

¹ Fonte de dados RAIS-Mte 2010.

² Fonte Contas Nacionais 2010 do IBGE obtido no IPEADATA.

³ Fonte CENSO DEMOGRAFICO-IBGE 2010.

⁴ Fonte Contas Nacionais 2010 do IBGE.

⁵ Puga (2009) faz uma discussão sobre as causas e magnitudes das economias de aglomeração.

⁶ 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.

grandes cidades (Glaeser e Maré 2001; Combes et. al. 2009; Rocha, et. al., 2011). Neste caso, é possível que firmas espacialmente concentradas sejam mais produtivas devido, em parte, as suas próprias características, e não apenas pelo fato de se beneficiarem das economias de aglomeração. Assim, de modo a apresentar evidências da importância das externalidades produtivas geradas pela concentração industrial para explicar as escolhas locais das firmas seria necessário isolar o efeito “seleção” do efeito “aglomeração” sobre a produtividade (Combes et. al. 2009).

Com este escopo de pesquisa, avanços têm sido observados na literatura empírica internacional, especialmente devido à existência e a uma maior facilidade de acesso aos dados ao nível das firmas. Boa parte desses estudos utiliza como base modelos de escolha discreta (Cheng e Stough, 2006; Arauzo e Manjon, 2004; Basile et. al., 2003; Baudewyns et. al., 2000), mas se diferenciam quanto ao nível de desagregação setorial e regional 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* (Balele et. al., 2003; Hansen 1987). A vantagem destes modelos, quando aplicado com o objetivo de obter evidências das economias de aglomeração para as escolhas locais das firmas, reside na possibilidade de minimizar os problemas de identificação que estão presentes nos estudos que utilizam dados agregados por indústrias, que consideram informações do emprego para construir tanto o indicador de concentração quanto as variáveis explicativas.⁷

Para o Brasil, destaca-se um único trabalho publicado realizado com microdados ao nível das firmas, 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 locais das firmas, e que a proximidade da grande São Paulo é um fator de atração. Porém, contrariando as expectativas o autor encontra indicações, para alguns dos modelos estimados, de que as firmas dos 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 investiga a influência de fatores locais, especialmente aqueles que devem gerar economias de aglomeração, para as escolhas locais⁸ das Indústrias de Alimentos e Bebidas das principais cidades brasileiras. A análise baseia-se nos microdados da Relação Anual de Informações Sociais (RAIS) do Ministério do Trabalho (Mte) para o ano de 2010, a qual compreendeu 5.196 estabelecimentos formais, e informações secundárias do Censo Demográfico do IBGE para o ano de 2010. Para tanto, seguindo a linha de investigação que utilizam modelos de escolha discreta, são estimados os modelos *logit condicional* e o *mixed logit*. Este último modelo apresenta a vantagem de flexibilizar a hipótese de independência das alternativas irrelevantes (IIA) imposta pelo *logit condicional* e, assim, é possível considerar a heterogeneidade das firmas⁹ sob a hipótese de que os fatores locais influenciam de forma diferenciada a produtividade das mesmas. Ressalta-se que esta estimação apenas é possível porque a base de dados utilizada, os microdados da RAIS do Ministério do Trabalho, é disponibilizado por firma e é possível identificar cada uma delas, como também a sua localização, através de um código único e imutável, o CNPJ.

Como as divisões de atividades da indústria de transformação se diferem quanto à intensidade dos fatores de produção, capital intensivo, trabalho ou recursos naturais, possivelmente as características locais das cidades, inclusive aquelas que podem gerar externalidades produtivas, devem influenciar de modo diferenciado as escolhas locais das firmas (LaFountain, 2005; Rocha et. al., 2013). Assim, escolheu-se como objetivo de análise da pesquisa, a Indústria de Alimentos e Bebidas, uma das indústrias mais importantes para a economia do país. No Brasil, entre as indústrias da transformação, a indústria de

⁷ Porém, reconhece-se os avanços realizados nesta linha de investigação em construir uma medida para mensurar o nível de concentração geográfica das indústrias, e a partir deles tentar explicá-los com base em modelos de regressão, conduzidos pioneiramente por Ellison e Glaeser (1994; 1999, 2007).

⁸ Apenas para simplificar a análise, como o termo “escolha locacional” é utilizado repetitivamente, por vezes atribui-se essa ação a Indústria ou Firma, embora sabe-se que a decisão da escolha é do (s) investidor (s). Deste modo, firma é utilizada como sinônimo de estabelecimento.

⁹ Aqui firma é utilizado como sinônimo de estabelecimento.

alimentos e bebidas (CNAE-95 a 2-dígitos), a qual representou em 2010 aproximadamente 20% do emprego gerado pela indústria da transformação segundo a RAIS-Mte, se destaca por se concentrar em maior potencial em áreas urbanas; 13% do emprego gerado estão nas capitais brasileiras. Além disso, dadas as características da cadeia produtiva da indústria de Alimentos e Bebidas, com importantes conexões para trás com os seus fornecedores de insumos *in natura* e para frente (com o mercado consumidor), acredita-se ser um bom referencial para se obter evidências das economias de aglomeração.

2. Distribuição Espacial dos Estabelecimentos Formais da Indústria de Alimentos e Bebidas no Brasil

A escolha de realizar o estudo para a indústria de alimentos e bebidas justifica-se pelo fato de ser uma das mais importantes indústrias do país. Segundo os resultados das contas nacionais do IBGE para o ano de 2005, da matriz de insumo-produto, a indústria de alimentos e bebidas apresentou um dos maiores graus de dependência de outros setores na economia enquanto consumidora de insumos¹⁰. Esse resultado sugere que o crescimento da indústria deve ter um grande impacto na economia do país.

Essa importância se dá no âmbito interno, mas também no externo, já que o setor se enquadra na cadeia produtiva do agronegócio, notadamente um dos setores mais dinâmicos e competitivos em termos mundiais que o Brasil possui. Em 2013, o faturamento (líquido de impostos indiretos), segundo dados da Associação Brasileira das Indústrias de Alimentos (ABIA), alcançou a marca de R\$ 484,7 bilhões, o que representa 10% do PIB Nacional e 21,8% do faturamento da indústria de transformação. Além disso, a indústria obteve forte expansão nos últimos anos, as vendas da indústria de alimentos no mercado interno no período de 2005 a 2013 passaram de R\$ 144,4 bilhões para R\$ 371,7 bilhões, um crescimento de 257% no período. Quanto ao comércio exterior, no ano de 2013, os produtos alimentícios industrializados representaram 17,8% do total das exportações nacionais. Pela ótica do mercado de trabalho, a importância é ainda mais evidente, pois segundo a RAIS, a indústria gerou, no ano de 2013, um pouco mais de 1,5 milhão de empregos formais (em torno de 1 milhão e 600 vínculos ativos, um aumento de 41 mil vínculos, em relação a 2012); e é composta por um pouco mais de 46 mil estabelecimentos, segundo maior número dentro da indústria de transformação, atrás apenas de confecção de artigos de vestuário e acessórios.

