

ECONOMIA do CEARÁ em DEBATE 2017

GOVERNO DO ESTADO DO CEARÁ
SECRETARIA DO PLANEJAMENTO E GESTÃO - SEPLAG
INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ - IPECE

ECONOMIA DO CEARÁ EM DEBATE 2017

GOVERNADOR DO ESTADO DO CEARÁ

Camilo Sobreira de Santana

Vice-Governadora do Estado do Ceará

Maria Izolda Cela de Arruda Coelho

SECRETARIA DO PLANEJAMENTO E GESTÃO – SEPLAG

Secretário

Francisco de Queiroz Maia Júnior

Secretário adjunto

Antônio Sérgio Montenegro Cavalcante

Secretário executivo

Júlio Cavalcante Neto

INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ – IPECE

Diretor Geral

Flávio Ataliba Flexa Daltro Barreto

Diretor de Estudos Econômicos - DIEC

Adriano Sarquis Bezerra de Menezes

Diretor de Estudos Sociais – DISOC

João Mário de França

Diretor de Estudos de Gestão Pública - DIGEP

Cláudio André Gondim Nogueira

Gerência de Estatística, Geografia e Informação - GEGIN

Mário Aragão

ECONOMIA DO CEARÁ EM DEBATE 2017

2017 v-1 - 2015

Flávio Ataliba Flexa Daltro Barreto, João Mário de França, Mário Aragão, Fátima Juvenal de Sousa (organizadores).

Anual

ISBN: 978-85-98664-40-8

1. Economia. 2. Ceará. I - Daltro, Barreto, Flávio, Ataliba, Flexa.

CDU 330 (813.1) (05)

304 páginas.

Copyright © 2018 - IPECE.

Os artigos apresentados neste livro são de inteira responsabilidade dos seus autores. As opiniões neles emitidas não exprimem, necessariamente, o ponto de vista do Instituto de Pesquisa e Estratégia Econômica do Ceará - IPECE.

GOVERNO DO ESTADO DO CEARÁ
SECRETARIA DO PLANEJAMENTO E GESTÃO - SEPLAG
INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ - IPECE

ECONOMIA DO CEARÁ EM DEBATE 2017

Organizadores

Flávio Ataliba Flexa Daltro Barreto

João Mário de França

Mário Aragão

Fátima Juvenal de Sousa

Conselho Editorial

Flávio Ataliba Flexa Daltro Barreto (IPECE / CAEN-UFC)

João Mário de França (IPECE / CAEN-UFC)

Carlos Alberto Manso (LEP-CAEN-UFC)

Leando de Almeida Rocco (CAEN-UFC)

Ricardo Brito Soares (CAEN-UFC)



INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ – IPECE

Av. General Afonso Albuquerque Lima, S/N - Cambéba - Edifício SEPLAG - Térreo
CEP 60.822-325 - Fortaleza / Ceará

Fones: (85) 3101-3496 | 3101-3521 - Fax: (85) 3101-3500

www.ipece.ce.gov.br

APRESENTAÇÃO

O livro *Economia do Ceará em Debate 2017* é uma coletânea dos onze artigos apresentados por ocasião do XIII Encontro Economia do Ceará em Debate, realizado pelo Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE), no dia 10 de novembro de 2017, no auditório da Secretaria do Planejamento e Gestão (SEPLAG).

O livro que apresentamos à sociedade é fruto do Encontro, bem como do trabalho interativo entre Governo do Estado e à academia, que contou com o envolvimento intenso dos colaboradores do IPECE. De forma a cobrir os principais temas de pesquisa em economia com foco no estado do Ceará, e conceber oportunidade para que estudiosos possam discutir os principais desafios da economia cearense, bem como possibilitar às autoridades públicas responsáveis pelo desenvolvimento do Estado, conhecer ideias inovadoras sobre as ações que têm efeitos diretos sobre a economia e a sociedade cearense.

Agradecemos o apoio recebido do Governo do Estado do Ceará, por intermédio da SEPLAG. Devemos igualmente gratidão ao Bradesco e Banco do Nordeste do Brasil pelo apoio financeiro para a premiação.

Flávio Ataliba Flexa Daltro Barreto
Diretor Geral do IPECE

SUMÁRIO

DETERMINANTES DO SOBREPESO E OBESIDADE ENTRE HOMENS E MULHERES NO CEARÁ

Natália Cecília de França, Diógenes Alberto de França 9

INVESTIGAÇÃO E JULGAMENTO DOS HOMICÍDIOS DOLOSOS NO MUNICÍPIO DE FORTALEZA

Denise Xavier Araújo de Oliveira, Andrei Gomes Simonassi 34

AGLOMERAÇÕES PRODUTIVAS NOS SETORES DE ATIVIDADE ECONOMICA DO CEARÁ: ANÁLISE TEMPORAL E ESPACIAL

Marcos Antônio Feijó Nagaki, Nicolino Trompieri Neto 63

DETERMINANTES DO RENDIMENTO ESCOLAR NAS ESCOLAS PÚBLICAS DE ENSINO MÉDIO DO ESTADO DO CEARÁ: UMA ANÁLISE DE PAINEL SOBRE O SPAECE DE 2008 A 2014

Vitor Borges Monteiro, Pablo Urano de Carvalho Castelar, João Benicio Sousa 93

UMA ANÁLISE ESPACIAL DA POBREZA MULTIDIMENSIONAL NOS BAIRROS DE FORTALEZA - CEARÁ

Ricardo Candéa Sá Barreto, Eli Izidro dos Santos, Ícaro Célio Santos de Carvalho..... 121

AVALIAÇÃO DE POLÍTICAS EDUCACIONAIS: UM ESTUDO DE CASO DA EDUCAÇÃO EM TEMPO INTEGRAL SOBRE O DESEMPENHO DOS ALUNOS DO ENSINO MÉDIO DO ESTADO DO CEARÁ PARA O ANO DE 2014

Antonio Bruno Fernandes Marcelino, Wellington Ribeiro Justo, Matheus Oliveira de Alencar..... 153

DILEMA EQUIDADE-EFICIÊNCIA PROPOSTA METODOLÓGICA E EVIDÊNCIAS EMPÍRICAS PARA O CEARÁ

Lyana Maria França da Costa Ribeiro, Christiano Modesto Penna 184

EDUCAÇÃO NO ESTADO DO CEARÁ: ANÁLISES UTILIZANDO O MÉTODO DO CONTROLE SINTÉTICO

Francisco Germano Carvalho Lúcio, Natanael Soares Leite 207

UMA ANÁLISE DOS DETERMINANTES SOCIOECONÔMICOS DA (IN)SEGURANÇA ALIMENTAR EM DOMICÍLIOS CEARENSES

Maria Hosana P. Moreira, Vitor Hugo Miro 230

ANÁLISE DA SIMULTANEIDADE ENTRE DESENVOLVIMENTO ECONÔMICO, TRANSPARÊNCIA MUNICIPAL E DESIGUALDADE DE RENDA PARA O CEARÁ

Márcio Aurélio Frota Pereira, Ronaldo de Albuquerque e Arraes 259

DETERMINANTES DA PARTICIPAÇÃO DOS CASAIS NO MERCADO DE TRABALHO: ANÁLISE PARA A REGIÃO NORDESTE E O ESTADO DO CEARÁ

Daniel Tomaz de Sousa, Maria Adreciana Silva de Aguiar, Francisco Germano Carvalho Lúcio, Priscila Silva Rodrigues 288

DETERMINANTES DO SOBREPESO E OBESIDADE ENTRE HOMENS E MULHERES NO CEARÁ

Natália Cecília de França*
Diógenes Alberto de França**

RESUMO

Sobrepeso e obesidade têm se tornado uma epidemia mundial ao longo dos últimos anos, configurando sério problema de saúde. Neste sentido, o presente estudo tem como objetivo contribuir com a literatura sobre tal agravo entre a população adulta cearense. Em primeiro lugar, fez-se uma avaliação dos determinantes do excesso de peso entre homens e mulheres cearenses por meio da estimação de um modelo Logit ordenado generalizado. Os resultados indicam, por exemplo, que idade, renda domiciliar per capita, consumo elevado de sal, consumo de frutas e mais horas assistindo televisão contribuíram para a prevalência do excesso de peso em ambos os sexos. Entre as mulheres, a escolaridade auxiliou na redução da obesidade, já entre os homens não se observou uma tendência clara. Na segunda etapa do trabalho, foram implementados diferentes algoritmos de mineração de dados na classificação antropométrica da população adulta residente no Ceará, cuja acurácia foi comparada com a do modelo Logit binário. A construção das árvores de decisão indicam que os principais determinantes da prevalência do excesso de peso entre homens e mulheres foram, respectivamente, a situação matrimonial e a condição de ocupação no mercado de trabalho. Ademais, os algoritmos de mineração de dados não apresentaram maior acurácia quando comparados com o modelo Logit binário.

Palavras-Chave: Obesidade, Determinantes, Classificação, Mineração de Dados.

DETERMINANTS OF OVERWEIGHT AND OBESITY BETWEEN MEN AND WOMEN IN CEARÁ

ABSTRACT

Overweight and obesity have become a worldwide epidemic over the past few years, configuring a serious health problem. In this sense, the present study aims to contribute

* Doutoranda em Economia, CAEN/UFC. nc.franca@hotmail.com. +1(305)4907631; +55(85)997664043.

** Tecnólogo em jogos Digitais, PUC MINAS. diogenesaf@yahoo.com.br. Av. da Universidade, 2700. 60020-181. Fortaleza – CE

with the literature on such an aggravation among the adult population of Ceará. Firstly, an evaluation of the determinants of overweight among men and women from Ceará was made through the estimation of a generalized ordered Logit model. The results indicate, for example, that age, per capita household income, high salt intake, fruit consumption and more hours watching television contributed to the prevalence of overweight in both sexes. Among women, schooling contributed to the reduction of obesity; among men, a clear trend was not observed. In the second stage of the work, different algorithms of data mining were implemented in the anthropometric classification of the adult population resident in Ceará, whose accuracy was compared with that of the Logit model. The construction of decision trees indicates that the main determinants of the prevalence of overweight among men and women were, respectively, the marital status and the occupation status in the labor market. In addition, the data mining algorithms were not more accurate when compared with the binary Logit model.

Keywords: Obesity, Determinants, Classification, Data Mining.

Classificação JEL: C81, I12.

1 INTRODUÇÃO

Sobrepeso e obesidade têm se tornado uma epidemia mundial nos últimos anos, além de se configurar um sério problema de saúde pública. Segundo dados da Organização Mundial de Saúde (OMS), a incidência mundial de obesidade dobrou desde 1980¹. Em 2014, 39% dos adultos com mais de 18 anos de idade estavam com sobrepeso e 13% eram obesos. O Brasil segue essa mesma tendência de aumento da obesidade. Resultados do relatório da Organização das Nações Unidas para Alimentação e Agricultura e da Organização Pan-americana de Saúde (2017) mostram que, em 2010, 17,8% da população brasileira era obesa; passando para 20% em 2014. Dentre os estados brasileiros, o Ceará, em 2013, foi o décimo em termos de prevalência de sobrepeso, conforme dados da Pesquisa Nacional de Saúde (PNS) realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Naquele ano, 57,1% dos cearenses acima de 18 anos de idade estava com sobrepeso e 18,7% era o percentual de obesos.

Excesso de peso corporal se configura como um dos fatores de risco às doenças crônicas não transmissíveis (DCNT)², responsáveis por 72,6% das mortes prematuras entre os brasileiros em 2013 (MS, 2015). A expansão da obesidade pode ter contribuído, por exem-

1. <http://www.who.int/mediacentre/factsheets/fs311/en/>. Acesso em 27 de setembro de 2017.

2 Exemplos de DCNT são doenças cardiovasculares, diabetes, câncer e doenças respiratórias crônicas.

plo, para o aumento da incidência de diabetes e hipertensão no Brasil na última década. Dados do Inquérito Telefônico de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas (Vigitel) indicam que, entre 2006 e 2016, o diagnóstico médico de diabetes no Brasil foi de 5,5% para 8,9%; e o de hipertensão, de 22,5% para 25,7%.

Além de ser fator de risco às DCNT, sobrepeso e obesidade também geram perdas econômicas. Os custos mundiais com obesidade representavam 2,8% do Produto Interno Bruto (PIB) global em 2012, de acordo com um estudo realizado pelo McKinsey Global Institute (2014). O mesmo estudo aponta que a obesidade era o terceiro maior encargo social no Brasil, com gastos que correspondiam a 2,4% do PIB nacional, ficando atrás apenas dos custos com mortes violentas e alcoolismo.

O recente boom da obesidade pode ser explicado pelo aumento da inatividade física, fatores genéticos e hormonais, crescimento econômico, urbanização, mudança nos hábitos alimentares reduzindo o consumo de alimentos saudáveis e aumentando o de alimentos ultraprocessados, etc. Dados do Vigitel mostram, por exemplo, que o consumo regular (ao menos 5 dias da semana) de feijão, alimento típico da dieta do brasileiro, caiu entre 2012 e 2016, passando de 67,5% para 61,3%.

Uma das medidas adotadas pelo Ministério da Saúde (MS) no enfrentamento da obesidade foi o lançamento, em 2014, do novo Guia Alimentar para a População Brasileira, que trás recomendações para uma alimentação saudável. Nesta mesma linha, o Governo Federal também lançou a campanha “Brasil Saudável e Sustentável” para mostrar as vantagens de uma alimentação saudável. O controle da obesidade também está presente entre as metas do “Plano de Ações Estratégicas para o Enfrentamento das Doenças Crônicas Não Transmissíveis (DCNT) no Brasil 2011 – 2022”, elaborado pelo MS em parceria com diversas outras instituições.

Neste sentido, este estudo visa contribuir com a literatura sobre os determinantes do sobrepeso e da obesidade entre a população adulta residente no Ceará. Para tal, foram estimados modelos Logit ordenados desagregados por sexo. A prevalência do excesso de peso aumentou com a idade e a renda domiciliar per capita em ambos os sexos. Alimentação saudável foi fator protetor contra obesidade entre homens e mulheres. Ainda em relação ao comportamento dos indivíduos, o consumo de frutas, o consumo excessivo de sal e o alto número de horas assistindo televisão contribuíram para o excesso de peso. Entre as mulheres, a escolaridade apresentou relação negativa com a prevalência do excesso de peso, o que não se observou entre os homens.

Ademais, foram utilizados algoritmos de mineração de dados na classificação antropométrica de homens e mulheres residentes no Ceará. Um bom algoritmo de classificação se mostra relevante na formulação de políticas públicas de forma mais eficiente e eficaz. De acordo com as árvores de decisão, a situação matrimonial foi o principal determinante do excesso de peso entre os homens; e a condição de ocupação no mercado de trabalho foi o principal determinante entre as mulheres. Os algoritmos utilizados neste trabalho não apresentaram ganhos em termos de acurácia quando comparados ao modelo Logit.

O presente estudo, além desta introdução é composto por uma seção sobre os determinantes da obesidade e outra que trata dos algoritmos de mineração de dados utilizados na classificação da obesidade. Mais adiante são mostrados os aspectos metodológicos, os resultados e as considerações finais encerram o trabalho.

2 DETERMINANTES SOCIOECONÔMICOS DA OBESIDADE

O entendimento sobre os determinantes do sobrepeso e da obesidade é muito importante ao ser capaz de gerar informações úteis para os formuladores de políticas públicas. Inatividade física, mudanças nos padrões de consumo, escolaridade, genética, dentre outros são fatores capazes de afetar a classificação antropométrica das pessoas.

Roemling e Qaim (2012) fizeram uma análise sobre a evolução da transição nutricional vivenciada na Indonésia entre 1993 e 2007. Os autores constataram que, seguindo a tendência mundial, os níveis de obesidade aumentaram no País, principalmente entre as mulheres. Além disso, seus resultados indicam que mudanças nos padrões alimentares e reduções nos níveis de atividade física foram fatores importantes para o aumento no índice de massa corporal (IMC) da população.

Ao fazer uma análise multinível dos determinantes da obesidade entre a população adulta chinesa em 2006, Zhang e Liao (2013) ressaltam que a renda tem impacto positivo na incidência de obesidade. Por outro lado, os autores destacam que escolaridade e ocupação manual se relacionam de maneira negativa com obesidade. Eles definiram como ocupação manual os cargos de fazendeiros, pescadores, caçadores, soldados e trabalhadores não qualificados. Em geral, pessoas alocadas em tais funções exercem algum grau de atividade física durante o trabalho, contribuindo, assim, para a menor chance de serem obesas em relação aos demais trabalhadores.

Em âmbito nacional, Andrade *et al.* (2015) constataram que a obesidade entre homens e mulheres residentes na área urbana de Belo Horizonte (BH) é influenciada por fato-

res distintos. Por exemplo, maior nível de escolaridade reduziu a prevalência de sobrepeso entre as mulheres e foi fator de risco entre os homens. Além disso, o consumo de refrigerantes dietéticos exerceu impacto positivo na prevalência de excesso de peso para ambos os sexos. Em outra análise sobre a população adulta de BH, Ward *et al.* (2015) também destacam a diferenciação dos determinantes da obesidade entre homens e mulheres. Segundo os autores, escolaridade se relacionou de maneira negativa com o IMC e a obesidade entre as mulheres, enquanto que as rendas do domicílio e da vizinhança tiveram impacto positivo entre os homens.

Ainda sobre a distinção dos fatores determinantes da obesidade entre os sexos, Siqueira *et al.* (2015) fizeram estudo sobre a obesidade abdominal na população com mais de 40 anos de idade residente no município de Cambé, Paraná. Os resultados indicam que a prevalência da obesidade se revelou bem mais acentuada entre as mulheres do que entre os homens. Fatores como idade, inatividade física no lazer e hipertensão arterial tiveram impacto positivo na obesidade em ambos os sexos. Por sua vez, classe econômica, tabagismo e diabetes se mostraram significativos apenas entre as mulheres.

Correia *et al.* (2011) analisaram os determinantes da obesidade entre as mulheres em idade reprodutiva residentes no Ceará. Por meio da utilização de análise multivariada, os autores salientam que fatores como ter idade acima de trinta anos, ser casada ou utilizar contraceptivos tiveram impacto positivo na obesidade. Ainda em relação às mulheres em idade reprodutiva, Costa *et al.* (2014) fizeram uma análise da evolução, entre 1997 e 2006, do excesso de peso e seus determinantes entre aquelas residentes em Pernambuco. Mais uma vez, a baixa escolaridade foi apontada como fator de risco para a prevalência de tal agravo. Idade e maior número de gestações também contribuíram positivamente para a obesidade.

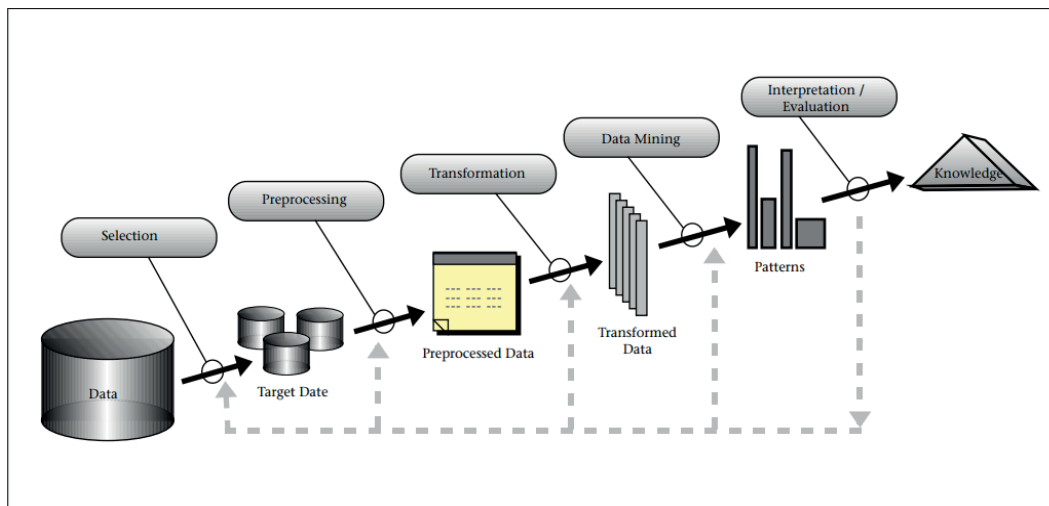
3 MINERAÇÃO DE DADOS NA CLASSIFICAÇÃO DA OBESIDADE

Atualmente, existe um enorme volume de informações disponíveis no dia-a-dia das pessoas. Torna-se, pois, necessária a obtenção de conhecimento útil em meio a tantas informações disponíveis. Fayyad, Piatetsky-Shapiro e Smyth (1996) definem o processo de Descoberta de Conhecimento nas Bases de Dados (do inglês *Knowledge Discovery in Databases* – KDD) como “um processo não trivial de identificação de novos padrões válidos, úteis e compreensíveis” nos dados.

A Figura 1 representa as etapas do processo de KDD segundo Fayyad, Piatetsky-

-Shapiro e Smyth (1996). Em primeiro lugar, selecionam-se os dados de interesse em meio às informações disponíveis. Tais dados passam por um tratamento que consiste nas etapas de pré-processamento e transformação. Posteriormente, são aplicadas ferramentas de mineração de dados para a extração de padrões e geração de conhecimento relevante.

Figura 1: Representação do processo de KDD



Fonte: Fayyad, Piatetsky-Shapiro e Smyth (1996).

A mineração de dados (data mining), uma das etapas do processo de KDD, é o processo de descobrir padrões em bases de dados (WITTEN & FRANK, 2005). Esses padrões devem constituir conhecimento útil capaz de gerar vantagens para quem os utiliza. Além disso, conforme Hand, Mannila e Smyth (2001), os padrões obtidos devem ser, preferencialmente, de fácil entendimento. Algoritmos como árvores de decisão, classificadores bayesianos, redes neurais artificiais e máquinas de vetores de suporte (SVM) têm sido utilizados na classificação da obesidade, principalmente obesidade infantil.

Zhang *et al.* (2009) fizeram uma comparação entre o modelo Logit e seis algoritmos de mineração de dados na classificação da obesidade infantil na Inglaterra entre 1988 e 2003. Os resultados indicam que, em geral, máquinas de vetores de suporte e classificadores bayesianos obtiveram um melhor desempenho na classificação do sobrepeso e da obesidade. Além disso, tais ferramentas mostraram uma maior acurácia em comparação ao modelo Logit.

Pochini, Wu e Hu (2014), ao construírem uma árvore de decisão e estimarem um modelo Logit, constataram que fazer atividade física e tomar café da manhã todos os dias reduziu a prevalência de obesidade entre adolescentes residentes nos Estados Unidos em 2011. Por sua vez, o consumo de cigarro e de bebidas doces elevou o risco de ser obeso.

Ainda sobre os principais fatores de risco da obesidade, Qendraj e Xhafaj (2015) construíram uma árvore de decisão a partir dos dados de Tirana Intermedical Centre para a população adulta. Os autores constataram que ser fumante foi o fator que mais contribuiu para a obesidade.

4 ASPECTOS METODOLÓGICOS

4.1 Base de dados e descrição das variáveis

As informações utilizadas neste estudo foram obtidas na Pesquisa Nacional de Saúde - PNS 2013, realizada pelo IBGE em convênio com o Ministério da Saúde. A Pesquisa abrange todo o território nacional e tem como objetivo gerar informações relevantes sobre as condições de saúde da população brasileira. Pelo fato de a PNS ser uma pesquisa amostral, considerou-se os devidos pesos para expansão da amostra nas análises.

As aferições de peso, altura, circunferência da cintura e pressão arterial, bem como a realização de exames laboratoriais foram feitas em adultos com idade superior a 18 anos selecionados aleatoriamente em cada domicílio com entrevista realizada³. Diante da natureza do presente estudo, apenas esses indivíduos foram selecionados para as análises, totalizando 6.223.101 pessoas residentes no Ceará.

O índice de massa corporal (IMC), uma medida bastante utilizada na classificação antropométrica, é obtido da seguinte forma:

$$IMC = \frac{\text{peso em quilogramas}}{(\text{altura em metros})^2} \quad (1)$$

A Tabela 1 mostra a classificação antropométrica segundo o IMC proposta pela OMS para a população adulta⁴.

Tabela 1: Classificação antropométrica segundo IMC para população adulta

Classificação	IMC (kg/m ²)
Baixo peso	< 18,5
Eutrofia	18,5 – 24,9
Sobrepeso	25 – 29,9
Obesidade	≥ 30

Seguindo essa classificação, 2,67% das pessoas consideradas neste estudo estavam abaixo do peso e foram, portanto, excluídas das análises. Foram feitos tratamentos adicionais na base de dados: (i) excluiu-se pessoas com mais de 60 anos de idade; (ii) desconsiderou-se pessoas cuja condição no domicílio fosse pensionista, empregado doméstico ou

3. Foi selecionado um adulto com mais de 18 anos de idade em cada domicílio.

4. <http://www.who.int/topics/obesity/en/>. Acesso em 27 de setembro de 2017.

parente de empregado doméstico; (iii) excluiu-se as observações que tinham alguma informação faltante nas variáveis relevantes. O número final de observações e sua distribuição por gênero é apresentado na Tabela 2.

Tabela 2: Total de observações e distribuição por gênero, Ceará, 2013

Gênero	Observações	Percentual
Masculino	1.971.217	58,70
Feminino	1.386.796	41,30
Total	3.358.013	100

Fonte: Elaboração própria com dados da PNS (2013).

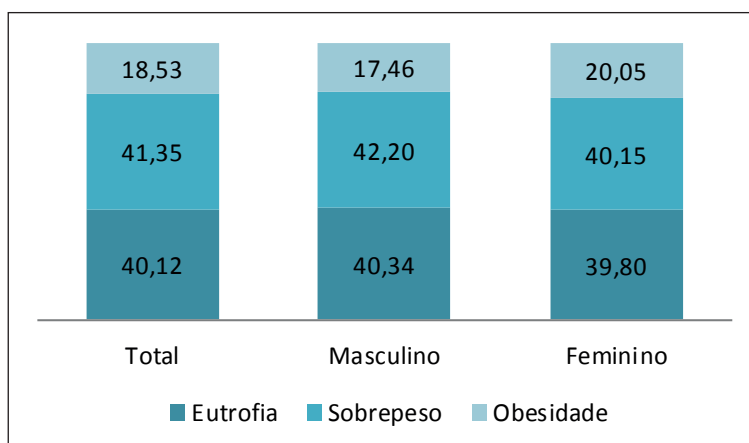
Uma vez excluídos os indivíduos com baixo peso, construiu-se uma variável categórica da seguinte forma:

$$y = \begin{cases} 0, & \text{eutrofia;} \\ 1, & \text{sobrepeso;} \\ 2, & \text{obesidade.} \end{cases} \quad (2)$$

Em que as categorias de são ordenadas conforme o IMC dos indivíduos.

Mais da metade da população adulta cearense estava acima do peso, sendo 41,35% e 18,53% o percentual de pessoas com sobrepeso e obesidade, respectivamente (Figura 2). A Figura 2 também mostra que a incidência de obesidade, mensurada pelo IMC, se mostrou ligeiramente superior entre as mulheres do que entre os homens.

Figura 2: População adulta cearense por classificação antropométrica segundo IMC (%), 2013



Fonte: Elaboração própria com dados da PNS (2013).

O IMC, apesar de ser um bom indicador, não é totalmente correlacionado com a gordura corporal (ABESO, 2009). Conforme destacado em ABESO (2009), suas principais limitações incluem o fato de não distinguir massa gordurosa de massa magra; e não

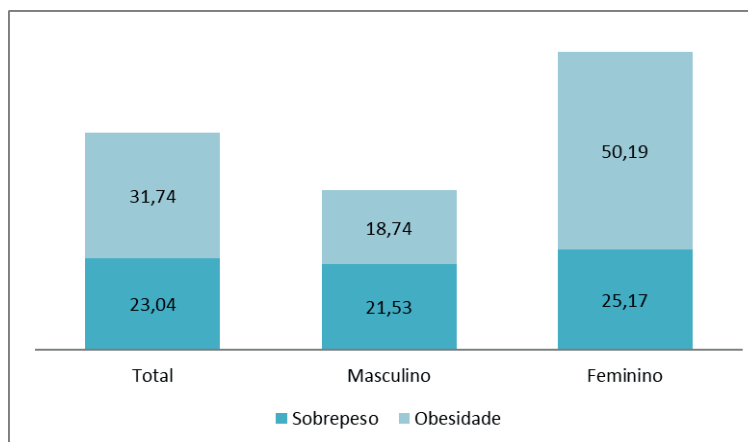
refletir, necessariamente, a distribuição da gordura corporal. A medida da circunferência abdominal, por sua vez, tem um alto grau de associação com a gordura corporal total. A OMS sugere pontos de corte para a classificação antropométrica segundo a circunferência abdominal, conforme mostra a Tabela 3.

Tabela 3: Classificação antropométrica segundo a circunferência abdominal

Classificação	Circunferência abdominal (cm)	
	Homem	Mulher
Sobrepeso	≥ 94	≥ 80
Obesidade	≥ 102	≥ 88

Assim como em Siqueira et al. (2015), os resultados da Figura 3 indicam que a obesidade abdominal se mostrou bem mais alarmante entre as mulheres (50,19%), sendo mais do que o dobro do percentual observado entre os homens (18,74%). Isto é um indicativo de que as mulheres estão muito mais sujeitas a complicações decorrentes da obesidade do que os homens, demandando, assim, uma atenção especial na formulação de políticas na área da saúde.

Figura 3: População adulta cearense segundo sobrepeso e obesidade abdominal (%), 2013



Fonte: Elaboração própria com dados da PNS (2013).

Com o intuito de verificar a relação entre as classificações antropométricas baseadas no IMC e na circunferência abdominal, calculou-se o coeficiente V de Cramer, que mede a relação entre variáveis categóricas. Foi encontrado um valor de 0,4577, indicando que ambas as classificações apresentam alto grau de relação (WITTE & WITTE, 2009). Assim sendo, as análises realizadas neste estudo serão sobre a classificação segundo o IMC em detrimento da circunferência abdominal ou uma combinação de ambas.

O Quadro 1 define as variáveis independentes utilizadas neste estudo, que contêm informações referentes ao indivíduo, ao domicílio de residência e à área de moradia. Todas

as variáveis foram categorizadas de forma a assumirem os valores zero e um.

Quadro 1: Definição das variáveis independentes

Variável	Definição
18 a 30 anos	1, se tem de 18 a 30 anos de idade; 0, caso contrário
31 a 40 anos	1, se tem de 31 a 40 anos de idade; 0, caso contrário
41 a 50 anos	1, se tem de 41 a 50 anos de idade; 0, caso contrário
51 a 60 anos	1, se tem de 51 a 60 anos de idade; 0, caso contrário
Branco	1, se branco(a); 0, caso contrário
Casado	1, se casado(a); 0, caso contrário
Direção*	1, se ocupação é classificada como “diretores e gerentes”, “profissionais das ciências e intelectuais”, “técnicos e profissionais de nível médio” ou “trabalhadores de apoio administrativo”; 0, caso contrário
Serviço*	1, se ocupação é classificada como “trabalhadores dos serviços, vendedores dos comércios e mercados” ou “ocupações elementares”; 0, caso contrário
Manual*	1, se ocupação é classificada como “trabalhadores qualificados, operários e artesãos da construção, das artes mecânicas e outros ofícios”, “operadores de instalações e máquinas e montadores” ou “membros das forças armadas, policiais e bombeiros militares”; 0, caso contrário
Ocupado	1, se ocupado(a); 0, caso contrário
Fundamental	1, se tem ensino fundamental completo ou médio incompleto; 0, caso contrário
Médio	1, se tem ensino médio completo ou superior incompleto; 0, caso contrário
Superior	1, se tem ensino superior completo; 0, caso contrário
Plano	1, se tem plano de saúde; 0, caso contrário
Rdpc1	1, se a renda domiciliar per capita está no primeiro quintil da distribuição; 0, caso contrário
Rdpc2	1, se a renda domiciliar per capita está no segundo quintil da distribuição; 0, caso contrário
Rdpc3	1, se a renda domiciliar per capita está no terceiro quintil da distribuição; 0, caso contrário
Rdpc4	1, se a renda domiciliar per capita está no quarto quintil da distribuição; 0, caso contrário
Rdpc5	1, se a renda domiciliar per capita está no quinto quintil da distribuição; 0, caso contrário
Feijão	1, se consome feijão ao menos 5 dias da semana; 0, caso contrário
Verdura e legume	1, se consome verduras e legumes ao menos 5 dias da semana; 0, caso contrário
Carne	1, se consome carne vermelha ao menos 4 dias da semana; 0, caso contrário
Frango	1, se consome frango ao menos 4 dias da semana; 0, caso contrário
Peixe	1, se consome peixe ao menos 4 dias da semana; 0, caso contrário
Fruta	1, se se consome fruta ao menos 5 dias da semana; 0, caso contrário
Refrigerante	1, se consome refrigerante ao menos 5 dias da semana; 0, caso contrário
Sal	1, se o se consumo de sal é alto ou muito alto; 0, caso contrário
Atividade	1, se praticou atividade física nos últimos 3 meses; 0, caso contrário
Tv	1, se assiste ao menos 4 horas de televisão por dia; 0, caso contrário
Bebida	1, se nunca bebeu; 0, caso contrário
Cigarro	1, se nunca fumou; 0, caso contrário
Urbano	1, se reside na zona urbana; 0, caso contrário
Fortaleza	1, se reside em Fortaleza; 0, caso contrário
Restormf	1, se reside na Região Metropolitana de Fortaleza exclusive Fortaleza; 0, caso contrário
Restouf	1, se reside nos demais municípios cearenses; 0, caso contrário

Fonte: Elaboração própria.

*Variável criada com base na Classificação de Ocupações para Pesquisas Domiciliares - CBO Domiciliar.

Na Tabela 4 são apresentadas algumas estatísticas descritivas das variáveis independentes, a saber, média e desvio padrão. Os resultados são apresentados para a população adulta total e também desagregados por gênero.

Em geral, mais de 30% das pessoas consideradas têm entre 18 e 30 anos, aproximadamente 25% se autodeclarou branco(a); 37%, casado(a); e 95%, ocupado(a). O perfil de ocupação se difere entre homens e mulheres: quase metade das mulheres está alocada na classe denominada serviço (47,63%); enquanto entre os homens, os cargos classificados como serviço e manual contabilizam uma participação relativa de cerca de 40%, cada.

Tabela 4: Estatísticas descritivas das variáveis independentes

Variáveis	Total		Masculino		Feminino	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
18 a 30 anos	0,3567	0,4790	0,3760	0,4844	0,3292	0,4699
31 a 40 anos	0,2781	0,4481	0,2692	0,4435	0,2908	0,4541
41 a 50 anos	0,2255	0,4179	0,2281	0,4196	0,2218	0,4154
51 a 60 anos	0,1397	0,3467	0,1267	0,3326	0,1583	0,3650
Branco	0,2524	0,4344	0,2566	0,4367	0,2466	0,4310
Casado	0,3764	0,4845	0,3651	0,4814	0,3924	0,4883
Direção	0,2385	0,4262	0,2057	0,4042	0,2852	0,4515
Serviço	0,4185	0,4933	0,3778	0,4848	0,4763	0,4994
Manual	0,3013	0,4588	0,3733	0,4837	0,1988	0,3991
Ocupado	0,9583	0,1999	0,9569	0,2031	0,9603	0,1953
Sem instrução	0,3640	0,4811	0,4077	0,4914	0,3018	0,4590
Fundamental	0,1638	0,3701	0,1762	0,3810	0,1462	0,3534
Médio	0,3566	0,4790	0,3296	0,4701	0,3950	0,4889
Superior	0,1155	0,3197	0,0864	0,2810	0,1569	0,3637
Plano	0,1897	0,3920	0,1850	0,3883	0,1963	0,3972
Rdpc1	0,1851	0,3884	0,1994	0,3995	0,1648	0,3710
Rdpc2	0,1932	0,3948	0,1837	0,3873	0,2067	0,4050
Rdpc3	0,2051	0,4038	0,2013	0,4010	0,2106	0,4077
Rdpc4	0,1755	0,3804	0,1478	0,3549	0,2150	0,4108
Rdpc5	0,2410	0,4277	0,2678	0,4428	0,2029	0,4021
Feijão	0,7444	0,4362	0,8147	0,3885	0,6446	0,4786
Verdura e legume	0,1184	0,3231	0,0871	0,2819	0,1630	0,3694
Carne	0,3321	0,4710	0,3937	0,4886	0,2445	0,4298
Frango	0,3369	0,4726	0,3303	0,4703	0,3463	0,4758
Peixe	0,0666	0,2494	0,0693	0,2540	0,0628	0,2426
Fruta	0,4115	0,4921	0,3654	0,4816	0,4770	0,4995
Refrigerante	0,2174	0,4124	0,2257	0,4181	0,2055	0,4041
Sal	0,1302	0,3365	0,1564	0,3632	0,0928	0,2902
Atividade	0,4071	0,4913	0,4433	0,4968	0,3556	0,4787
Tv	0,1145	0,3184	0,1059	0,3077	0,1266	0,3325
Bebida	0,5598	0,4964	0,4385	0,4962	0,7323	0,4428
Cigarro	0,6766	0,4678	0,6281	0,4833	0,7455	0,4356
Urbano	0,8323	0,3736	0,8211	0,3832	0,8482	0,3588
Fortaleza	0,3308	0,4705	0,3157	0,4648	0,3523	0,4777
Restormf	0,2347	0,4238	0,2272	0,4190	0,2455	0,4304
Restouf	0,4345	0,4957	0,4572	0,4982	0,4022	0,4903

Fonte: Elaboração própria com dados da PNS 2013.

A maioria dos homens não tem instrução ou possui o ensino fundamental incompleto (40,77%). O grupo educacional de maior destaque entre as mulheres é o que corresponde ao ensino médio completo ou superior incompleto. (39,50%). O percentual de mulheres que consumia verduras e legumes em pelo menos 5 dias da semana (16,30%) foi praticamente o dobro do percentual de homens (8,71%).

4.2 Modelo Logit ordenado

A primeira etapa deste trabalho consiste em analisar os determinantes do sobrepeso e da obesidade na população adulta cearense em 2013. A variável dependente é constituída das seguintes categorias ordenadas segundo o IMC: eutrofia, sobrepeso e obesidade. Desta forma, foi estimado um modelo Logit ordenado, que não faz nenhuma suposição sobre a distribuição de probabilidade das variáveis independentes (GREENE, 2012). Tal modelo, em sua versão padrão, é construído a partir de uma variável latente, y^* , definida como segue:

$$y^* = x'\beta + \varepsilon \quad (3)$$

Em que, x é o vetor de variáveis explicativas; β , o vetor de parâmetros a ser estimado; ε , os fatores não observados; e y^* , a variável latente, constituída por j categorias, é não observada. Neste modelo, a variável y é observada:

$$\begin{aligned} y &= 0 \text{ se } y^* \leq 0 \\ y &= 1 \text{ se } 0 \leq y^* \leq \mu_1 \\ y &= 2 \text{ se } \mu_1 \leq y^* \leq \mu_2 \\ &\vdots \\ y &= J \text{ se } \mu_{j-1} \leq y^* \end{aligned} \quad (4)$$

Em que, os μ 's são parâmetros desconhecidos satisfazendo:

$$0 < \mu_1 < \mu_2 < \dots < \mu_{j-1} \quad (5)$$

No caso em que y apresenta três categorias, como no presente estudo, a função de distribuição de probabilidade pode ser escrita como:

$$\begin{aligned} \text{Prob}(y = 0 | x) &= F(-x'\beta) \\ \text{Prob}(y = 1 | x) &= F(\mu_1 - x'\beta) - F(-x'\beta) \\ \text{Prob}(y = 2 | x) &= 1 - F(\mu_1 - x'\beta) \end{aligned} \quad (6)$$

Assim como nos modelos Logit binários, os efeitos marginais das variáveis explicativas x não são exatamente os coeficientes estimados. Para as três categorias da variável dependente, os efeitos marginais são dados por:

$$\begin{aligned}
\frac{\partial \text{Prob}(y=0|x)}{\partial x} &= -f(-x'\beta)\beta \\
\frac{\partial \text{Prob}(y=1|x)}{\partial x} &= [f(-x'\beta) - f(\mu_1 - x'\beta)]\beta \\
\frac{\partial \text{Prob}(y=2|x)}{\partial x} &= f(\mu_1 - x'\beta)\beta
\end{aligned} \tag{7}$$

Generalizando (7), somente os sinais das mudanças em $\text{Prob}(y=0|x)$ e $\text{Prob}(y=J|x)$ não são ambíguos em relação ao sinal dos coeficientes.

Em sua versão padrão, o modelo ordenado pressupõe que os coeficientes estimados, β , se mantêm constantes ao longo de todas as categorias da variável dependente (LONG & FREESE, 2006; ABREU *et al.*, 2009). No entanto, essa hipótese pode ser violada em todas ou em algumas das categorias de y , dando origem, respectivamente, ao modelo ordinal generalizado e ao modelo ordenado de chances proporcionais parciais. No modelo ordinal generalizado, os β 's são diferentes para todas as categorias da variável dependente. Por sua vez, o modelo ordenado de chances proporcionais parciais permite que os β 's coincidam em algumas categorias e diverjam em outras.

A escolha do modelo adequado se deu pela aplicação do teste de Brant (1990), cuja hipótese nula é a versão padrão do modelo ordenado. Desta forma, o modelo ordinal generalizado é o mais indicado quando a hipótese nula do teste é rejeitada para todos os coeficientes. Quanto é rejeitada em alguns coeficientes e em outros não, o modelo de chances proporcionais parciais é o mais adequado.

4.3 Algoritmos de mineração de dados para classificação

O segundo objetivo deste trabalho consiste em comparar o desempenho do modelo Logit binário com diferentes algoritmos de mineração de dados na classificação antropométrica da população adulta cearense. Para tal, a variável dependente foi recodificada gerando a variável a seguir:

$$z = \begin{cases} 0, & \text{eutrofia;} \\ 1, & \text{sobrepeso ou obesidade.} \end{cases} \tag{8}$$

O desempenho dos diferentes algoritmos de classificação utilizados neste trabalho será comparado pela acurácia global de cada modelo:

$$\text{Acurácia}_{\text{global}} = \frac{\text{número de casos classificados corretamente}}{\text{número total de casos}} \tag{9}$$

Para a implementação dos algoritmos, a base de dados é dividida em dois conjuntos:

treinamento e teste. Os modelos são desenvolvidos no primeiro conjunto e testados no segundo. Como os dados utilizados neste estudo são categóricos, os classificadores do tipo Árvore de Decisão e Bayesianos são os modelos mais indicados (Dugan *et al.*, 2015).

Seja S uma base de dados composta por s observações⁵. Assuma que a variável a ser classificada é composta por m valores distintos definindo m classes C_i , $i = 1, \dots, m$. Defina s_i como o número de observações que pertence à classe C_i . Cada amostra de dados X tem n atributos⁶, que podem ser expressos vetorialmente como $X = (x_1, x_2, \dots, x_n)$. Cada elemento x_i é um valor nominal ou contínuo que corresponde com o atributo A_i ($i = 1, \dots, n$). Suponha A_i que o atributo tem q valores distintos $\{a_{i1}, a_{i2}, \dots, a_{iq}\}$, e A_i pode particionar a base de dados original S em q subconjuntos, $\{S_1, S_2, \dots, S_q\}$, em que S_j contém dados com o valor a_{ij} em A_i . Então, seja s_{ij} o número de observações da classe C_i no subconjunto S_j .

4.3.1 Árvores de decisão

Este algoritmo gera um modelo em estrutura de árvore, que é uma estrutura hierárquica, composta por arestas consistindo de diferentes atributos e possui três tipos de nós: nó raiz, nós intermediários e nós finais representando uma classe. A árvore é construída com base no ganho de informação. O procedimento é descrito de forma breve a seguir. Em primeiro lugar, deve-se calcular a seguinte medida de informação:

$$I(s_1, s_2, \dots, s_m) = -\sum_{i=1}^m p_i \log_2(p_i) \quad (10)$$

$$p_i = \frac{s_i}{s} \quad (11)$$

Em que, p_i é a probabilidade de uma observação aleatória pertencer à classe C_i .

Em segundo lugar, calcula-se a entropia $C(A_i)$, que é a informação esperada baseada na partição pelo atributo A_i :

$$E(A_i) = \sum_{j=1}^q \frac{s_{1j} + s_{2j} + \dots + s_{mj}}{s} I(s_{1j}, \dots, s_{mj}) \quad (12)$$

$$I(s_{1j}, \dots, s_{mj}) = \sum_{i=1}^m p_{ij} \log_2(p_{ij}) \quad (13)$$

$$p_{ij} = \frac{s_{ij}}{|S_j|} \quad (14)$$

⁵ Esta seção baseia-se em Zhang et al. (2009).

⁶ Na linguagem de mineração de dados, atributos correspondem às variáveis independentes.

Em que, $|S_j|$ é o número de observações no subconjunto S_j .

Logo, a informação obtida em ramificar a árvore em s é dada por:

$$E(A_i) = \sum_{j=1}^q \frac{s_{1j} + s_{2j} + \dots + s_{mj}}{s} I(s_{1j}, \dots, s_{mj}) \quad (12)$$

O atributo A_i com o maior ganho de informação é escolhido como nó raiz e suas arestas são formadas de acordo com os diferentes valores de a_{ij} , $J = 1, \dots, q$. Este processo é repetido até que todas as observações pertençam à mesma classe, que rotulará o nó final. Ou seja, quanto mais alto na árvore, mais informativo é o atributo. Neste trabalho será implementado o algoritmo C4.5 (Quinlan, 1993), que é um dos mais aplicados na classificação de obesidade (Abdullah et al., 2016).

4.3.2 Classificadores Bayesianos

Os classificadores Bayesianos fazem a classificação com base na probabilidade de dada observação pertencer à determinada classe. Redes bayesianas (Bayesian networks) e Naïve Bayes são os dois modelos fundamentais nesse grupo de classificadores:

“Naïve Bayes algorithms assume that the effect that an attribute plays on a given class is independent of the values of other attributes. In practice, dependencies often exist among attributes; hence Bayesian networks are graphical models, which, unlike naïve Bayesian classifiers, can describe joint conditional probability distributions.” (ZHANG *et al.*, 2009).

Dada uma amostra de dados, X , cuja classe será determinada, o algoritmo Naïve Bayes classifica determinada observação como pertencente à classe C_i se, e somente se:

$$P(C_i \setminus X) > P(C_j \setminus X), \text{ para } 1 \leq j \leq m, \neq i \quad (16)$$

Em que a probabilidade a posteriori $P(C_i \setminus X)$, $1 \leq i \leq m$, pode ser determinada pelo Teorema de Bayes:

$$P(C_i \setminus X) = \frac{P(X \setminus C_i)P(C_i)}{P(X)} \quad (17)$$

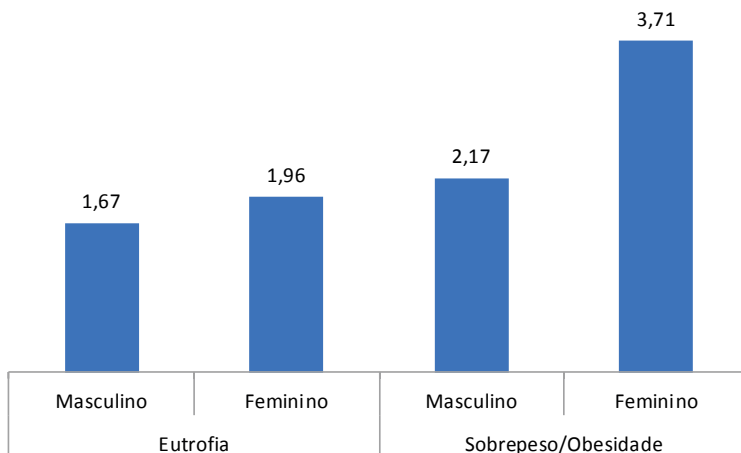
5. ANÁLISE DE RESULTADOS E DISCUSSÕES

5.1 Prevalência de DCNT por classificação nutricional segundo IMC

Nesta seção, a população adulta cearense foi classificada em duas categorias segundo o IMC: eutrofia e sobrepeso/obesidade. Os resultados apresentados a seguir revelam que a prevalência do diagnóstico de algumas DCNT se mostrou maior entre a parcela da população adulta cearense acima do peso. Ademais, foi um problema que afetou as mulheres de forma mais acentuada.

A prevalência do diagnóstico de diabetes entre as mulheres acima do peso foi 3,71% (Figura 4) contra 1,96% entre as mulheres com peso normal. Já entre os homens, tais rubricas representaram, respectivamente, 2,17% e 1,67%.

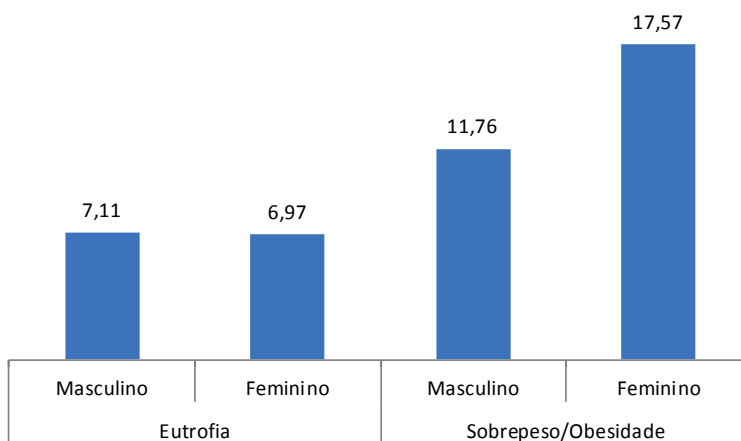
Figura 4: Prevalência do diagnóstico de diabetes por classificação nutricional e gênero (%), Ceará, 2013



Fonte: Elaboração própria com dados da PNS 2013.

A incidência de pessoas diagnosticadas com pressão alta também se mostrou superior entre as pessoas acima do peso (Figura 5). Neste grupo da população cearense, entre as mulheres 17,57% foram diagnosticados com pressão alta e entre os homens, 11,76%.

Figura 5: Prevalência do diagnóstico de pressão alta por classificação nutricional e gênero (%), Ceará, 2013

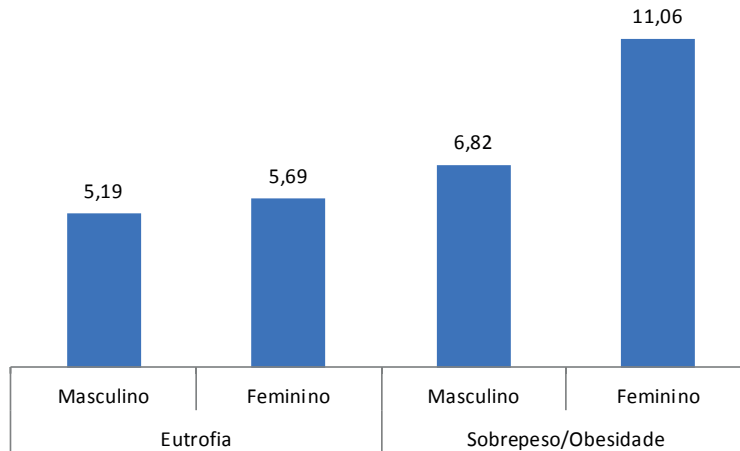


Fonte: Elaboração própria com dados da PNS 2013.

A Figura 6 indica que a prevalência do diagnóstico de colesterol alto foi semelhante

entre homens e mulheres com peso normal, sendo muito mais elevada entre as mulheres acima do peso (11,06%) do que entre os homens (6,82%).

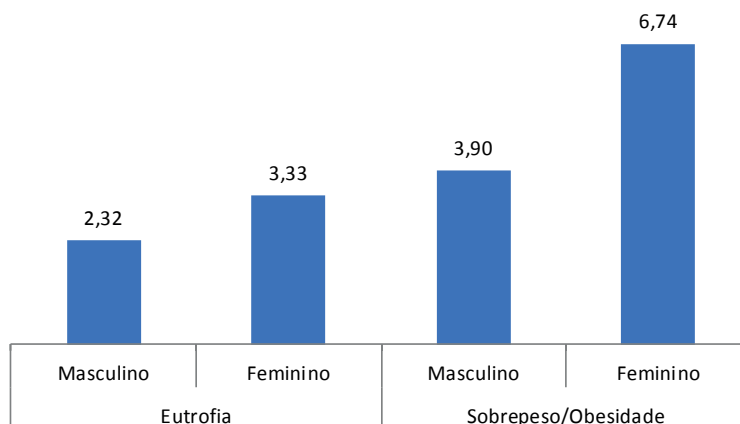
Figura 6: Prevalência do diagnóstico de colesterol alto por classificação nutricional e gênero (%), Ceará, 2013



Fonte: Elaboração própria com dados da PNS 2013.

A parcela de mulheres acima do peso diagnosticadas com reumatismo foi de 6,74% contra 3,33% das mulheres com peso normal. Já o percentual de homens eutróficos diagnosticados com reumatismo foi 2,32%, passando para 3,90% dos homens com sobrepeso ou obesos.

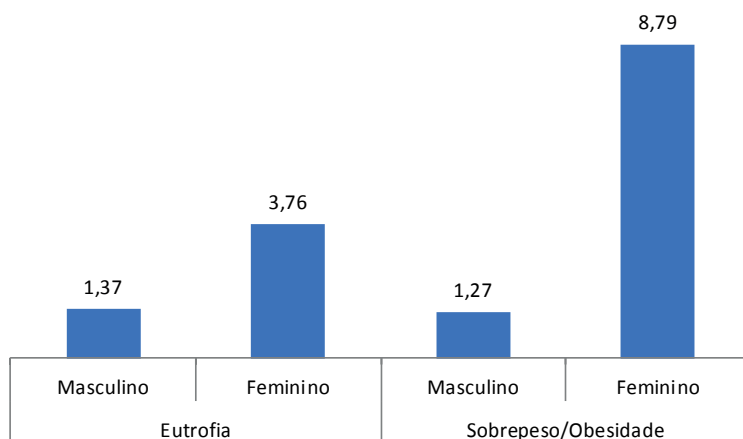
Figura 7: Prevalência do diagnóstico de reumatismo por classificação nutricional e gênero (%), Ceará, 2013



Fonte: Elaboração própria com dados da PNS 2013.

Por fim, a prevalência do diagnóstico de depressão entre os homens foi semelhante entre aqueles com peso normal e acima do peso. No entanto, 8,79% das mulheres acima do peso foram diagnosticadas como sofrendo depressão, enquanto apenas 3,76% das mulheres de peso normal o foram.

Figura 8: Prevalência do diagnóstico de depressão por classificação nutricional e gênero (%), Ceará, 2013



Fonte: Elaboração própria com dados da PNS 2013.

Os resultados desta seção corroboram o fato de sobrepeso e obesidade serem fatores de risco para a ocorrência de DCNT. Assim sendo, e diante da atual epidemia da obesidade, a adoção de medidas que atuem na prevenção e no combate dessa doença é muito importante. Neste sentido, a análise dos fatores determinantes do sobrepeso e da obesidade se mostra relevante ao gerar informações úteis para os formuladores de políticas públicas.

5.2 Determinantes do sobrepeso e da obesidade

A análise do impacto dos determinantes do sobrepeso e da obesidade na população adulta cearense se deu por meio da estimação de um modelo Logit ordenado. Os dados foram desagregados por gênero para a realização das estimações. O teste de Brant indicou que o modelo ordinal generalizado foi o mais apropriado tanto entre os homens quanto entre as mulheres. Os efeitos marginais das variáveis explicativas são apresentados na Tabela 5.

Sobrepeso e obesidade aumentaram com a idade para ambos os sexos. Entre os homens, o fato de ser branco se mostrou fator de risco para sobrepeso e obesidade, enquanto que entre as mulheres observou-se resultado oposto (mulheres brancas têm probabilidade de terem peso normal 11,3 p.p. maior do que mulheres não brancas). Os homens casados

têm uma chance de terem sobrepeso 11,1 p.p. maior do que homens não casados. Mulheres casadas também se mostraram mais suscetíveis a terem problemas de sobrepeso e obesidade.

Em relação ao tipo de ocupação, homens empregados nos cargos classificados como direção e serviços tem mais chance de serem obesos em comparação àqueles trabalhando na classe denominada manual. De fato, o crescente sedentarismo de muitas formas de trabalho é apontado pela OMS como um fator de risco para obesidade. Por sua vez, mulheres alocadas nas classes direção e serviço apresentaram uma maior probabilidade de terem peso normal quando comparadas com aquelas na classe manual. A condição na força de trabalho teve impacto divergente entre homens e mulheres. Homens (mulheres) ocupados(as) têm uma chance de terem peso normal 36p.p. maior (menor) do que àqueles(as) não ocupados(as).

A educação teve relação inversa com a obesidade entre as mulheres, enquanto que entre os homens não se observou uma relação clara. Este resultado corrobora o de Ward et al. (2015), ao encontrarem relação negativa entre IMC e escolaridade para mulheres e padrão não estabelecido entre homens na população adulta de Belo Horizonte.

Tabela 5: Efeitos marginais estimados para classificação antropométrica segundo IMC no Ceará desagregado por gênero, 2013 *continua*

Coeficientes	Masculino			Feminino		
	Eutrofia	Sobrepeso	Obesidade	Eutrofia	Sobrepeso	Obesidade
31 a 40 anos	-0,176*** (0,000938)	0,0792*** (0,00102)	0,0969*** (0,000874)	-0,164*** (0,00104)	0,135*** (0,00114)	0,0290*** (0,000997)
41 a 50 anos	-0,230*** (0,000949)	0,173*** (0,00110)	0,0573*** (0,000941)	-0,274*** (0,00104)	0,247*** (0,00124)	0,0277*** (0,00111)
51 a 60 anos	-0,158*** (0,00123)	0,0813*** (0,00141)	0,0766*** (0,00125)	-0,170*** (0,00134)	0,0927*** (0,00152)	0,0771*** (0,00141)
branco	-0,0516*** (0,000904)	0,0300*** (0,000897)	0,0216*** (0,000623)	0,113*** (0,00111)	-0,0618*** (0,00106)	-0,0515*** (0,000745)
casado	-0,102*** (0,000859)	0,111*** (0,000857)	-0,00902*** (0,000571)	-0,0541*** (0,000972)	0,0146*** (0,000981)	0,0395*** (0,000739)
direcao	-0,0639*** (0,00115)	-0,00172 (0,00123)	0,0656*** (0,000980)	0,142*** (0,00168)	-0,110*** (0,00164)	-0,0322*** (0,00109)
servico	-0,107*** (0,000834)	0,0210*** (0,000863)	0,0861*** (0,000676)	0,146*** (0,00129)	-0,112*** (0,00129)	-0,0334*** (0,000872)
ocupado	0,362*** (0,000639)	-0,322*** (0,00195)	-0,0394*** (0,00196)	-0,363*** (0,00220)	0,248*** (0,00207)	0,115*** (0,00114)
fundamental	-0,0590*** (0,00111)	0,0954*** (0,00109)	-0,0364*** (0,000730)	-0,0672*** (0,00152)	0,156*** (0,00146)	-0,0892*** (0,000801)
medio	-0,126*** (0,00106)	0,111*** (0,00106)	0,0148*** (0,000757)	0,0480*** (0,00128)	0,0639*** (0,00125)	-0,112*** (0,000825)
superior	0,182***	-0,105***	-0,0771***	0,0794***	-0,00769***	-0,0718***

Tabela 5: Efeitos marginais estimados para classificação antropométrica segundo IMC no Ceará desagregado por gênero, 2013 *conclusão*

	(0,00183)	(0,00174)	(0,000785)	(0,00201)	(0,00185)	(0,00104)
rdpc2	-0,140***	0,0380***	0,102***	-0,0143***	-0,0624***	0,0767***
	(0,00106)	(0,00107)	(0,00107)	(0,00147)	(0,00165)	(0,00143)
rdpc3	-0,0355***	-0,0313***	0,0668***	-0,0270***	-0,0677***	0,0946***
	(0,00115)	(0,00105)	(0,00100)	(0,00149)	(0,00168)	(0,00143)
rdpc4	-0,190***	0,119***	0,0714***	-0,0621***	0,0644***	-0,00235
	(0,00104)	(0,00108)	(0,00108)	(0,00150)	(0,00160)	(0,00123)
rdpc5	-0,206***	0,0692***	0,136***	-0,0704***	-0,0135***	0,0839***
	(0,000996)	(0,000298)	(0,000832)	(0,00154)	(0,00170)	(0,00150)
plano	-0,0967***	0,00216	0,0945***	-0,0263***	0,0451***	-0,0188***
	(0,00115)	(0,00118)	(0,000929)	(0,00135)	(0,00129)	(0,000891)
feijao	-0,0158***	0,0777***	-0,0619***	0,0992***	-0,0630***	-0,0363***
	(0,00101)	(0,000968)	(0,000751)	(0,000989)	(0,000986)	(0,000762)
verdura e legume	-0,00142	0,0258***	-0,0244***	0,0322***	-0,0133***	-0,0189***
	(0,00147)	(0,00145)	(0,000826)	(0,00118)	(0,000517)	(0,000667)
carne	-0,0150***	0,0469***	-0,0318***	0,0748***	-0,133***	0,0579***
	(0,000798)	(0,000786)	(0,000512)	(0,00113)	(0,00118)	(0,000939)
frango	0,0169***	0,0378***	-0,0547***	-0,0188***	0,0389***	-0,0201***
	(0,000831)	(0,000819)	(0,000527)	(0,000987)	(0,00100)	(0,000731)
peixe	0,0738***	-0,0668***	-0,00707***	0,159***	-0,102***	-0,0574***
	(0,00155)	(0,00150)	(0,00101)	(0,00217)	(0,00209)	(0,00119)
fruta	-0,0276***	-0,0225***	0,0501***	-0,0910***	0,0216***	0,0694***
	(0,000807)	(0,000806)	(0,000569)	(0,000944)	(0,000966)	(0,000740)
refrigerante	0,0416***	-0,0298***	-0,0118***	-0,0231***	-0,0145***	0,0376***
	(0,000941)	(0,000909)	(0,000620)	(0,00117)	(0,00118)	(0,000923)
sal	0,0232***	-0,0936***	0,0705***	0,0198***	-0,134***	0,114***
	(0,00109)	(0,00106)	(0,000863)	(0,00167)	(0,00166)	(0,00146)
atividade	0,00261**	0,0517***	-0,0543***	-0,0602***	0,0458***	0,0144***
	(0,000833)	(0,000842)	(0,000554)	(0,000991)	(0,00100)	(0,000737)
tv	-0,0228***	0,0269***	-0,00414***	-0,0977***	0,163***	-0,0649***
	(0,00123)	(0,00120)	(0,000801)	(0,00129)	(0,00131)	(0,000816)
bebida	0,104***	-0,109***	0,00560***	0,142***	-0,0691***	-0,0725***
	(0,000812)	(0,000817)	(0,000552)	(0,00102)	(0,00105)	(0,000846)
cigarro	-0,0798***	0,0737***	0,00607***	0,0812***	-0,0528***	-0,0284***
	(0,000913)	(0,000877)	(0,000591)	(0,00115)	(0,00117)	(0,000861)
urbano	0,0000273	-0,0336***	0,0336***	0,0447***	-0,0253***	-0,0194***
	(0,00107)	(0,00103)	(0,000667)	(0,00133)	(0,00141)	(0,00111)
fortaleza	0,0536***	-0,0459***	-0,00770***	-0,105***	0,00746***	0,0977***
	(0,000963)	(0,000922)	(0,000610)	(0,00113)	(0,00122)	(0,00103)
restormf	0,0125***	0,0299***	-0,0424***	-0,0954***	-0,0649***	0,160***
	(0,00101)	(0,000995)	(0,000592)	(0,00113)	(0,00132)	(0,00119)

Fonte: Elaboração própria com dados da PNS 2013.

***Significativo a 1%.

De um modo geral, maiores níveis de renda domiciliar *per capita* se mostraram como fator de risco para obesidade entre homens e mulheres. Por exemplo, homens no quintil mais rico da distribuição da renda domiciliar per capita têm uma chance de serem obesos 13,6p.p. maior do que aqueles no quintil mais pobre. Já entre as mulheres, esse aumento na chance de ser obesa corresponde a 8,39p.p. Ter plano de saúde foi fator de risco para obesidade (9,45p.p. a mais de chance em relação à categoria base) entre os homens, e foi fator de risco para sobrepeso (4,51p.p. a mais de chance em relação à categoria base) entre as mulheres.

Em relação à alimentação, o consumo regular de feijão, verdura e legume reduziu as chances dos homens serem obesos. No entanto, também reduziu a chance de terem peso normal, tendo efeito positivo sobre sobrepeso. Já entre as mulheres, tais variáveis foram fator de proteção contra sobrepeso e obesidade. O consumo de carne vermelha foi fator de risco para sobrepeso entre os homens (4,69p.p. a mais de chance em relação à categoria base). Já entre as mulheres foi o consumo de frango que aumentou as chances de terem sobrepeso em 3,89p.p. em comparação à categoria base. O consumo de peixe foi fator protetor nos dois sexos, sendo o efeito muito maior entre as mulheres na chance de ter peso normal (7,38p.p. para os homens e 15,9p.p. para as mulheres). O consumo de refrigerante foi fator de risco para obesidade entre as mulheres (3,76p.p. a mais de chance de serem obesas) e fator protetor de peso normal entre homens (4,16p.p. a mais de chance de terem peso normal).

A obesidade aumentou com o consumo de frutas e de sal em ambos os sexos. Além disso, passar mais de 4 horas por dia assistindo TV contribuiu para o aumento no sobrepeso em 2,69p.p. entre os homens e em 16,30p.p. entre as mulheres. O fato de não consumir bebidas alcoólicas aumentou a probabilidade de homens e mulheres terem peso normal. Por sua vez, nunca ter fumado foi fator de risco para sobrepeso e obesidade entre os homens, e contribuiu para o aumento da chance das mulheres terem peso normal.

Quanto ao local de moradia, homens residindo em Fortaleza e na Região Metropolitana de Fortaleza apresentaram menores chances de serem obesos em relação aos homens residindo nos demais municípios cearenses. Resultado oposto se observou entre as mulheres. De um modo geral, os resultados dessa seção mostraram que o excesso de peso é afetado de maneira diferente entre homens e mulheres.

5.3 Acurácia dos algoritmos de classificação

Os algoritmos de classificação foram implementados nos dados desagregados por gênero. A divisão da base de dados entre os conjuntos de treinamento e teste seguiu a *cross-validation* de 10 passos. Ou seja, a amostra foi dividida em 10 partes iguais e em cada

rodada, nove destas partes foi o conjunto de treinamento e uma, o conjunto de teste. Após as dez repetições deste exercício, foram obtidos a média e o desvio-padrão da acurácia de cada modelo. Os resultados são apresentados na Tabela 6 a seguir.

Tabela 6: Acurácia dos algoritmos na classificação antropométrica segundo gênero (%), 2013

Gênero	Logit	Árvore de decisão	Naïve Bayes	Rede bayesiana
Masculino	59,96	56,67	58,57	58,61
Feminino	66,85	62,99	62,49*	62,54*

Fonte: Elaboração própria com dados da PNS 2013.

*Acurácia inferior ao modelo Logit (nível de significância de 5%).

Os modelos tiveram desempenho similar entre os homens, com aproximadamente 60% dos casos classificados corretamente (Tabela 6). Por sua vez, entre as mulheres, o modelo Logit teve desempenho superior ao dos classificadores bayesianos e similar ao da árvore de decisão.

A construção das árvores de decisão indicou que, entre os homens, a situação matrimonial foi o principal fator que contribuiu para o excesso de peso, já entre as mulheres foi a condição de ocupação no mercado de trabalho. Mais uma vez, sobrepeso e obesidade são afetados de maneiras distintas por seus determinantes entre homens e mulheres.

6. CONSIDERAÇÕES FINAIS

Este estudo tem como objetivo contribuir para a literatura de Economia da Saúde no Ceará. Mais especificamente, visa favorecer o debate sobre sobrepeso e obesidade na população adulta cearense. Tal agravo, que representa fator de risco para as DCNT e gera outros efeitos perversos na qualidade de vida das pessoas, tem se tornado uma epidemia mundial ao longo dos últimos anos.

As informações utilizadas neste trabalho foram obtidas na PNS 2013 realizada pelo IBGE. A classificação antropométrica, feita com base no IMC, seguiu a proposição da OMS para a população adulta. As pessoas com baixo peso foram excluídas das análises, de forma que foram consideradas apenas aquelas com peso normal, sobrepeso e obesidade. Em vista da natureza da variável dependente, foram estimados modelos Logit ordenados com o intuito de analisar os determinantes do excesso de peso entre homens e mulheres residentes no Ceará. Dentre as variáveis explicativas, têm-se informações sobre o indivíduo, o domicílio e a região de moradia. Os resultados mostraram peculiaridades nos determinantes do excesso de peso entre os dois sexos. A prevalência de sobrepeso e obesidade

aumentou com a idade e com a renda domiciliar per capita entre homens e mulheres. Assim como em Ward et al. (2015), a escolaridade apresentou relação negativa com a prevalência de excesso de peso entre as mulheres, não tendo padrão muito claro entre os homens. Alimentação saudável, representada pelo consumo regular de feijão, verduras, legumes e peixe, reduz a obesidade em ambos os sexos. Ademais, o consumo de frutas, o consumo excessivo de sal e número de horas assistindo televisão foram fatores de risco para excesso de peso entre homens e mulheres.

Na segunda parte do trabalho, foram utilizados algoritmos de mineração de dados na classificação antropométrica da população adulta cearense. A construção das árvores de decisão mostrou que, novamente, a prevalência de excesso de peso é afetada de maneira diferente entre homens e mulheres. Entre os homens, a situação matrimonial foi o principal fator que contribuiu para sobrepeso e obesidade, já entre as mulheres foi a condição de ocupação no mercado de trabalho. A utilização de algoritmos de mineração de dados não apresentou ganhos em termos de acurácia quando comparados ao modelo Logit. Em trabalhos futuros, pode ser feita a comparação do modelo Logit com modelos de redes neurais artificiais e máquinas de vetores de suporte, por exemplo.

De um modo geral, os resultados indicam que a prevalência de excesso de peso é afetada de maneira diferente entre homens e mulheres, o que deve ser levado em consideração para uma melhor formulação de políticas para o enfrentamento deste agravo. Tais políticas devem ser compostas por medidas de diversas áreas, como saúde e educação, tendo em vista o perfil multifacetado dos determinantes do excesso de peso.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ABDULLAH, F. S. *et al.* **Data Mining Techniques for Classification of Childhood Obesity Among Year 6 School Children.** In: International Conference on Soft Computing and Data Mining. Springer, Cham, 2016. p. 465-474.

ABESO - Associação Brasileira para o Estudo da Obesidade e da Síndrome Metabólica. **Diagnósticos brasileiros de obesidade 2009/2010.** 3.ed. Itapevi, SP : AC Farmacêutica, 2009.

ABREU, M. N. S.; SIQUEIRA, A. L.; CAIAFFA, W. T. **Regressão logística ordinal em estudos epidemiológicos.** Rev Saúde Pública 2009; 43(1):183-194.

ANDRADE, R. G. et al. **Excesso de peso em homens e mulheres residentes em área urbana: fatores individuais e contexto socioeconômico.** Cad. Saúde Pública, Rio de Janeiro, 31 Sup:S1-S11, 2015.