Com o intuito de apresentar uma caracterização geral da distribuição geográfica da indústria de alimentos e bebidas (código 15, segundo CNAE 95), a Tabela 1 apresenta a participação no emprego gerado e no total de estabelecimentos para as principais capitais do país. Essa primeira análise apenas dá uma indicação de como a indústria está distribuída regionalmente: 11 cidades do país, para as quais o estudo é realizado, concentraram mais de 10% de todo o emprego e mais de 11% de todos os estabelecimentos.

Tabela 1. Percentual do Emprego e Estabelecimentos da indústria de alimentos e bebidas nas principais capitais brasileira (2010)

Cidade	Emprego	Estabelecimentos
Belém	0,41%	0,45%
Fortaleza	0,96%	1,36%
Recife	0,71%	1,13%
Salvador	0,33%	0,64%
Belo Horizonte	0,57%	1,29%
Rio de Janeiro	1,47%	1,12%
São Paulo	2,77%	2,79%
Curitiba	0,94%	0,95%
Porto Alegre	0,64%	0,43%
Goiânia	1,05%	0,70%
Brasília	1,02%	0,68%
Demais Cidades	89,15%	88,48%

¹⁰ Segundas as informações da Matriz de Insumo-Produto das Contas Nacionais do ano de 2005 do IBGE medido pelos coeficientes técnicos.

Brasil 100,0% 100,0%

Fonte: Elaborado pelos autores com Base nos dados da Rais-MTE 2010.

A figura 1 dá uma dimensão da distribuição geográfica do emprego gerado pela indústria de alimentos e bebidas no ano de 2010 entre os municípios brasileiros. A parte escura representa a concentração de empregos na indústria.



Figura 1. Distribuição do Número de Empregos Formais da Indústria de Alimento e Bebidas
Fonte: RAIS-Mte 2010.

Por sua vez, nota-se de acordo com a Tabela 2 que a maioria dos estabelecimentos localizados nessas capitais se enquadra na faixa de Micro e Pequena empresa (definidas aqui como as que possuem até 20 empregados). As empresas de porte médio, definidas como aquelas que possuem entre 20 e 240 empregados, vem em seguida com uma participação variando de 9% à 22%.

Tabela 2. Percentual, por tamanho, dos Estabelecimentos da Indústria de Alimentos e Bebidas nas principais capitais brasileiras (2010)

Cidades	Micro e pequena	Media	Grande
Belém	78,33%	19,70%	1,97%
Fortaleza	79,67%	18,21%	2,11%
Recife	82,58%	15,85%	1,57%
Salvador	89,97%	9,00%	1,04%
Belo Horizonte	85,49%	13,99%	0,51%
Rio de Janeiro	78,74%	17,91%	3,35%
São Paulo	74,92%	22,85%	2,23%
Curitiba	85,21%	12,91%	1,88%
Porto Alegre	82,82%	15,46%	1,72%
Goiânia	81,30%	17,65%	1,05%
Brasília	83,73%	14,97%	1,30%
Demais Cidades	86,40%	11,26%	2,34%

Fonte: Elaborado pelos autores com Base nos dados da Rais-MTE 2010.

A Tabela 2 indica que o setor analisado é consideravelmente composto por micro e pequenas empresas. A maioria dos estabelecimentos é de pequeno porte, sendo mais representativo na capital Salvador, com uma participação de quase 90% no total das indústrias localizadas na cidade, quando comparado com as cidades consideradas na análise. Ao comparar o tamanho das indústrias nas cidades selecionadas com as demais cidades, observa-se que há uma maior participação de médios estabelecimentos nas cidades consideradas na análise (com médio em torno de 16%), a exceção é Salvador (9% tem porte médio). Cabe destacar ainda a cidade de São Paulo, que possui a maior participação de estabelecimentos de médio e grande porte no total localizado na cidade (em torno de 25%), quando comparado com as cidades consideradas na análise (percentual médio em torno de 17%)

A Tabela 3 apresenta a participação do emprego da indústria de alimentos e bebidas no emprego industrial na cidade (E_{ij} / E_J) e a participação dos estabelecimentos da Indústria de Alimentos e bebidas no total de estabelecimentos industriais na cidade (Est_{ij} / Est_J). De acordo com estes dados, mais uma vez verifica-se a importância da indústria para as cidades selecionadas

Tabela 3. Percentual do Emprego e Estabelecimento da Indústria de Alimentos e bebidas, por cidade (2010)

Cidades	%	%
	Emprego E_{ij} / E_J	Estabelecimento Est_{ij} / Est_J
Belém	39%	27%
Fortaleza	17%	12%
Recife	31%	24%
Salvador	17%	17%
Belo Horizonte	12%	12%
Rio de Janeiro	13%	9%
São Paulo	8%	5%
Curitiba	15%	10%
Porto Alegre	14%	11%
Goiânia	21%	10%
Brasília	33%	18%
Demais municípios	21%	15%

Fonte: Elaboração própria com base na Rais (2010).

O destaque aqui está para as cidades de Belém, Recife e Brasília, onde a referida indústria representa mais de 30% do emprego gerado por toda a indústria da transformação, e por 27%, 24% e 18% dos estabelecimentos industriais formais, respectivamente.

Buscando uma medida relativa para mensurar a importância da Indústria de Alimentos e Bebidas para cada cidade, que pondere pelo tamanho do setor industrial da cidade, calculou-se o Quociente Locacional (QL), obtido pela razão entre a participação da cidade no emprego gerado pela Indústria de Alimento e Bebidas nacionalmente (E_{ij} / E_I) e a participação da cidade no emprego gerado por toda a indústria de transformação (E_J / E). Analisando estes dados, constata-se que para quatro cidades o QL é acima da unidade, indicando que a indústria de alimentos e bebidas é a mais importante para essas cidades quando comparada com toda a indústria.

Tabela 4. Quociente Locacional da Indústria de Alimentos e Bebidas para as principais capitais brasileiras (2010)

Cidades	E_{ij} / E_I	E_j / E	QL
Belém	0,004	0,002	1,956
Fortaleza	0,010	0,011	0,868
Recife	0,007	0,005	1,531
Salvador	0,003	0,004	0,842
Belo Horizonte	0,006	0,009	0,622
Rio de Janeiro	0,015	0,022	0,659
São Paulo	0,028	0,070	0,400
Curitiba	0,010	0,013	0,739
Porto Alegre	0,004	0,006	0,724
Goiânia	0,007	0,006	1,070
Brasília	0,007	0,004	1,650

Fonte: Elaboração Própria

Nota: E_{ij} – emprego da indústria de alimentos e bebidas (i) da cidade j; E_I – total do emprego da indústria de alimentos; E_j – emprego total da indústria da transformação da cidade j; E – emprego total da indústria da transformação do Brasil.