- BRANT, R. **Assessing proportionality in the proportional odds model for ordinal logistic regression**. *Biometrics*, v. 46, n. 4, p. 1171–1178, 1990.
- CORREIA, L. L. *et al.* **Prevalência e determinantes de obesidade e sobrepeso em mulheres em idade reprodutiva residentes na região semiárida do Brasil**. *Ciência & Saúde Coletiva*, 16(1):133-145, 2011.
- COSTA, E. C. *et al.* **Evolução do excesso de peso e fatores associados em mulheres de 10 a 49 anos em Pernambuco, Nordeste, Brasil**. *Rev. Nutr.*, Campinas, 27(5):513-524, set./out., 2014.
- DUGAN, T. M. *et al.* **Machine learning techniques for prediction of early childhood obesity**. *Appl Clin Inform* 2015; 6: 506–520.
- FAYYAD, U; PIATETSKY-SHAPIRO, G; SMYTH, P. **From Data Mining to Knowledge Discovery in Databases**. American Association for Artificial Intelligence, 1996.
- GREENE, W. H. **Econometric Analysis**. 7. ed. New Jersey: Prentice Hall, 2012.
- HAND, D.; MANNILA, H.; SMYTH, P. **Principles of Data Mining**. The MIT Press, 2001.
- LONG, J. S.; FREESE, J. **Regression Models for Categorical Dependent Variables Using Stata**, 2. ed. Stata Press, 2006.
- MCKINSEY GLOBAL INSTITUTE. **Overcoming obesity: An initial economic analysis**. 2014.
- MINISTÉRIO DA SAÚDE (BR). Secretaria de Vigilância em Saúde. Departamento de Vigilância de Doenças e Agravos não Transmissíveis e Promoção da Saúde. **Vigitel Brasil 2014: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico**. Brasília: Ministério da Saúde; 2015.
- Organização das Nações Unidas para Alimentação e Agricultura e Organização Pan-americana de Saúde. **Panorama de la seguridad alimentaria y nutricional**. Santiago, 2017.
- POCHINI, A.; WU, Y.; HU, G. **Data Mining for Lifestyle Risk Factors Associated with Overweight and Obesity among Adolescents**. In: *Advanced Applied Informatics (IIAIA-AI)*, 2014 IIAI 3rd International Conference on. IEEE, 2014. p. 883-888.
- QENDRAJ, D.; XHAFAJ, E. **Evaluating risk factors of being obese, by using id3 algorithm in Weka software**. *European Scientific Journal*, ESJ, v. 11, n. 24, 2015.
- ROEMLING, C.; QAIM, M. **Obesity trends and determinants in Indonesia**. *Appetite*, v. 58, n. 3, p. 1005-1013, 2012.

SIQUEIRA, D. G. B. et al. **Diferenças entre sexos nos determinantes da obesidade abdominal em adultos de 40 anos ou mais: estudo de base populacional.** Rev. Nutr., Campinas, 28(5):485-496, set./out., 2015.

WARD, J. et al. **Associação de fatores socioeconômicos com o índice de massa corporal, obesidade, atividade física e dieta em Belo Horizonte, Minas Gerais, Brasil: Estudo Saúde em Beagá.** Cad. Saúde Pública, Rio de Janeiro, 31 Sup:S182-S194, 2015.

WITTEN, I. H.; FRANK, E. **Data mining : practical machine learning tools and techniques.** 2. Ed. (Morgan Kaufmann series in data management systems) 2005.

WITTE, R. S.; WITTE, J. S. **Statistics.** 9. Ed. Willey, 2009.

ZHANG, L.; LIAO, T. F. **Multilevel Effects of Socioeconomic Status and Income Inequality on Obesity among Chinese Adults.** In: <http://paa2013.princeton.edu/papers/132519>. Acesso em 20 de julho de 2017.

ZHANG, S. et al. **Comparing data mining methods with logistic regression in childhood obesity prediction.** Information Systems Frontiers, v. 11, n. 4, p. 449-460, 2009.

INVESTIGAÇÃO E JULGAMENTO DOS HOMICÍDIOS DOLOSOS NO MUNICÍPIO DE FORTALEZA

Denise Xavier Araújo de Oliveira*

Andrei Gomes Simonassi**

RESUMO

O estudo faz uma análise da trajetória de apuração dos crimes de homicídios dolosos consumados e registrados na cidade de Fortaleza no ano de 2014, focando em uma das causas que contribuem para a manutenção ou incremento desse crime: a impunidade. Ao longo dessa trajetória observa-se que inúmeros casos registrados não resultaram em imposição da pena prevista em lei, o que estimulou a investigação deste fenômeno a partir da aplicação de técnicas de análise de sobrevivência às fases do processo de investigação. Conclui-se que o maior entrave à apuração de crimes está já na fase de investigação policial. Não obstante, constata-se que do total de 1.896 crimes, apenas 18 casos foram julgados (0,95% do total). Outro ponto de especialização do estudo consiste na divisão dos casos segundo a ocorrência ou não de prisão (e a sua modalidade). Ao todo, apenas 1,5% dos casos foram iniciados com prisão em flagrante, porém representaram 67% dos casos julgados, o que demonstra a necessidade de elaboração de novas metodologias de trabalho no sentido de efetivar, dentro dos limites legais, a captura do autor ainda em situação de flagrante delito.

Palavras-Chave: Fluxo da Justiça Criminal. Homicídio Doloso. Análise de Sobrevivência.

ABSTRACT

The study analyzes the trajectory of investigation of the crimes of felony murder committed in the city of Fortaleza, practiced in 2014, focusing on one of the causes that contribute to the maintenance or increase of this crime: impunity. Throughout this trajectory, we observed that many cases registered had no imposition of the penalty provided by the Brazilian law, which stimulate an investigation by applying survival analysis techniques on the steps of the investigation process. We find that just 18 cases in the overall of 1.896 crimes were judged (0,95% of the total). Other special point of the investigation is

* Denise Xavier Araújo de Oliveira. CAEN/UFC. denisexaraujo@gmail.com. (85)99926.4044.

** Andrei Gomes Simonassi. CAEN/UFC. agsimonassi@gmail.com.

related to regarding a division of the cases according to the occurrence or not of arrest (and its modality). In this case, 1.5% of cases were initiated with flagrant arrest, but denoted 67% of the cases tried, which demonstrates the need to elaborate new work methodologies in order to achieve, within legal limits, the capture of the author still in flagrante delicto situation.

Keywords: Flow of Criminal Justice. Willful Homicide. Survival Analysis.

1 INTRODUÇÃO

O combate à criminalidade tem sido um dos objetos de maior atenção na agenda dos formuladores de política no Brasil nos últimos anos. Com efeito, mesmo com um ambiente de elevada instabilidade política e de crise fiscal recente, o aumento da violência de forma generalizada no país compõe o grupo dos maiores desafios aos gestores públicos.

Os livros básicos de economia tratam esse problema já com requisitos teóricos bem básicos como um simples diagrama de oferta e demanda, quando em um exemplo o gestor público teria duas opções para o combate ao comércio de drogas ilegais, quais sejam o aumento do poder de polícia ou o investimento em educação. A partir deste simples exemplo, o que se observa é que enquanto o investimento em educação ataca a atividade ilegal a partir da retração da demanda, desestimulando a prática destes atos pelo menor preço final atribuído ao produto ilícito, o aumento do poder de polícia impacta na retração da oferta do bem neste mercado, mas o que acaba elevando o preço de equilíbrio e, por conseguinte, estimula a prática da atividade ilícita que é a venda de drogas ilícitas.

O que se busca com este estudo é investigar um terceiro componente determinante do combate à violência que os simples livros textos de economia não consideraram em suas análises: a efetividade do judiciário enquanto inibidor da prática da atividade criminosa. De fato, tanto a impunidade jurídica em si quanto à percepção de impunidade sentida pela sociedade pode ser considerada como um incentivo para o aumento, ou, no mínimo, para a continuidade das ações criminosas. Segundo Becker (1968), antes de ingressar em uma atividade ilícita, assim como em qualquer outro tipo de atividade legal, o indivíduo analisa os custos e benefícios de praticá-la. Somente quando os benefícios superam os custos ele decide por iniciar a empreitada delituosa. A impunidade reduz um desses custos que é a probabilidade de aprisionamento e condenação, tornando, portanto, essa atividade mais vantajosa.

No ordenamento jurídico-penal brasileiro, existe uma gama de delitos que diferem

entre si tanto pelo bem jurídico lesado, quanto pela gravidade da lesão ao bem jurídico, ou ainda por sua forma de execução, dentre outros aspectos. Os ritos de apuração também são diversos, assim como a pena prevista ou a natureza jurídica do crime, o que influencia diretamente o tempo de julgamento e imposição da pena.

Neste contexto, qualquer análise da criminalidade de forma agregada implica incorrer em imprecisões pela opção por uma análise superficial. A especial lesividade do homicídio e a sua crescente incidência na sociedade brasileira justificam a sua análise em separado e a investigação da eficiência na apuração e no julgamento dos delitos de homicídio é um dos fatores mais importantes para compreensão da importância da variável que representa o fluxo da Justiça Criminal enquanto instrumento de combate a esse problema social.

O homicídio (juntamente com os outros delitos dolosos contra a vida, dos quais ele é o principal expoente) recebe um tratamento diferenciado no nosso ordenamento jurídico, a começar pelo próprio rito de apuração (Rito do Tribunal do Júri), definido pelos artigos 406 a 497 do Código de Processo Penal, com redação dada pela Lei nº 11.689/09. Não obstante, este tipo de crime, diferentemente do que ocorre com vários outros (como furtos, roubos, receptação, apropriação indébita, ofensas à honra, violação à propriedade intelectual, etc.), é um dos menos sujeitos ao subregistro, fato que torna mais robusta uma análise quantitativa desta variável. Isso se dá principalmente em casos de homicídio consumado (morte da vítima) em decorrência da maior dificuldade de desaparecimento dos vestígios materiais da infração, notadamente o corpo de delito (que, no homicídio, consiste no próprio corpo da vítima).

Tendo em vista a peculiaridade dessa modalidade de crime, busca-se contribuir com subsídios à formação de políticas eficazes no combate à violência generalizada nas cidades brasileiras a partir de uma análise do fluxo do sistema de apuração e julgamento dos crimes de homicídio consumados em um estudo de caso restrito ao ano de 2014 na cidade de Fortaleza, por sua representatividade como a 5ª maior cidade do país em população¹ e pelo destaque dado à violência na capital cearense em nível nacional.

Em termos da representatividade deste problema na economia cearense, tem-se que em 2014 as despesas realizadas pelo Governo do Estado do Ceará na área de segurança pública² corresponderam a 7,8% do total do orçamento. Gastou-se, portanto, nesta área o equivalente à soma das despesas com assistência social, cultura, direitos da cidadania, urbanismo, habitação, saneamento, gestão ambiental e organização agrária, daí a grande

1. 2,6 milhões de habitantes. Base IBGE 2016.

2. Dados do Fórum Brasileiro de Segurança Pública, Anuário do ano de 2016.

necessidade de identificar os maiores gargalos na área de forma a melhor gerir os gastos públicos no setor.

Para tanto, propõe-se uma divisão do fluxo da Justiça Criminal em fases e realiza-se o cálculo dos tempos médios de cada fase. Serão utilizados dois cálculos distintos de média. O primeiro será o de média simples e o segundo uma estimativa do tempo médio através de técnicas de análise de sobrevivência (que levam em consideração informações censuradas, ou seja, que não terminaram determinada fase). Serão ainda estimadas as funções de sobrevivência para cada fase, utilizando o método de Kaplan-Meier, e calculada a probabilidade de um caso de homicídio “terminar” determinada fase, ou seja, prosseguir para a fase seguinte ao longo do fluxo da justiça. Deste modo, em última instância, tem-se ainda um mapeamento do Sistema de Justiça Criminal na apuração e julgamento do crime de homicídio a partir deste estudo de caso para a capital cearense, explicitando os pontos de entrave ao insucesso ou à morosidade para conclusão da investigação e o julgamento desse tipo de crime.

Além desta introdução, o presente estudo é dividido em mais três seções. A seção dois explicita o funcionamento da Justiça Criminal na investigação e no julgamento dos crimes e traz também um resumo da literatura sobre o tema e uma breve explicação sobre as prisões cautelares no Brasil. A seção três é dedicada ao estudo de análise de sobrevivência com estimação da função de sobrevivência (Kaplan-Meier) e, por fim, a seção 4 expõe as conclusões do estudo.

2 FLUXO DO SISTEMA DE JUSTIÇA CRIMINAL

O sistema de Justiça Criminal pode ser conceituado, do ponto de vista estrutural, como o conjunto de órgãos que são acionados quando do cometimento de um crime. São eles: i) a Polícia Militar, a Polícia Civil, o Ministério Público e o Poder Judiciário que serão considerados neste estudo, por serem os responsáveis para apurar, processar e julgar os delitos tipificados no Código Penal ou nas Leis Penais Extravagantes, segundo as diretrizes do Código de Processo Penal ou de Leis Processuais Penais Especiais (Ribeiro, 2010)³ e ainda ii) os órgãos de Execução Penal (Secretarias Estaduais de Justiça, Centros de Privação Provisória da Liberdade, Penitenciárias, etc)⁴ em caso de prisão em flagrante ou de

3. Apesar de não ter função tipicamente apuratória, Polícia Militar foi incluída nesse rol por razões empíricas: tanto em virtude do relevante trabalho que presta junto ao local de crime, como também porque, na maioria dos casos, ela é de fato o primeiro órgão público a ter contato direto com a infração.

4. Vale advertir que a Defensoria Pública e a OAB não foram incluídas no conceito proposto, apesar de exercerem funções essenciais à Justiça e de contribuírem enormemente para a legitimidade da prestação jurisdicional. A razão da sua não inclusão se deu porque não são instituições sobre as quais pesa, de forma direta, o dever público de promover a eficiência e a eficácia dos atos de apuração dos delitos.

prisão preventiva, ou ainda, quando do cumprimento da pena.

No processo de apuração de cada crime de homicídio doloso como em quase todo delito, nem todos os casos conseguem passar por todas as fases de apuração previstas até o julgamento final. Por isso, é também importante mensurar a quantidade dos delitos que não chegam a perfazer toda a trajetória de apuração, identificando onde ocorrem as maiores perdas, bem como o porquê de elas ocorrerem (Rifiotis et al, 2010). Essas perdas fazem com que o fluxo possua a estrutura de funil, com muitos casos ocorridos/registrados e cada vez menos casos investigados, denunciados e julgados. Ou seja, há uma filtragem dos casos ao longo do tempo.

A análise científica da impunidade deve englobar também a questão do tempo decorrido entre o crime e o seu desfecho. Em outras palavras, não basta a conclusão do processo apuratório (quando se impõe a sanção penal), é necessário ainda que o julgamento ocorra em um tempo suficientemente curto para que se tenha a percepção de que a justiça foi cumprida. Por outro lado, a trajetória ou as etapas de apuração não poderão ser aceleradas ou decotadas a ponto de serem suprimidos direitos ou garantias de qualquer das partes envolvidas no conflito.

A seguir, será apresentada uma explanação sobre a trajetória de investigação de um crime de homicídio, seguida por uma revisão da literatura sobre o tema e, por fim, uma subseção que aborda os tipos de prisões cautelares existentes no Brasil.

2.1 Trajetória de Investigação e Julgamento dos Crimes

A maioria dos estudos que analisam o tempo de processamento dos crimes de homicídio leva em consideração apenas a fase policial, a data do oferecimento da denúncia e a data da sentença condenatória. Essa simplificação da análise, desconsiderando o tempo despendido entre cada movimentação processual, gera uma perda de informação sobre os fatos e os atores responsáveis pela famigerada “morosidade processual”. À medida que se aprofunda a investigação dos atos de apuração dos delitos, observa-se que algumas subetapas ou subfases assumem especial relevo (e poderiam escapar a um olhar meramente panorâmico).

O tratamento dos delitos englobados no espaço amostral deste estudo (homicídios registrados no Município de Fortaleza no ano de 2014) revelou a necessidade de se alçar à categoria de fase autônoma do processo apuratório o tempo gasto pela Justiça para a efetivação da citação dos acusados em processo penal. Deveras, em alguns casos, como nos

de réus presos, essa etapa foi realizada com razoável facilidade, mas, em outros, como nas ações penais contra muitos réus em liberdade, essa etapa nem sequer chegou a ser superada, dado que o insucesso na localização dos citandos implicou o travamento/suspensão da própria ação penal.

Na esteira desse raciocínio, buscando-se, tanto quanto possível, suprir eventuais lacunas que distanciariam a análise matemática da realidade prática (praxe forense), este estudo trabalhará com seis fases que perfazem os principais estágios percorridos pelos órgãos que compõe o sistema de Justiça Criminal, desde o cometimento do crime de homicídio até o seu julgamento, assim divididas:

1. Fase Policial: inicia-se a partir do momento do cometimento do crime e encerra-se com a última remessa dos autos ao Poder Judiciário.
2. Fase de Denúncia: Compreende o período entre a última remessa do inquérito e o recebimento da denúncia pelo Poder Judiciário.
3. Fase de Citação: Tem início com o recebimento da denúncia termina quando da 1ª audiência de instrução.⁵
4. Fase de Instrução: Corresponde ao tempo entre a primeira audiência de instrução e a emissão da decisão de pronúncia⁶.
5. Fase de Recurso contra a Pronúncia e Julgamento: Tempo entre a decisão de pronúncia e a data da sessão de julgamento de Júri.

2.2 Literatura sobre Fluxo de Justiça Criminal no Brasil

Tendo em conta a literatura brasileira, pode-se dizer que a partir da década de 80 cientistas sociais têm se dedicado mais detidamente sobre o estudo do fluxo do Sistema de Justiça Criminal e sobre o seu grau de eficiência. O trabalho considerado pioneiro foi o de Coelho (1986). Nesse estudo, o autor utilizou dados oficiais do Serviço de Estatística, Demografia, Moral e Política para o Estado do Rio de Janeiro entre 1942 e 1967 e concluiu, tomando o último ano como representativo, que, para os crimes de roubo, furto, homicídio e estelionato, apenas 16% dos indiciados foram condenados.

5. Como um processo pode ter vários réus e cada um pode ser citado em uma data diferente, optou-se por considerar que a fase de citação compreende o tempo entre o recebimento da denúncia e a data da 1ª audiência de instrução. Dessa forma, restam englobados, inclusive, os desdobramentos entre as várias formas de citação que podem ser utilizadas, além da citação pessoal, que é a mais comum.

6. Por razões metodológicas, optou-se por não tomar como marco temporal de encerramento da fase de instrução criminal as datas de oferecimento das alegações finais pelas partes, uma vez que estas são extremamente variáveis.

Cano e Duarte (2010), além de analisarem o fluxo de processos de homicídio doloso e dos processos de roubo para a cidade do Rio de Janeiro entre 2000 e 2007, buscaram também explicar os fatores que estariam associados a uma maior probabilidade de se alcançar a sanção penal. Para o primeiro objetivo, utilizaram uma abordagem transversal, com dados de Registro de Ocorrências Policiais e de Sentenças, e concluíram que apenas 8% dos homicídios e 3% dos roubos, em média, culminaram em uma sanção penal para pelo menos um dos acusados.

Uma grande contribuição desse estudo foi a de fazer a divisão entre os crimes cuja apuração se iniciou com uma prisão em flagrante e os crimes cuja apuração se iniciou tão somente com a comunicação da ocorrência à Polícia. Da totalidade dos casos de flagrante de crimes de homicídio, 26% haviam sido julgados, em comparação com 6% daqueles que se iniciaram por registro de ocorrência. Nos casos de crimes de roubo, 68% das apurações iniciadas por prisão em flagrante haviam chegado a uma sanção penal, contra 33% daqueles que não eram flagrantes.

Em relação ao segundo objetivo do estudo, utilizando uma metodologia longitudinal retrospectiva com dados obtidos na Vara de Execuções Penais, concluíram que crimes com maior número de réus e vítimas e cometidos com o uso de arma de fogo, provavelmente pela maior repercussão social e gravidade, são os que têm menor probabilidade de ficarem impunes. De acordo com o estudo, essa probabilidade aumenta em crimes cometidos na rua e contra mulheres.

Ribeiro (2010) utilizou dados longitudinais ortodoxos provenientes da Fundação Seade para os crimes de homicídio no Estado de São Paulo, entre 1991 e 1998, e concluiu que quase a totalidade dos Boletins de Ocorrência havia se transformado em Inquérito Policial, mas apenas 22% tornaram-se Ações Penais, 14% chegaram à fase de sentença, e 8%, à condenação. O estudo constatou ainda que o tempo médio de processamento de um crime (tempo entre a data do crime e a sentença) era de 2,8 anos.

A referida autora estimou ainda três modelos logit para identificar as chances de: um caso passar da Polícia para o Judiciário; de chegar à fase de sentença, dado que o processo foi iniciado; e de o réu ser condenado, dado que o processo chegou à fase de sentença. As variáveis explicativas foram: sexo, cor, escolaridade do acusado e se a apuração havia ou não sido iniciada com a prisão em flagrante.

Os modelos apontaram que réus negros e jovens seriam aqueles com maior probabilidade de serem condenados. A autora concluiu que as características físicas dos acusados

seriam mais importantes para explicar a chance de um processo ser esclarecido e sentenciado com condenação que as características processuais. Entretanto, há de se ter cautela com esse tipo de conclusão. No estudo, os modelos apresentados possuem baixo poder de explicação, o que, em nosso sentir, indica que características processuais relevantes tenham sido omitidas, as quais poderiam definir, com maior precisão e objetividade, a chance de um processo chegar à fase de condenação.

Misse e Vargas (2007) buscaram comparar o fluxo da Justiça Criminal para os crimes contra o patrimônio e para os crimes de homicídio na cidade do Rio de Janeiro em duas séries históricas: de 1953 a 1957, utilizando uma abordagem longitudinal ortodoxa, e de 1997 a 2001, utilizando uma metodologia transversal. Para poder fazer a comparação entre os dois períodos, os autores calcularam, para o primeiro, as taxas de: elucidação policial, esclarecimento e condenação.

Na série mais antiga, os homicídios eram, em média, mais elucidados e resultavam em maior número de condenações do que as observadas em crimes de roubo. Essas taxas, para o crime de homicídio, permaneceram praticamente constantes nas duas séries. Entretanto, para o crime contra o patrimônio, essas taxas apresentaram uma queda significativa na série mais atual.

Silva (2010) trouxe um novo enfoque ao categorizar os crimes de homicídio a partir de informações de três características do crime: relação entre vítima e agressor, circunstâncias do crime (se em casa, via pública, bar etc.) e motivação atribuída pelo Promotor quando do oferecimento da denúncia. As informações foram coletadas a partir de 254 denúncias de homicídios dolosos consumados e tentados entre os anos de 2003 e 2005 de Belo Horizonte. Foram definidas nove categorias: conflitos cotidianos, questões de drogas, vingança, motivos financeiros, conflitos amorosos, trabalho policial, “bala perdida”, homicídio simples e outros.

Em relação à análise de fluxo, a autora fez, assim como proposto neste estudo, uma nova divisão intra-processual e dividiu as fases entre a denúncia e a pronúncia, entre a pronúncia e a sessão de Júri e entre o Júri e a execução da pena. A autora analisou as referidas denúncias (oferecidas entre os anos de 2003 e 2005) nos anos de 2007 e 2009 e concluiu que, enquanto no ano de 2007, 20% dos casos não tinham sido pronunciados e 66,6% não haviam sido julgados; em 2009, esses valores se reduziram para 5% e 40%, respectivamente. De acordo com a amostra da autora, o tempo médio decorrido entre a data do fato até o seu julgamento foi de 3,4 anos, e a fase mais demorada dos processos foi entre a denúncia e a pronúncia.

Mais do que uma simples verificação das médias dos tempos processuais, Vargas et al (2010) inovaram e utilizaram técnicas de análise de sobrevivência para mensurar, de forma mais apurada, a morosidade processual e a probabilidade de um processo ser sentenciado. Essa técnica também será utilizada no presente estudo. As autoras utilizaram a mesma base de dados coletada por Vargas (2000) para os crimes de estupro e calcularam que o tempo médio de processamento desde a data do registro da ocorrência até o julgamento passou de 932 dias (média simples) para 1263 dias (análise de sobrevivência). As autoras concluíram ainda que 80% dos casos foram sentenciados após 2000 dias e que a probabilidade de ser sentenciado em 500 dias era de 15%.

Outro ponto que merece destaque no trabalho em comento foi a divisão da análise dos casos de processos com réus presos e de processos com réus soltos. Isso porque o Código de Processo Penal estabelece prazos diferenciados de acordo com a situação do réu. Utilizando a técnica de análise de sobrevivência, as autoras calcularam que o tempo médio de processamento para o réu solto foi quase três vezes maior que para o réu preso (1716 e 623 dias respectivamente).

3 ESTUDO DE CASO A PARTIR DOS HOMICÍDIOS EM FORTALEZA EM 2014

A capacidade de que os órgãos que compõem a Justiça Criminal dispõem para processar e julgar os crimes de homicídio e a velocidade com que o fazem são de fundamental importância para a sensação de que a justiça foi feita, ou, caso contrário, para que seja difundida no meio social a nefasta percepção ou sensação de impunidade.

Nessa senda, esta seção apresenta uma análise dos tempos entre as fases da justiça criminal citadas na subseção 2.1, sendo tal análise feita de forma global e também subdividida em processos com réus presos e soltos. Antes, porém, será feita uma explanação sobre a base de dados do estudo, seguida de uma evidência empírica dos dados. Em seguida, será apresentada a metodologia e os resultados do estudo.

3.1 O Processo de Filtragem para Obtenção da Base de Dados

O primeiro passo dado para a construção da base de dados foi a coleta, através do sítio eletrônico da Secretaria de Segurança Pública e Defesa Social do Ceará, SSPDS/CE, dos nomes de todas as vítimas de CVLI na cidade de Fortaleza em todo o ano de 2014⁷, quando foram registradas 1.989 vítimas de CVLI⁸.

7 O ano de 2014 foi escolhido, pois é o mais antigo que a SSPDS/CE divulga o nome das vítimas de CVLI.

8 Ressalte-se que, como a base de dados foi construída a partir do nome das vítimas que vieram a óbito, ela só engloba os homicídios dolosos consumados.

Como o objetivo do vertente estudo restringe-se à análise dos crimes de homicídio dolosos, foram excluídos os casos de latrocínio, lesão corporal seguida de morte, as vítimas não identificadas, aqueles casos que posteriormente foram identificados como sendo de outra comarca e os de competência das Varas de Infância e Juventude, totalizando 1.906 vítimas e 1.896 casos de homicídio.⁹

A partir dos nomes das vítimas, foi feita uma consulta junto ao site do Tribunal de Justiça do Estado do Ceará, por meio do sistema SAJ. Esta consulta gerou três resultados possíveis: i) Não encontrado: a vítima não figura como parte em nenhum processo, ou seja, o Inquérito Policial, se aberto, ainda não foi remetido ao Judiciário¹⁰; ii) Central de Inquéritos – autos físicos: o inquérito já foi remetido ao Judiciário (mais precisamente ao órgão de protocolo e distribuição do Poder Judiciário) e, portanto, recebeu um número próprio dentro do sistema de processos judiciais; mas, além disso, o resultado Central de Inquéritos – autos físicos do sistema SAJ significa também que o órgão de protocolo e distribuição já encaminhou os autos do inquérito ao Ministério Público¹¹ e iii) Processos eletrônicos – autos virtuais: referentes àqueles casos em que o Promotor de Justiça emitiu alguma manifestação que deva ser apreciada pelo Juiz de Direito, a exemplo de um parecer pela quebra de sigilo telefônico, pela busca e apreensão domiciliar, ou, até mesmo, um parecer pelo arquivamento do feito ou uma denúncia-crime; vale dizer: todas as manifestações diversas do mero requerimento de novas diligências ou dilações de prazo de conclusão do inquérito; é importante advertir que, nas manifestações sujeitas à apreciação de um Juiz, os autos físicos são integralmente digitalizados; por outro lado, nas manifestações que implicam apenas o retorno dos autos à Delegacia de origem, não há digitalização¹².

No resultado “ii”, correspondente à informação do SAJ Central de Inquéritos – autos físicos, o sistema eletrônico do Tribunal de Justiça não monitora mais as várias tramitações do feito físico entre o Ministério Público e a Polícia Civil. Tal situação comumente perdura no tempo. O sistema SAJ somente será alimentado novamente, chegando-se ao resultado iii quando for emitida pelo Promotor alguma manifestação sujeito ao crivo do Juiz do Direito.

9. Quando o mesmo fato resultou em mais de uma morte, foi considerada apenas uma vítima, a mais velha. Dessa forma, a base de dados é constituída por casos de homicídio (mesmo processo) e não de vítimas.

10. Existe a possibilidade de o processo não ser encontrado por um erro de digitação por parte da Polícia ou por parte do Tribunal de Justiça. Embora essa seja uma hipótese provável para poucos casos ela fica difícil de ser sustentada se a quantidade de processos Não Encontrados for considerável.

11. A Central de Inquéritos é um órgão integrante da estrutura do Ministério Público destinado a viabilizar a chamada tramitação direta dos inquéritos policiais. Na prática, as Delegacias de Polícia Civil de Fortaleza remetem seus inquéritos diretamente a essa Central. Assim, os autos aportam no Ministério Público sem a necessidade da intervenção de um Juiz.

12. Os inquéritos policiais instaurados atualmente, notadamente partir do ano de 2016, já são digitais (meio virtual). Nesses casos, não há mais a tramitação física do feito policial.

Dos processos eletrônicos foram extraídas informações sobre a data de remessa do inquérito ao Judiciário, a data do oferecimento e a do recebimento da denúncia, bem como sobre as datas da primeira e da última audiência, da pronúncia, da sessão de Júri. Todas as informações foram coletadas entre abril e julho de 2016, ou seja, entre 16 a 30 meses (em média, quase dois anos) após o cometimento do crime.

Em cada fase, serão apresentadas as informações referentes ao total de processos e, após, os casos serão divididos em três categorias: Flagrante, aqueles oriundos de prisão em flagrante; Preso Não Flagrante, aqueles em que o acusado/réu foi preso preventivamente em algum momento da investigação ou da ação penal ou ainda que se tenha notícia (dentro do processo) que ele se encontra preso nos autos de outro processo; e Solto, aqueles em que o acusado/réu se encontra solto, mesmo que tenha sido expedido mandado de prisão preventiva contra ele (não cumprido).

3.2 Evidência Empírica

A partir dos 1.896 casos de homicídio dolosos consumados divulgados pelas SSSPS/CE e considerados após o processo de filtragem para esse estudo, apenas 1.517, o que representa 80% do total, foram remetidos ao Poder Judiciário. Portanto, 20% dos homicídios que chegaram ao conhecimento da Polícia Civil, após dois anos (em média) do cometimento crime, ou não tiveram o Inquérito Policial instaurado, ou este, apesar de aberto, além de não ter sido concluído, não foi encaminhado, nem mesmo por uma vez, ao Judiciário. Assim, como não se tem nenhuma informação sobre os processos que não foram remetidos, a base de dados do estudo será composta apenas dos 1.517 casos que chegaram ao conhecimento do Poder Judiciário.

A Tabela 1 mostra o total de casos em cada fase, bem como a quantidade de processos que finalizaram cada uma delas. São apresentados os dados para o total de casos e subdivididos segundo a ocorrência ou não de prisão (e o tipo) ao longo da investigação e da Ação Penal.

Observando o total de casos, percebe-se que a fase que concentra a maior porcentagem de casos não concluídos é a Fase Policial. Dos 1.517 casos que foram remetidos ao menos uma vez ao Poder Judiciário quase 90% não tiveram a investigação concluída, ou seja, não conseguiram reunir elementos suficientes acerca da autoria e materialidade do crime que possibilitassem o oferecimento da denúncia-crime por parte do Promotor de Justiça. A partir dessas informações, independente de como o fluxo se comporte na fase judicial, pode-se afirmar que a maior filtragem de processos, e, conseqüentemente, a

maior causa da impunidade, é a fase de investigação policial. Resultados semelhantes são encontrados por Ribeiro (2010) e Vargas et al (2010). No primeiro estudo, apenas 22% dos inquéritos de homicídio foram remetidos e, no segundo, apenas 55% dos inquéritos de estupro instaurados prosseguiram para as demais fases.

Tabela 1: Total de Casos em Cada Fase

	Total Casos	Concluíram a fase	% Não Concluíram		Total Casos	Concluíram a fase	% Não Concluíram
<i>Policial</i>	1517	183	87,9%	<i>Instrução</i>	102	46	54,9%
Flagrante	28	28	0%	Flagrante	23	17	26,1%
Preso Não Flag	73	72	1,4%	Preso Não Flag	53	21	60,4%
Solto	1416	83	94,1%	Solto	26	8	69,2%
<i>Denúncia</i>	183	164	10,4%	<i>Rec. Pron. e Julg</i>	46	18	60,9%
Flagrante	28	27	3,6%	Flagrante	17	10	41,2%
Preso Não Flag	72	72	0%	Preso Não Flag	21	8	61,9%
Solto	83	65	21,7%	Solto	8	0	100%
<i>Citação</i>	164	102	37,8%	<i>Pós-Julgamento</i>	1517	18	1,2%
Flagrante	27	23	14,8%	Flagrante	28	10	35,7%
Preso Não Flag	72	53	26,4%	Preso Não Flag	73	8	11%
Solto	65	26	60%	Solto	1416	0	0%

Fonte: elaboração própria.

* Dados para o município de Fortaleza no ano de 2014.

Ademais, constata-se que dos 183 inquéritos com investigação concluída dos crimes de 2014 em Fortaleza, 164 foram denunciados (89,6%). Dentre os casos que não concluíram a fase, 17 foram arquivados (9,29%) e 2 receberam decisão de declínio de competência (1,09%). Dentre os motivos para o arquivamento, 9 casos foram por morte do indiciado, 3 por legítima defesa e 5 por restarem exauridas as possibilidades de investigação sem que chegasse a nenhum indiciamento. Nos dois primeiros casos e nos casos de declínio de competência não há o que se falar em matéria de impunidade. O primeiro (morte do agente) é resolvido com a declaração de extinção da punibilidade, prevista no artigo 107, I, do Código Penal; o segundo (de legítima defesa) representa a situação mais comum dentre as hipóteses de exclusão da ilicitude do fato, previstas no artigo 23 do Código Penal; e, por derradeiro, no terceiro caso, não há impunidade, pois o fato ainda será apreciado, só que por outro juízo. O filtro da impunidade refere-se, então, apenas aos 5 casos arquivados por falta de indiciamento, que representam 0,26% do total de crimes, 0,33% do total de crimes remetidos e 2,73% dos crimes com investigação concluída. Assim, não há o que falar sobre filtro de impunidade nessa fase, e sim sobre uma possível morosidade processual. Ressalte-se que estudos transversais, que consideram apenas o número de arquivamentos, sem

atentar para as suas razões, obtêm resultados viesados.

Na fase de Citação, dos 164 casos denunciados, a maioria foi finalizada e apenas 38% dos casos permaneceram nesta fase. Da leitura dos processos, percebeu-se que é muito comum só se conseguir efetuar a citação pessoal do réu quando este é preso, seja por força de um mandado de prisão preventiva, seja porque cometeu outro delito posterior e veio a ser preso em flagrante. Outrossim, dos 102 casos que iniciaram a Fase de Instrução, pouco mais da metade permaneceu nessa fase em 2014.

Muito embora o Código de Processo Penal determine que a audiência de instrução seja una, e, embora não seja raro encontrar processos que atendam esse quesito, forçoso é convir que bem mais comum é encontrar processos que se arrastam na tentativa de localizar as testemunhas ouvidas na Fase Policial. É comum também que as testemunhas, ao serem ouvidas em juízo, alterem, total ou parcialmente, o teor de seus depoimentos prestados perante a Autoridade Policial Civil, principalmente por medo de represálias dos próprios acusados ou de comparsas destes. Isso dificulta muito o trabalho acusatório, haja vista que os elementos de convicção produzidos na Fase Policial precisam ser confirmados em juízo para que possam justificar uma decisão condenatória.

Em relação aos processos que tiveram decisão de pronúncia, do total de 46 apenas 18 foram julgados, ou seja, 61% dos casos ainda esperam decisão do recurso contra a decisão de pronúncia ou que chegue a data designada para a sessão de julgamento.

Quando se divide os casos em Flagrante, Preso Não Flagrante e Solto percebe-se que em todas as fases os casos com réu solto apresentam maior porcentagem de casos que não conseguiram terminar a fase. Merecem destaque a Fase Policial e a Fase de Recurso Contra a Decisão de Pronúncia e Julgamento. No primeiro caso, enquanto todos os casos iniciados com prisão em flagrante e 98,6% dos casos de Preso Não Flagrante tiveram a investigação concluída, no caso de réu solto esse valor foi de apenas 5,9%. No que tange ao julgamento dos casos, quase 60% dos casos oriundos de prisão em flagrante foram julgados, o que não ocorreu com nenhum processo de réu solto.