Finalmente, os Quocientes Locacionais calculados com base na participação de cada grupo de atividades no emprego gerado e no total de estabelecimentos da indústria de alimentos e bebidas, dá indicações da magnitude da diversidade produtiva da Indústria de Alimentos e Bebidas observada no Brasil. (ver tabela A2)

Observa-se, a princípio, o quanto as cidades se diferenciam ao nível de especialização da indústria de alimentos e bebidas, tomando como referência para a análise a comparação relativa entre os QLs. A cidade de Belém tem os grupos de atividade 151 (Abate e Preparação de Produtos de Carne e de Pescado), 152 (Processamento, Preservação e Produção de Conservas de Frutas, Legumes e Vegetais) e 153 (Produção de Óleos e Gorduras Vegetais e Animais) como os mais importantes, o que decorre, certamente, das vantagens naturais que a região da cidade possui. Em Fortaleza, têm-se os grupos 152 e 159 (Fabricação de Bebidas) com as maiores participações relativas no emprego gerado pela indústria. Na cidade de Recife, por sua vez se destaca o grupo 156 (Fabricação e Refino de Açúcar), condizente com a cultura da cana-de-açúcar no estado. A cidade de Salvador também apresenta destaque no grupo 151, além do grupo 155 (Moagem, Fabricação de Produtos Amiláceos e de Rações Balanceadas para Animais).

A cidade do Rio de Janeiro tem como atividade mais representativa o Grupo 151, principalmente devido as atividades de Pesca. Já a cidade São Paulo, percebe-se uma maior diversidade das atividades da indústria, de acordo com este indicador, que pode ser observado por uma maior homogeneidade dos valores dos quocientes locacionais dos grupos de atividades.

Em Curitiba, o grupo 158 (Fabricação de Outros Produtos Alimentícios) é o mais importante em termos de participação relativo no emprego gerado pela indústria local, em comparação com os demais. Para a cidade de Porto Alegre, o destaque é para a fabricação de bebidas (grupo 159); em Goiânia tem-se o Grupo 152 e, em Brasília, o grupo 157 (Torrefação e Moagem de Café).

Em resumo, comparando os grupos de atividades considerando o QL calculado a partir da variável “estabelecimento”, o grupo de atividade de Abate e Preparação de Produtos de Carne e de Pescado (CNAE 151) é a atividade com maior participação relativa para cinco das onze cidades consideradas na análise.

3. Modelos de Escolha Discreta: Escolha Locacional das Firmas

Baseando-se em modelos de escolha locacional que adotam como base os pressupostos neoclássicos para a função de produção (Carlton, 1979; Glaeser e Mare, 2001) assume-se que a firma escolhe a cidade j para se instalar de forma a maximizar o seu lucro, dados os atributos locais das cidades e os preços dos

fatores de produção, capital e trabalho. Desse modo, o problema de maximização de lucro das firmas (Glaeser e Mare, 2001) pode ser descrito na forma:

$$\pi_{ij} = A_j K^\delta L^{1-\delta} - W_j L - R K \quad (1)$$

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

Apenas para simplificação, sem alterações nas principais conclusões que se quer obter para o objetivo da pesquisa, assim como Hansen (1987), assume-se 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:

$$\pi_{ij} = A_j L^\delta K^{1-\delta} - W_j L \quad (2)$$

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

$$L^* = \left(\frac{W_j}{A_j \delta K^{1-\delta}} \right)^{\frac{1}{\delta-1}} \quad (3)$$

Substituindo L^* na função lucro tem-se o lucro ótimo na forma:

$$\pi^* = (1-\delta) A_j^{1/(1-\delta)} \left(\frac{W_j}{\delta} \right)^{-\delta/(1-\delta)} K \quad (4)$$

Da equação (4), observa-se que o lucro ótimo correlaciona-se positivamente com A_j , que inclui os fatores locais que influenciam positivamente a produtividade 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, a firma/investidor i escolhe a localidade j para se instalar entre as J opções de escolha de forma a maximizar os lucros π_{ij} que assume a seguinte forma linear:

$$\pi_{ij} = X_j \beta + \varepsilon_{ij}, \quad i = 1, \dots, N \text{ e } j = 1, \dots, J \quad (5)$$

X_j é um vetor $1 \times M$ das características locais, β é o vetor dos parâmetros, ε_{ij} 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' ($\pi_{ij'}$), é maior do que o que o investidor teria em j'' ($\pi_{ij''}$), de modo que a probabilidade do investidor i escolher a localidade j' para se instalar é dada por:

$$p[\pi_{ij'} > \pi_{ij''}] = p[(X_{j'} \beta + \varepsilon_{ij'}) > (X_{j''} \beta + \varepsilon_{ij''})] \text{ onde } j' \neq j'' \quad (6)$$

os resíduos têm distribuição do tipo “valor extremo” $F(\varepsilon_{ij}) = \exp(-e^{-\varepsilon_{ij}})$

McFadden (1974) mostra que se os erros da função π_{ij} são identicamente e independentemente

distribuídos, a probabilidade da região j' ser escolhida é:

$$L_{ij} = \text{Prob}(Y_i = j') = \frac{\exp(\pi_{ij'})}{\sum_{j=1}^J \exp(\pi_{ij})} \quad (7)$$

Y_i 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:

$$LL = \sum_i \sum_j I_{i,j} \ln(L_{i,j}) \quad (8)$$

o indicador I_{ij} é igual a um se a firma i escolhe a localidade j' ; caso contrário, I_{ij} 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)¹¹ (Wooldridge, 2002; Greene, 2003), 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 Independência das Alternativas Irrelevantes (IIA) for rejeitada, outros modelos que relaxam essa hipótese devem ser estimados com o mesmo propósito da pesquisa.

Neste caso, é possível que os coeficientes do modelo (1) variem sobre a população, compatível com a situação em que os fatores locais (variáveis explicativas) influenciam de forma diferenciada as escolhas locais dos investidores/firmas. Como consequência, a IIA derivada das suposições feitas para o termo de erro (ε_{ij}) do modelo da equação (5) passa a ser flexibilizada (Train, 2009). Neste caso, o vetor dos coeficientes das variáveis explicativas, que aqui mensuram atributos da região j , passa a não ser mais fixo e variam de acordo com as preferências locais dos investidores/firmas, sendo decomposto em dois termos: $\beta = b + \eta_i$ (Revelt e Train, 1998, Revelt e Train, 1999), com o termo " η_i " captando a heterogeneidade das preferências das firmas.

O modelo com as características acima descritas é conhecido na literatura como "*mixed logit*"¹², e pode ser representado a partir da equação (1) na forma:

$$\pi_{ij} = X_j \beta + \varepsilon_{ij}, \text{ com } \beta = b + \eta_i \quad (9)$$

Ou,

$$\pi_{i,j} = x_j b + x_j \eta_i + \varepsilon_{ij} \quad (10)$$

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

Para o caso da presente pesquisa e recorrendo a descrição do processo de simulação em Train (2009), 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 A hipótese da IIA pode ser testada a partir do teste de Hausman.

¹² Aqui, assim como em Train (2009), o termo *mixed logit* está sendo utilizando para generalizar a função de distribuição de β , a qual pode assumir diversas forma funcional. No caso particular do *logit condicional*, a distribuição de β é dita "degenerativa".

$$S_i(\beta_i) = \prod L_{ij}(\beta_i) = P_r\{(\pi_{i1} > \pi_{ij}) \cap (\pi_{i2} > \pi_{ij}) \cap \dots \cap (\pi_{i11} > \pi_{ij})\} = \prod_{j=1}^{11} \frac{\exp(\pi_{ij})}{\sum_{j=1}^J \exp(\pi_{ij})} \quad (11)$$

$S_i(\beta_i)$ 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 “ β ” para cada estabelecimento é desconhecido, sendo necessário integrar a probabilidade do *logit* condicional sobre todos os possíveis valores de “ β ” e então obtém a probabilidade de escolha do *mixed logit*.

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

$$P_i(\theta) = \int S_i(\beta_i) f(\beta_i | \theta) d\beta_i \quad (12)$$

$f(\beta_i | \theta)$ é a função densidade. A função densidade pode tomar diferentes distribuições a depender das características das variáveis em estudos (Train, 2009).

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, 2009)

O modelo *Mixed Logit*, assim como o modelo *logit* condicional, tem importantes aplicações em diversas áreas de estudo, destacam-se aqui as pesquisas que buscam modelar as preferências dos consumidores por atributos de um bem qualquer, levando em consideração que esses fatores podem ter diferentes impactos sobre a função de utilidade dos consumidores. Como exemplo, citam-se os estudos de demanda por automóvel (Cardell e Dunbar, 1980); os que buscam valorar as amenidades locais (Rocha e Matos, 2013); e, mais recentemente, destacam-se as pesquisas que investigam a importância das economias de aglomeração para explicar as escolhas locais das firmas (Basile et. al., 2008; Defever, 2012).

4. Modelo Empírico e Dados

Seguindo a literatura¹³ que objetiva obter evidências da importância dos fatores locais para a escolha espacial das indústrias a partir de modelos de escolha discreta, o *logit* condicional e o *mixed logit* são estimados com dados ao nível da firma. A vantagem de utilizar este tipo de dados reside no fato da possibilidade de incluir no modelo controles para as características da firma que são correlacionadas com a sua produtividade e, assim, poderia contaminar o “efeito aglomeração”, o ganho da produtividade devido a fatores externos as firmas. Além disso, é importante destacar que modelos como este são menos sujeitos a problemas de endogeneidade gerada pelo fato da causalidade reversa entre os regressores e a variável dependente, já que é pouco provável que uma firma individualmente venha a ter um impacto significativo nas variáveis locais, como no nível de emprego ou produção. Este problema deve ser mais recorrente em modelos que utilizam como variável dependente indicadores agregados regionalmente. Ressalta-se que esta estimação apenas foi possível devido a existência dos microdados da RAIS-MTe do Ministério do Trabalho, instituição que disponibilizou os dados por firma para a presente pesquisa.¹⁴

O modelo estimado segue em sua especificação geral:

¹³ ARAUZO-CAROD (2010) faz um *survey* dos modelos e estudos empíricos que tem como foco de análise a influência dos fatores locais para a escolha espacial das firmas com base em modelos de escolha discreta.

¹⁴ A base de dados pode ser requerida por qualquer pesquisador ao Ministério do Trabalho mediante um Termo de Compromisso que exige, entre outras condições, que os dados não sejam utilizados para outro propósito que não seja a pesquisa. Para maiores informações de acesso aos dados consultar o link <http://acesso.mte.gov.br/portal-pdet/home/>.

$$y_{ij} = X_i \beta + Z_j \delta + \varepsilon_{ij} \quad (13)$$

A variável dependente, y_{ij} , assume valor igual a “1” se a firma “i” está localizada na cidade “j”; o vetor X_i inclui as características controles das firmas, tais como, tamanho e nível médio da escolaridade dos trabalhadores. No vetor Z_j são inclusos as características das cidades que podem explicar a localização das firmas e são definidos logo a seguir. E, ε_{ij} , é o termo de erro do modelo.

De acordo com o modelo teórico que dá suporte a equação (13), a estimativa para o coeficiente de cada variável explicativa informa sobre a preferência dos investidores/firmas pelo respectivo atributo local, e esta informação, sendo corretamente identificada, pode ser utilizada para auferir sobre a importância que cada variável tem sobre a função lucro das firmas. A este respeito, cabe ressaltar que a proposta deste artigo é de apontar evidências que sejam consistentes com a existências de economias de aglomeração geradas pela concentração da atividade produtiva não tendo, portanto, a pretensão de mensurar a magnitude destas.¹⁵

Além da escolha das variáveis explicativas – descritas a seguir com as suas respectivas justificativas, uma atribuição do pesquisador –, que deve ter um respaldo na literatura já publicada, também é preciso escolher o nível de desagregação geográfica para o estudo. A estimação dos modelos *logit* condicional e *mixed logit* pode não ser viável se o conjunto de opções de escolha for muito grande. Assim, o pesquisador enfrenta um *trade off* entre a escolha do nível de agregação regional a ser considerado na pesquisa e o grau de heterogeneidade que quer captar com os dados. Como a presente pesquisa tem como objetivo obter evidências das economias de localização e urbanização, dados ao nível das cidades são necessários, e deste modo optou-se por realizar o estudo para as principais capitais do país, as que apresentavam um maior grau de concentração destas indústrias, conforme já relatado.

A base de dados utilizada na pesquisa são os microdados derivados do Relatório Anual de Informações Sociais (RAIS)¹⁶ do Ministério do Trabalho e do Emprego, relatório este que todas as firmas formais são obrigadas a responder no final de cada ano. É um registro administrativo anual que permite identificar a localização dos estabelecimentos e dos trabalhadores formalmente inseridos no mercado de trabalho, assim como, característica dos trabalhadores, da ocupação, do vínculo, entre outras informações. A cobertura da base deve abranger, portanto, a população dos estabelecimentos e trabalhadores formais do país, correspondendo a uma espécie de censo do mercado de trabalho formal brasileiro.

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; Goiânia e Brasília. O recorte setorial é para a indústria da transformação de Alimentos e Bebidas (CNAE 15), divisão de atividade a 2-dígitos de acordo com a Classificação Nacional de Atividades Econômicas (CNAE).

Quanto as variáveis explicativas consideradas na análise, o quadro 1 descreve cada uma delas com sua respectiva fonte e sinal esperado encontrados nos estudos empíricos citados a seguir.

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 número 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

¹⁵ Um modo alternativo de se obter evidências das economias de aglomeração geradas por firmas espacialmente concentradas é através de estudos de casos que busquem mensurar o impacto de choques exógenos, como eventos naturais, sobre a concentração da atividade produtiva, um campo ainda pouco explorado em estudos de economia regional e urbana. Combes e Gobillon (2014) fazem uma discussão dos estudos com este escopo de pesquisa.