Por fim, analisando o denominado “Pós-Julgamento” (ver definição em 2.1), fica evidente a grande impunidade aos autores de crimes de homicídio no município de Fortaleza. Em média dois anos após o cometimento do crime, dos 1.517 casos de homicídios apenas 18 foram julgados, o que representa 1,2% do total. A exceção se encontra nos crimes iniciados com prisão de flagrante, com 35% dos casos julgados no mesmo período.

Em resumo, em todo o ano de 2014, dos 1.517 casos de homicídio, houve apenas 28

flagrantes, o que representa 1,85% do total. Cuida-se de um número extremamente baixo. Esse resultado demonstra a necessidade de elaboração de novas metodologias de trabalho para os policiais em campo.

3.3 Metodologia para um Exercício Empírico¹³

Essa seção trará a metodologia que será utilizada para análise do tempo em que cada fase citada na subseção 2.1 se desenvolve. Devido à peculiaridade das informações para os casos de homicídio, com muitos crimes cometidos e pouco julgados, o que faz com que a estrutura do fluxo de processamento tenha o formato de um funil, a utilização de média simples, ou outras medidas similares que ignoram esta estrutura de dados, caso utilizadas, levarão a resultados viesados, ou, no mínimo, resultados incapazes de sintetizar o comportamento do fenômeno.

É necessário então utilizar uma técnica que considere todos os casos, tendo ele alcançado determinada fase ou não, para a computação dos tempos médios. Essa técnica é a de Análise de Sobrevivência, cuja aplicação nesse tipo de dados iniciou-se com Vargas et al (2010).

O primeiro passo para o cálculo desta média é a estimação da Função de Sobrevivência, que atribui a cada tempo uma probabilidade de um caso de homicídio permanecer (ou sobreviver) em determinada fase. A seção 3.3.1 apresenta a metodologia geral para o cálculo dessa função e a seção 3.3.2 mostra como utilizar esta técnica para dados judiciais.

3.3.1. Função de Sobrevivência e Estimador *Kaplan-Meier*¹⁴

Muito utilizado em estudos da área médica, na Análise de Sobrevivência, a variável de interesse é o tempo até a ocorrência de determinado evento, T . Como geralmente esse tempo remete-se a algo negativo, ele é denominado tempo de falha (ou tempo de sobrevivência).

Essa abordagem apresenta como principal vantagem a de ser robusta à presença de censura. Esta pode ocorrer quando se perde o acompanhamento de uma determinada

13. A metodologia descrita neste capítulo seguirá Colosimo e Giolo (2006) e Favero e Belfiore (2015)

14. Embora se tenha conhecimento que a grande maioria dos estudos que utilizam essa metodologia associa à análise da Função de Sobrevivência o Modelo de Regressão de Cox, optou-se neste estudo pela não utilização do referido modelo. Isso porque o Modelo de Cox serve para avaliar o efeito de covariadas no evento de interesse (nesse caso, o término de cada fase). Uma das principais suposições do modelo é que indivíduos (casos de homicídio) de grupos diferentes apresentam funções da taxa de sobrevivência proporcionais entre si. No presente caso, a única variável citada no estudo que poderia ter influencia na variável de interesse é justamente a variável indicativa de grupos, o que prejudica a análise.

observação ou quando o estudo termina antes que o evento de interesse aconteça, porém, é importante destacar que, para o cálculo das estatísticas, são utilizados todos os dados, inclusive os censurados.

Seja δ_i uma função indicadora de falha ou censura e o tempo de falha para as observações i ($i=1, 2, 3, \dots$), os dados podem ser representados pelo par (t_i, δ_i) da seguinte forma: $\delta_i=1$, se t_i é um tempo de falha ou $\delta_i = 0$, se t_i é um tempo censurado.

A função de sobrevivência, $S(t)$, definida como a probabilidade de uma observação não falhar até determinado tempo t , é uma forma de representar o tempo de sobrevivência, variável aleatória contínua e não negativa, que pode ser descrita forma: $S(t)=P(T \geq t)$ e, em termos de função de sobrevivência em termos da função acumulada, $S(t)=P(T > -t)=1-F(t)$.

Pode-se representar o tempo de sobrevivência através da função falha ou função risco (*hazard function*), $\lambda(t)$, que é a probabilidade que o evento de interesse ocorra em um determinado intervalo de tempo, Δt , dado que ele ainda não ocorreu.

$$\lambda(t) = \frac{S(t) - S(t + \Delta t)}{\Delta t S(t)}$$

A função de sobrevivência pode ser estimada utilizando estimador não-paramétrico de *Kaplan-Meier* (1958). Na ausência de censura, ele pode ser definido como a razão entre o número de observações que falharam até o tempo t e o número total de observações.

Na presença de censura, suponha que as falhas (evento de interesse) ocorram no tempo t_j com $j=1, 2, \dots, k$ e seja d_j o número de falhas. Então, considerando n_j o número de observações que ainda não sofreram o evento em t_j , $S(t)$ pode ser escrito em termos de probabilidade condicional da seguinte forma:

$$\hat{S}(t) = \prod_{j:t_j < t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j < t} \left(1 - \frac{d_j}{n_j} \right)$$

Pode-se ainda dividir as observações em grupos, ou estratos, distintos e comparar as suas curvas de sobrevivência. Neste estudo, assim como feito na seção 3, dividiremos os casos em três grupos, de acordo com a presença ou não de prisão (e o tipo): *Flagrante*, *Preso Não-Flagrante* e *Solto*.

Para a comparação das curvas de sobrevivência será utilizado o teste de log-rank (Mantel, 1966). Esse teste analisa os eventos que foram observados em cada grupo e os compara com os seus valores esperados, sob a hipótese nula de que o risco é o mesmo para todos os grupos, ou seja, $H_0: \lambda_1(t) = \lambda_2(t) = \dots = \lambda_k(t)$, com igual ao número de estratos. A hipótese alternativa é que pelo menos uma função é diferente da outra, ou seja, no caso de

três ou mais estratos, que é a realidade deste estudo, rejeitar a hipótese nula não significa dizer que todas as funções de sobrevivência diferem significativamente entre si, basta que apenas uma se comporte assim. Ressalte-se ainda que a hipótese é feita para todo o tempo de sobrevivência e que, portanto, pode haver trechos de duas ou mais curvas semelhantes, e, mesmo assim, elas diferirem entre si (Carvalho *et al*, 2011). Como temos três estratos (Flagrante, Preso Não-Flagrante e Solto), adicionalmente, será realizado o teste *log-rank* (Cox-Mantel) para a comparação das funções de sobrevivência em pares.

3.3.2 Análise de Sobrevivência e Dados Judiciais

Para utilizar a técnica de Análise de Sobrevivência, é necessário definir o tempo inicial do estudo, a escala de medida do tempo, o evento de interesse (falha) e o que constitui a censura.

A escala de medida do tempo é em dias e serão calculadas as médias intra-fases. Assim, para a primeira fase, o tempo inicial é o cometimento do crime e, para cada fase subsequente, o tempo inicial é o fim da fase anterior. Por sua vez, o evento de interesse será o fim da respectiva fase, da seguinte forma: Fase Policial: remessa dos fólhos policias ao Poder Judiciário; Fase de Denúncia: recebimento da denúncia; Fase de Citação: primeira audiência de instrução; Fase de Instrução: decisão de pronúncia; Fase de Recurso contra a Pronúncia e Julgamento: sessão de julgamento do júri. Será ainda calculado o tempo total para o julgamento, cujo tempo inicial é o cometimento do crime e o evento de interesse é a sessão de julgamento do júri.

Por sua vez, as censuras são aqueles processos que ainda não terminaram determinada fase. Foram excluídos dessa análise os processos classificados como “Não Encontrados”, pois não pode haver censura no tempo inicial. O espaço amostral do estudo são os 1.517 casos remetidos ao Poder Judiciário e, assim como na sessão anterior, os casos foram divididos segundo a existência ou não de prisões cautelares durante o curso do processo, totalizando três grupos: Flagrante, Preso Não-Flagrante e Solto.

Em síntese, serão estimadas as funções de sobrevivência para cada fase e para o período total entre o cometimento do crime e o seu julgamento (denominado Pós-Julgamento), utilizando o estimador de Kaplan-Meier, e, a partir dessa função serão estimados os tempos médios de cada uma das fases citadas na seção 2.1.

3.4 Resultados

3.4.1 Tempo Médio Estimado

Nesta seção serão apresentados os resultados dos tempos médios que um caso de homicídio leva para terminar determinada fase da investigação ou da Ação Penal. A Tabela 3 mostra duas estimativas de tempo médio para todas as fases propostas, para o total de casos e dividindo-os segundo a ocorrência (e o tipo) de prisão. A primeira é a média simples que, como visto em 3.3, tende a subestimar os resultados e a segunda utilizando técnicas de Análise de Sobrevivência, que utiliza em seu cálculo todos os casos, tendo eles concluído a fase ou não, e que apresenta resultados mais fidedignos.

Tabela 2: Estimativa do tempo médio de cada fase

	Média Simples	Análise de Sobrevivência		Média Simples	Análise de Sobrevivência
<i>Policial</i>	231,99	492,06	<i>Instrução</i>	169,58	399,95
Flagrante	13,28	13,28	Flagrante	228,14	280,45
Preso Não Flag.	241,71	248,92	Preso Não Flag.	126,94	407,77
Solto	290,96	557,13	Solto	157,81	493,63
<i>Denúncia</i>	67,26	101,49	<i>Rec. Pron. e Julg.</i>	152,89	331,96
Flagrante	79,33	79,44	Flagrante	159,4	-
Preso Não Flag.	53,49	53,49	Preso Não Flag.	147	-
Solto	94,12	157,93	Solto	-	-
<i>Citação</i>	252,99	268,47	<i>Pós-Julgamento</i>	499,28	852,5
Flagrante	169,3	208,69	Flagrante	479,6	-
Preso Não Flag.	252,98	254,38	Preso Não Flag.	523,88	-
Solto	318,48	338,33	Solto	-	-

Fonte: elaboração própria.

* Dados para o município de Fortaleza no ano de 2014.

Em uma primeira leitura da tabela, percebe-se que os valores obtidos via média simples são sempre iguais ou inferiores aos estimados via Análise de Sobrevivência, sendo que a igualdade ocorre nos casos de ausência de censura. Quanto maior a quantidade de casos censurados (que não terminam a fase), maior a diferença entre as médias.

Focando a análise primeiramente no total de casos em cada fase (sem a subdivisão por tipo de prisão), nota-se que a fase mais demorada é a fase policial. Demora-se, em média, 1 ano e 4 meses apenas para concluir a investigação de um crime de homicídio e se iniciar a etapa judicial, sendo essa fase uma das grandes responsáveis pela sensação de impunidade.

Em sequência, tem-se a fase de instrução. A demora na conclusão desta fase deve-se principalmente a dificuldades de localizar as testemunhas ouvidas na fase policial. A lentidão da fase de recurso contra pronúncia e julgamento é devida, a existência de

recurso e tempo para o seu julgamento da fase de Citação a dificuldade de localizar o réu. A fase de denúncia mostrou-se relativamente rápida.

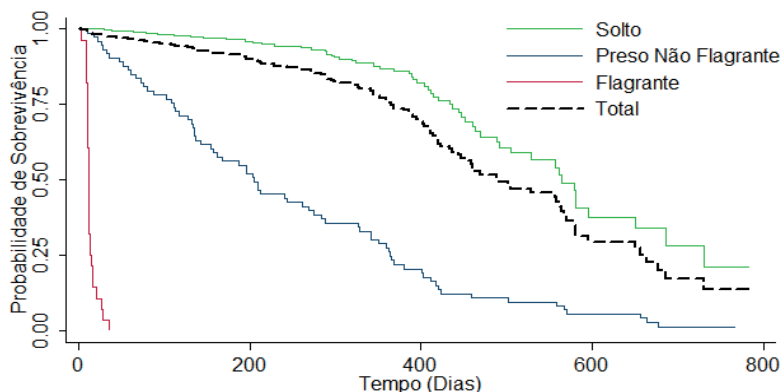
Quando se analisa o tempo médio estimado entre o cometimento do crime e o seu julgamento (Pós-julgamento), percebe-se que esse valor foi de 852 dias, ou seja, aproximadamente 2 anos e quatro meses. Comparando com outros estudos, Ribeiro (2010) concluiu que, em média, o julgamento do crime de homicídio no Estado de São Paulo demora 2,8 anos. Esse valor é de 3,4 anos para a cidade de Belo Horizonte, segundo Silva (2010). Por sua vez, Adorno e Izumino (2007), com base em estudos de Adorno (1991 e 1994) e Izumino (1998), concluíram que, para os crimes dolosos contra a vida no Estado de São Paulo, a morosidade padrão é de, em média, dois anos.

Analisando os tempos estimados intra-fases de acordo com a subdivisão segundo a ocorrência ou não de prisão percebe-se os tempos médios estimados dos casos de Flagrante e Preso Não Flagrante são sempre inferiores dos casos de réus Soltos. A maior diferença é observada justamente na fase policial em que todos os casos de Flagrante tiveram a investigação concluída com tempo médio de 13 dias.

3.4.2 Funções de Sobrevivência

O Gráfico 1 mostra a estimativa da função de sobrevivência da Fase Policial, utilizando o estimador de Kaplan-Meier, para o total de casos e subdividido segundo a ocorrência (e o tipo) de prisão. No gráfico abaixo, o eixo y, indica a probabilidade de sobrevivência. No presente estudo, “sobreviver” significa permanecer em determinada fase, ou seja, algo indesejável. Para a comparação das curvas utiliza-se a regra de que quanto mais à esquerda ela estiver mais rápido o evento de interesse (fim de cada fase) ocorrerá.

Gráfico 1: Função de Sobrevivência para Fase Policial



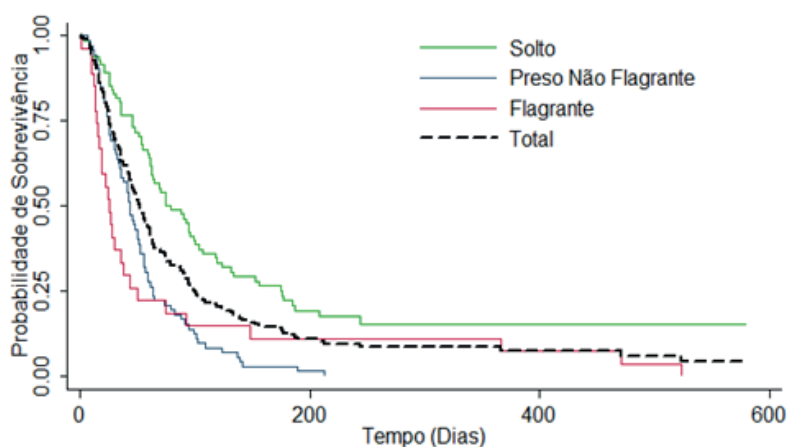
Fonte: elaboração própria. (*) Dados para o município de Fortaleza no ano de 2014. (**) De acordo com o teste de log-rank, rejeita-se a hipótese nula de igualdade das funções de sobrevivência (Ver Apêndice).

Percebe-se que as funções de sobrevivência referentes ao total de casos e de réu Solto se comportam de maneira bastante similar, pois, conforme argumentado em 3.2, a grande maioria dos casos transcorre sem a imposição (ou o cumprimento) de alguma prisão cautelar. Extrai-se do gráfico acima que estas duas funções de sobrevivência iniciam-se côncavas e assim permanecem até aproximadamente 450 dias após o cometimento do crime, quando passa a ser convexa. Assim, até atingir o ponto de inflexão, a probabilidade de sobrevivência, ou seja, de não atingir o evento de interesse (que na Fase Policial é a conclusão das investigações) se mantém bastante elevada, em torno de 70% para réu solto e 57% para o total de casos. Após esse período passa a decrescer a taxas crescentes e 785 dias após o crime, essas probabilidades passam a ser, respectivamente de 21% e 14%.

A função de sobrevivência referente aos casos iniciados com prisão em Flagrante, por sua vez, decresce rapidamente apenas alguns dias após a data de cometimento do crime. Após 10 dias a probabilidade de um caso não ter a investigação concluída é de 61%. Com 20 dias esse valor já passa a ser de 14% e com 28 dias a probabilidade tende a 0%. A função de sobrevivência dos casos Preso Não Flagrante, por sua vez, se comporta de maneira intermediária entre as curvas de Flagrante e Solto.

O Gráfico 2 a estimativa da função de sobrevivência da Fase Denúncia utilizando o estimador de Kaplan-Meier, para o total de casos e segundo a divisão proposta. Essa função foi estimada utilizando apenas os 183 casos que concluíram a Fase Policial, conforme dados da Tabela 1.

Gráfico 2: Função de Sobrevivência para Fase de Denúncia



Fonte: elaboração própria. (*) Dados para o município de Fortaleza no ano de 2014. (**) O teste de log-rank (Ver Apêndice) rejeita a hipótese nula de igualdade entre as curvas. Na comparação entre os pares, não se pode rejeitar a hipótese nula de igualdade entre as funções de Preso não Flagrante e Flagrante.

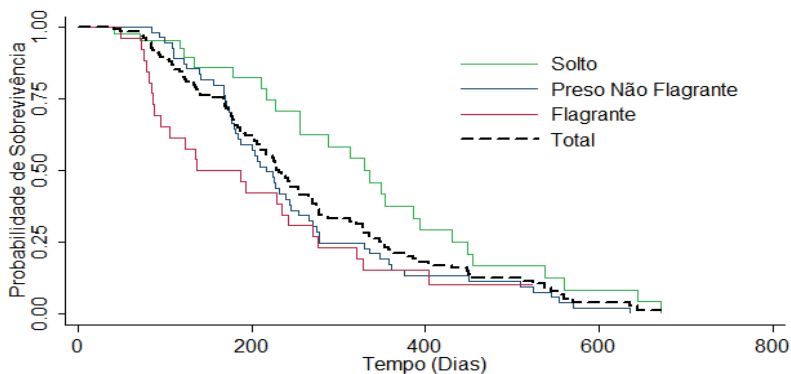
Percebe-se que a função de sobrevivência para o total de casos decai de forma bastante acentuada até aproximadamente 100 dias após o seu início (conclusão das investigações). Isso indica que essa fase transcorre com relativa rapidez, com muitos casos terminando logo e poucos que ainda perduram muitos dias, mas com baixa probabilidade de sobrevivência (permanecer na fase).

A curva de casos com réu solto, como esperado, está mais a direita durante toda a sua extensão. Por sua vez, as curvas de Flagrante e Preso não Flagrante possuem características semelhantes. Ambas iniciam-se com uma acentuada queda e se estabilizam com baixa probabilidade de sobrevivência. Uma das possíveis explicações para o andamento processual semelhante nesses dois casos é que muitas vezes o pedido de prisão preventiva é feito juntamente com o oferecimento da denúncia-crime por parte do Promotor de Justiça e, caso o Mandado seja expedido e cumprido com celeridade, os dois casos estariam no mesmo patamar de prioridade. A probabilidade que a denúncia seja recebida pelo Juiz 200 dias após a conclusão das investigações é de 99% para os casos de Preso Não Flagrante e de 89% para os casos de Flagrante.

O Gráfico 3 mostra o comportamento da função de sobrevivência, estimada via *Kaplan-Meier*, para os 164 casos que iniciaram a Fase de Citação, para o total de casos e de acordo com a subdivisão proposta. Esta fase é, em geral, bastante morosa.

Analisando a função de sobrevivência para o total de casos, nota-se que a probabilidade que esta fase termine em 90 dias é de apenas 9%; em 180, de 33%, e, em 365 dias esse valor sobe para 79%. Isso indica que, para a maioria dos casos, a citação é realizada entre seis meses e um ano a partir do recebimento da denúncia.

Gráfico 3: Função de Sobrevivência para Fase de Citação

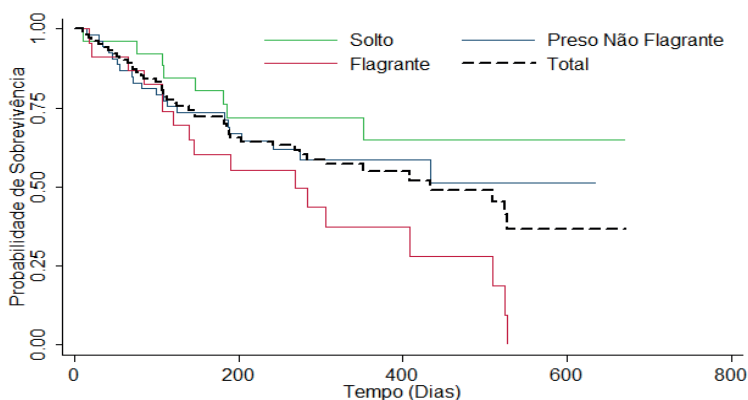


Fonte: elaboração própria. (*) Dados para o município de Fortaleza no ano de 2014. (**) O teste geral, rejeita a hipótese nula de igualdade das funções, indicando que pelo menos uma difere significativamente das demais. Na comparação aos pares, rejeita-se a hipótese nula de igualdade da função Solto com as demais, porém não se rejeita a igualdade entre as funções Flagrante e Preso Não Flagrante (Ver Apêndice).

Em relação a subdivisão dos casos em Flagrante, Preso Não Flagrante e Solto, um ano após o recebimento da denúncia, a probabilidade que a fase de citação termine, ou, em outras palavras, que seja realizada a 1ª audiência de instrução, é, respectivamente, de 85%, 85% e 62%. As três funções de sobrevivência apresentam comportamento semelhante, com a função Solto predominantemente à esquerda e a função Flagrante, à direita.

O Gráfico 4 mostra a função de sobrevivência para o total de casos e segundo a ocorrência (e o tipo) de prisão para os 102 casos que iniciaram a Fase de Instrução.

Gráfico 4: Função de Sobrevivência - Fase de Instrução



Fonte: elaboração própria. (*) Dados para o município de Fortaleza no ano de 2014. (**) O teste de log-rank, rejeita a hipótese nula de igualdade entre as três funções. Na análise aos pares, verifica-se que a função de sobrevivência relativa aos casos de Flagrante se apresenta significativamente diferente das demais e que não se pode rejeitar a hipótese de igualdade entre as funções Preso não Flagrante e Solto (Ver Apêndice).

Para o total de casos, percebe-se no gráfico que, em relação as fases anteriores, a função de sobrevivência se apresentam de forma mais aplainada, o que indica uma maior dispersão. Por exemplo, diferente da Fase de Denúncia, em que há uma acentuada queda no início da curva, o que indica que uma grande quantidade de processos consegue terminar nos primeiros dias, neste caso, a probabilidade de um processo terminar em 200 dias não difere drasticamente da probabilidade de terminar em 400 dias. A título de ilustração, essas probabilidades são, respectivamente, 34% e 45%.

Em relação a divisão proposta, nota-se que até aproximadamente 200 dias as três funções apresentam comportamento semelhante. Após, a função de sobrevivência relativa aos casos de Flagrante decresce até a probabilidade de a Fase de Instrução não terminar tender a zero em 527 dias após o seu início. As funções de Preso não-Flagrante e Solto, por sua vez, com respectivamente se estabilizam com probabilidade de não terminar a fase superiores a 50%.

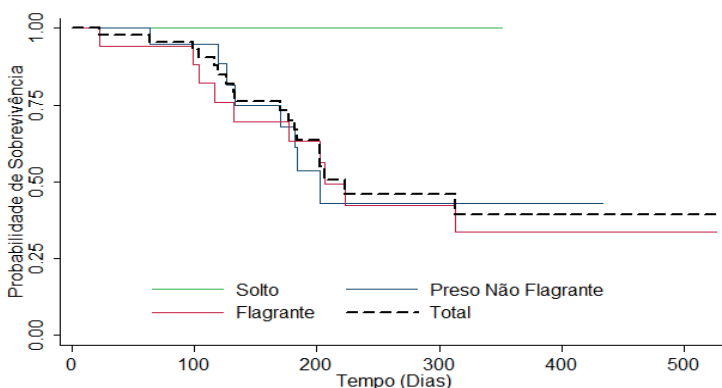
O Gráfico 5 mostra as funções de sobrevivência, estimada via Kaplan-Meier, para o

total dos casos e segundo a subdivisão proposta, para os 46 casos que iniciaram a Fase de Recurso contra a Pronúncia e Julgamento.

A função de sobrevivência Solto é representada por uma linha horizontal, com probabilidade de sobrevivência (processo não terminar) igual a um, por não ter nenhum processo julgado (evento de interesse) no período analisado. Por conseguinte, a função de sobrevivência para o total de casos, que nas fases anteriores se apresentava bastante similar a função de Solto, neste caso se aproxima do formato das funções Flagrante e Preso Não-Flagrante.

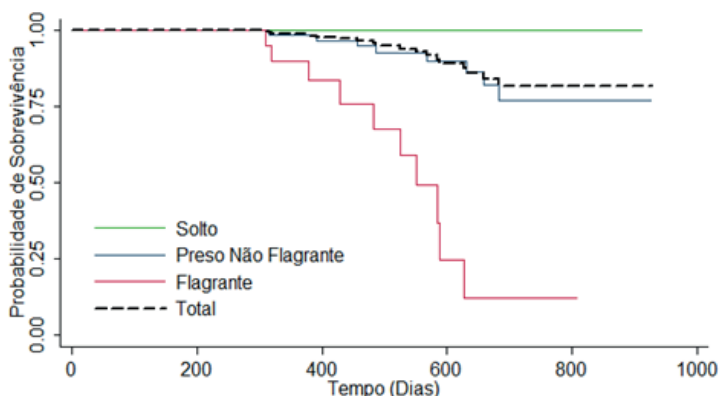
De forma geral, percebe-se que até aproximadamente 100 dias, as três funções possuem comportamento que sugere elevada probabilidade de não terminar a fase (ser julgado). Por conseguinte, observa-se uma forte queda até aproximadamente 200 dias, quando volta a se estabilizar. A título de exemplo, as probabilidades de o processo ser julgado em 100 e 200 dias são, respectivamente, de 9% e 36%.

Gráfico 5: Função de Sobrevivência: Fase de Recurso contra a Pronúncia e Julgamento



Fonte: elaboração própria. (*) Dados para o município de Fortaleza no ano de 2014. (**) De acordo com o teste de log-rank, não se pode rejeitar a hipótese de igualdade das funções de sobrevivência. Esse resultado é confirmado pela análise em pares (Ver Apêndice).

Conforme exposto, após a análise intra-fase, será realizada uma análise geral, denominada Pós-Julgamento, cujo tempo inicial é o cometimento do crime e o evento de interesse o julgamento do processo. O Gráfico 6 mostra a função de sobrevivência para o total de casos e subdivididos segundo a ocorrência ou não de prisão. Como poucos processos alcançaram o evento (foram julgados), a probabilidade de sobrevivência (processo não terminar) se mantém alta durante todo o domínio.

Gráfico 6: Função de Sobrevivência – Pós-julgamento

Fonte: elaboração própria. (*) Dados para o município de Fortaleza no ano de 2014. (**) De acordo com o teste de log-rank, rejeita-se a hipótese nula de igualdade das funções de sobrevivência e, na comparação em pares todas as funções se mostraram diferentes entre si (Ver Apêndice).

Mais uma vez, a função referente aos processos com réus Soltos se apresenta como uma linha horizontal com probabilidade de sobrevivência igual a um por não ter tido nenhum processo julgado. Até aproximadamente 300 dias todas as três funções apresentam probabilidade de sobrevivência (não terminar) igual a um. Isso indica que por mais célere que transcorra a investigação e a Ação Penal a chance que se tenha um resultado condenatório até cerca de dez meses após o do crime é quase zero.

Após esse período, a função de sobrevivência Flagrante decresce de forma acentuada e se mantém bem mais a esquerda que as demais. Isso indica que os processos que se iniciam com prisão em Flagrante são aqueles com maior probabilidade de terminarem e com maior rapidez. 629 dias após o cometimento do crime a probabilidade de o processo dessa natureza não ter sido julgado é de apenas 12%.

Por sua vez, as funções de sobrevivência para o total de casos e de Preso não Flagrante passam a decrescer, só que a primeira de forma bem menos expressiva, se estabilizando com altas probabilidades de não ter o processo julgado. A título de exemplo, 900 dias após o crime, essa probabilidade é de, respectivamente, 82% e 77%.

Para sintetizar as informações das funções de sobrevivência acima, a Tabela 4 mostra os valores dos quartis para cada fase, para o total de casos de acordo com a subdivisão proposta, e as respectivas probabilidades de um caso concluir a fase em cada um dos quartis.

Tabela 4: Probabilidade de Sobrevivência por Quartis da Distribuição

	Q1	Q2 (dias)	Q3	Q1 (Probabilidade de passar de Fase)	Q2	Q3
<i>Policial</i>	56	109	210	3,1%	5,2%	11,0%
Flagrante	9	10	13	17,9%	39,3%	75,0%
Preso Não Flag	111	205	364	26,0%	50,7%	75,3%
Solto	57	108	208	0,7%	1,9%	4,5%
<i>Denúncia</i>	25	50	96	23,6%	48,3%	73,8%
Flagrante	13	24	43	22,1%	48,1%	74,0%
Preso Não Flag	24	42	61	25,0%	48,6%	73,6%
Solto	36	74	152	23,2%	47,6%	72,1%
<i>Citação</i>	71	152	255	2,3%	23,7%	58,7%
Flagrante	84	137	276	23,1%	50,0%	76,9%
Preso Não Flag	99	174	254	3,5%	29,5%	65,7%
Solto	45	107	244	2,0%	4,5%	29,1%
<i>Instrução</i>	120	228	353	23,6%	35,6%	45,0%
Flagrante	107	190	321	26,1%	44,7%	62,7%
Preso Não Flag	112	225	336	24,6%	35,6%	41,5%
Solto	147	313	431	19,6%	28,1%	35,3%
<i>Rec. Pron. e Julg.</i>	107	177	206	9,4%	30,0%	49,6%
Flagrante	116	202	313	24,0%	43,7%	66,2%
Preso Não Flag	112	170	188	5,3%	32,0%	46,4%
Solto	76	108	181	0,0%	0,0%	0,0%
<i>Pós-Julgamento</i>	60	123	249	0,0%	0,0%	0,0%
Flagrante	320	400	540	10,0%	16,4%	40,9%
Preso Não Flag	419	540	668	3,2%	7,2%	17,7%
Solto	58	113	217	0,0%	0,0%	0,0%

Fonte: elaboração própria. (*) Dados para o município de Fortaleza no ano de 2014.

A Tabela 4 deve ser interpretada da seguinte forma: para o total dos casos da fase Policial, 25% dos 1.517 casos que iniciaram essa fase foram censurados ou passaram para a fase de Denúncia em até 56 dias após o seu início, com probabilidade de passar de fase (em 56 dias) de 3,1%. Nesse mesmo raciocínio, 75% dos casos foram censurados ou passaram de fase em até 210 dias após o início, com probabilidade de passar de fase de apenas 11%.

De forma geral, mesmo analisando o terceiro quartil, ou seja, para 75% dos casos que iniciam cada fase, percebe-se que para as fases a probabilidade de passar fase não passa dos 50% (com exceção das fases de Denúncia e Citação).

Quando se subdivide os casos segundo a ocorrência de prisão, ainda focando-se no terceiro quartil, nota-se que as maiores probabilidades são encontradas nos casos iniciados com prisão em flagrante. Nos casos Solto, com exceção da fase de Denúncia, a maior probabilidade observada foi de apenas 35%.

Ressalte-se que o Pós-Julgamento tem probabilidade zero em todos os quartis, pois pelo menos 75% de todos os casos foram censurados em até 249 e o processo que foi julgado mais rápido durou 310 dias, ou seja, todos os quartis estão localizados em um trecho em que a função de sobrevivência assume o valor constante igual a 1.

4 CONSIDERAÇÕES FINAIS

Considerando a necessidade da elaboração e implementação de políticas públicas alternativas eficazes no combate à criminalidade nas cidades brasileiras, este estudo considera para uma “análise fora da caixa” o combate à sensação de impunidade principalmente por parte do criminoso e o respectivo ganho de sinergia que a inserção do poder judiciário pode trazer aos tradicionais mecanismos já previstos nos livros básicos de economia.

Como objeto da investigação para um estudo de caso tem-se a trajetória de apuração dos 1.896 casos de homicídios dolosos, praticados e registrados na cidade de Fortaleza no ano de 2014, na modalidade consumada.

Observa-se inicialmente que ao longo do fluxo da justiça criminal o sistema vai perdendo a sua eficácia e, portanto, sistema vai assumindo assim a forma de funil em termos de registros de crimes e efetiva penalidade desde a Fase Policial até a Fase de Recurso contra a Pronúncia e Julgamento.

O tempo em que cada um dessas fases se desenvolveu, nos casos em apreço, também foi abordado e medido, por ser um dos aspectos mais impactantes da percepção social da impunidade. Foram utilizados dois diferentes métodos para o seu cálculo: o de Média Simples e o de Análise de Sobrevivência, este último incorporando em seu cálculo informações referentes àqueles casos que ainda não conseguiram concluir determinada fase (informação censurada).

A diferença entre os dois métodos em alguns casos foi bastante significativa. Na Fase Policial, por exemplo, considerando apenas os 183 casos que tiveram a investigação concluída, o tempo médio de conclusão foi de 232 dias, contra 492 dias quando se considerou as informações censuradas (Análise de Sobrevivência). Essa disparidade entre os valores das médias obtidas utilizando duas diferentes metodologias repete-se em diversas fases, sempre com a estatística de Análise de Sobrevivência apresentado valores superiores.

Outro aspecto que merece destaque é o tempo médio entre o cometimento do crime e o seu julgamento: apenas 18 casos foram julgados até julho de 2016. Utilizando média simples, constatou-se que o tempo médio entre o crime e o julgamento foi de 16,6 meses, o

que é considerado baixo na literatura. Quando se incorporam ao cálculo da média informações sobre os tempos dos processos que ainda não foram julgados, o tempo médio para o julgamento passa a ser de 28,4 meses, valor bem mais próximo daqueles obtidos em outros Estados.

A estimação da Função de Sobrevivência seguindo o método de Kaplan-Meier permitiu ainda o cálculo da probabilidade de um processo terminar determinada fase em um determinado tempo. Neste aspecto, destaca-se que: i) em exatamente 1 ano após o crime, a probabilidade que um processo tenha a sua investigação concluída (Fase Policial) é de 23% que em 1 ano e meio passa a ser de 54%; ii) a probabilidade de um processo terminar a Fase de Denúncia em 60 dias é de 57%, mas em 90 dias aumenta para 70%, indicando a rapidez desta fase; iii) a Fase de Citação é, em geral, morosa, quando é de apenas 8,7% a chance de que esta fase termine em 90 dias, de 33% em 180 dias e de 79% em 365 dias a partir do recebimento da denúncia; iv) há uma grande dispersão nos tempos da Fase de Instrução; v) uma vez pronunciado, as chances de um processo ser julgado (Fase de Recurso contra a Pronúncia e Julgamento) em 100 e 200 dias são, respectivamente, de 9% e 45%.

As evidências do estudo corroboram ainda o argumento de que a efetivação da prisão em flagrante traz efeitos extremamente positivos para celeridade na duração do processo e aumenta a chance de condenação (reduzindo a impunidade). O oposto também é verdadeiro, isto é, não sendo caso de flagrante, uma série de implicações negativas sobrevêm ao fluxo do sistema de Justiça Criminal.

Foi exposto, no corpo do vertente estudo, que os resultados evidenciaram a necessidade de elaboração de novas metodologias de trabalho para os policiais em campo. Uma das metodologias que pode ser sugerida é o acionamento simultâneo de dois tipos de equipes da PM: uma para o isolamento de local de crime e outra para concentrar esforços exclusivamente na captura do autor. No tocante ao Poder Judiciário, percebe-se que uma gestão melhor racionalizada poderia reverter em importantes mudanças no cenário da impunidade na medida em que agilizaria o andamento dos processos. Além dessas medidas, urge que seja criada uma base de dados oficial interligando todos os órgãos que compõem o sistema de Justiça Criminal, com dados de ocorrências policiais, inquéritos, denúncias, sentenças, e execuções penais, organizados longitudinalmente, como sugerem quase todos os artigos da área.

5 REFERÊNCIAS BIBLIOGRÁFICAS

ADORNO, Sérgio, 1994. **Cidadania e Administração da Justiça Criminal**. In: Diniz, E. et al (org.) O Brasil no rastro da crise. São Paulo, Brasília: HUCITEC, ANPOCS/IPEA, pp. 304-327.

ADORNO, Sérgio; IZUMINO, Wânia Pasinato, 2007. **A Justiça no Tempo, o Tempo da Justiça**. Tempo Social, revista de sociologia da USP, v. 19, n.2, nov. 2007.

BECKER, G. S, 1968. **Crime and Punishment: An Economic Approach**. Journal of Political Economy, nº 76, pp. 169-217.

BRASIL. **Código de Processo Penal**. Disponível em: www.planalto.gov.br. Acesso em 17/08/2016.

_____. Ministério da Fazenda. Secretaria do Tesouro Nacional. **Sistema de Informações Contábeis e Fiscais do Setor Público Brasileiro – SICONFI**. Disponível em www.siconfi.tesouro.gov.br. Acesso em: 09/10/2016.

CANO, I., Duarte, T. L., 2010. **A mensuração da impunidade no sistema de justiça criminal do Rio de Janeiro**. Segurança Justiça e Cidadania, ano 2, n. 4, SENASP, Ministério da Justiça.

CARVALHO et al. (2011). **Análise de Sobrevivência: teoria e aplicações em saúde**. Rio de Janeiro. Editora Fiocruz. 2ª ed.

CEARÁ. **Tribunal de Justiça do Estado do Ceará**. Disponível em: www.tjce.jus.br. Acesso em: 21/09/16.

_____. **Secretaria de Segurança Pública e Defesa Social do Estado do Ceará**. Disponível em: www.sspds.ce.gov.br. Acesso em: 21/09/2016.

COLOSIMO, Enrico A., GIOLO, Suely R., 2006. **Análise de Sobrevivência Aplicada**. São Paulo: Blucher.

COELHO, E. C., 1986. **A administração da Justiça Criminal no Rio de Janeiro: 1942-1967**. Dados – **Revista de Ciências Sociais**. Rio de Janeiro, vol. 29, n.1.

DATASUS. **Sistema de Informação de Mortalidade**. Ministério da Saúde. Disponível em: <http://www2.datasus.gov.br>. Acesso em 17/08/2016.

FAVERO, Luiz Paulo; BELFIORE, Patricia. 2015. **Métodos Quantitativos com Stata: Procedimentos, Rotinas e Análise de Resultados**, Elsevier Brasil.

IBGE (2016). **Instituto Brasileiro de Demografia e Estatística**. Disponível em: [www.](http://www.ibge.gov.br)

ibge.gov.br. Acesso em: 21/09/2016.

KAPLAN, E. L.; MEIER, P. (1958). **Nonparametric Estimation from Incomplete Observations**. Journal of the American Statistical Association, 53, 457-81.

MANTEL, N. (1966). **Evaluation of Survival data and Two New Rank Order Statistics Arising in its Consideration**. Cancer Chemotherapy Reports, 50, 163-170.

MISSE, Michel., VARGAS, Joana Domingues. 2009. A produção decisória do sistema de justiça criminal no Rio de Janeiro ontem e hoje: um estudo preliminar. **Revista Brasileira de Ciências Criminais**, n. 77: 237-260.

RIBEIRO, Ludmila. 2010. A Produção decisória do sistema de justiça criminal para o crime de homicídio: análise dos dados do estado de São Paulo entre 1991 e 1998. Dados – **Revista de Ciências Sociais**, Rio de Janeiro, vol. 53, nº 1, 2010. p. 159-93.

RIFIOTIS, Theophilos; VENTURA, Andresa Burigo; CARDOSO, Gabriela Ribeiro. 2010. Reflexões críticas sobre a metodologia do estudo do fluxo de justiça criminal em caso de homicídios dolosos. **Revista de Antropologia**, 53, 2, pp. 689-714.

SILVA, Klarissa Almeida. 2010. O papel dos tipo de homicídio dolosos na construção social da incriminação dos sujeitos pelos promotores de justiça: Belo Horizonte, processos com andamento entre 2007 e 2009. **Dilemas: Revista de Estudos de Conflitos e Controle Social**. Vol.3, n. 8, pp. 101-123

VARGAS, Joana Domingues. 2000. **Crimes Sexuais e Sistema de Justiça**. São Paulo, IBCCRIM.

VARGAS, Joana D.; MAGALHÃES, Ismênia B. de; RIBEIRO, Ludmila M. L. 2010. Tempo da Justiça: Metodologia de Tratamento do Tempo e da Morosidade Processual na Justiça Criminal. **Segurança, Justiça e Cidadania: Pesquisas Aplicadas em Segurança Pública**. Brasília: Secretaria Nacional de Segurança Pública, Ministério da Justiça, pp. 45-72.

APÊNDICE

Tabela 5: Teste *log-rank* de igualdade de Funções de Sobrevivência por fases

<i>Fase Policial</i>		<i>Fase de Denúncia</i>		<i>Fase de Citação</i>	
$\lambda^2 =$	2622,92	$\lambda^2 =$	29,01	$\lambda^2 =$	8,45
Prob> $\lambda^2 =$	0,00	Prob> $\lambda^2 =$	0,00	Prob> $\lambda^2 =$	0,02

<i>Fase de Instrução</i>		<i>Fase Rec. Pron. Julg.</i>		<i>Pós-Julgamento</i>	
$\lambda^2 =$	7,97	$\lambda^2 =$	2,55	$\lambda^2 =$	87,85
Prob> $\lambda^2 =$	0,02	Prob> $\lambda^2 =$	0,28	Prob> $\lambda^2 =$	0,00

Fonte: elaboração própria

Tabela 6: Teste *log-rank* (Cox-Mantel) de igualdade de Funções de Sobrevivência por fase

<i>Fase Policial</i>				<i>Fase de Instrução</i>				
	λ^2	Prob> λ^2			λ^2	Prob> λ^2	λ^2	Prob> λ^2
	Solto		Preso N Flag		Solto		Preso N Flag	
Preso N Flag	189,98	0,00			Preso N Flag	0,91	0,34	
Flagrante	2123,41	0,00	131,13	0,00	Flagrante	7,82	0,01	3,94
								0,047

<i>Fase de Denúncia</i>				<i>Fase Recurso contra Pronúncia e Julgamento</i>				
	λ^2	Prob> λ^2			λ^2	Prob> λ^2	λ^2	Prob> λ^2
	Solto		Preso N Flag		Solto		Preso N Flag	
Preso N Flag	27,48	0,00			Preso N Flag	2,20	0,14	
Flagrante	13,32	0,00	0,017	0,90	Flagrante	2,66	0,103	0,01
								0,94

<i>Fase de Citação</i>				<i>Pós-Julgamento</i>				
	λ^2	Prob> λ^2			λ^2	Prob> λ^2	λ^2	Prob> λ^2
	Solto		Preso N Flag		Solto		Preso N Flag	
Preso N Flag	5,78	0,02			Preso N Flag	12,29	0,00	
Flagrante	7,00	0,01	0,93	0,33	Flagrante	91,13	0,00	25,77
								0,00

Fonte: elaboração própria

AGLOMERAÇÕES PRODUTIVAS NOS SETORES DE ATIVIDADE ECONOMICA DO CEARÁ: ANÁLISE TEMPORAL E ESPACIAL

Marcos Antônio Feijó Nagaki*

Nicolino Trompieri Neto**

RESUMO

O presente trabalho tem por objetivo identificar e analisar a dinâmica da distribuição espacial dos setores de atividades econômicas previstos na determinação do Produto Interno Bruto (PIB) - Agropecuária, Indústria e Serviços, e dos subsetores da Administração Pública e do Comércio no estado do Ceará para o período 2002-2014. Portanto, investigar-se-á a evolução dos agrupamentos (*clusters*), bem como a relação espacial existente entre a concentração do estoque de empregos formais e o Valor Adicionado Bruto (VAB) de tais atividades. Para tanto, adotou-se como procedimento metodológico a Análise Exploratória de Dados Espaciais (AEDE) aplicada à concentração setorial, cuja mensuração se dera mediante o cálculo do Índice de Concentração Normalizado (*ICn*).

Palavras-Chave: Economia Cearense; Valor Adicionado Bruto; Índices de Concentração; Análise Exploratória de Dados Espaciais.

ABSTRACT

The objective of this work is to identify and analyze the dynamics of the spatial distribution of the sectors of economic activity foreseen in the determination of the Gross Domestic Product (GDP) - Agriculture, Industry and Services, and of the sub-sectors of Public Administration and Commerce in the state of Ceara for the period 2002-2014. Therefore, we will investigate the evolution of clusters, as well as the spatial relationship between the stock concentration of formal jobs and the Gross Value Added (GVA) of these sectors. For this purpose, was used the method of Exploratory Analysis of Spatial Data (AEDE) applied to the sectorial concentration, which was measured by means of the Normalized Concentration Index (ICn).

Keywords: Cearense Economy; Gross Value Added; Concentration Index; Exploratory Analysis of Spatial Data.

*Bacharel em Informática pela Universidade de Fortaleza – UNIFOR. Mestrando em Economia no Setor Público (CAEN/UFC). Contato: 85-996626044. E-mail: marcosfn@gmail.com.

**Instituto de Pesquisa e Estratégia Econômica do Ceará – IPECE. Universidade de Fortaleza – UNIFOR. E-mail: nicolino.trompieri@unifor.br.

1 INTRODUÇÃO

O estado do Ceará compreende um total de 184 municípios, os quais apresentam características geoambientais, socioeconômicas, culturais e de redes de fluxos municipais distintas e, por conseguinte, foram distribuídos em quatorze Regiões de Planejamento pela Secretaria do Planejamento e Gestão (SEPLAG) de acordo com as suas semelhanças. Tal divisão, incorporada ao planejamento estatal por meio do Plano Plurianual (PPA), no qual constam todos os investimentos e ações continuadas do Estado, atende aos objetivos de descentralização das políticas, de integração da ação governamental nos territórios estaduais, de estímulo e facilitação da participação social e do fortalecimento das vocações regionais, tendo como fim último à redução das desigualdades regionais.

Assim, uma vez que as estratégias governamentais se baseiam em aspectos territoriais, torna-se imprescindível considerar o efeito espacial entre os municípios, visto que o estudo do comportamento dos aglomerados industriais e de trabalho per si não seja suficiente para retratar de forma fidedigna a dinâmica econômica regional e os impactos oriundos da organização espacial inter-firmas. Dessa forma, a análise espacial se insere como ferramental estatístico imprescindível a adequada compreensão de tais concentrações, já que pondera as interdependências e heterogeneidades regionais.

Vale salientar que, conforme Marshall (1982), as aglomerações contribuem na criação de um ambiente favorável às empresas que ali coexistem, isto é, a concentração espacial de firmas pode fornecer aos produtores vantagens competitivas, as quais não seriam possíveis caso eles estivessem atuando em lugares distantes uns dos outros. Portanto, pode-se afirmar que a proximidade geográfica proporciona o surgimento de externalidades, pecuniárias e tecnológicas, favorecendo a existência de *linkages*, entre produtores e usuários e, conseqüentemente, cria um ambiente propício à existência de *spillovers* tecnológicos (CROCCO *et al.*, 2003).

Notadamente em países em desenvolvimento como o Brasil, os aglomerados informais são representados predominantemente por micro e pequenas empresas, e se caracterizam, em geral, por qualificação profissional inadequada, infraestrutura precarizada, recursos financeiros insuficientes e pouca ou nenhuma inovação. Todavia, ressalvada a importância do setor informal presente na economia cearense, optou-se neste estudo por considerar apenas o estoque de empregos formais, cujas motivações serão explicitadas em momento oportuno.

Portanto, as aglomerações produtivas existentes ou potenciais, identificadas através

das relações formais de emprego, permitem subsidiar as políticas governamentais de apoio ao desenvolvimento regional, visto que exercem impacto direto na geração de empregos, renda e, conseqüentemente, contribuem no desenvolvimento econômico regional.

Segundo dados do Ministério do Trabalho e Emprego (MTE), considerando todos os setores econômicos, o estado do Ceará gerou 1,55 milhões de empregos formais em 2014, correspondendo a uma massa salarial de R\$ 2,6 bilhões em dezembro do mesmo ano. Desse total, os empregos do Estado se concentraram, principalmente, no setor de serviços (74%), no setor da indústria (24%) e no setor agropecuário (2%).

Em relação ao Valor Adicionado Bruto¹ (VAB), segundo dados do Instituto Brasileiro de Geografia e Estatística (IBGE) de 2014, o setor de serviços representou a maior participação no PIB do Estado correspondendo a 75,64% da riqueza gerada. Em seguida, o setor da indústria registrou participação de 19,16% na economia cearense e, finalmente, o setor da agropecuária respondendo por 5,20% do PIB.

Desse modo, este trabalho pretende realizar um estudo exploratório sobre a influência dos grandes setores econômicos cearenses: agricultura, indústria e serviços, examinando, a partir da mensuração da concentração setorial, o fenômeno espacial do estoque de empregos formais, bem como a sua correlação com o VAB. Vale salientar que em relação ao setor de serviços, os subsetores da administração pública e do comércio serão estudados individualmente, dada a relevância econômica de ambos.

Procurou-se, portanto, analisar a distribuição espacial da concentração empregatícia dos grandes setores econômicos entre os 184 municípios cearenses nos anos 2002, 2008 e 2014. O objetivo é descobrir se uma atividade econômica de um dado município influencia a formação do estoque de emprego nos seus vizinhos. Isso significaria que uma determinada concentração setorial seria capaz de influenciar o grau de empregabilidade desse setor nas regiões adjacentes. Este estudo visa, também, compreender o tipo de relação existente entre a variável endógena (concentração empregatícia) e a variável exógena (valor adicionado bruto) de cada setor.

Dessa forma, o trabalho encontra-se distribuído em quatro seções além desta introdução. A segunda seção descreve a metodologia e o tratamento dos dados empregados, ilustrando a aplicação da Análise Exploratória dos Dados Espaciais (AEDE) ao Índice de Concentração Normalizado (ICn); na terceira seção, os resultados empíricos são apresentados; e na última seção, são feitas as considerações finais.

1. VAB é a contribuição ao produto interno bruto pelas diversas atividades econômicas, obtida pela diferença entre o valor bruto de produção e o consumo intermediário absorvido por essas atividades.

2 METODOLOGIA E TRATAMENTO DOS DADOS

Os estudos que abordam aglomerados produtivos locais, normalmente, ignoram tendências de longo prazo ao focar apenas nas aglomerações existentes. Ao desprezar a dinâmica da estrutura produtiva, tal abordagem fragiliza a consecução de políticas públicas voltadas ao desenvolvimento econômico regional, visto que as aglomerações em expansão podem afetar às já existentes. Assim, quando as especificidades regionais relativas à estagnação, declínio, expansão e dinamismo setorial são levados em consideração, a análise se torna muito mais precisa.

Sugerido inicialmente por Crocco *et al.* (2006), o Índice de Concentração Normalizado (ICn) é uma estatística que permite a identificação das aglomerações produtivas reais e potenciais. Porém, indicadores de concentração como o *ICn*, Índice de Gini, Quociente Locacional (*QL*), Hirschman Herfindal modificado (*HHm*), entre outros, não mensuram os aglomerados dadas as características geográficas, sendo considerados não espaciais, ou seja, unidades geográficas vizinhas ou distantes são tratadas de forma idêntica ao ignorar as dependências espaciais existentes entre elas e, conseqüentemente, podendo gerar resultados subestimados da aglomeração espacial com um todo.

Para a realização deste estudo, utilizou-se como fonte de dados a relação Anual das Informações Sociais (RAIS) disponibilizada no sítio do Ministério do Trabalho e Emprego (MTE). Trata-se de um relatório anual a ser preenchido e enviado ao MTE por todas as pessoas jurídicas do país. Vale salientar que a ampla maioria dos trabalhos relacionados ao tema, pontua diversas considerações sobre o uso dessa base de dados, por exemplo, o registro exclusivo das relações formais de trabalho e a autoclassificação (por meio do formulário de preenchimento on-line) são algumas das fragilidades dessa metodologia de coleta de dados.

Sobre a primeira restrição, ao abordar apenas as relações formais de trabalho, pode-se resultar em valores subestimados do emprego, ou seja, análises baseadas unicamente no estoque de empregos formais tendem a viesar os resultados, uma vez que as relações informais de trabalho não se distribuem de forma homogênea no espaço, no tempo e nem entre os setores (CAMPOS, 2004). A segunda problemática se refere ao autopreenchimento do formulário on-line da RAIS, o qual possibilita a ocorrência de erros cadastrais e, por conseguinte, reduz a confiabilidade das informações registradas.

Todavia, diversos outros aspectos positivos legitimam o uso da RAIS para analisar a distribuição espacial do emprego de forma segura, por exemplo, o nível detalhado de desagregação geográfica e setorial dos dados (SUZIGAN *et al.*, 2003), a obrigatoriedade e

a automatização do cadastro on-line, ademais, o caráter censitário permite que eventuais erros cadastrais sejam diluídos em setores ou regiões com grande quantidade de empregados.

Dessa forma, calculou-se o *ICn* do estoque de empregos dos setores componentes do PIB para os 184 municípios cearense nos anos 2002, 2008 e 2014. Em seguida, aplicou-se a Análise Exploratória de Dados Espaciais (AEDE) a fim de extrair medidas de autocorrelação global e local para identificar padrões de regimes espaciais ou clusters, os quais examinam aspectos como associação espacial e heterogeneidade espacial.

Destaca-se que cada uma das estatísticas descritas possui vantagens e desvantagens a se considerar. Assim, a intenção de unir os índices locacionais à AEDE foi superar as limitações de cada metodologia caso fossem realizadas isoladamente. O resultado dessa abordagem conjunta auxilia não só na identificação das potências aglomerações produtivas, mas, também, na demonstração das relações espaciais existente entre elas.

2.1 Índice de Concentração Normalizado

De acordo com Crocco *et al.* (2006), o *ICn* é formado pela junção de três outros índices: *QL*, *HHm* e Participação Relativa (PR), sendo responsável por localizar aglomerações produtivas locais e em especial às potenciais.

O Quociente Locacional (*QL*) é adequado para ilustrar a especialização em certo setor ou atividade específica e representa a razão entre duas estruturas econômicas: concentração relativa de um setor ou atividade em um dado município comparativamente à participação deste mesmo setor ou atividade no espaço definido como referência.

Segundo Haddad *et al.* (1985), a representação algébrica para o *QL* é dada por:

$$QL = \left(\frac{E_j^i}{E_j} \right) / \left(\frac{E_{CE}^i}{E_{CE}} \right)$$

Onde E_j^i = e emprego do setor ou atividade *i* no município *j*; E_j = emprego total no município *j*; E_{CE}^i = emprego do setor ou atividade *i* no Ceará; E_{CE} = emprego total no Ceará.

Para Haddad *et al.* (1985 p. 232), um município possuiria especialização em determinado setor ou atividade caso o seu *QL* fosse superior a 01 (um). Caso contrário, a especialização do município no setor ou atividade seria inferior à especialização do Ceará no referido setor ou atividade. Vale frisar que as análises de aglomerações produtivas locais baseadas apenas no *QL* estão sujeita a erros, posto que o quociente tende a sobrevalorizar o peso de um determinado setor para regiões pequenas e com nível de emprego reduzido e estrutura produtiva pouco diversificada. Por outro lado, o quociente tende a subvalorizar a importância de determinados setores em regiões com uma estrutura produtiva bem

diversificada, mesmo que este setor possua peso significativo no contexto da economia de referência.

Com o propósito de mitigar esse problema, utilizou-se o índice Hirschman-Herfindahl modificado (*HHm*) a fim de quantificar o peso do setor ou atividade do município em relação ao estado do Ceará. Conforme Crocco et al. (2006), o *HHm* estuda a comparação do peso do setor ou atividade *i* do município *j* com o setor ou atividade *i* do Ceará na estrutura produtiva cearense.

$$HHm = \left(\frac{E_j^i}{E^i} \right) - \left(\frac{E_j}{E_{CE}} \right)$$

O terceiro indicador denominado de Participação Relativa (*PR*) capta a importância relativa do setor em relação ao total do setor no estado do Ceará. Portanto, a relevância da atividade *i* no município *j* mediante o total de emprego do setor no Ceará é representada. Esse índice pode variar entre zero e um, sendo que quanto maior o grau de importância do setor em relação ao Estado, mais próximo de 01 (um) será o indicador.

$$PR = \left(\frac{E_j^i}{E_{CE}^i} \right)$$

A partir dos valores conhecidos do *QL*, do *HHm* e da *PR*, é possível calcular um único indicador de concentração setorial denominado de Índice de Concentração Normalizado (*ICn*), conforme ilustrado na equação abaixo:

$$ICn_{ij} = \theta_1 QLn_{ij} + \theta_2 HHm_{ij} + \theta_3 PR_{ij}$$

Onde os valores dos θ indicam os pesos de cada um dos indicadores para cada setor específico. Para a obtenção desses valores utilizou-se uma das técnicas de análise multivariada e, provavelmente, a mais antiga e conhecida dentre elas, a Análise de Componentes Principais (ACP) (JOLLIFFE, 2002). Essa metodologia consiste em converter um conjunto de valores de uma série de dados em um novo conjunto de valores de uma série de dados linearmente não correlacionada, ordenada pelo maior poder de explicação, denominada de componentes principais.

Dessa forma, a primeira etapa para o cálculo dos pesos via ACP consiste na obtenção dos autovalores da matriz de correlação. Cada autovalor corresponde a uma explicação na variância β_1 , β_2 e β_3 em que o somatório deles é igual à variância total dos componentes e, portanto, à variância total dos indicadores locais. A Tabela 2.1.1 apresenta os autovalores ou variância (e sua acumulação) dos três componentes principais.

Tabela 2.1.1 - Autovalores da matriz de correlação

Componente	Autovalor	Variância Explicada pelo Componente	Variância Explicada Total
1	λ_1	β_1	β_1
2	λ_2	β_2	$\beta_1 + \beta_2$
3	λ_3	β_3	$\beta_1 + \beta_2 + \beta_3 = (100\%)$

Fonte: Elaboração própria com base em Crocco *et al.* (2006).

Em seguida, calcula-se a participação relativa de cada um dos indicadores em cada um dos componentes para encontrar a importância das variáveis nos componentes. Para isso, toma-se como base a matriz de coeficientes ou os autovetores da matriz de correlação (Tabela 2.1.2) e toma a função módulo dos autovetores associados a cada componente. Posteriormente, efetua-se a divisão entre cada autovetor, em valor absoluto, pelo somatório dos autovetores absolutos de seu componente (Tabela 2.1.3).

Tabela 2.1.2 - Autovetores da matriz de correlação

Índice de Concentração	Componente 1	Componente 2	Componente 3
<i>QL</i>	α_{11}	α_{12}	α_{13}
<i>HHm</i>	α_{21}	α_{22}	α_{23}
<i>PR</i>	α_{31}	α_{32}	α_{33}

Fonte: Elaboração própria com base em Crocco *et al.* (2006).

Tabela 2.1.3 - Autovetores da matriz de correlação recalculados

Índice de Concentração	Componente 1	Componente 2	Componente 3
<i>QL</i>	$\alpha'_{11} = \frac{ \alpha_{11} }{ \alpha_{11} + \alpha_{21} + \alpha_{31} }$	$\alpha'_{12} = \frac{ \alpha_{12} }{ \alpha_{12} + \alpha_{22} + \alpha_{32} }$	$\alpha'_{13} = \frac{ \alpha_{13} }{ \alpha_{13} + \alpha_{23} + \alpha_{33} }$
<i>HHm</i>	$\alpha'_{21} = \frac{ \alpha_{21} }{ \alpha_{11} + \alpha_{21} + \alpha_{31} }$	$\alpha'_{22} = \frac{ \alpha_{22} }{ \alpha_{12} + \alpha_{22} + \alpha_{32} }$	$\alpha'_{23} = \frac{ \alpha_{23} }{ \alpha_{13} + \alpha_{23} + \alpha_{33} }$
<i>PR</i>	$\alpha'_{31} = \frac{ \alpha_{31} }{ \alpha_{11} + \alpha_{21} + \alpha_{31} }$	$\alpha'_{32} = \frac{ \alpha_{32} }{ \alpha_{12} + \alpha_{22} + \alpha_{32} }$	$\alpha'_{33} = \frac{ \alpha_{33} }{ \alpha_{13} + \alpha_{23} + \alpha_{33} }$

Fonte: Elaboração própria com base em Crocco *et al.* (2006).

Finalmente, o peso final de cada índice de concentração resulta da soma dos produtos dos α'_{ij} pelo seu autovalor correspondente, conforme Tabela 2.1.4. Onde os α'_{ij} representam o peso que cada variável assume dentro de cada componente e que os autovalores (β_s) fornecem a variância dos dados associada ao componente.

Tabela 2.1.4 – Pesos dos índices de concentração para o *ICn*

Índice de Concentração	Peso
<i>QL</i>	$\theta_1 = \alpha'_{11}\beta_1 + \alpha'_{12}\beta_2 + \alpha'_{13}\beta_3$
<i>HHm</i>	$\theta_2 = \alpha'_{21}\beta_1 + \alpha'_{22}\beta_2 + \alpha'_{23}\beta_3$
<i>PR</i>	$\theta_3 = \alpha'_{31}\beta_1 + \alpha'_{32}\beta_2 + \alpha'_{33}\beta_3$

Fonte: Elaboração própria com base em Crocco *et al.* (2006).

Ao calcular os pesos² dos indicadores locais dos setores econômicos, observou-se uma ligeira variação da proposição feita por Crocco *et al.* (2003) e encontrados no estudo de Paula (2008), em que estariam próximos de 1/3. Ainda assim, o cálculo do *ICn* via ACP é útil para garantir o melhor ajuste das cargas dos índices.

2.2 Análise Exploratória de Dados Espaciais

A AEDE se baseia nos aspectos espaciais presentes em bases de dados e considera tanto a dependência espacial quanto a heterogeneidade espacial. Princípio básico da análise espacial, conhecida como Primeira Lei da Geografia, Tobler (1970) argumenta que: tudo é relacionado com tudo, mas coisas mais próximas estão mais relacionadas que coisas distantes.

A AEDE procura retratar a distribuição espacial, localizar os padrões de associação espacial, verificar a presença de diferentes regimes espaciais, identificar fenômenos geográficos com extrema irregularidade (não estacionariedade) e determinar *outliers*.

O uso de variáveis espacialmente densas ou intensivas são imprescindíveis à AEDE (HAINING, 2003). Variáveis desse tipo são divididas por algum indicador de intensidade (variáveis *per capita* ou por área, ou ainda, as divididas pela quantidade de capital ou trabalho). Desse modo, o *ICn* sendo uma proporção, torna-se adequado para este estudo.

A fim de implementar a AEDE com o intuito de descrever a distribuição espacial dos setores do PIB, faz-se uso do I de Moran, cujo teste é o mais comumente aplicado em análises de aglomerações produtivas (FESER & ISSERMAN, 2005).

Através da estatística *I* de Moran obtém-se a indicação formal do grau de associação linear entre os vetores de valores observados no tempo *t* (z_t) e a média ponderada dos valores da vizinhança, os *lags* espaciais (Wz_t), e S_0 representa a soma dos elementos da matriz *W*. Os valores de *I* maiores (ou menores) do que o valor esperado $E(I) = -1/(n-1)$

2. Os pesos para os setores da indústria, dos serviços, da agropecuária, da administração pública, do comércio e dos serviços+comércio foram calculados com base no programa *EViews9* e se encontram à disposição com os autores.

demonstram a existência de autocorrelação positiva (ou negativa).

Segundo Anselin (1995), a estatística de Moran é descrita como:

$$I_t = \left(\frac{n}{S_0} \right) \left(\frac{z_t' W z_t}{z_t' z_t} \right) t = 1, \dots, n$$

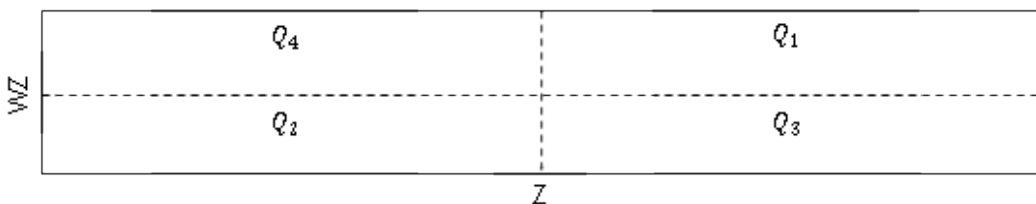
Neste trabalho, a execução do I de Moran foi realizada através da matriz de contiguidade *Queen* de primeira ordem, isto é, os municípios contendo um ou mais vértices comuns foram considerados adjacentes.

Conforme afirma Gonçalves (2005), a estatística *I* de Moran é uma medida global e, portanto, indica a tendência geral de agrupamento de dados, porém, é incapaz de descrever a estrutura de correlação espacial regional. Assim, a observação da existência de clusters espaciais locais se derá por meio do diagrama de dispersão de Moran e do Indicador de Associação Espacial Local (LISA).

O diagrama de dispersão de Moran é uma ferramenta utilizada para interpretar a estatística *I* de Moran (ANSELIN, 1996). Visa identificar a tendência geral de associação por meio da correlação linear entre os valores da variável de interesse defasados espacialmente W_z e os valores observados em cada unidade espacial z (ANSELIN, 1995). Ainda, o coeficiente *I* de Moran é dado pela inclinação da curva de regressão de W_z contra z e essa apresenta o grau de ajustamento.

A representação gráfica do valor da estatística *I* de Moran para cada região em análise é ilustrada na Figura 2.2.1. O primeiro quadrante Q_1 , Alto-Alto (AA), indica regiões com altos valores para a variável em análise cercada por regiões que, também, contenham valores acima da média para a mesma variável. O segundo quadrante, Baixo-Baixo (BB), representa as regiões e vizinhos com baixo valor para a variável analisada. O terceiro quadrante Q_3 , Alto-Baixo (AB), indica regiões com valores acima da média cercadas por vizinhos de baixo valor. Por fim, o quarto quadrante Q_4 , Baixo-Alto (BA), ilustra as localidades com baixo valor cercadas por vizinhos de alto valor.

Figura 2.2.1 – Representação gráfica de Moran



Sugerido inicialmente por Anselin (1995), o Moran local permite observar os padrões locais de associação linear estatisticamente significativa ao mensurar a contribuição individual de cada região na estatística I de Moran global, ou seja, objetiva avaliar o padrão local de autocorrelação espacial.

O Moran local é representado pela seguinte expressão:

$$I_i = \frac{(y_i - \bar{y}) \sum_j w_{ij} (y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2 / n}$$

Valores de $I_i > 0$ indicam a existência de uma concentração espacial positiva ou *clusters* de valores similares (Alto-Alto ou Baixo-Baixo). Por outro lado, $I_i < 0$ apontam a presença de uma concentração espacial negativa ou *clusters* de valores distintos (Alto-Baixo ou Baixo-Alto). Vale salientar, que similar ao gráfico de dispersão de Moran global, o mapa de *cluster* LISA indica as correlações espaciais locais significantes por meio dos quadrantes AA, AB, BA e BB.

É possível, também, analisar o fenômeno da autocorrelação espacial para variáveis distintas, ou seja, a descoberta do grau de relacionamento entre os valores de uma variável observada em uma região com os valores de outra variável observada nas regiões adjacentes é fornecida pelo estudo do I de Moran bivariado. As mesmas ferramentas e análises existentes para o estudo de uma variável também se aplicam neste caso. Isso quer dizer que além de permitir a obtenção de um coeficiente de autocorrelação espacial global bivariado, é possível medir a associação local entre as variáveis do modelo em estudo através do mapa de *cluster* bivariado. Assim, essa estatística mensura o grau de associação linear existente entre o valor de uma variável em uma determinada região e a média de outra variável nas regiões vizinhas.

Para implementar a análise multivariada, utilizou-se o ICn do estoque de empregos como variável endógena e o Valor Adicionado Bruto (VAB), a preços correntes, da agropecuária, indústria, serviços e administração pública como variável exógena, disponíveis no sítio do IBGE. A ressalva que se faz, neste caso, ao implementar a estatística bivariada, é a exigência de que as duas variáveis em exame correspondam a exatamente os mesmos setores e subsetores. Todavia, essa obrigação não foi atendida para a atividade do comércio, inviabilizando a sua investigação. Para o setor de serviços, foi necessário remover a atividade da administração pública, uma vez que a estimativa do VAB para esse setor exclui a referida atividade.

3 RESULTADOS E DISCUSSÃO³

3.1 Setor da Indústria

A distribuição do IC_n desse setor no estado do Ceará indicou que o número de municípios especializados com valores acima da média estadual permaneceu igual em 2002 e 2008 com 58 localidades, aumentando em seguida para 63 em 2014. Em postos de trabalho ocupados, isso representou um total de 188.740, 265.551 e 361.063 respectivamente, o que correspondeu a um incremento de 36% no último ano.

A estimação do I de Moran sobre o IC_n dos 184 municípios cearenses (Tabela 3.1.1) indicou a presença de autocorrelação espacial positiva com valores estatisticamente significantes no período. Visualmente, isso pode ser observado através da inclinação positiva da reta nos diagramas de dispersão de Moran (Figura 3.1.1).

Tabela 3.1.1 – Índice de Moran global da indústria

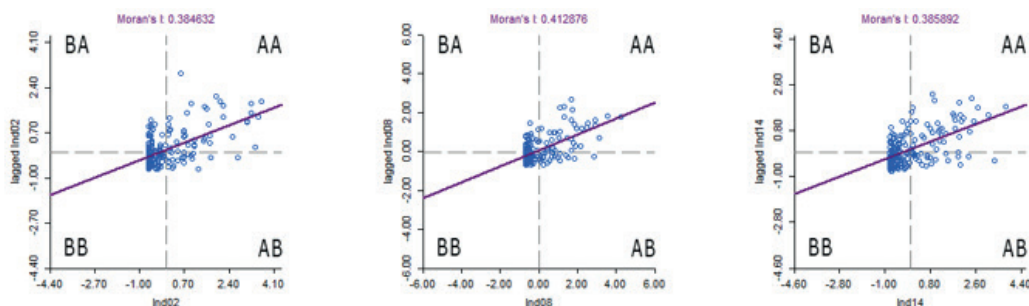
Anos	I de Moran Global	*Significância
2002	0,3846	0,001
2008	0,4128	0,001
2014	0,3858	0,001

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Em 2002, 71% dos municípios cearenses possuíam relações espaciais positivas, dos quais 21% se situavam no quadrante Alto-Alto e 50% no Baixo-Baixo. Em 2008, houve uma redução nesse número para 67% do total, desses, 19% estavam no Alto-Alto e 48% no Baixo-Baixo. Em 2014, a porcentagem cresceu para 69%, dos quais 23% estavam no Alto-Alto e 46% no Baixo-Baixo.

Figura 3.1.1 – Diagrama de dispersão de Moran da indústria



Fonte: Elaboração própria com base no programa *GeoDa 1.10*.

3. Os mapas de desvio padrão, os mapas de cluster LISA bivariado e os mapas de significância dos setores desta seção se encontram à disposição com os autores, dado que não foi possível incluí-los pela limitação do número máximo de páginas exigido.

A Tabela 3.1.2 consolida os dados das aglomerações Alto-Alto exibidas na Figura 3.1.2 em anexo, e sugere que havia uma concentração na região da Grande Fortaleza com transbordamentos para o Litoral Oeste/Vale do Curu, Maciço de Baturité, Litoral Leste e Vale do Jaguaribe em 2002. Em 2008, esse cluster se deslocou em direção ao oeste do Estado, ampliando a sua presença no Litoral Oeste/Vale do Curu, enquanto desapareciam os municípios das regiões Litoral Leste e Vale do Jaguaribe do período anterior. Em 2014, a região da Grande Fortaleza e o Litoral Oeste/Vale do Curu se consolidaram com essa relação espacial quase na totalidade, ao ponto de gerar *spillover* em Forquilha no Sertão de Sobral. No mesmo ritmo, os empregos gerados no último ano no Ceará aumentaram em 37% em relação ao período anterior. Vale destacar que, em 2002, o regime Baixo-Baixo que era marcante no Sertão dos Crateús, porém com pequena presença no Sertão de Sobral e no Cariri, intensificou-se em 2014, abarcando grande parte da região Sertão dos Crateús e perpassando a Serra da Ibiapaba, o Sertão de Sobral e, mais ao sul, quase integralmente o Sertão dos Inhamuns, ultrapassando, ainda, as fronteiras do Centro Sul e do Cariri.