¹⁶ Maiores informações sobre esta base constam em <http://www.mte.gov.br/pdet>

Renda Média das Famílias	Somatório da renda de todas as famílias dividido pelo número de famílias nas principais capitais	(+)	Censo 2010
Setor Agrícola: Quociente Locacional $QL_{ij} = \frac{\frac{E_{ij}}{E_i}}{\frac{E_{.j}}{E_{..}}}$	Participação dos estabelecimentos do município j na indústria i ($\frac{E_{ij}}{E_i}$) dividido pela participação em todas as indústrias ($\frac{E_{.j}}{E_{..}}$)	(+)	RAIS 2010
Setor Alojamento e Alimentação: Quociente Locacional $QL_{ij} = \frac{\frac{E_{ij}}{E_i}}{\frac{E_{.j}}{E_{..}}}$	Participação dos estabelecimentos do município j na indústria i ($\frac{E_{ij}}{E_i}$) dividido pela participação em todas as indústrias ($\frac{E_{.j}}{E_{..}}$)	(+)	RAIS 2010
Medida de Concorrência $MC = \left(\frac{Est_{ij}}{Est_i} \right) / \left(\frac{E_{ij}}{E_i} \right)$	Participação da cidade i no total dos estabelecimentos da indústria de alimentos e bebidas ($\frac{Est_{ij}}{Est_i}$) dividido pela Participação da cidade i no emprego total da indústria de alimentos e bebidas ($\frac{E_{ij}}{E_i}$)	(-)	RAIS 2010
Índice de <i>Herfindahl</i> $HERF = \sum_{i=15}^{37} (PART)^2$	Somatório da participação no emprego de cada divisão de atividades da indústria da transformação elevado ao quadrado	(-)	RAIS 2010
Participação no Emprego (PART)	Participação da indústria i no emprego da cidade j	(+)	RAIS 2010
TamanhoXEscolha Locacional	Número de Empregos por Firma multiplicado pela <i>dummy</i> da escolha locacional	(-)	RAIS 2010
Escolaridade(Anos)XEscolha Locacional	Nível de Instrução Médio dos trabalhadores da firma multiplicado pela <i>dummy</i> da escolha locacional	(-)	RAIS 2010

Fonte: Elaboração Própria

As variáveis *proxies* para o custo de produção, o salário médio pago na indústria e o 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 apresentem os menores custos de produção, considerando os demais fatores constantes. Nesse sentido, boa parte das evidências empíricas para outros países, de fato, obtêm indicações que corroboram a suposição 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) obtêm indicações para as indústrias tradicionais e de bens intermediários do Estado de São Paulo que os estabelecimentos industriais são atraídos para se instalarem nas cidades com os maiores salários. Com respeito ao custo da mão-de-obra alocada na indústria, é razoável supor que as menos mecanizadas, de menor intensidade tecnológica ou mais intensivas em mão-de-obra, sejam atraídas para as cidades que apresentem um menor salário médio para os trabalhadores alocados no mesmo grupo de atividades, sendo este fator uma vantagem competitiva. Quanto ao custo do aluguel, Figueiredo *et al.* (2002) e Hansen (1987) encontram evidências empíricas que este é um fator de repulsão das indústrias.

Espera-se que os indicadores do potencial do mercado, aqui captados através do salário médio da cidade e da renda média das famílias da cidade, apresentem sinais positivos indicando que as firmas são atraídas a se instalarem nas cidades onde a população possivelmente tenha mais condições financeiras de consumir uma maior variedade e qualidade de produtos alimentícios. Coughlin et. al. (1991) encontram evidências de que as indústrias são atraídas para as cidades Norte-Americanas que apresentam uma maior renda per capita.

As variáveis locais que buscam captar a influência das economias de localização e urbanização para a escolha locacional das firmas devem atuar como fatores de aglomeração da atividade produtiva (Hansen, 1987; Head et al.; 1995; Figueiredo et al., 2002; Arauzo e Manjon, 2004). Seguindo a literatura aqui citada,

a participação da indústria de alimentos e bebidas no total do emprego da cidade objetiva mensurar as economias de localização, os ganhos produtivos obtidos pela interação da indústria com outras especializadas na mesma atividade produtiva. O índice de *Hirschman-Herfindahl* (H-H), que mensura o grau da diversidade produtiva da cidade, é utilizado como *proxy* das economias de urbanização. Quanto menor o índice, maior é a diversidade produtiva da cidade. A suposição é que as firmas podem se beneficiar por estarem localizadas em cidades que apresentam uma maior diversidade produtiva, as quais teriam ganhos de produtividade pela interação com firmas de diferentes grupos de atividades (Jacobs, 1969). Diferentemente das economias de localização, que se tem um consenso na literatura quando a sua mensuração, para as economias de urbanização, há diferentes especificações. Tem-se, por exemplo, os estudos que utilizam como variável *proxy* a concentração do emprego por km² (Hansen; 1987), os que consideram o índice de diversidade produtiva de H-H (Galinari et. al., 2007; Dalbertoa et. al, 2013), entre outros indicadores.

Como controles para as conexões de mercado para trás, acesso a insumos, considerou-se o QL dos estabelecimentos do setor agrícola da cidade (QL), setor este que segundo a matriz de insumo-produto do IBGE do ano 2005 apresentou o maior coeficiente técnico como fornecedor de insumos para a Indústria de Alimentos e Bebidas. Utilizando a mesma ideia, o QL dos estabelecimentos de serviços de alojamento e alimentação considerado no modelo empírico pretende captar as conexões de mercado para frente, já que este é o setor que apresentou maior conexão de demanda segundo coeficiente técnico da referida matriz de insumo-produto.

Para contrabalancear esses fatores, que presumivelmente atuariam como fatores de aglomeração das indústrias, incorporou-se no modelo uma medida de concorrência sugerida por Glaeser et. al. (1992), a razão entre a participação da cidade *i* no total dos estabelecimentos da indústria de alimentos e bebidas e a participação no emprego total gerado pela indústria de alimentos e bebidas.

Os controles para as características do estabelecimento, o tamanho (pessoal ocupado) e o nível de escolaridade médio dos trabalhadores, são interagidos com as variáveis *dummies* de escolha locacional, de modo a ser possível a estimação dos modelos de escolha discreta que exige que as variáveis variem pelas *js* alternativas de localização. Porém, tais variáveis têm importantes interpretações. Araujo e Manjon (2004) utilizando estas mesmas variáveis obtêm indicações para cidades da Espanha de que os estabelecimentos maiores tendem a se localizarem 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 entre as variáveis *dummies* de localização 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 A1 apresenta as informações das variáveis explicativas por cidade.

5. Resultados

A tabela 7 apresenta os resultados dos modelos estimados com base na equação 13, a partir do qual é possível obter indicações da influência de atributos locais/características das cidades, potencialmente gerados de economias de localização e urbanização, para as escolhas locacionais das indústrias de alimentos e bebidas; se atuam como fator de atração ou repulsão desses estabelecimentos. Quanto a interpretação dos coeficientes, no modelo aqui estimado um coeficiente positivo indica que a variável é um fator de aglomeração das indústrias consideradas no estudo, e se for negativo, um fator de dispersão. No primeiro caso, haveria indicações de que as indústrias estariam mais propensas a se localizarem nas cidades que apresentam uma maior dotação do atributo local, já que atuaria incrementando a produtividade da indústria, se configurando, portanto, em uma força de aglomeração.

Inicialmente a equação 13 é estimada considerando a especificação mais simples para o modelo de escolha discreta, o *logit condicional*, o qual assume que a função de distribuição dos parâmetros do modelo, β , é “degenerada”. A partir dos resultados do *logit* condicional, a hipótese de coeficientes fixos conhecida

como Hipótese da Independência das Alternativas Irrelevantes (IIA) foi testada através do teste de *Hausman* e foi rejeitada ao nível de confiança de 5%¹⁷.

Desse modo, procedeu-se as estimações flexibilizando a forma funcional da distribuição dos parâmetros do modelo *logit* condicional, e considerou-se a distribuição normal para a maioria dos parâmetros. Neste caso, são estimados a média (β) e o desvio padrão (σ) que descrevem a distribuição das preferências locais das firmas por cada atributo local que segue uma distribuição normal. A significância estatística dos coeficientes estimados do desvio (ver segunda parte da tabela 7) apenas reforçam a importância de considerar a heterogeneidade da importância dos fatores locais para a escolha locacional dos investidores/firmas.

No primeiro modelo (coluna 1) são considerados os atributos locais que devem se correlacionar positivamente com a produtividade das firmas (QL Agrícola; QL Alojamento e Serviços; Índice de H-H; Participação no Emprego), e a variável *proxy* para concorrência, fator que deve contrabalancear as forças que atuam a favor da aglomeração das indústrias. Um segundo modelo (coluna 2) adiciona-se controles para os custos de produção (Salário Manufatura e Aluguel) e para o potencial de mercado (renda das famílias). E, buscando controlar pelas características produtivas internas as firmas, um terceiro modelo foi estimado (coluna 3) incorporando variáveis do nível de escolaridade médio dos trabalhadores (*grinstru*) e tamanho da firma (*tamestab_city*) interagido com a *dummy* da escolha locacional (*city*) de forma a ser possível estimar o modelo, dado que o mesmo exige variabilidade das variáveis entre as alternativas de escolhas.

Quanto ao ajuste do modelo, a maioria das variáveis apresentam coeficientes significantes a menos de 5%, e os coeficientes mostram-se conjuntamente significantes a 5%. Ressalta-se, porém, a importância da correta especificação do modelo, dado os diferentes resultados obtidos quando não se controla pelas características individuais das firmas, modelo 1 e modelo 2.

Focando, assim, a análise nos resultados do terceiro modelo, o com melhor especificação¹⁸, há indicações que tanto as economias de localização, quanto as de urbanização, são importantes para explicar as escolhas locais das indústrias de alimentos e bebidas consideradas na análise. Ou seja, quanto maior for a diversidade produtiva da cidade e a concentração relativa de empregos da indústria de alimentos e bebidas, maior será a probabilidade de a cidade ser escolhida pelo investidor.

Os resultados também sugerem que as indústrias são atraídas para se instalarem nas cidades onde o setor agrícola tem uma maior importância relativa, sugerindo que as conexões para trás estão correlacionadas positivamente com a probabilidade de escolha locacional das indústrias de alimentos e bebidas. Esses resultados corroboram com o fato da indústria ser dependente de uma cadeia de produção agrícola que geram altos custos de transporte e, portanto, a proximidade das indústrias aos seus fornecedores de matéria-prima deve atuar como um fator de atração dessa atividade. Corroborando a este resultado, porém com outro propósito de pesquisa, Rocha et. al. (2011) mostra que o modelo de vantagens naturais se adequa melhor para explicar a aglomeração da indústria de alimentos e bebidas.

Já os resultados para a concentração relativa de estabelecimentos de alojamentos e alimentação, considerado aqui como *proxy* das conexões de demanda para frente, apresentou coeficiente negativo e estatisticamente significativo, o que não é coerente com a suposição de que as indústrias de alimentos e bebidas teriam maiores vantagens produtivas por estarem instaladas em cidades com um maior percentual de estabelecimentos de alojamentos e alimentação, relativamente aos demais setores. Porém, há indicações de que a renda média da população da cidade, variável que também capta as conexões de demanda para frente, o acesso ao mercado, é um fator de aglomeração dessas indústrias.

Contrabalanceando esses efeitos, a variável concorrência, como esperado, é um fator de dispersão das indústrias consideradas na análise.

As variáveis que buscam mensurar os custos de produção das firmas, salário médio da indústria e o preço do aluguel apresentaram coeficientes estatisticamente significantes a menos de 5%. Contrariando a

¹⁷ Resultados encontram-se no apêndice.

¹⁸ O teste da razão da verossimilhança, aplicado ao modelo 1 versus o modelo 3, $2(11851,1-12019,2) = 336,2$, com 23 graus de liberdade e com um nível crítico a 5% (35,17), rejeitou-se o modelo 1 em relação ao modelo 3. Assim como, o modelo 1 versus o modelo 3, $2(11851,1-11967,6) = 233$, com 20 graus de liberdade e com um nível crítico a 5% (31,41), rejeitou-se o modelo 2 em relação ao modelo 3.

teoria, os resultados indicam que as firmas da amostra considerada na análise são atraídas a se estabelecerem naquelas cidades que pagam um salário médio maior para os trabalhadores da indústria de alimentos e bebidas. 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 a disponibilidade de trabalhadores mais qualificados. Neste caso, 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 relacionado a salários mais elevados. O sinal negativo do custo do aluguel, conforme esperado, indica que este é um fator de repulsão das indústrias.

Ademais, as variáveis de interação incluídas no modelo assinalam que as indústrias de alimentos e bebidas são atraídas para se instalarem na cidade de São Paulo e, possivelmente, os ganhos produtivos seriam maiores para as indústrias de grande porte, dado que os coeficientes observados são negativos com base na categoria de referência (São Paulo).

Tabela 7. Resultados do Modelo *Mixed Logit* para a Indústria de Alimentos e Bebidas
(Variável Dependente – Escolha Locacional)

Variáveis	Modelo (1)	Modelo (2)	Modelo (3)
Média			
Aluguel		-0,012 *** (0,001)	-0,020 *** (0,006)
Salário Manufatura		0,001 *** (0,000)	0,003 *** (0,000)
Renda das Famílias		0,001 *** (0,000)	0,001 *** (,000)
QL Agrícola	-15,040 *** (1,182)	1,629 * (0,839)	17,797 *** (3,456)
QL Alojamento e Hotelaria	2,516 *** (0,315)	0,318 (0,383)	-5,944 *** (2,184)
Índice H-H	-10,385 *** (2,506)	3,548 ** (1,675)	-80,60 *** (28,57)
Participação Emprego	-4,182 *** (0,679)	-7,358 *** (0,540)	15,563 *** (3,837)
Concorrência	0,910 *** (0,066)	-0,520 *** (0,106)	-0,973 ** (0,484)
Tamanho*Belém			-0,191 (0,219)
Tamanho *Fortaleza			-0,274 (0,337)
Tamanho* Recife			-0,334 *** (0,127)
Tamanho * Salvador			-0,338 *** (0,067)
Tamanho* Belo Horizonte			-0,117 *** (0,036)
Tamanho*Rio de Janeiro			0,016 (0,035)
Tamanho*Curitiba			-0,174 *** (0,040)
Tamanho*Porto Alegre			-0,242 *** (0,048)

Tamanho*Goiânia	-0,357 ***
	(0,119)
Tamanho*Brasília	-0,425 ***
	(0,162)
Escolaridade*Belém	-0,770 **
	(0,348)
Escolaridade *Fortaleza	-1,298 *
	(0,651)
Escolaridade* Recife	-0,035
	(0,151)
Escolaridade * Salvador	0,338 ***
	(0,085)
Escolaridade* Belo Horizonte	-0,199 ***
	(0,051)
Escolaridade*Rio de Janeiro	-0,337 ***
	(0,048)
Escolaridade*Curitiba	0,160 ***
	(0,053)
Escolaridade*Porto Alegre	0,098 **
	(0,047)
Escolaridade*Goiânia	-0,084
	(0,140)
Escolaridade*Brasília	-0,203
	(0,211)