Em todas as observações deste setor, houve predominância absoluta da região da Grande Fortaleza associada positivamente no Alto-Alto; inicialmente, possuía 14 cidades com 140.519 trabalhadores, passou para 13 cidades com 185.409 empregados em 2008 e, pulou para 16 municípios com 251.163 postos de trabalho em 2014, representando um incremento de 35% no estoque de empregos. Nessa região, os municípios de Horizonte e Maracanaú obtiveram *ICn* acima de dois desvios padrão da média nos três períodos e, com a mesma intensidade, as cidades de Pindoretama, São Gonçalo do Amarante e Pacajus apareceram em 2014, indicando serem localidades fortemente especializadas e, portanto, geradoras de *spillovers*. A segunda região mais relevante deste setor foi o Litoral Oeste/Vale do Curu; com apenas 2 cidades dependentes espacialmente em 2002, saltou para 4 em 2008 e ampliou para 7 em 2014, ocasionando um incremento de 110% no nível de empregos no último ano. Vale frisar que Itapajé foi a única cidade dessa região com *ICn* acima de dois desvios padrão da média em 2008, porém, Uruburetama passou a figurar ao lado desse município no período seguinte. O Sertão de Sobral que não apareceu nos dois primeiros anos, surgiu em 2014 com a cidade de Forquilha, a qual possuía baixa especialização e, portanto, indicando localidade recebedora dos efeitos de transbordamento. Inversamente, o Litoral Leste e o Vale do Jaguaribe não indicarem representantes no último período. Seguindo essa mesma tendência, o Maciço de Baturité evidenciou situação retracionista em 2014, visto que reduziu em uma cidade nesse ano, além de apresentar uma queda de 84% no estoque de empregos.

Tabela 3.1.2 – Evolução das aglomerações Alto-Alto da indústria

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	Δ% (AB)
Grande Fortaleza	14	13	16	140.519	185.409	251.163	35%
Litoral Leste	1	-	-	246	-	-	-
Litoral Oeste/Vale do Curu	2	4	7	1.527	4.705	9.901	110%
Maciço de Baturité	2	2	1	602	978	154	-84%
Vale do Jaguaribe	1	-	-	829	-	-	-
Sertão de Sobral	-	-	1	-	-	267	∞
Ceará	20	19	25	143.723	191.092	261.485	37%

Fonte: Elaboração própria.

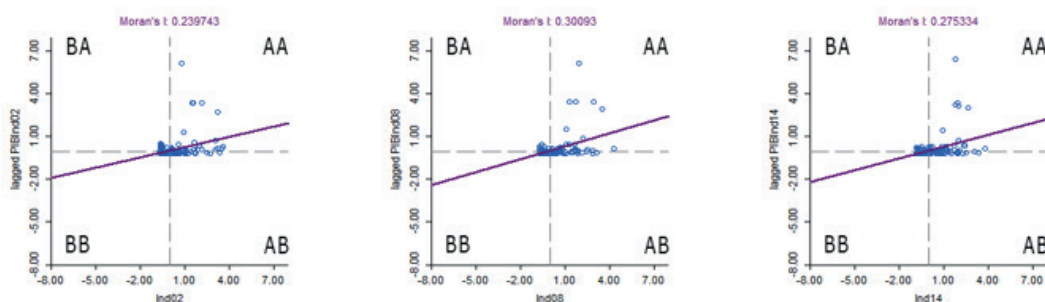
A Tabela 3.1.3 exibe os valores do I de Moran bivariado, onde se demonstra haver evidência de correlação direta entre o *ICn* e o VAB do setor industrial. A inclinação positiva da reta nos diagramas de dispersão de Moran bivariado da Figura 3.1.3 representa graficamente tal afirmativa.

Tabela 3.1.3 – Índice de Moran bivariado da indústria

Anos	<i>I</i> de Moran Global	*Significância
2002	0,2397	0,001
2008	0,3009	0,001
2014	0,2753	0,001

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Figura 3.1.3 – Diagrama de dispersão de Moran bivariado da indústria

Fonte: Elaboração própria com base no programa *GeoDa* 1.10.

A Tabela 3.1.4 consolida as informações das aglomerações Alto-Alto, onde são exibidos os municípios com altos *ICn* circundados por vizinhos com altos valores do VAB. Os resultados revelaram um cluster formado por cidades Alto-Alto concentrado na região da Grande Fortaleza em todos os anos, quais sejam: Maracanaú, Maranguape, Pacatuba, Caucaia, Eusébio, Itaitinga e Fortaleza. Importante frisar que houve um crescimento de 34% no estoque de empregos, ainda que Pindoretama tenha deixado de constar nessa região em

2014, apontando para uma maior concentração empregatícia. Observando o Sertão de Sobral, a cidade de Forquilha apresentou transbordamento apenas nas duas primeiras análises.

Tabela 3.1.4 – Evolução das aglomerações bivariadas Alto-Alto da indústria

Regiões	Número de Municípios			Número de Vínculos Ativos			$\Delta\%$ (AB)
	2002	2008	2014	2002	2008 (A)	2014 (B)	
Grande Fortaleza	8	8	7	123.516	173.867	233.394	34%
Sertão de Sobral	1	1	-	478	415	-	-100%
Ceará	9	9	7	123.994	174.282	233.394	34%

Fonte: Elaboração própria.

3.2 Setor de Serviços

O cálculo do *ICn* desse setor exigiu a remoção dos subsetores do comércio e da administração pública, visto que tais subsetores foram explorados em capítulo próprio. Assim, a distribuição do *ICn* revelou que o número de municípios especializados com valores acima da média estadual sofreu pouca alteração no período. Em 2002 constavam 56 localidades, em 2008 esse número caiu para 55 cidades especializadas e, em 2014, aumentou para 57 municípios, o que correspondeu a um incremento de 4% nesse ano. De forma mais intensa, em postos de trabalho ocupados, verificou-se um crescimento importante, foram 221.327, 301.934 e 480.121 novos empregos criados respectivamente, ocasionando incremento de 59% no último ano.

A estimação do *I* de Moran sobre o *ICn* dos 184 municípios cearenses (Tabela 3.2.1) indicou ausência de autocorrelação espacial em 2002, já que a pseudo-significância da variável em questão apresentou probabilidade não significativa. Todavia, em 2008 e 2014, os resultados revelaram indício de autocorrelação espacial positiva com valores estatisticamente significantes. Tais cenários podem ser observados através da inclinação da reta nos diagramas de dispersão de Moran (Figura 3.2.1).

Tabela 3.2.1 – Índice de Moran global dos serviços

Anos	<i>I</i> de Moran Global	*Significância
2002	0,0234	0,227
2008	0,2464	0,001
2014	0,1698	0,002

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Em 2008, 57% dos municípios cearenses possuíam relações espaciais positivas, dos quais 16% se situavam no quadrante Alto-Alto e 41% no Baixo-Baixo. Em 2014, houve uma redução nesse número para 53% do total, desses, 15% estavam no Alto-Alto e 38% no Baixo-Baixo.

Figura 3.2.1 – Diagrama de dispersão de Moran dos serviços

Anos	<i>I</i> de Moran Global	*Significância
2002	0,0234	0,227
2008	0,2464	0,001
2014	0,1698	0,002

Fonte: Elaboração própria com base no programa GeoDa 1.10.

A Tabela 3.2.2 consolida os dados das aglomerações Alto-Alto ilustrada na Figura 3.2.2 em anexo, e sugere um intenso crescimento em 2008 e uma leve redução em 2014 no número de municípios com essa associação. Em 2002, eram 3 localidades com essa relação, sendo distribuídos igualmente nas regiões Centro Sul, Grande Fortaleza e Litoral Leste. Em 2008, essa associação saltou para 7 cidades, apenas uma delas representada no Maciço de Baturité e o restante na região da Grande Fortaleza. Em 2014, a quantidade de cidades diminuiu para 6, mantendo as duas regiões do período anterior como protagonistas dessa dependência espacial. Apesar da redução no número de cidades no último ano, o Estado sinalizou expansão de 54% no número de empregados no mesmo período.

As regiões Centro Sul e Litoral Leste só exibiram relação espacial Alto-Alto no primeiro período observado, sumindo nos anos seguintes. Contudo, a região da Grande Fortaleza concentrou a maior quantidade de municípios espacialmente relacionados, bem como expôs incremento acentuado de 54% no estoque de empregos no último período; Em 2002, somente a cidade de Eusébio figurava nessa região, mas a partir de 2008 surgiram novos municípios ao lado dessa localidade, quais sejam: Fortaleza, Aquiraz, Caucaia, Itaitinga e Maracanaú. Em 2014, porém, apenas Aquiraz não gerou relação espacial. Possivelmente o transbordamento nessa região em 2014, decorreu-se principalmente pelos municípios de Fortaleza e Euzébio, mediante a elevada especialização dessas duas cidades, cujo ICn se encontrava acima de dois desvios padrão da média. A região do Maciço de Baturité, embora não tenha indicado cluster em 2002, revelou a cidade de Pacoti em 2008 com efeito transbordamento. Em 2014, essa cidade cedeu lugar para a cidade de Barreira, em grande parte por ter tido especialização pouco expressiva. Vale salientar, que essa região ampliou fortemente a geração de empregos, da ordem de 74% no último ano.

Tabela 3.2.2 – Evolução das aglomerações Alto-Alto dos serviços

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	$\Delta\%$ (AB)
Centro Sul	1	-	-	117	-	-	-
Grande Fortaleza	1	6	5	15.497	260.786	401.160	54%
Litoral Leste	1	-	-	303	-	-	-
Maciço de Baturité	-	1	1	-	65	113	74%
Ceará	3	7	6	15.917	260.851	401.273	54%

Fonte: Elaboração própria.

Conforme explicitado na metodologia, a análise do I de Moran bivariado para esse setor necessitou de ajustes a fim de se adequar a forma como os dados do valor adicionado dos setores são disponibilizados. Isto é, visto que as informações referentes às estimativas do VAB apresentam os dados dos serviços excluindo apenas a administração pública do rol de atividades que o compõem, foi necessário, também, excluir somente a referida atividade para o cálculo do ICn do setor.

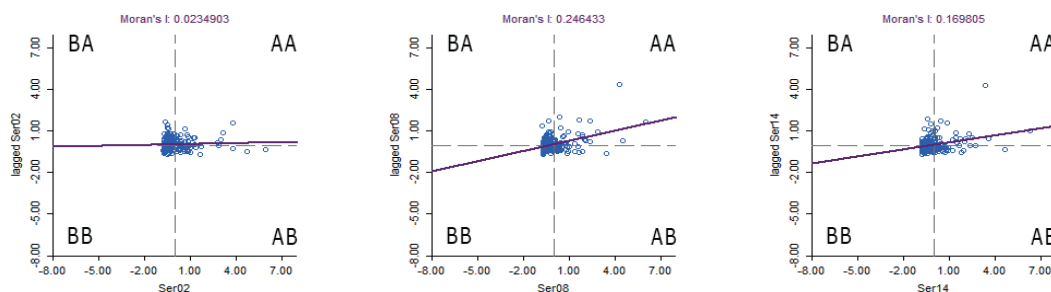
Assim, a Tabela 3.2.3 exibe os valores do I de Moran bivariado, onde se demonstra haver evidência de correlação direta entre o ICn e o VAB do setor de serviços. A inclinação positiva da reta nos diagramas de dispersão de Moran bivariado da Figura 3.2.3 representa graficamente tal afirmativa.

Tabela 3.2.3 – Índice de Moran bivariado dos serviços

Anos	I de Moran Global	*Significância
2002	0,1215	0,005
2008	0,1690	0,003
2014	0,1360	0,002

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Figura 3.2.3 – Diagrama de dispersão de Moran bivariado dos serviços

Fonte: Elaboração própria com base no programa *GeoDa 1.10*.

A Tabela 3.2.4 consolida as informações dos municípios com altos valores do ICn circundados por vizinhos com altos valores do VAB. As observações indicaram a existência de um cluster do tipo Alto-Alto situado nas regiões da Grande Fortaleza e do Cariri que se perdurou durante todo o período. No estado do Ceará, o número de municípios com transbordamento cresceu gradativamente, registrando 5, 7 e 8 cidades em 2002, 2008 e 2014 respectivamente. O crescimento do número de vínculos empregatícios acompanhou essa tendência, ampliando em 57% no último ano.

A região do Cariri teve o município de Barbalha constando nas três observações, enquanto que a cidade de Crato entrou para a lista das localidades com relação Alto-Alto em 2014. É impressionante perceber o elevado avanço no estoque de empregos nessa região no último ano, visto que apresentou incremento de 534% no número de trabalhadores, o qual se deu principalmente ao surgimento de Crato.

A região da Grande Fortaleza revelou os municípios de Fortaleza e Euzébio com forte especialização no setor e com relação Alto-Alto em todas as observações, indicando serem localidades geradoras de *spillovers* em seus vizinhos. Em 2014, com especialização inferior às dessas localidades, mas não menos importante, as cidades: Itaitinga, Caucaia, Maracanaú e Maranguape, figuraram, também, com dependência espacial positiva.

Tabela 3.2.4 – Evolução das aglomerações bivariadas Alto-Alto dos serviços

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	$\Delta\%$ (AB)
Cariri	1	1	2	1.576	2.147	13.611	534%
Grande Fortaleza	4	6	6	257.029	379.706	587.375	55%
Ceará	5	7	8	258.605	381.853	600.986	57%

Fonte: Elaboração própria.

3.3 Setor da Agropecuária

A distribuição do ICn desse setor indicou um crescimento gradativo no número de municípios especializados com valores acima da média estadual. Em 2002 constavam 34 localidades especializadas no setor, em 2008 passou para 38 e, em 2014 constou 41 municípios. Em postos de trabalho ocupados, isso representou um total de 10.002, 20.804 e 21.229 respectivamente, o que correspondeu a um incremento de 2% no último ano.

O I de Moran para o ICn dos 184 municípios cearenses (Tabela 3.3.1) revelou a presença de autocorrelação espacial positiva com valores estatisticamente significantes no

período. Tal afirmativa pode ser confirmada através da inclinação positiva da reta nos diagramas de dispersão de Moran (Figura 3.3.1).

Tabela 3.3.1 – Índice de Moran global da agropecuária

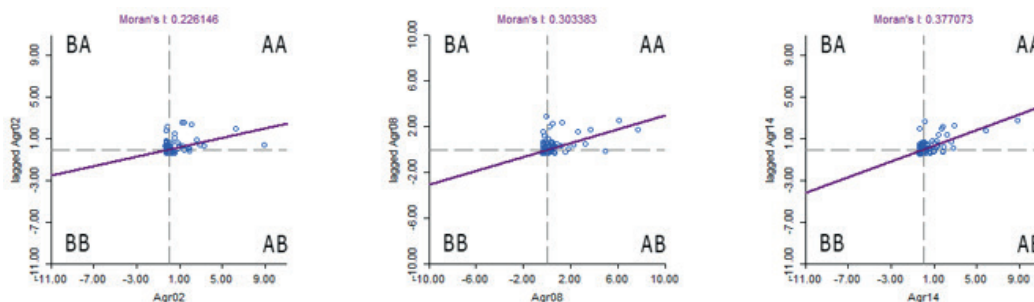
Anos	<i>I</i> de Moran Global	*Significância
2002	0,2261	0,001
2008	0,3033	0,001
2014	0,3770	0,001

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Em 2002, 79% dos municípios cearenses possuíam relações espaciais positivas, dos quais 11% se situavam no quadrante Alto-Alto e 68% no Baixo-Baixo. Em 2008, houve uma elevação para 81% do total, desses, 14% estavam no Alto-Alto e 67% no Baixo-Baixo. Em 2014, a porcentagem caiu para 74%, dos quais 14% estavam no Alto-Alto e 60% no Baixo-Baixo.

Figura 3.3.1 – Diagrama de dispersão de Moran da agropecuária



Fonte: Elaboração própria com base no programa *GeoDa 1.10*.

A Tabela 3.3.2 consolida os dados das aglomerações Alto-Alto ilustradas na Figura 3.3.2 em anexo, e sugere a existência de um agrupamento que sofreu pouca alteração desde 2002, o qual se manteve concentrado na região do Litoral Leste e Vale do Jaguaribe. Especificamente, o número de municípios com essa associação espacial, que era igual a 7 em 2002 e 2008, elevou-se em uma unidade em 2014, mesmo que o estoque de empregos formais no Estado tenha caído 2% na última observação. Em contrapartida, o padrão Baixo-Baixo demonstrou um avanço significativo no lado oeste do Estado, em que o primeiro ano revelou uma presença marcante em parte das regiões Sertão dos Crateús, Sertão dos Inhamuns, Cariri e, com menor influência, no sertão de Sobral e Vale do Jaguaribe. Tal cluster provocou transbordamento em regiões adjacentes ao longo do tempo,

onde o último ano indicou um espalhamento que englobou grande parte da região do Sertão dos Crateús e do Sertão dos Inhamuns, avançando, ainda, no Sertão de Sobral, Litoral Oeste/Vale do Curu e Cariri.

As cidades de Icapuí e Aracati na região do Litoral Leste apresentaram *ICn* acima de dois desvios padrão da média em todos os períodos, isto é, são localidades com extrema especialização no setor e potencialmente geradoras de spillovers. É importante destacar que essa região elevou em 77% o número de vínculos empregatícios em 2014. Embora incipiente em 2002, o Vale do Jaguaribe evoluiu em termos de dependência espacial durante o período, em grande parte ocasionada pela elevada especialização das cidades de Quixeré e Limoeiro do Norte, as quais demonstraram *ICn* acima de dois desvios padrão da média nas duas últimas averiguações. Entretanto, tal região apresenta indícios de estagnação ou declínio no setor, haja vista a quantidade de municípios relacionados positivamente ter se mantido constante nos dois últimos períodos e o grau de empregabilidade em 2014 ter reduzido em 39%. Por fim, Ibiapina na Serra da Ibiapaba constou nas duas primeiras análises, porém desapareceu na última.

Tabela 3.3.2 – Evolução das aglomerações Alto-Alto da agropecuária

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	$\Delta\%$ (AB)
Litoral Leste	5	3	5	2.346	3.378	5.971	77%
Serra da Ibiapaba	1	1	-	54	63	-	-100%
Vale do Jaguaribe	1	3	3	231	7.135	4.367	-39%
Ceará	7	7	8	2.631	10.576	10.338	-2%

Fonte: Elaboração própria.

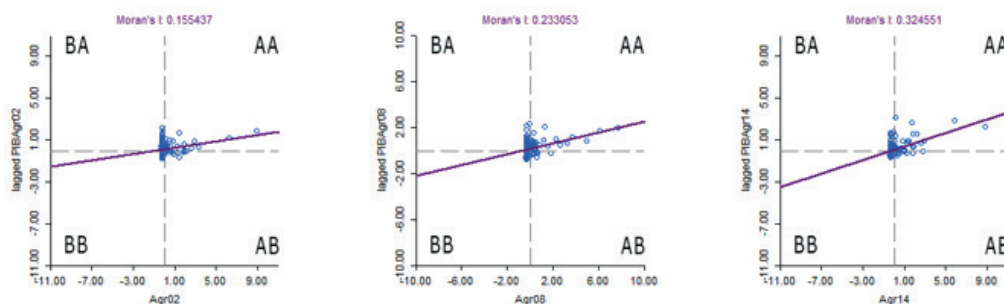
A Tabela 3.3.3 exibe os valores do *I* de Moran bivariado, onde se demonstra haver evidência de correlação direta entre o *ICn* e o VAB do setor agropecuário. A inclinação positiva da reta nos diagramas de dispersão de Moran bivariado da Figura 3.3.3 ilustra graficamente tal afirmativa.

Tabela 3.3.3 – Índice de Moran bivariado da agropecuária

Anos	<i>I</i> de Moran Global	*Significância
2002	0,1554	0,001
2008	0,2330	0,001
2014	0,3245	0,001

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Figura 3.3.3 – Diagrama de dispersão de Moran bivariado da agropecuária

Fonte: Elaboração própria com base no programa *GeoDa 1.10*.

A Tabela 3.3.4 consolida os dados dos municípios com altos *ICn* circundados por vizinhos com altos valores do VAB. Em 2002, há sinal da existência de pequenos clusters Alto-Alto nas regiões Litoral Oeste/Vale do Curu, Litoral Leste e Vale do Jaguaribe, os quais aumentaram sua representatividade em 2014, excetuando a primeira região que desapareceu. Destaca-se que essa relação adentrou o leste da região da Grande Fortaleza em 2014 e, no estado do Ceará, o número de municípios com transbordamento nesse setor dobrou, ao passo que a quantidade de novos empregados se elevou em 65%. Nesse mesmo ano, o padrão Baixo-Baixo, ou seja, cidade pouca especializada circundada por vizinhos com baixo VAB, foi representado por: Sobral, Mucambo e Pacujá no Sertão de Sobral; Canindé no Sertão de Canindé e Aiuaba no Sertão dos Inhamuns.

A quantidade de municípios com transbordamentos e o nível de vínculos empregatícios do Litoral Leste cresceram significativamente no período. De apenas Fortim em 2002, essa região ampliou para cinco cidades em 2014: Icapuí, Aracati, Beberibe, Jaguaruana e Fortim, com destaque para as três primeiras que registraram elevada especialização no setor agropecuário. O Litoral Oeste/Vale do Curu, porém, obteve pouca expressividade nessa análise, já que não indicou representantes nos dois últimos períodos. O Vale do Jaguaribe, embora tenha indicado uma redução de 39% no estoque de empregos em 2014, conquistou o segundo lugar do Ceará no quesito postos de trabalho criados, sendo representado pelas cidades: Quixeré, Limoeiro do Norte e Russas. A região da Grande Fortaleza, representada apenas pelo município de Cascavel em 2008, demonstrou avanço importante neste setor, a qual dobrou o número de cidades e incrementou em 205% o estoque de empregos no período seguinte. Vale salientar que esse elevado percentual não se deveu ao aparecimento da cidade de Pindoretama, mas sim ao aumento súbito no estoque de empregos de Cascavel naquele ano.

Tabela 3.3.4 – Evolução das aglomerações bivariadas Alto-Alto da agropecuária

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	Δ% (AB)
Litoral Leste	5	3	5	2.346	3.378	5.971	77%
Serra da Ibiapaba	1	1	-	54	63	-	-100%
Vale do Jaguaribe	1	3	3	231	7.135	4.367	-39%
Ceará	7	7	8	2.631	10.576	10.338	-2%

Fonte: Elaboração própria.

3.4 Setor da Administração Pública (APU)

A distribuição do ICn desse setor indicou que o número de municípios especializados com valores acima da média estadual cresceu 9% em 2008, mas, em 2014 retornou praticamente ao patamar de 2002, constando 104 localidades, ou seja, o número de cidades em 2002, 2008 e 2014, nesta ordem, foi respectivamente; 105, 114 e 104. Em postos de trabalho ocupados, isso representou um total de 55.686, 95.424 e 91.423 respectivamente, o que significou uma queda de 4% no nível de empregos em 2014.

O I de Moran para o ICn dos 184 municípios cearenses (Tabela 3.4.1) indicou a presença de autocorrelação espacial positiva com valores estatisticamente significantes no período. Visualmente, isso pode ser observado através da inclinação positiva da reta nos diagramas de dispersão de Moran (Figura 3.4.1).

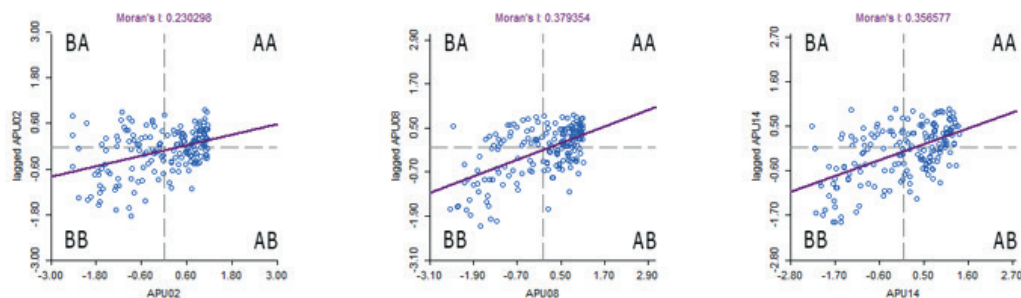
Tabela 3.4.1 – Índice de Moran global da APU

Anos	I de Moran Global	*Significância
2002	0,2302	0,001
2008	0,3793	0,001
2014	0,3565	0,001

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Em 2002, 61% das cidades cearenses possuíam relações espaciais positivas, dos quais 35% se situavam no quadrante Alto-Alto e 26% no Baixo-Baixo. Em 2008, esse número permaneceu constante, porém, 39% estavam no Alto-Alto e 22% no Baixo-Baixo. Em 2014, subiu para 62%, dos quais 34% estavam no Alto-Alto e 28% no Baixo-Baixo.

Figura 3.4.1 – Diagrama de dispersão de Moran da APU

Fonte: Elaboração própria com base no programa *GeoDa 1.10*.

A Tabela 3.4.2 consolida as aglomerações Alto-Alto exibidas na Figura 3.4.2 em anexo, e sugere um deslocamento do centro do Estado em direção ao oeste, isto é, em 2002 indicava *clusters* isolados nas regiões Sertão de Sobral e Vale do Jaguaribe e, de forma mais concentrada, no Sertão de Canindé e no Sertão dos Crateús. Em 2008, tal aglomerado manteve representação no Sertão de Sobral e no Sertão dos Crateús, contudo, a maior participação se concentrou no Sertão dos Inhamuns perpassando o Centro Sul e o Cariri. Por fim, em 2014, houve forte transbordamento nas cidades da região do Sertão dos Inhamuns, atravessando o Centro Sul e o Cariri, os quais ampliaram, também, a sua participação. De forma sintética, percebe-se um crescimento gradual no número de municípios positivamente relacionados no Alto-Alto; em 2002, 2008 e 2014, nesta ordem, obtiveram respectivamente, 5; 7 e 13 municípios, ao passo que o estoque de empregos formais no Estado apresentou evolução de 70% em 2014.

Devido ao elevado número de cidades com relação Baixo-Baixo detectado nas análises, cabe destacar que, em 2014, tal associação foi marcante na região da Grande Fortaleza com 14 municípios, no Litoral Leste e Vale do Jaguaribe com 3, no Cariri com 2 e no Litoral Oeste/Vale do Curu e Maciço de Baturité com uma cidade.

As regiões Sertão de Canindé e Vale do Jaguaribe só exibiram padrão Alto-Alto no primeiro ano, sumindo nos períodos seguintes. O Sertão de Sobral, representado unicamente pelo município de Pacujá, manteve-se com relação espacial positiva nos três períodos, porém apontou redução de 31% no número de empregos no último ano. As localidades Monsenhor Tabosa e Nova Russas no Sertão dos Crateús mostraram efeito transbordamento em 2002, contudo, somente Nova Russas permaneceu em 2008, desaparecendo em 2014. Na região Centro Sul, Saboeiro foi a única cidade a apresentar, concomitantemente, associação positiva e forte especialização em 2008 e, possivelmente, acarretou *spillover* em Catarina no último ano, que indicou, também, forte especialização. Em 2008, o Sertão dos Inhamuns indicou Aiuaba com elevada especialização ao lado de Pambu e Tauá. Es-

ses municípios permaneceram com relação positiva em 2014, os quais acabaram por gerar transbordamento para Arneiroz que obteve o maior valor de *ICn* dessa região. Por fim, a região do Cariri demonstrou a maior evolução desse setor, de apenas Várzea Alegre em 2008, ampliou para 6 cidades em 2014, representando um incremento de 224% no número de trabalhadores e, ainda, apontando extrema especialização em Tarrafas. Há evidências de que o espalhamento dessa região se deu mediante a pressão das regiões adjacentes Sertão dos Inhamuns e Centro Sul.

Tabela 3.4.2 – Evolução das aglomerações Alto-Alto da APU

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	$\Delta\%$ (AB)
Sertão de Canindé	1	-	-	1.201	-	-	-
Sertão de Sobral	1	1	1	379	558	385	-31%
Sertão dos Crateús	2	1	-	1.379	1.125	-	-100%
Vale do Jaguaribe	1	-	-	286	-	-	-
Cariri	-	1	6	-	1.581	5.124	224%
Centro Sul	-	1	2	-	905	1.886	108%
Sertão dos Inhamuns	-	3	4	-	3.547	5.713	61%
Ceará	5	7	13	3.245	7.716	13.108	70%

Fonte: Elaboração própria.

A Tabela 3.4.3 exibe os valores do *I* de Moran bivariado, onde se demonstra haver evidência de correlação inversa (autocorrelação negativa) entre o *ICn* e o VAB do setor da administração pública. A inclinação negativa da reta nos diagramas de dispersão de Moran bivariado da Figura 3.4.3 demonstra graficamente tal afirmativa.

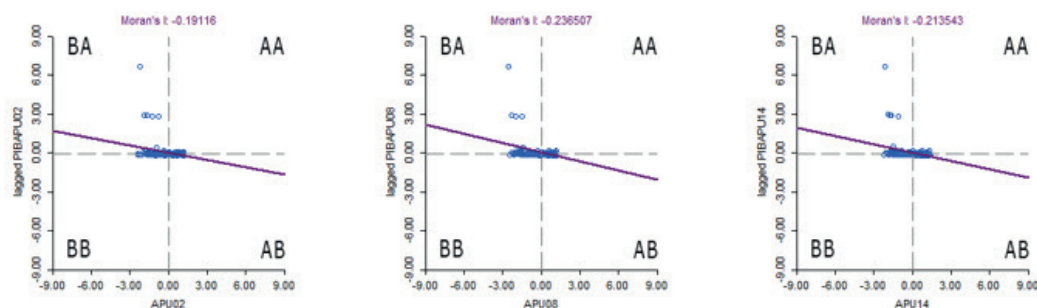
Tabela 3.4.3 – Índice de Moran bivariado da APU

Anos	<i>I</i> de Moran Global	*Significância
2002	-0,1911	0,001
2008	-0,2365	0,001
2014	-0,2135	0,001

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Figura 3.4.3 – Diagrama de dispersão de Moran bivariado da APU



Fonte: Elaboração própria com base no programa GeoDa 1.10.

A Tabela 3.4.4 e a Tabela 3.4.5 consolidam as informações dos municípios com altos (baixos) valores do *ICn* circundados por vizinhos com altos (baixos) valores do VAB. Os resultados mostraram a existência de um pequeno cluster do tipo Alto-Alto que não figurou em 2002, mas surgiu em 2008 e em 2014. Tal *cluster* esteve situado na região do Litoral Oeste/Vale do Curu e era representado pelo município de Miraima. Os estudos indicaram também a existência de pequenos clusters do tipo baixo-baixo nos três períodos. Situados no Cariri e no Sertão de Canindé, somente esse último se manteve em 2014, o qual pertencia à cidade de Canindé.

Tabela 3.4.4 – Evolução das aglomerações bivariadas Alto-Alto da APU

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	$\Delta\%$ (AB)
Litoral Oeste/Vale do Curu	-	1	1	-	590	598	1%

Fonte: Elaboração própria.

Tabela 3.4.5 – Evolução das aglomerações bivariadas Baixo-Baixo da APU

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	$\Delta\%$ (AB)
Cariri	2	1	-	1.087	780	-	-100%
Sertão de Canindé	1	1	1	2.088	2.767	2.744	-1%
Ceará	3	2	1	3.175	3.547	2.744	-23%

Fonte: Elaboração própria.

3.5 Setor do Comércio

A distribuição do *ICn* desse setor no estado do Ceará revelou que o número de municípios especializados com valores acima da média estadual cresceu gradativamente. Em 2002 constavam 62 localidades, em 2008 esse número saltou para 64 cidades especializadas e, em 2014 constou 73 municípios, o que correspondeu a um incremento de 14% nesse ano. Acompanhando esse ritmo de expansão, a quantidade de postos de trabalho gerados representou um total de 102.027, 160.954 e 256.654 respectivamente, ocasionando elevação de 59% no último ano.

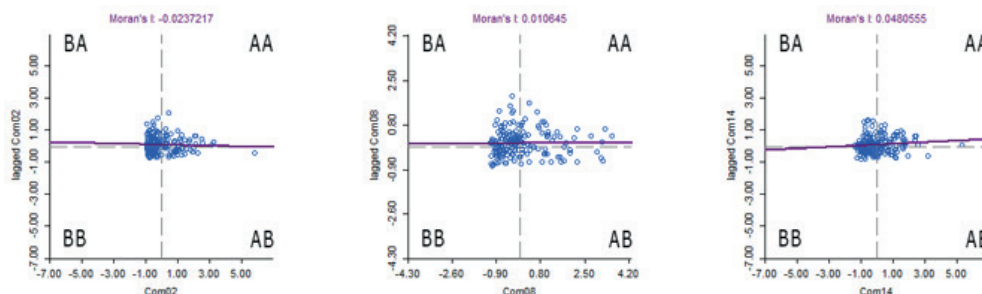
A estimação do *I* de Moran para o *ICn* dos 184 municípios cearenses (Tabela 3.5.1) indicou ausência de autocorrelação espacial nos três anos investigados, já que a pseudo-significância da variável em questão apresentou probabilidade não significativa. A pouca inclinação da reta nos diagramas de dispersão de Moran da Figura 3.5.1 corrobora tal afirmativa.

Tabela 3.5.1 – Índice de Moran global do setor do comércio

Anos	<i>I</i> de Moran Global	*Significância
2002	-0,0237	0,3730
2008	0,0106	0,3540
2014	0,0480	0,1260

*Valor da pseudo-significância baseada em 999 permutações.

Fonte: Elaboração própria.

Figura 3.5.1 – Diagrama de dispersão de Moran do setor do comércio

Fonte: Elaboração própria com base no programa GeoDa 1.10.

É importante frisar que a simples indicação de ausência de autocorrelação global não afasta a possibilidade de haver padrões de associação local, isto é, não necessariamente os padrões globais de associação espacial estarão alinhados com os padrões locais. Por conseguinte, faz-se necessário a investigação do Moran local.

A Tabela 3.5.2 consolida os dados das aglomerações Alto-Alto ilustrada na Figura 3.5.2 em anexo, e sugere uma evolução gradual de transbordamentos. Em 2002, pequenos clusters estavam divididos pelas regiões do Cariri, Centro Sul e Grande Fortaleza e, em 2008, o Sertão Central foi incluído nessa listagem. Todavia, enquanto a Grande Fortaleza reduziu a participação em 2014, a região do Vale do Jaguaribe despontou em toda a sua faixa oeste, ao mesmo tempo em que o Centro Sul e o Sertão Central ampliaram a presença. Em suma, de apenas 6 municípios associados positivamente em 2002, saltou para 7 em 2008 e ampliou para 9 em 2014. Tal incremento se refletiu de forma acentuada na geração de empregos no Ceará, apresentando elevação de 62% no último ano.

O aglomerado localizado na região Centro Sul sofreu pouca alteração no período, em 2002 figuravam apenas as cidades de Cedro e Orós, em 2008 somente essa última permaneceu e, em 2014, as cidades de Óros e Quixelô indicaram transbordamento. Vale salientar que essa região apresentou incremento de 208% no número de postos de trabalho criados no último ano. Na região da Grande Fortaleza, apenas Caucaia apareceu em 2002, mas a cidade de Maracanaú ampliou a presença dessa região em 2008, porém somente essa

última localidade permaneceu em 2014. É importante frisar que a geração de empregos proporcionada por Maracanaú em 2014, superou em 51% o somatório da quantidade de empregos das duas cidades do período anterior. Em 2008, Piquet Carneiro era a única cidade do Sertão Central com efeito transbordamento, todavia, em 2014, registraram ao lado dessa cidade, as localidades Senador Pompeu e Dep. Irapuan Pinheiro, alcançando juntas, aumento de 1.010% no estoque de empregos. Por fim, o Vale do Jaguaribe que não possuía representante em 2002 e 2008, figurou no último período com os municípios: Jaguaribe, Jaguaretama e Morada Nova.