	Desvio Padrão		
QL Agrícola	29,677 ***	-1,890	-1,261
	(1,560)	(3,077)	(1,568)
QL Alojamento e Hotelaria	-1,489	-0,855	-0,685
	(1,167) ***	(2,604)	(2,125)
Índice de Herfindahl	20,957	3,652	175,98 ***
	(3,031)	(3,926)	(43,29)
Participação Emprego	16,452 ***	11,118 ***	-3,693
	(0,803)	(0,901)	(2,303)
Concorrência	0,137	0,884 ***	0,072
	(0,226)	(0,194)	(0,272)
Log-Verossimilhança	-12019,2	-11967,6	-11851,1
LR chi2	477,92 ***	101,77 ***	36,28 ***
N	57156	57156	57156

Nota: * coeficiente significativo à 10%; ** significativo a 5%;***significativo a 1%

Com base nos resultados e considerando as características intrínsecas das Indústrias de Alimentos e Bebidas do Brasil – pois possivelmente há diferenças do grau da intensidade tecnológica entre os países –, há indicações de que os estabelecimentos considerados na pesquisa têm ganhos produtivos por se situarem próximos aos seus fornecedores. Os resultados também são coerentes com o a existência de economias de localização e urbanização geradas por firmas espacialmente concentradas.

Vale salientar que estudos que buscam obter evidências da importância das economias de aglomeração geradas pela concentração especial da atividade produtiva e/ou pela diversidade da economia são relativamente recentes e têm avançando no sentido de identificar as fontes destas forças aglomerativas.

Um primeiro avanço foi a adaptação dos modelos de escolha discreta para o contexto da localização das firmas, e a possibilidade de flexibilizar o modelo *logit* condicional de modo a captar a influência da heterogeneidade das firmas. Porém, ainda é preciso deixar claro as limitações presentes nesses modelos. A estimação pode não ser viável se o conjunto de opções de escolha for muito grande. Outra limitação está na possibilidade de se omitir alguma variável importante para a análise e ter-se o termo de erro correlacionado com a (s) variável (s) explicativa (s).

A este respeito, caso os coeficientes sejam corretamente identificados, poder-se-ia tecer afirmações sobre o peso que cada atributo local teria sobre a função lucro da firma, e a partir dessas informações construir *rankings* para as cidades de acordo com um índice que mensurasse a qualidade das amenidades locais para as firmas. Este é um desafio futuro dessa pesquisa.

6. Considerações Finais

A presente pesquisa buscou obter evidências da importância dos fatores locais, especialmente aqueles que geram externalidades produtivas para as firmas, já consagrados na literatura, para explicar a escolha locacional das indústrias de Alimentos e Bebidas nas principais cidades brasileiras. A investigação empírica baseou-se em modelos de escolha discreta, *Logit* condicional e *Mixed Logit*, estimados a partir dos microdados da RAIS-Mte para o ano de 2010.

Os resultados indicaram para importância de flexibilizar a hipótese da independência das alternativas irrelevantes para captar a heterogeneidade das preferências locacionais das firmas e, assim, o modelo *mixed logit* foi estimado com respaldo no teste de Hausman. De acordo com os resultados, há evidências de que as cidades com uma maior concentração relativa de indústrias de alimentos e bebidas, assim como, com uma maior diversidade industrial atuam como polo de atração das industriais de Alimentos e Bebidas, sugerindo que as economias de localização e urbanização são importantes para explicar as escolhas locacionais dessas indústrias. Além disso, há indicações de que as conexões de mercado para trás são importantes para explicar a distribuição espacial das indústrias, e são forças que atuam a favor da concentração industrial. Por outro lado, cidades que apresentam um ambiente de maior concorrência para as indústrias seriam preteridas pelas indústrias em análise.

A respeito dos custos de produção, os resultados sugerem que o custo do aluguel atua como fator de repulsão da atividade produtiva, mas as indústrias seriam atraídas a se instalarem nas cidades que apresentam uma maior média salarial para a indústria de alimentos e bebidas.

Os resultados também apontam que, mesmo após estes controles, as indústrias de alimentos e bebidas são mais propensas em se instalarem na cidade de São Paulo, em comparação com as demais cidades incluídas na análise.

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 a estimação do modelo com dados em painel para um nível de desagregação regional maior com o objetivo de obter efeitos dinâmicos da distribuição espacial das indústrias. Além disso, pretende-se avançar no sentido de estabelecer *rankings* da qualidade dos fatores locais para as localidades em estudo.

As evidências obtidas para a influência dos fatores locais, potencialmente geradores de economias de localização e urbanização nas escolhas locacionais dos investidores/indústrias, ao menos indicam que é importante se fortalecer no país as políticas voltadas para a consolidação de arranjos produtivos industriais, aproveitando as vantagens produtivas das cidades com vista à elevação da produtividade no país. Cabe assinalar que este tipo de política, de um certo modo, vai no sentido contrário as políticas de incentivos fiscais que simplesmente buscam atrair firmas para localidades que não apresentam vantagens produtivas específicas para a atividade produtiva da indústria contemplada pela política; compatível com a situação em que a cidade não teria ganhos produtivos com a instalação da (s) indústria (s), nem tão pouco a indústria teria maiores vantagens produtivas, senão aquela oriunda de uma isenção fiscal.

Referências Bibliográficas

ARAUZO-CAROD, JOSEP-MARIA, Daniel LIVIANO-SOLIS, and Miguel MANJÓN-ANTOLÍN. "Empirical Studies in Industrial Location: An Assessment of Their Methods and Results*." *Journal of Regional Science* 50.3 (2010): 685-711.

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.

BAUDEWYNS, Didier. La localisation intra-urbaine des firmes: une estimation logit multinomiale. ULB-Université Libre de Bruxelles, 1999.

BERRY, S., LEVINSOHN, J., PAKES, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841-890.

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

CARLTON, Dennis W. 1979. "Why New Firms Locate Where They Do: An Econometric Model," in W. Wheaton (ed.), *Interregional Movements and Regional Growth*. Washington: The Urban Institute.

CAROD, Josep Maria Arauzo; ANTOLÍN, Miguel C. Manjón. Firm size and geographical aggregation: an empirical appraisal in industrial location. *Small Business Economics*, v. 22, n. 3-4, p. 299-312, 2004.

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.

COMBES, Pierre-Philippe; DURANTON, Gilles; GOBILLON, Laurent. The identification of agglomeration economies. *Journal of Economic Geography*, v. 11, n. 2, p. 253-266, 2011.

COMBES, Pierre-Philippe; GOBILLON, Laurent. The empirics of agglomeration economies. 2014.

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

CROZET, Matthieu; MAYER, Thierry; MUCCHIELLI, Jean-Louis. How do firms agglomerate? A study of FDI in France. *Regional Science and Urban Economics*, v. 34, n. 1, p. 27-54, 2004.

DALBERTOA, Cassiano Ricardo; STADUTOB, Jefferson Andronio Ramundo. Uma Análise das Economias de Aglomeração e seus Efeitos sobre os Salários Industriais Brasileiros. *Rev. Econ. Contemp*, p. 539-569.

DEFEVER, F. (2012). The spatial organization of multinational firms. *Canadian Journal of Economics/Revue canadienne d'économie*, 45(2), 672-697.

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. Geographic concentration in US manufacturing industries: a dartboard approach. *National Bureau of economic research*, 1994.

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

FEIJO, Carmem Aparecida; CARVALHO, Paulo Gonzaga M.; RODRIGUEZ, Maristella Schaeffers. 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.

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

GALINARI, R.; CROCCO, M. A.; LEMOS, M. B.; BASQUES, M. F. D. O efeito das economias de aglomeração sobre os salários industriais: uma aplicação ao caso brasileiro. *Revista de Economia Contemporânea*, v. 11, n. 3, p. 391-420, set./dez. 2007.

GLAESER, Edward L. et al. Growth in Cities. *Journal of Political Economy*, p. 1126-1152, 1992.

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.

LAFOUNTAIN (2005). Where do firms locate? Testing competing models of agglomeration. *Journal of Urban Economics*, 58:338–366.

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.

MELITZ, Marc J. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, v. 71, n. 6, p. 1695-1725, 2003.

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

PUGA, Diego. The Magnitude and Causes of Agglomeration Economies. *Journal of Regional Science*, v. 50, n. 1, p. 203-219, 2010.

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

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

REVELT D. e Train K. (1998) "Mixed Logit with Repeated Choices of Appliance Efficiency Levels." *Review of Economics and Statistics*, v. LXXX, no. 4, p. 647-657.

REVELT, D. and K. Train (1999) “Customer-Specific Taste Parameters and Mixed Logit” University of California at Berkeley, Working Paper no E00-274.

ROCHA, Roberta de Moraes; NETO, Raul da Mota Silveira; GOMES, Sonia Maria Fonseca Pereira Oliveira. Maiores cidades, maiores habilidades produtivas: ganhos de aglomeração ou atração de habilitados? uma análise para as cidades brasileiras. *REN. Revista economica do nordeste*, v. 42, n. 4, p. 675-695, 2011.

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

ROSENTHAL, Stuart S.; STRANGE, William C. Geography, industrial organization, and agglomeration. *review of Economics and Statistics*, v. 85, n. 2, p. 377-393, 2003.

SILVA, Magno Vamberto Batista ; SILVEIRA NETO, Raul da Mota . Determinantes da localização industrial no brasil e geografia econômica: evidências para o período pós-real. In: X Encontro Regional de Economia, 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.

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

TRAIN, Kenneth E. Discrete choice methods with simulation. Cambridge university press, 2009.
 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

Tabela A1. Valores Observados das Variáveis Explicativas para as 11 Cidades - 2010

CIDADES	Aluguel	Renda das Famílias	Salário Manufatura	QL Agrícola	QL Alojamento e Hotelaria	Índice H-H	Concorrência	Participação Emprego
Belém	437,10	1542,67	2951,16	0,11	0,60	1,09	0,19	0,07
Fortaleza	284,00	998,02	2771,98	0,03	0,63	1,41	0,25	0,14
Recife	382,31	2093,19	3487,90	0,05	0,69	1,59	0,13	0,08
Salvador	384,99	2088,69	2843,45	0,08	0,82	1,94	0,10	0,09
Belo Horizonte	514,22	2068,20	4480,00	0,12	0,70	2,28	0,08	0,08
Rio de Janeiro	599,00	3042,95	4131,57	0,02	0,74	0,76	0,08	0,08
São Paulo	594,48	2959,50	4388,05	0,03	0,77	0,99	0,08	0,09
Curitiba	551,05	1974,69	4594,80	0,04	0,76	0,99	0,09	0,06
Porto Alegre	571,51	1970,19	4742,65	0,06	0,73	1,49	0,08	0,07
Goiânia	406,80	1178,60	4035,64	0,27	0,59	1,51	0,13	0,10
Brasília	583,26	1636,77	5397,10	0,23	0,85	1,49	0,15	0,07

Fonte: Elaboração própria.

Tabela A2. Quocientes Locacionais para os grupos da indústria de alimentos e bebidas para as principais capitais brasileiras (2010).

CIDADES	151		152		153		154		155		156		157		158		159	
	Estab	Emp	Estab	Emp	Estab	Emp	Estab	Emp	Estab	Emp	Estab	Emp	Estab	Emp	Estab	Emp	Estab	Emp
Belém – PA	6.95	0.46	3.22	1.40	1.66	7.81	1.19	0.50	0.65	0.66	0.62	0.00	0.41	0.20	1.04	1.12	1.33	3.80
Fortaleza – CE	0.49	0.05	1.25	4.75	0.95	1.23	0.89	0.80	0.81	2.35	0.00	0.00	0.09	0.01	2.03	1.86	1.14	1.98
Recife – PE	1.37	0.97	0.89	0.38	0.64	0.04	0.69	0.68	0.46	0.43	4.98	0.11	0.36	2.67	2.34	1.92	1.60	1.89
Salvador – BA	3.23	0.12	1.92	0.65	2.08	0.36	2.11	0.74	0.62	5.23	0.77	0.02	1.28	0.59	2.41	1.77	1.02	0.91
Belo Horizonte – MG	2.42	0.40	1.32	0.49	0.40	0.00	1.81	1.05	0.51	0.32	0.45	0.00	1.19	0.44	3.17	2.74	1.78	1.92
Rio de Janeiro – RJ	6.83	0.44	0.59	0.19	0.31	0.30	0.55	0.73	0.24	0.37	0.34	0.09	0.63	0.88	0.99	2.29	1.23	2.97
São Paulo – SP	1.53	0.22	0.64	0.65	1.31	2.30	1.25	2.36	0.44	0.82	0.54	0.02	0.33	0.62	1.17	2.42	0.60	1.03
Curitiba – PR	0.08	0.28	0.36	0.06	0.24	0.00	0.35	0.13	0.39	0.79	0.27	0.00	0.35	2.12	1.56	3.02	0.57	1.84
Porto Alegre – RS	3.39	0.09	1.33	0.48	1.58	0.39	0.66	0.58	0.76	0.53	0.00	0.00	0.39	0.21	2.13	2.39	1.46	4.02
Goiânia – GO	3.71	0.53	1.16	6.70	0.65	0.00	1.16	1.91	1.15	1.10	0.73	0.00	0.73	0.17	1.88	1.49	0.85	1.26
Brasília – DF	1.38	1.23	1.01	0.22	0.00	0.00	0.95	0.57	0.81	0.93	0.37	0.02	0.87	3.53	1.99	1.23	0.62	2.73

Fonte: Elaboração própria com base nos dados da Rais (2010).

Nota: E_{ij} – emprego da indústria do grupo de atividade (i) da cidade j; E_i – total do emprego do grupo de atividade (i); E_j – emprego total da indústria de alimentos da cidade j ; E – emprego total da indústria de alimentos do Brasil