Tabela 3.5.2 – Evolução das aglomerações Alto-Alto do setor do comércio

Regiões	Número de Municípios			Número de Vínculos Ativos			
	2002	2008	2014	2002	2008 (A)	2014 (B)	$\Delta\%$ (AB)
Centro Sul	2	1	2	160	85	262	208%
Cariri	3	3	-	534	963	-	-100%
Grande Fortaleza	1	2	1	1.251	7.869	11.903	51%
Sertão Central	-	1	3	-	59	655	1.010%
Vale do Jaguaribe	-	-	3	-	-	1.721	∞
Ceará	6	7	9	1.945	8.976	14.541	62%

Fonte: Elaboração própria.

A análise do *I* de Moran bivariado para essa atividade não foi passível de execução, haja vista as informações referentes às estimativas do VAB do setor do comércio não apresentarem os dados de forma desagregada, ou seja, os dados pertinentes a essa atividade estão consolidados às informações do setor de Serviços.

4 CONSIDERAÇÕES FINAIS

A metodologia adotada neste trabalho objetivou identificar e analisar a evolução de potenciais aglomerados produtivos locais, bem como a sua relação com o valor adicionado bruto dos setores de atividade econômica. Para tanto, observou-se tais arranjos mediante o uso de indicadores locacionais, principalmente o *ICn*, em seguida verificou-se o entorno relacionado ao setor produtivo para cada município por meio da Análise Exploratória de Dados Espaciais (AEDE). Dessa forma, foi possível garantir uma análise mais fidedigna à realidade dos fatos, uma vez que se considerou a influência da concentração localizada sobre as atividades do mesmo ramo nos demais municípios adjacentes.

O uso da Análise de Componentes Principais (*ACP*) para assegurar a melhor justeza dos pesos dos índices de concentração (*QL*, *HHm* e *PR*) _utilizados na elaboração do *ICn* dos setores econômicos _mostrou-se útil ao superar as limitações internas de cada índice, ainda que a sua carga relativa tenha variado ligeiramente da proporção de 1/3, valor comu-

mente aplicado em estudos que examinam aglomerados potenciais.

A aplicação da AEDE demonstrou que o estado do Ceará evoluiu em termos de concentração e dependência espacial em todos os setores e subsetores investigados. O setor industrial se concentrou em torno da região da Grande Fortaleza e se desenvolveu a partir do Litoral Leste em direção ao Litoral Oeste/Vale do Curu. O setor de serviços que se encontrava disperso pelo Estado no início da análise, convergiu-se na região da Grande Fortaleza ao final do período. A agropecuária manteve a sua predominância nas cidades do Litoral Leste e do Vale do Jaguaribe em todas as observações. A administração pública aglomerou-se nas cidades do Sertão dos Inhamus, Centro Sul e Cariri. Finalmente, o setor do comércio terminou por se fixar no Vale do Jaguaribe com transbordamentos para o Sertão Central e o Centro Sul.

Assim, uma vez que os resultados deste estudo estabeleceram um panorama temporal e espacial dos grandes setores econômicos cearenses através de observações empíricas, entende-se ser possível apoiar o governo no desenvolvimento de ações e programas mais assertivos para a formulação de políticas públicas que modifiquem positivamente situações sociais, econômicas, culturais e políticas.

REFERÊNCIAS

- ANSELIN, L. Local indicators of spatial association – LISA. **Geographical Analysis**. V 27 (2), April. p 93-115. 1995.
- ANSELIN, Luc et al. Spatial analytical perspectives on GIS. **The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association**. Taylor and Francis: London, p. 111-125, 1996.
- CAMPOS, Luís Henrique Romani de. **O Processo de Aglomeração Produtiva em Pernambuco**. Recife, 2004. 165 p. Tese de Doutorado em Economia. Universidade Federal de Pernambuco.
- CROCCO, Marco Aurelio et al. **Metodologia de identificação de arranjos produtivos locais potenciais: uma nota técnica**. Belo Horizonte: UFMG/Cedeplar, 2003. (Texto para discussão; 191).
- CROCCO, Marco Aurélio et al. **Metodologia de identificação de aglomerações produtivas locais**. Nova economia, v. 16, n. 2, p. 211-241, 2006.
- FESER, Edward; ISSERMAN, Andrew. **Clusters and rural economies in economic and geographic space**. 2005.

GONÇALVES, E. **A Distribuição Espacial da Atividade Inovadora Brasileira: Uma Análise Exploratória**. Texto para discussão CEDEPLAR N° 246, Belo Horizonte: UFMG, 2005. 33p.

HADDAD, Paulo Roberto et al. **Economia regional: teorias e métodos de análise**. Fortaleza: BNB, 1985.

HAINING, Robert P. **Spatial data analysis: theory and practice**. Cambridge University Press, 2003.

IBGE – Instituto Brasileiro de Geografia e Estatística. Disponível em: <https://ww2.ibge.gov.br/home/estatistica/economia/pibmunicipios/2014/default_base.shtm>. Acesso em: mai. 2017.

IPECE – Instituto de Pesquisa e Estratégia Econômica do Ceará, Ceará em Mapas Informações georreferenciadas e especializadas para os 184 municípios cearenses. Disponível em: <<http://www2.ipece.ce.gov.br/atlas/capitulo1/11.htm>>. Acesso em: 10 mai. 2017.

JOLLIFFE, I.T. **Principal component analysis**. 2.ed. New York, Springer-Verlag, 2002.

MARSHALL, Alfred. **Princípios de economia: tratado introdutório**. Trad. Revista Rômulo de Almeida, Ottomy Stauch. São Paulo: Abril Cultural, p. 231-239, 1982.

RAIS – Relação Anual de Informações Sociais. Disponível em: <<http://bi.mte.gov.br/bgcaged/rais.php>>. Acesso em: mai. 2017.

PAULA, Teófilo Henrique Pereira. **Arranjo Produtivo Local de Rochas Ornamentais do Estado do Espírito Santo: delimitação e análise dos níveis de desenvolvimento dos municípios constituintes**. Revista Econômica do Nordeste, v. 39, n. 1, p. 66-83, 2008.

RODRIGUES, Marcos Aurelio et al. **Identificação e análise espacial das aglomerações produtivas do setor de confecções na região Sul**. Economia Aplicada, v. 16, n. 2, p. 311-338, 2012.

SUZIGAN, Wilson et al. **Coefficientes de Gini locacionais–GL: aplicação à indústria de calçados do Estado de São Paulo**. Nova Economia, v. 13, n. 2, p. 39-60, 2003.

TOBLER, W. **A computer movie simulating urban growth in the Detroit Region**. Economic Geography. v.46, n.2, p.234-240. 1970.

ANEXOS⁴

Figura 3.1.2 – Mapa de *Cluster* LISA da indústria

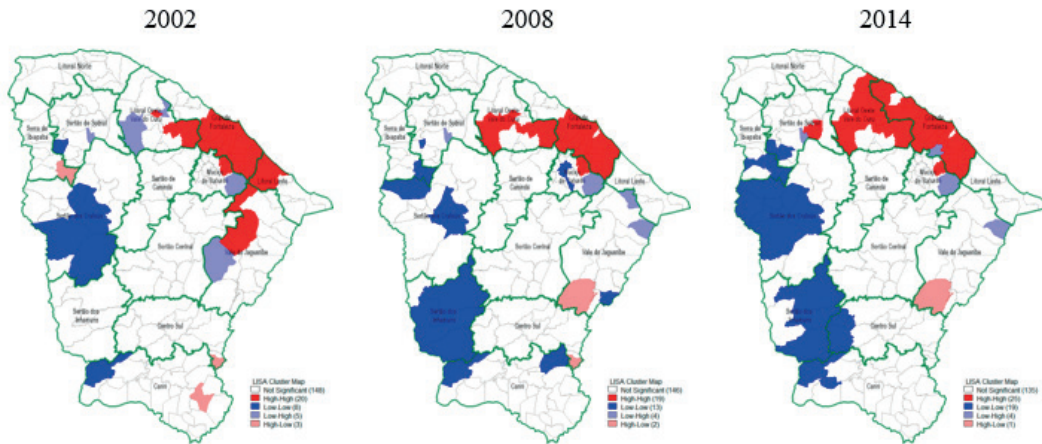


Figura 3.2.2 – Mapa de *Cluster* LISA do setor de serviços

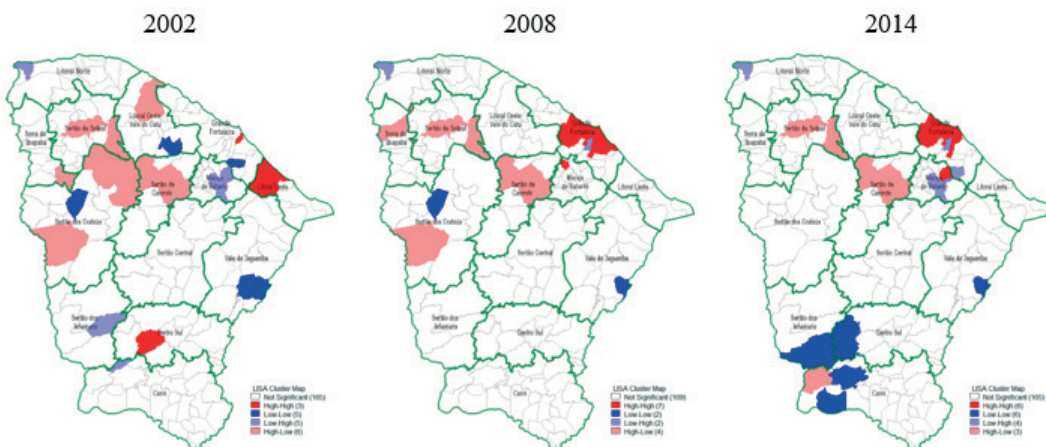
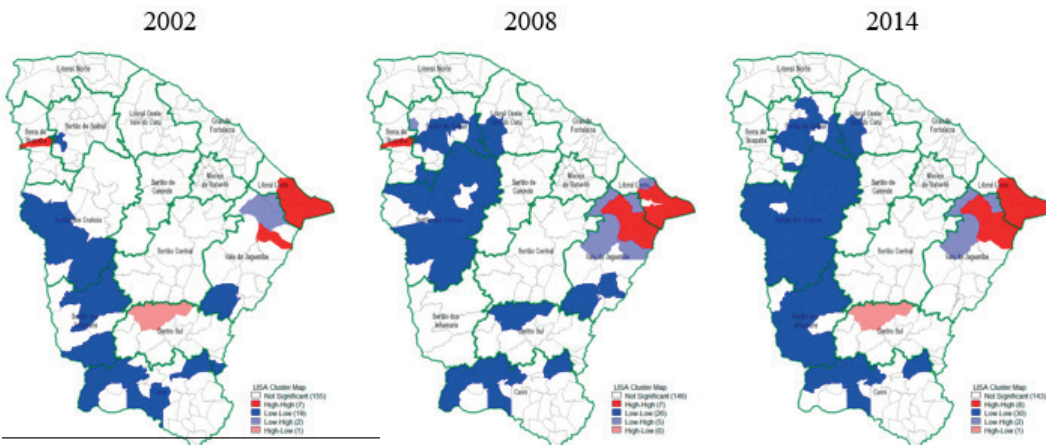


Figura 3.3.2 – Mapa de *Cluster* LISA da agropecuária



4 Fonte: Elaboração própria com base no programa GeoDa 1.10.

Figura 3.4.2 – Mapa de *Cluster* LISA da administração pública

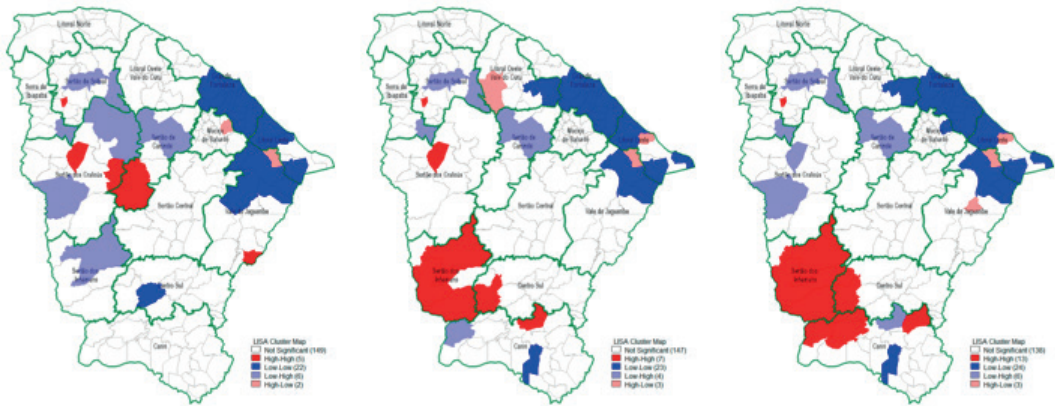
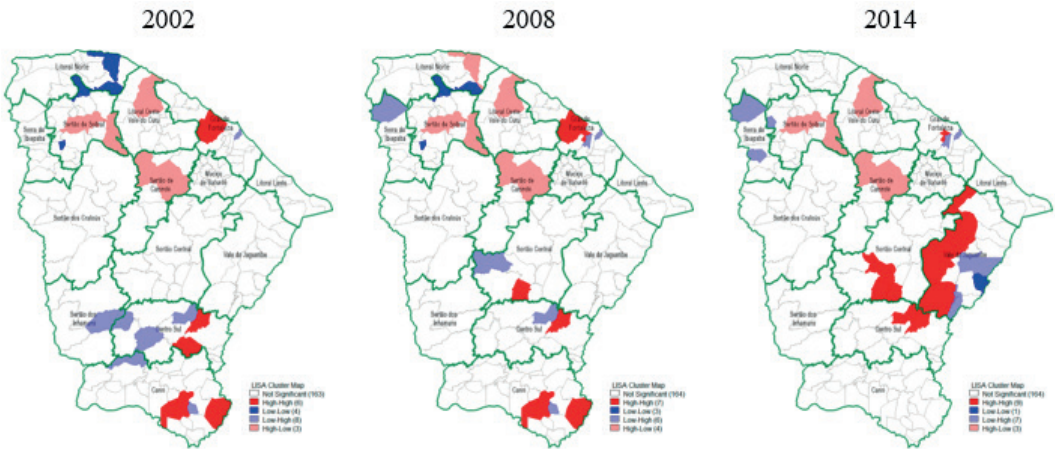


Figura 3.5.2 – Mapa de *Cluster* LISA do setor do comércio



DETERMINANTES DO RENDIMENTO ESCOLAR NAS ESCOLAS PÚBLICAS DE ENSINO MÉDIO DO ESTADO DO CEARÁ: UMA ANÁLISE DE PAINEL SOBRE O SPAECE DE 2008 A 2014

Vitor Borges Monteiro*

Pablo Urano de Carvalho Castelar**

João Benicio Sousa***

RESUMO

Este trabalho investiga os determinantes da nota de proficiência em língua portuguesa e matemática, extraídas do SPAECE, para o 3º ano do ensino médio de 441 escolas do Estado do Ceará, entre os anos de 2008 e 2014. O painel balanceado foi estimado controlando os Efeitos. A análise indica que número de docentes, número de alunos transportados e infraestrutura impactam positivamente na nota de língua portuguesa. Já a quantidade de alunos da escola, o percentual de docentes com nível superior, números de casos de dengue, taxa de mortalidade infantil e densidade demográfica impactam negativamente a nota. Os resultados para proficiência em matemática divergem da nota de português apenas em duas variáveis. Enquanto a distorção para a idade certa não apresentou significância para o exame de língua portuguesa, foi estatisticamente significativa e positivo para a nota de matemática. A infraestrutura, que foi positiva e significa para a explicar a nota em língua portuguesa, foi estatisticamente insignificante no modelo da nota de matemática. Um resultado marcante encontrado em todos os quantis e para os dois modelos, foi o sinal negativo e significativo das variáveis aprovados e abandono. O sinal negativo da variável aprovados, como discutido ao longo do texto, pode ser resultado do afrouxamento da regras e exigências aos alunos na busca da escola por bons indicadores.

Palavras-Chave: SPAECE; Regressão quantílica; Dados em Painel, Ensino Médio.

ABSTRACT

This work investigates the determinants of the proficiency score in Portuguese and Mathematics exams, extracted from the SPAECE, for the last year of high school in 441

* Professor Doutor do Curso de Bacharelado em Finanças da UFC. vitorborges@ufc.br. (85)996381459.

** Professor Doutor do Curso de Bacharelado em Finanças da UFC. pcastelar@ufc.br.

*** Graduando do Curso de Bacharelado em Finanças da UFC. beniciojb@hotmail.com.

schools in the State of Ceará, between 2008 and 2014. A balanced data panel was estimated to control for fixed effects. The analysis indicates that the number of teachers, the number of students transported and the presence of infrastructure positively impact the Portuguese note. However, the number of high school teachers, the number of dengue cases, the infant mortality rate and the demographic density have a negative impact on that grade. The results for proficiency in mathematics differ from the ones in Portuguese only in two variables. While age distortion did not present significance for the Portuguese exam, it was statistically significant and positive for the mathematics note. The infrastructure was shown to be positive to explain the note in Portuguese, however, it was statistically insignificant in the model of the mathematics note. A striking result found in all quantiles and for both models was the negative and significant sign of the approved students and number of abandonments variables. The negative sign of the approved variable, as discussed throughout the text, may be a result of loosening the rules for the school to obtain good indicators.

Keywords: SPAECE, Quantile regression; Panel Data, high school.

1 INTRODUÇÃO

Foi formado um consenso na sociedade ao longo do tempo sobre o papel relevante da educação para o desenvolvimento de um indivíduo e conseqüentemente da comunidade a qual pertença, quer seja local, regional, nacional ou até mundial, haja vista a globalização à qual estamos inseridos. A oportunidade do acesso ao conhecimento faz com que o ser humano desenvolva seu potencial, contribuindo em paralelo para o desenvolvimento da sociedade como um todo. O recente trabalho de Soares (2017), por exemplo, estima o valor de mercado da educação pública comparando as proficiências de alunos das redes estadual e privada, e constata que educação infantil e os níveis fundamental e médio são altamente progressivos, e os gastos aumentam o bem-estar.

Com isso, cada vez mais a sociedade tem exigido que a Educação seja o pilar primordial no desenho das políticas públicas, de maneira que essa priorização venha a ser representada por aportes maiores de investimento e, por conseguinte, visem maior qualidade do ensino. Nesse sentido, uma das maneiras de analisar essa qualidade é através de avaliações de conhecimento aplicadas ao corpo discente.

Apesar da quase universalização do acesso ao ensino fundamental e médio ser uma conquista significativa e meritória de comemoração, vários indicadores educacionais apontam a existência de problemas associados à qualidade da educação no Brasil. Em 2013, em

levantamento feito pelo movimento Todos Pela Educação¹, com dados da PNAD (Pesquisa Nacional por Amostra de Domicílio), verificou que pouco mais da metade dos jovens brasileiros (54,3%) concluiu o ensino médio até os 19 anos, idade considerada adequada para o término da educação básica e apenas 40% de jovens entre 15 e 17 anos frequentam o ensino médio.

A meta do governo brasileiro é atingir em 2021 os índices máximos definidos para o Ensino Básico². Para acompanhar o processo, o Ministério da Educação realiza, a cada dois anos, uma avaliação denominada Prova Brasil, voltada para as áreas de Língua Portuguesa e Matemática. São coletados dados sobre a capacidade dos alunos em leitura e resolução de problemas. Além dos testes, os alunos respondem um questionário com questões sociais e culturais relacionadas ao contexto no qual o aluno está inserido. A análise das provas dos alunos não contribui por si só para a melhoria do desempenho, é necessária a identificação das causas das dificuldades dos alunos e, com base nesse conhecimento, estabelecer formas de intervenção para a melhoria da qualidade do ensino.

Para atingir essa meta, torna-se necessário avaliar de forma acurada a condição atual das escolas do país, em especial as públicas, responsáveis por 90% das matrículas. Apesar da literatura identificar diversos fatores para determinar o rendimento escolar, como background familiar, mercado de trabalho atraente, turmas homogêneas, reforço escolar, aspectos pedagógicos, infraestrutura etc podem exercer influência significativa sobre a qualidade da educação. A pouca atenção dada ao tema é paradoxal, especialmente em face da riqueza de dados disponíveis. Informações sobre as condições materiais da escola são sistematicamente levantadas pelo Censo Escolar realizado pelo Inep/MEC, mas vêm sendo apresentadas de forma consolidada apenas nas sinopses anuais do Censo produzidas por aquele órgão.

Este estudo visa preencher essa lacuna com uma análise do impacto de itens de infraestrutura das escolas sobre rendimento escolar no Estado do Ceará. Além disso, o trabalho traz como contribuição diferenciada a utilização da metodologia apresentada por Koenker (2004), ainda incipiente em estudos que abordam o cenário brasileiro.

O Estado do Ceará realiza exames de avaliações anuais através do SPAECE, o presente trabalho visa investigar os resultados obtidos pelas escolas públicas estaduais do Ceará nos exames SPAECE de 2008 a 2014 e analisar a relação existente entre esses

1. <http://www.todospelaeducacao.org.br/educacao-na-midia/indice/31405/pnad-2013-confirma-lentidao-na-melhoria-da-educacao/>. Acesso em 05/09/2017.

2. Como apresentado na nota técnica http://fne.mec.gov.br/images/pdf/notas_tecnicas_pne_2011_2020.pdf. Acesso em 05/09/2017.

resultados e diversos indicadores estatísticos educacionais e socioeconômicos. As principais variáveis analisadas neste trabalho são infraestrutura da escola, indicadores socioeconômicos do município, número de docentes, qualificação do docente, tamanho da escola, distorção da idade certa, alunos transportados da zona rural para urbana etc.

O presente trabalho lançará uma investigação empírica de dados em painel com 441 escolas no período informado, utilizando técnicas de controles de efeito fixo, efeitos aleatórios e regressões quantílicas. Além desta introdução, segue uma revisão de literatura, metodologia e resultados e conclusão.

2 REVISÃO DE LITERATURA

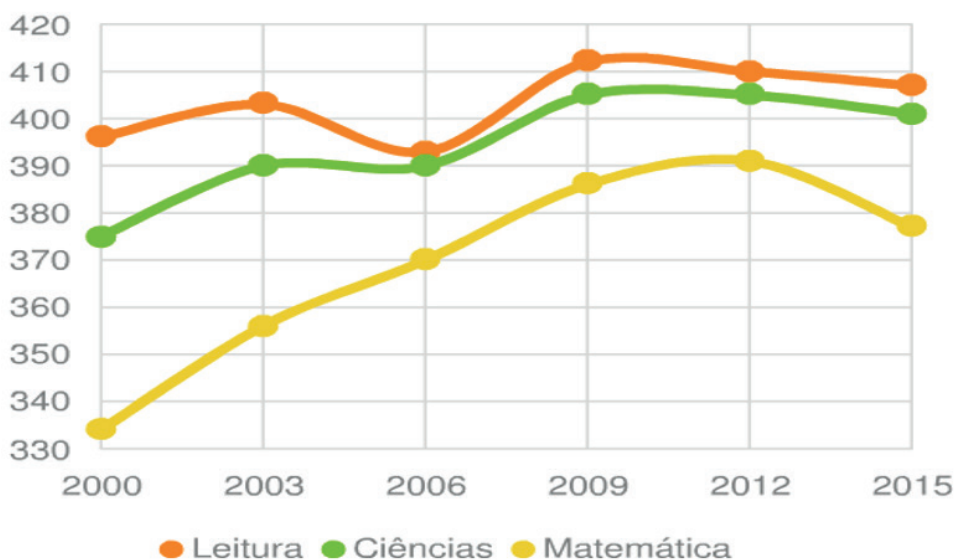
2.1 Os Sistemas de Avaliações no Brasil

Atualmente existem cinco grandes eixos de avaliações em larga escala no Brasil, todas sob a responsabilidade da Diretoria de Avaliação da Educação Básica (DAEB), vinculada ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). São elas: o Programa Internacional de Avaliação de Alunos (PISA); o Exame Nacional do Ensino Médio (ENEM); o Exame Nacional para Certificação de Competência de Jovens e Adultos (ENCCEJA); a Provinha Brasil; o Sistema Nacional de Avaliação da Educação Básica (SAEB/Prova Brasil).

O INEP é uma autarquia federal vinculada ao Ministério da Educação (MEC), cujo objetivo é promover estudos, pesquisas e avaliações sobre o Sistema Educacional Brasileiro com a intenção de subsidiar a formulação e implementação de políticas públicas para a área educacional. O INEP realiza levantamentos estatísticos e avaliativos em algumas etapas da educação básica, assim como na modalidade de educação de jovens e adultos. (SAEB/Prova Brasil).

O Programme for International Student Assessment (PISA) consiste em um programa de avaliação internacional, desenvolvido pelos países participantes da Organização para a Cooperação do Desenvolvimento Econômico (OCDE). Participam os 34 membros da OCDE e mais 36 convidados, totalizando 70 países. A avaliação é realizada a cada três anos e tem como objeto o desempenho de alunos na faixa etária de 15 anos, a partir do 7º ano, nas áreas de Leitura, Matemática e Ciências. O Brasil passou a participar do exame a partir do ano de 2000 e os resultados encontram-se na figura 1.

Figura 1- Evolução do desempenho do Brasil na Avaliação PISA por área de ensino



Fonte: Centro de Referências em Educação Integral

Note que o desempenho brasileiro evoluiu entre os anos de 2000 e 2009, porém reverteu a tendência para os exames posteriormente. Na avaliação de 2015, dos 70 países participantes, o Brasil ficou na posição 63º em Ciências, 65º em Matemática e 59º em Leitura.

O ENEM consiste em um exame anual, individual e de caráter voluntário. (BRASIL, 2008). Para Castro (1999), além de o aluno poder realizar uma autoavaliação de seus conhecimentos, ele obtém parâmetros para a continuidade dos estudos ou para o ingresso no mercado de trabalho. Importante ressaltar que a implantação desta avaliação contribuiu para discussões acerca de mudanças na forma de ingresso no Ensino Superior, cooperando com a implementação de uma seleção nacional unificada e, conseqüentemente, potencializando oportunidades. As quatro provas objetivas, de 45 questões cada, consistem em quatro eixos (Ciências Humanas, Ciências da Natureza, Linguagens e Matemática) mais uma redação.

No ano de 2012, o Estado do Ceará ficou em 25º lugar no ENEM, dentre os 27 Estados da federação mais o Distrito Federal, considerando o agregado de todas as redes. Vale ressaltar que no mesmo ano, quatro escolas privadas do Ceará ficaram entre as dez melhores do país, evidenciando o abismo existente entre o ensino médio público e privado no Estado. Essas evidências reforçam a importância do problema de pesquisa proposto neste trabalho

O ENCCEJA (Exame Nacional para Certificação de Competências de Jovens e

Adultos) é um exame nacional, cujo objetivo principal é “[...] avaliar as competências e habilidades básicas de jovens e adultos que não tiveram acesso aos estudos ou não puderam continuá-los na idade própria.” (BRASIL, 2008, p. 15). O participante do ENCCEJA se inscreve por áreas (Ensino Fundamental ou Ensino Médio), sendo possível obter certificação em cada prova. As certificações são feitas diretamente nas secretarias municipais e estaduais de educação que aderirem ao exame e equivale à conclusão da etapa de ensino.

A *Provinha Brasil* foi criada em consonância com o objetivo do Plano de Desenvolvimento da Educação (PDE) do MEC de viabilizar ações que contribuam para equidade e a qualidade da educação pública brasileira (BRASIL, 2008, p. 15). Tal prova está voltada para alunos que estão no início do 2º ano do Ensino Fundamental e, por enquanto, só existe avaliação com enfoque na alfabetização e ao início do letramento, mas o INEP planeja uma Provinha voltada para o pensamento lógico matemático.

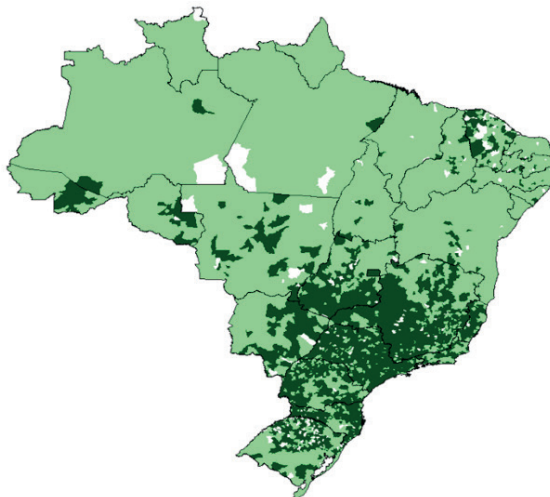
O Sistema de Avaliação da Educação Básica - SAEB e a Prova Brasil são dois exames complementares. As primeiras discussões sobre a criação e implementação das avaliações em larga escala no Brasil ocorreram no período entre 1985 e 1986. (BRASIL, 2008). De acordo com Castro (1999), o Brasil, e igualmente outros países em desenvolvimento, naquela época, necessitavam de estratégias para promover a melhoria da qualidade do ensino ofertado. Nesse sentido, o Saeb foi instituído em 1990, composto por um conjunto de avaliações externas, de caráter amostral, em larga escala e tem como principal objetivo realizar um diagnóstico da educação básica brasileira. O primeiro exame em 1990 consistia avaliar somente a rede pública do Ensino Fundamental, as séries contempladas para serem avaliadas nas áreas de Língua Portuguesa, Matemática e Ciências eram a 1ª, a 3ª, a 5ª e a 7ª séries.

Em 2005 o Saeb foi reestruturado³, e o exame começou a ser realizado por duas avaliações: Avaliação Nacional da Educação Básica (Aneb) e Avaliação Nacional do Rendimento Escolar (Anresc), conhecida como Prova Brasil. A Aneb permaneceu com os procedimentos da avaliação amostral (atendendo aos critérios estatísticos de no mínimo 10 estudantes por turma) das redes públicas e privadas, com foco na gestão da educação básica que vinha sendo realizado pelo Saeb. A Anresc (Prova Brasil) passou a avaliar de forma censitária testes que avaliam proficiência em Língua Portuguesa e Matemática, a cada dois anos, as escolas que atendessem aos critérios de no mínimo 30 estudantes matriculados na última etapa dos anos iniciais (4ª série/5º ano) ou dos anos finais (8ª série/9º ano) do Ensino Fundamental das escolas públicas.

3. Portaria Ministerial nº 931, de 21 de março de 2005

Ao ampliar a amostragem do SAEB, a Prova Brasil, além de exibir os desempenhos da federação, das regiões e dos estados, ampliou a verificação para as cidades e escolas. O 3º ano do Ensino Médio, que vinha sendo avaliado pelo Saeb desde 1997, também passou a compor a Prova Brasil em 2005, porém, com dados divulgados apenas por Estado. A Prova Brasil também é parte integrante do Índice de Desenvolvimento da Educação Básica (IDEB). Este índice é uma forma de mensuração que associa os resultados da Prova Brasil com as informações sobre o rendimento escolar dos alunos, sua taxa de aprovação, e determina a qualidade da educação baseado nesses fatores. (FERNANDES, 2007). Vale ressaltar que a escola só possui seu resultado individual divulgado se pelo menos 50% dos alunos realizarem o exame.

Figura 2: RESULTADOS DO SAEB 2015 - Proficiências médias por Município - 5º ano do Ensino Fundamental – Língua Portuguesa



Fonte: Relatório INEP Resultados (2016).

Reportado para uma análise para o Estado do Ceará, note na figura 2 o destaque para vários municípios cearenses que obtiveram notas médias acima da média nacional na proficiência de língua portuguesa, no 5º ano do Ensino Fundamental, no ano de 2015, representados pela área mais escura no mapa. Resultados semelhantes também são evidenciados para o 9º ano e para os exames de proficiência em matemática. Vale destacar várias políticas estaduais que contribuíram para o Estado alcançar este patamar, como: i) Leis de Incentivo à Educação no Ceará, que transfere recursos do ICMS vinculada a proporção de gasto com educação e ao atingimento de resultados em avaliações externas, ver avaliação desta política em Carneiro e Irffi (2017) e; ii) O Programa de Alfabetização na Idade Certa (PAIC), que passou a servir de elemento catalisador de políticas públicas a partir de 2007, através de exitosa experiência na cidade de Sobral, como discutido em Lavor (2012).

2.2 SPAECE - Sistema Permanente de Avaliação da Educação Básica do Ceará

O exame anual SPAECE é desenvolvido pela Secretaria da Educação (SEDUC) do estado do Ceará, em parceria com o Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF). Essa Avaliação abrange as escolas públicas das redes estadual e municipal do Estado, avaliando os alunos da Educação Básica, desde as etapas de Alfabetização até o Ensino Médio.

Do ano de sua criação, em 1992, até os dias atuais, o SPAECE fornece subsídios para formulação, reformulação e monitoramento das políticas educacionais, vislumbrando a oferta de um ensino de qualidade a todos os alunos da rede pública do Ceará. Para isso, a cada edição, são aplicados testes de desempenho e questionários contextuais, que possibilitam extrair dados, visando traçar um panorama da qualidade da educação dos alunos.

Realizada de forma censitária e universal, essa avaliação abrange as escolas estaduais e municipais, utilizando testes, com itens elaborados pelos professores da Rede Pública, tendo como orientação os Parâmetros Curriculares Nacionais (PCN) do Ministério da Educação (MEC) e os Referenciais Curriculares Básicos (RCB) da SEDUC. São aplicados, também, questionários contextuais, investigando dados socioeconômicos e hábitos de estudo dos alunos, perfil e prática dos professores e diretores.

Por considerar a importância da avaliação como instrumento eficaz de gestão, a SEDUC amplia, a partir de 2007, a abrangência do SPAECE, incorporando a avaliação da alfabetização e expandindo a avaliação do Ensino Médio para as três séries de forma censitária. Desta forma, o SPAECE passa a ter três focos: Avaliação da Alfabetização - SPAECE-Alfa (2º ano). Avaliação do Ensino Fundamental (5º e 9º anos) e Avaliação do Ensino Médio (1a, 2a e 3a séries).

2.3 Fatores Determinantes do Desempenho Escolar

Na literatura nacional, encontra-se uma vasta quantidade de trabalhos, por exemplo, Barros *et al.* (2001), Albernaz, Ferreira e Franco (2002), Felício e Fernandes (2005), Alves e Soares (2007), Sátyro e Soares (2007), Jacobson (2010) e Almeida (2014) que avaliam os principais fatores relacionados ao resultado da educação. Vários são os fatores levantados por estes autores como possíveis determinantes do rendimento do aluno, tais como aspectos socioeconômicos, aspectos de características familiares ou background familiar, mercados de trabalho atraentes, escolaridade dos professores, a qualidade da infraestrutura física, turmas homogêneas, trabalho pedagógico intencional e sistematizado, hábito de leitura, reforço escolar etc.

No estudo desenvolvido por Barros *et al.* (2001), a despeito das elevadas taxas de retorno à educação, o Brasil apresenta indicadores educacionais abaixo dos padrões internacionais. Ainda mais preocupante é o fato de que o subinvestimento em capital humano é tanto mais acentuado quanto mais pobre é a família. Segundo os autores, dado que pessoas menos escolarizadas serão com maior probabilidade pobres no futuro, essa natureza diferenciada do investimento em educação leva à transmissão intergeracional da pobreza.

Para analisar o fraco desempenho escolar, o trabalho utiliza o universo de indivíduos entre 11 e 25 anos de idade que vivem em áreas urbanas das regiões Nordeste e Sudeste, com dados da PNAD e PPV. Quatro tipos de determinantes são investigados: a disponibilidade e qualidade dos serviços educacionais, a atratividade do mercado de trabalho local, a disponibilidade de recursos familiares (financeiros e não-financeiros) e o volume de recursos da comunidade em que o indivíduo vive. Os resultados apontam que características familiares têm impacto no rendimento escolar, principalmente a escolaridade da mãe. Mercados de trabalho atraentes também parecem estar relacionados à menor escolaridade.

Ressalte-se, no entanto, o estudo quanto à qualidade dos professores, onde os autores observam que a escolaridade dos professores da segunda etapa do ensino fundamental contribui mais para a melhora do desempenho educacional que a escolaridade dos professores do ensino médio. Mais importante: a escolaridade média dos professores do ensino médio pareceu estar negativamente relacionada com a escolaridade dos jovens entre 11 e 25 anos. Essa relação negativa concentra-se particularmente entre os habitantes da região Nordeste.

Já o trabalho de Albernaz, Ferreira e Franco (2002) estima uma função de produção educacional para o ensino fundamental brasileiro, com base nos dados do Sistema de Avaliação da Educação Básica (SAEB) para as 8as séries em 1999. Utilizando modelos hierárquicos lineares, de forma a evitar vieses de agregação, a estimação sugere que cerca de oitenta por cento da variância em desempenho médio entre as escolas deve-se a diferenças na composição socioeconômica de seus alunos. Como em outros estudos internacionais, a rede a que a escola pertence está relacionada de forma importante com o desempenho de seus alunos. Não obstante, e em contraste com a maior parte da literatura internacional, outras variáveis escolares, tais como a escolaridade dos professores e a qualidade da infraestrutura física, também afetam o desempenho dos estudantes.

Para Felício e Fernandes (2005), a relação entre a qualidade da escola e o desempenho escolar deve-se a uma grande diversidade de fatores, mas, de modo geral, o impacto da escola é de relevância pequena especialmente quando comparada à parcela explicada pelo *background* familiar. Os autores utilizam dados do SAEB 2001 referentes às 4as séries do

Ensino Fundamental do Estado de São Paulo. A decomposição da desigualdade de notas em uma parte explicada pela escola e outra atribuída ao background familiar, resulta em um intervalo de valores possíveis para a proporção da desigualdade de desempenho explicada pelas diferenças entre as escolas (efeito escola). Os resultados indicam que o efeito escola pode explicar entre 0 e 28,4% da desigualdade total de notas de Língua Portuguesa e entre 8,7 e 34,4% para Matemática.

Segundo Alves e Soares (2007), o efeito-escola no desempenho acadêmico dos alunos de 28 turmas de sete escolas públicas da mesma região administrativa de Belo Horizonte, que atendem a comunidades vizinhas relativamente homogêneas do ponto de vista socioeconômico, foi mensurado através da identificação da parte do progresso dos alunos que pode ser atribuída às escolas. A pesquisa coletou dados longitudinais junto a todas as turmas de 5ª série do Ensino Fundamental no início e no final do ano letivo. Os alunos responderam a testes de Português (leitura) e Matemática e preencheram um questionário sobre suas características demográficas, socioeconômicas e culturais, itens sobre o percurso escolar e hábitos de estudo. Também foram realizadas entrevistas com profissionais das escolas e com familiares dos alunos com o objetivo de contextualizar os resultados obtidos com os instrumentos quantitativos. Os resultados revelam que o efeito-escola se estrutura pela forma como as escolas organizam os alunos em turmas. A formação das turmas por nível de habilidade impacta não apenas o nível de desempenho dos alunos, mas também as diferenças nas taxas de progresso observadas durante o estudo. Os critérios de formação de turmas contribuem para que pequenas diferenças entre os alunos se transformem em grandes diferenças entre os grupos e acirrem a estratificação escolar de uma forma mais ampla.

Conforme Sátyro e Soares (2007), que se baseia no Censo Escolar para descrever as condições de infraestrutura, tanto física como de recursos humanos, das escolas do ensino fundamental brasileiro, o acesso a serviços básicos como água, eletricidade e esgotamento sanitário; dependências escolares; existência de biblioteca ou sala de leitura; infraestrutura de comunicação e informação; além de formação de professores, que não é infraestrutura, mas é de importância fundamental para a qualidade do ensino, todos são itens que podem contribuir para o rendimento do aluno. No caso das dependências escolares e das bibliotecas e salas de leitura, os autores foram forçados a construir um índice sintético que resume a grande variedade de informações disponíveis no Censo Escolar de cada ano.

O primeiro resultado é que as condições materiais do ensino melhoraram muito entre 1997 e 2005, embora os resultados escolares, tanto em termos de repetência como em termos de aprendizado, não tenham mudado muito neste mesmo período. O segundo é que não há grandes diferenças de infraestrutura ou formação docente entre as escolas estaduais e privadas, embora existam grandes diferenças entre estas duas e as municipais.

Este é um resultado curioso, já que não há grandes diferenças de resultado de rendimento escolar entre municipais e estaduais, mas há, sim, grandes diferenças entre estas duas e as privadas. Estes dois resultados colocam em dúvida o impacto de melhorias de infraestrutura ou de formação docente sobre os resultados escolares. O último resultado vai no sentido oposto. As escolas rurais sofrem tanto com condições precárias, quanto com resultados educacionais muito baixos até hoje, o que sugere que talvez uma política de incremento de infraestrutura de escolas rurais possa ter impactos sobre repetência ou aprendizado.

Prédios e instalações adequadas, existência de biblioteca escolar, espaços esportivos e laboratórios, acesso a livros didáticos, materiais de leitura e pedagógicos, relação adequada entre o número de alunos e o professor na sala de aula e maior tempo efetivo de aula, por exemplo, possivelmente melhorem o desempenho dos alunos. Um dos poucos exemplos de estudo nesta linha é o trabalho realizado por Castro e Fletcher (1986), em que os autores analisam amostra de 600 escolas em todo o território nacional para retratar a situação da rede física escolar, a disponibilidade de equipamentos necessários à atividade pedagógica e o funcionamento da escola. A pergunta que orientou o estudo foi: “O que vem a ser a escola pública de primeiro grau frequentada pelos brasileiros em 1985?”. Os autores concluem que a infraestrutura das escolas tem impacto na qualidade de ensino, assim como ressaltam a importância das reprovações no rendimento discente no Brasil.

O mesmo resultado é encontrado por Lee *et al.* (2004), onde é encontrado que uma melhor infraestrutura têm impactos positivos no aprendizado dos alunos, particularmente para o Brasil, Portugal e Estados Unidos. Os autores estudam, através da metodologia multinível (HLM) escolas secundárias a nível mundial utilizando dados do PISA para o ano de 2000. Para o caso específico do Brasil, são analisadas as performances de 4.893 alunos de 15 anos de idade em 324 escolas secundárias. Além da correlação entre presença de infraestrutura e rendimento dos alunos dos testes padronizados, os autores também destacam a importância da ausência e evasão escolar.

Particularmente sobre a importância da infraestrutura, cabe ressaltar o trabalho de Santos *et al.* (2014), que faz uma análise das infraestruturas das escolas brasileiras através de dados do Censo Escolar de 2011 utilizando o método KDD. Os resultados apontam uma infraestrutura precária em número considerável de escolas no Brasil. Caso, de fato, haja correlação entre o rendimento escolar e a presença e qualidade da infraestrutura, tal perspectiva é preocupante.

Jacobson (2010) oferece uma investigação acerca dos fatores determinantes do desempenho de alunos de 4ª. série do ensino básico na Prova Brasil. Os dados foram colhidos por meio de entrevistas semi-estruturadas e análise dos planos e diários dos docentes de

seis escolas paraenses que apresentaram os maiores e menores índices em Língua Portuguesa na avaliação realizada em 2007. Das três escolas com piores resultados, em duas delas houve falhas na aplicação da prova, inclusive com troca de gabaritos por parte dos aplicadores.

Todavia, o fator preponderante para o baixo desempenho dos alunos é que eles não tinham os conhecimentos exigidos. Por outro lado, nas escolas com índices mais altos, verificamos um trabalho pedagógico sistematizado de ensino dos conteúdos. Além disso, os professores demonstraram maior conhecimento que o primeiro grupo sobre a avaliação e os resultados alcançados por sua escola. Os resultados confirmam a importância do contexto socioeconômico em que os alunos estão inseridos, da formação dos professores e também do processo de aplicação das avaliações em larga escala. No entanto, o mais forte determinante do desempenho das unidades escolares é o trabalho pedagógico intencional e sistematizado dos conteúdos de Língua Portuguesa.

Por fim, Almeida (2014) busca identificar os principais determinantes dos resultados escolares dos alunos, em especial, daqueles discentes de escolas públicas que apresentam as piores e as melhores proficiências em testes padronizados na primeira e segunda fase do ensino fundamental no Brasil. A pesquisa traz como elementos diferenciados nesta temática de estudo no país, a análise dos efeitos de programas de reforço de aprendizagem e do hábito de leitura dos discentes sobre o desempenho escolar.

Para atender a tais objetivos, são utilizados os microdados da Prova Brasil 2011 e a regressão quantílica, método econométrico que permite compreender os efeitos das variáveis explanatórias sobre o desempenho escolar dos alunos para os n -ésimos quantis da distribuição condicional. Os resultados centrais da pesquisa apontam que a existência de programas de reforço escolar não se mostra determinante para os resultados dos alunos, sobretudo, para aqueles do 5o ano do ensino fundamental. Além do mais, a falta de hábito de leitura do estudante é um fator que contribui para a redução do seu desempenho escolar.

3 METODOLOGIA E RESULTADOS

Esta seção tem por objetivo descrever, analisar e discutir os resultados. A seção 3.1 abordará a base de dados, a seção 3.2 descreve o modelo econométrico e a análise dos resultados do modelo.

3.1 Base de Dados

A fim de analisar a evolução da Educação no Ceará nos últimos anos, faz-se uso das notas na avaliação de competências e habilidades em Língua Portuguesa e Matemática, aplicadas através do exame SPAECE. Essa avaliação abrange as escolas públicas das redes

estadual e municipais do Estado, avaliando os alunos da Educação Básica, desde as etapas de Alfabetização até o Ensino Médio. Para a presente pesquisa, foram utilizados dados de 2008 a 2014 dos alunos de 3º ano de Ensino Médio da rede estadual. Em 2014, dos 97.654 alunos do ensino médio aptos a realizarem a prova, em todo o Ceará, 92,10% efetivamente realizaram. Para analisar os determinantes de rendimento escolar faz-se necessário, antes de qualquer análise, compreender o comportamento desta variável. O banco de dados dispõe 441 escolas do ensino médio da rede pública estadual, do total de 522, para os anos de 2008 a 2014, totalizando 3087 observações.

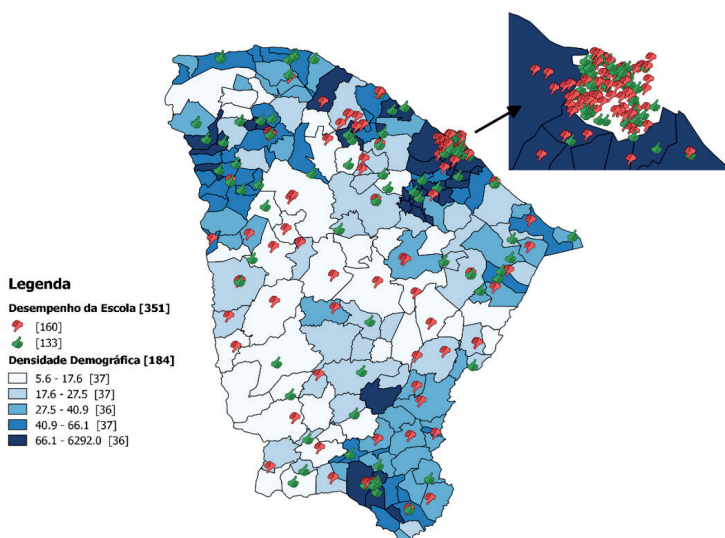
Tabela 1: Média das Proficiências de Língua Portuguesa e Matemática por ano

Ano	Média Nota de Língua Portuguesa	Média Nota de Matemática	Observações
2008	234.7522	247.5291	441
2009	250.2596	258.9714	441
2010	258.8000	257.8623	441
2011	257.6883	261.6065	441
2012	248.9823	257.6988	441
2013	253.3071	262.4005	441
2014	258.9368	260.2611	441
TOTAL	251.8181	258.0471	3087

Fonte: Elaboração dos autores.

Algumas foram excluídas em virtude do equilíbrio do painel. Observe na Tabela 1 a média de notas das 441 escolas da amostra, por ano. A média global de todo período aponta um desempenho maior na nota de matemática. As notas de ambas as proficiências crescem no período analisado, entretanto, a nota de português cresce mais rapidamente

Figura 3: Mapa Interativo entre Densidade Populacional dos municípios e Desempenho na nota de proficiência em Língua Portuguesa. Período 2008 a 2014

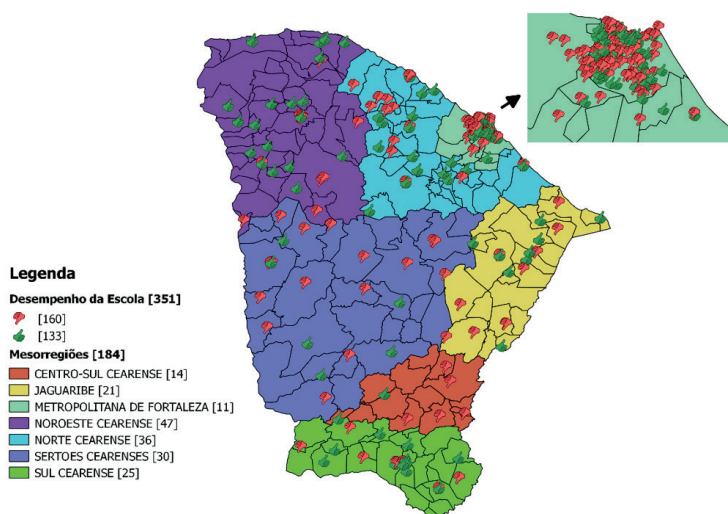


Fonte: Elaboração dos autores.

A Figura 3 mostra cinco *quantis* de densidade populacional (hab/km²) dos 184 municípios cearenses, e a localização de 351 escolas. Algumas escolas não foram localizadas pela geocodificação do software utilizado. O desempenho da escola representa a nota média da escola no período de 2008 a 2014 na avaliação língua portuguesa. O polegar vermelho indica que a escola obteve desempenho em língua portuguesa inferior a 251,8181 e o polegar verde indica que a escola obteve desempenho superior a 251,8181. Os efeitos da densidade demográfica sobre o rendimento escolar podem estar relacionados a diversos fatores positivos e negativos. Como aspectos negativos, destacam-se os grandes problemas urbanos como violência, tempo de deslocamento para o colégio etc. Como aspecto positivo destaca-se o acesso a serviços públicos de biblioteca, internet etc.

Nos 37 municípios com densidade demográfica inferior a 17,6, representados pelo primeiro quantil, existem 24 escolas, das quais 62,5% destas obtiveram nota inferior à média. Nos 36 municípios com densidade demográfica superior a 66,1, representados pelo último quantil, existem 201 escolas, das quais 55,35% das escolas obtiveram nota inferior a média. Logo, existe uma predominância de escolas inferiores à média no menor e maior quantil, entretanto, o percentual de escolas com boas notas aumenta com o aumento da densidade.

Figura 4: Mapa Interativo entre Mesorregiões dos municípios e Desempenho na nota de proficiência em Língua Portuguesa. Período 2008 a 2014.



Fonte: Elaboração dos autores.

Pode-se observar no mapa da Figura 4 que a mesorregião Noroeste do Estado possui uma concentração de escolas com bons resultados, se destacando frente as demais mesorregiões. Os resultados do SPAECE foram confrontados com indicadores estatísticos disponibilizados através da CEGED (Célula de Estudos, Gestão de Dados e Disseminação

de Informações Educacionais) da Secretaria de Educação do Ceará, descritos na tabela 1, a fim de analisar as relações entre desempenho escolas e características das escolas e fatores socioeconômicos do município.

Tabela 2: Descrição das variáveis, quanto a Fonte e Nível e Agregação

Variável	Fonte	Nível de Agregação	Descrição
LP	SPAECE SEDUC/CE	Escola	Nota da proficiência de Língua Portuguesa para o 3º ano do Ensino Médio.
MT	SPAECE SEDUC/CE	Escola	Nota da proficiência de Matemática para o 3º ano do Ensino Médio.
DV LP	SPAECE SEDUC/CE	Escola	Desvio padrão da nota de português
DV MT	SPAECE SEDUC/CE	Escola	Desvio padrão da nota de matemática
APROV	CEGED- SEDUC/CE	Escola	Percentual de alunos aprovados total no Ensino Médio
ABN	CEGED- SEDUC/CE	Escola	Percentual de abandono escolar total no Ensino Médio
DISTORCAO	CEGED- SEDUC/CE	Município	Percentual de alunos com distorção da idade certa
TRANSPORTE	CEGED- SEDUC/CE	Município	Percentual de alunos transportados da zona rural para a zona urbana
MATRICULADOS	CEGED- SEDUC/CE	Escola	Quantidade total de alunos matriculados no Ensino Médio
DOCENTES	CEGED- SEDUC/CE	Escola	Quantidade total de professores no Ensino Médio
DOC SUP	CEGED- SEDUC/CE	Escola	Percentual de professores com ensino superior
POP	IPECE	Município	População total do município
AREA	IPECE	Município	Área total do município
TMI	IPECE	Município	Taxa de mortalidade infantil
DENGUE	DATASUS	Município	Casos de Dengue registrados.
LOCALIZAÇÃO	Elaboração própria	Escola	1 se a escola está na zona urbana e 0 caso contrário

Fonte: Elaboração dos autores

Além dos indicadores já disponibilizados, foi criado um indicador de infraestrutura para representar a nota de infraestrutura das escolas. A composição dessa nota se dá pela soma de dez indicadores de infraestrutura (sala de diretoria, cozinha, biblioteca, sala de professores, sala de leitura, laboratório de informática, laboratório de ciências, banheiro, quadra de esportes, acesso à internet). Caso a escola possua o item analisado será computado 1(um) ponto e em caso negativo será computado 0 (zero) ponto. Assim, essa nota criada tem variação de 0(zero) a 10(dez). Note na tabela 2 que algumas variáveis não possuem nível de agregação por escola. Por exemplo, todas as escolas de Fortaleza recebem a mesma pontuação número de casos de dengue em Fortaleza. Essas informações variam apenas nos municípios.

3.2 Análise Econométrica

O modelo de dados em painel caracteriza-se pela existência de dados *cross-section* e dados longitudinais, representados no presente estudo por dados de 441 escolas de Ensino Médio no Estado do Ceará para os anos de 2008 a 2014. A principal característica da modelagem em painel consiste na presença de uma heterogeneidade não observada nos termos *cross-section*, que apresenta correlação com o termo de erro. A heterogeneidade não observada, invariante no tempo e distinta entre os termos *cross-section*, pode ser caracterizada

por aspectos cognitivos de cada escola. O processo de estimação consiste na eliminação deste termo de heterogeneidade não observada através de um modelo em diferenças ou a modelagem deste termo através de efeitos fixos, com inclusão de variáveis *dummy*, ou a modelagem de efeitos aleatórios.

$$y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it} \quad (1)$$

Onde $\varepsilon_{it} = \mu_i + v_{it}$

A equação (1) apresenta a formalização de um modelo de dados em painel, a variável dependente y_{it} representa as notas da escola i , no ano t , no SPAECE. Para tanto, são estimados dois modelos, representando as notas de proficiência em língua portuguesa e matemática. O vetor de variáveis explicativas X_{it} representa o conjunto de variáveis que serão testadas como variáveis que influenciam as notas examinadas no SPAECE, e são apresentadas na tabela 2. O termo de erro do painel é representado por um efeito fixo μ_i e um componente idiossincrático v_{it} . O teste de Hausman é utilizado para testar a hipótese nula de efeitos aleatórios.

A regressão quantílica apresenta riqueza quanto à análise dos resultados devido ao impacto da variável explanatória sobre a explicada poder ser verificada em diferentes níveis de quantis da variável dependente. Os pioneiros no estudo de regressões quantílicas (RQ) foram Koenker e Bassett (1978), e o modelo consiste na minimização de uma função objetivo com restrições de parâmetros que controlam a força de regularização e o quantil de interesse. Outra vantagem da regressão quantílica consiste na sua utilização em situações nas quais ocorrem problemas de multicolinearidade e de dimensionalidade decorrentes da quantidade, geralmente superior, de parâmetros a serem estimados em relação ao número de observações, e os modelos podem ser usados para caracterizar toda a distribuição condicional de uma variável resposta dado um conjunto de regressores. Koenker (2004) introduziu uma nova abordagem para estimar regressões quantílicas para dados em painel.

Ajustar um modelo de regressão quantílica para dados em painel consiste atualmente em utilizar os seguintes modelos disponíveis: (i) o método de estimação de efeitos fixos penalizados (FE) proposto por Koenker (2004) e (ii) o método dos efeitos aleatórios correlacionados (CRE) proposto pela primeira vez por Abrevaya e Dahl (2008) e elaborado por Bache *et al.* (2013).

O estimador de FE baseia-se na minimização de uma soma ponderada de K funções de objetivo de regressão de quantile comuns correspondentes a uma seleção de valores de K de τ , com pesos específicos de tau especificados pelo usuário. Os coeficientes de inclinação desta função objetiva são dependentes de τ , enquanto que os coeficientes correspondentes

aos efeitos fixos são assumidos como independentes de τ . O vetor de coeficientes de efeitos fixos é penalizado por um termo de penalidade λ , diminuindo assim esses coeficientes em direção a zero.

Ressalta-se que no modelo FE o termo α_i é o efeito fixo que capta outras características do indivíduo que não são observáveis, mas que certamente têm impacto nas notas de avaliação. Como exemplo destas características, podemos citar as habilidades cognitivas e não-cognitivas das escolas. Além disso, α_i é independente dos quantis, ou seja, $\alpha_i(\tau) \equiv \alpha_i$, hipótese adotada por Koenker (2004).

Assim, para inferir o modelo para diversos quantis, simultaneamente, consideramos os estimadores que resolvem o seguinte problema penalizado (FE) proposto por Koenker (2004):

$$\min_{(\alpha, \beta)} \sum_{k=1}^q \sum_{t=1}^T \sum_{i=1}^n w_k \rho_{\tau_k} (y_{it} - \alpha_i - x'_{it} \beta(\tau_k)) + \lambda \sum_{i=1}^n |\alpha_i| \quad (3)$$

em que $\rho_{\tau}(u) = u(1 - I(u < 0))$ é uma função ponderação linear piecewise, u o resíduo e $I()$ uma função indicador, conforme Koenker e Basset (1978). O termo λ é um parâmetro que penaliza o efeito fixo α na função-objetivo, caso este se distancie de um valor comum (no nosso caso seria n zero, ou seja, conforme a equação acima seria $\lambda \sum_{i=1}^n |\alpha_i - 0|$). Esta classe de estimadores penalizados, por meio do λ , melhora a performance não apenas do efeito fixo, mas também melhora a performance da estimativa do β em termos de variabilidade (KOENKER, 2004, 2005). A escolha deste parâmetro λ , segundo Koenker (2004) e Lamarche (2010), permanece ainda como um problema em aberto nesta literatura. Os pesos w_k ponderam o efeito relativo de cada um dos $q \{\tau_1, \dots, \tau_q\}$ quantis na estimação dos parâmetros. Adota-se pesos constantes iguais a $1/q$ (LAMARCHE, 2010). Para obter os erros padrão decide-se utilizar uma técnica de reamostragem, baseada em um *bootstrap* para dados em painel. Assim, o presente trabalho baseia-se na estratégia de Lamarche (2010), que utiliza um bootstrap de painel descrito em Cameron e Trivedi (2005, p. 708), mas aplicado para regressões quantílicas.

A seguir, nas tabelas 3 e 8, seguem o resultado da estimação para proficiência de língua portuguesa e matemática, respectivamente. Para ambos os modelos foi rejeitada a hipótese nula de efeitos aleatórios, com valores Chi-quadrado, respectivamente, de 248,8349 e 191,1492. Utilizar-se-á o modelo de regressão quantílica para dados em painel nos quantis .10, .25, .50, .75 e .90, com lambda igual 5.

Tabela 3: Resultados da Estimação por Efeito Fixo (EF) para proficiência em Língua Portuguesa

Variável	OLS	EA	EF	RQ 0.1	RQ 0.25	RQ 0.50	RQ 0.75	RQ 0.9
LOCALIZACAO	5.725301***	6.853517**	5.202207	6.07293*	7.07378**	5.88504*	3.39290	3.76012
DISTORCAO	0.338471*	0.480620***	0.553224***	0.36158***	0.37632***	0.35511***	0.35088***	0.23733***
MATRICULADOS	-0.003718***	-0.007954***	-0.015548***	-0.00557**	-0.00173	-0.00070	-0.00060	-0.00234
APROV	-0.264599***	-0.057771	0.023785	-0.28118***	-0.23111***	-0.32824***	-0.27466***	-0.29412***
ABN	-1.210543***	-0.757524***	-0.516477***	-1.00276***	-1.02118***	-1.14673***	-1.20233***	-1.30999***
DOCENTES	0.170083***	0.263986***	0.336742***	0.28750***	0.18294***	0.05830	0.05835	0.02315
DOC_SUP	-0.038179	-0.099515***	-0.141651***	-0.04972	-0.09795***	-0.07345**	-0.07186*	-0.02838
POP/AREA	0.000746***	0.000374**	-0.035204***	0.00078***	0.00078***	0.00060***	0.00055***	0.00090***
DENGUE	-0.003143***	-0.002962***	-0.005296***	-0.00435***	-0.00441***	-0.00345***	-0.00208***	-0.00199*
INFRA	1.099404***	1.051940***	0.839164***	1.59973***	0.79813**	1.20601***	1.07860***	1.18406***
TRANSPORTE	-0.000439	0.001284**	0.010230***					
TMI	-0.017709	-0.060890*	-0.106494***					
CONSTANTE	266.4764***	245.2846***	317.3320***	240.85617 ***	255.66527 ***	273.19561***	281.43434 ***	291.92516***
R-quadrado	0.2001	0.1550	0.5970					
Crítério de Schwarz	8.2257		8.6942					
Durbin-Watson	0.9397	1.3669	1.6596					

(*), (**) e (***) representam significância estatística aos níveis de 1%, 5% e 10% respectivamente.

A tabela 3 apresenta os resultados quando a variável dependente representa a nota de proficiência em Língua Portuguesa. Observa-se que o modelo de OLS é mais parcimonioso do que o modelo de EF, devido à não inclusão de variáveis *dummies* para controlar a heterogeneidade não observada no modelo de EF, esse resultado pode ser observado pelo critério de Schwarz. Entretanto, o modelo de EF apresenta um maior poder de explicação pela estatística R² e não apresenta sinais de correlação serial dos resíduos pela estatística Durbin-Watson. Vale ressaltar que a estatística do teste de Hausman 248,8349 rejeita a hipótese nula de efeitos aleatórios.

Nos resultados para EF, o intercepto foi estatisticamente significativo e positivo. Aponta que independente das variáveis explicativas do modelo, haverá uma nota média para as escolas captada pelo intercepto, note que o intercepto tem valor crescente nos quantis.

A distorção, mensurada como o percentual de alunos fora da idade certa no município na qual a escola encontra-se inserida, apresenta sinal positivo e estatisticamente significativo. Intuitivamente, o sinal esperado para esta variável seria negativo, logo, o resultado contrariou as expectativas iniciais. Para entender esse resultado, vale o exame da tabela 4. Note que a medida que aumenta as classes de distorção, em média a nota de português reduz, o que é esperado. Entretanto, da classe [20, 30) para [30, 40) há um leve aumento da nota, e como a maioria das observações encontram-se nesses estratos, contaminou o resultado do parâmetro. Esse resultado é replicado para todos os quantis na análise de regressão quantílica.

Tabela 4: Média de Nota em Língua Portuguesa estratificada por classes Distorção

Média de Distorção	Média da Nota de Língua Portuguesa.	obs
[0, 10)	256.9110	22
[10, 20)	252.5093	786
[20, 30)	251.4645	1124
[30, 40)	251.6358	1133
[40, 50)	249.4749	22
Todos	251.8181	3087

Fonte: Elaborada pelos autores.

A localização não apresenta significância estatística para o modelo de FE, algo que pode ser justificado pela baixa variabilidade dessa informação, haja vista que mais de 90% das escolas, nessa amostra de 441 escolas, estão na zona Urbana. Entretanto, no recorte quantílico, pode-se observar que o fato de estar na zona urbana, aumenta a nota das escolas nos quantis 0.1, 0.25 e 0.5, ou seja, para as 50% piores escolas da amostra. Quanto aos resultados em língua portuguesa, o efeito zona urbana é positivo.

A variável *matriculados* representa a quantidade de alunos total matriculados na escola, essa variável geralmente é utilizada na literatura como *proxy* do efeito tamanho da escola. Para o modelo de FE, esta variável apresenta sinal negativo e estatisticamente significativo, apontando que quanto maior a escola, menor a nota. Isso possivelmente deve-se ao efeito do envolvimento dos docentes, servidores e diretoria com a vida acadêmica do aluno, pois quanto maior escola, mais difícil o acompanhamento individualizado. Quando a análise é feita para os quantis, esse resultado só pode ser verificado para as 10% piores escolas.

A variável *aprovados* é definida como o percentual de alunos aprovados no terceiro ano do ensino médio. A variável possui impacto insignificante sobre a nota de língua portuguesa para o modelo de FE. Quando a análise é feita para os quantis, observa-se um impacto negativo e significativo para todos os cortes de quantis. Como o resultado mais um vez mostrou-se inesperado, torna-se interessante apresentar uma tabela descritiva da variável, reportada pela tabela 5.

Tabela 5: Média de Nota em Língua Portuguesa estratificada por classes Aprovados

Média de Aprovados	Média da Nota de Língua Portuguesa.	obs
[20, 40)	256.3311	6
[40, 60)	245.0255	179
[60, 80)	248.7039	1194
[80, 100)	254.7142	1707
[100, 120)	215.2885	1
Todos	251.8181	3087

Fonte: Elaborada pelos autores.

Note que as 6 (seis) escolas com menores índices de aprovação possuem uma nota média superior às demais. A exigência pode refletir positivamente nas notas de exames de

avaliação, já as escolas que relaxam nas exigências aos seus alunos, para produzir bons índices, acabam penalizadas nos exames de avaliação.

Já o abandono escolar, que independe da vontade da escola, apresenta-se com sinal negativo e significativo, dentro do esperado. Quanto maior o percentual de abandono no terceiro ano do ensino médio, menor a nota da escola. A literatura⁴ sobre as causas do abandono e evasão escolar é bastante vasta, apresentando diversos motivos, como efeito *background* familiar, efeito professor, efeito escola, violência etc.

O efeito escola representa aquelas que não possuem uma gestão de acompanhamento adequado que visam uma motivação do aluno, aquelas escolas que perdem mais alunos são as mesmas que acabam por produzir os piores resultados em língua portuguesa. O resultado é verificado para o modelo de FE e todos os recortes de quantis de notas. Note também uma magnitude do impacto crescente nos quantis, ou seja, as melhores escolas sentem mais o efeito do abandono do que as piores escolas.

Quanto ao impacto dos docentes, observa-se um efeito positivo e significativo, ou seja, quanto maior a quantidade de docentes, maiores serão as notas. Esse resultado é verificado para o modelo de FE e corroborado apenas para os quantis 0.1 e 0.25, ou seja, a presença maior dos professores na nota dos alunos é significativa para as 25% piores escolas.

A variável *doc_superior* representa o percentual de docentes com nível superior na escola, variável que apresenta sinal negativo e significativo para o modelo de FE. Esse resultado parece contra-intuitivo, entretanto, analisando a estatística descritiva, observa-se uma média de 91,67% de docentes com nível superior entre as 3087 observações, com desvio padrão de 9,74%, ou melhor, das 441 escolas em análise, 392 possuem um percentual acima de 80% de docentes com nível superior e 155 escolas com 100% de docentes com nível superior. Esse percentual elevado gera uma homogeneidade entre os professores, o que pode causar um efeito acomodação. Análises como a de Louzano *et al.* (2010), por exemplo, justificam a baixa atratividade da profissão e seu efeito acomodação dada à falta de incentivos na carreira. De toda forma, tal resultado pode ser associado ao encontrado em Barros *et al.* (2001) discutido anteriormente, onde também encontra-se o resultado surpreendente de relação inversa entre qualidade do professor e rendimento escolar.

Outro dado importante é que as escolas com percentuais de docentes com nível superior entre 40% e 60%, possuem a maior média de nota do grupo, 254,15, contra uma média de 250,27 para aquelas que possuem 100% de docentes com nível superior, como pode ser observado na tabela 6. Isso indica que as escolas mais heterogêneas podem gerar uma competição sadia entre os professores e isso justificaria o sinal negativo. Observando

4. Para uma revisão dessa literatura, sugere-se, por exemplo, Castelar, Monteiro e Lavor (2012).

as regressões quantílicas, pode-se concluir que esse “efeito competição” é significativo para as escolas dos quantis centrais, não tendo efeito nem para as 10% piores e nem as 10% melhores. Ainda assim, certamente é uma constatação preocupante que carece de maior investigação.

Tabela 6: Média da Nota de Língua portuguesa por classes de estratificação de docentes com nível superior (DOC_SUP)

DOC_SUP	Média da Nota de Língua Portuguesa	Obs.
0% a 19,9%	208.1195	1
20* a 39,9%	250.9836	3
40% a 59,9%	254.1466	27
60% a 79,9%	252.8914	311
80% a 99,9%	252.6233	1657
100%	250.2696	1088
Todos	251.8181	3087

Fonte: Elaborada pelos autores

A variável densidade demográfica, dada pelo tamanho da população dividido pela área do município na qual a escola está inserida, possui sinal negativo e estatisticamente significativo no modelo de FE. Quanto maior for a densidade, menores serão as notas. Pode-se apontar pontos positivos e negativos da densidade demográfica sobre o rendimento escolar: como ponto positivo, pode-se elencar o acesso a serviços públicos e privados em geral, como acesso à internet móvel, bibliotecas públicas, museus etc., além de que os melhores professores tendem a se instalar em centros urbanos. Já os pontos negativos da densidade demográfica são os problemas urbanos relacionados à criminalidade.

Note na tabela 7 que a grande maioria dos municípios possuem baixa densidade populacional e há uma queda da nota de português da classe [0, 2000) para a classe [6000, 8000) e um aumento de nota da classe [6000, 8000) para a classe [8000, 10000). Entretanto, pelo volume de observações, o impacto negativo se impõe. Comparando o modelo de efeito fixo com a análise quantílica, pode-se observar o seguinte: Quando a heterogeneidade não observada é modelada por efeitos fixos, o impacto da densidade demográfica é negativo. Quando o modelo é estimado por regressões quantílicas, dada a hipótese de independência do efeito fixo nos quantis, o impacto da densidade é positivo para todos os quantis.

Tabela 7: Média da Nota de Língua portuguesa por classes de estratificação de Densidade Populacional (Pop/Area)

POP/AREA	Média da Nota de Língua Portuguesa	Obs.
[0, 2000)	251.8526	2261
[6000, 8000)	250.9017	590
[8000, 10000)	253.7783	236
todos	251.8181	3087

Fonte: Elaborada pelos autores.

A quantidade de casos de dengue ocorridos e registrados no município na qual a escola está inserida, representada pela variável *dengue*, possui sinal negativo e significativo, ou seja, quanto mais casos de dengue ocorrem no município, afeta negativamente a nota de língua portuguesa. Obviamente que a dengue, como qualquer doença, impede o aluno de estudar e conseqüentemente diminui o seu rendimento. Esse resultado é corroborado pela análise em todos os quantis, com impacto absoluto maior sobre as piores escolas.

A taxa de mortalidade infantil (TMI), apesar de não atingir diretamente os alunos do ensino médio, representa uma proxy da qualidade da saúde do município na qual a escola está inserida. Estimada apenas para o modelo de FE⁵, apresenta de forma significativa que quanto maior a mortalidade infantil, pior a saúde em geral do município e, conseqüentemente, impacta negativamente na nota.

A variável Transporte indica a quantidade total de alunos, do município na qual a escola está inserida, transportados da zona rural para a zona urbana. Também estimado apenas por FE, indica que quanto mais alunos transportados maior a nota, representado pelo parâmetro positivo e estatisticamente significativo da variável.

A variável Infra varia de 0 a 10, indica a existência de 10 itens de infraestrutura, onde nota 10 indica que a escola possui todos os itens. Para o modelo de Efeito Fixo, a infraestrutura apresenta sinal positivo e significativo, indicando que quanto melhor a infraestrutura da escola, melhor os resultados em língua portuguesa. A variável também apresenta parâmetro significativo e positivo em todos os quantis em análise, com destaque para o impacto absoluto maior sobre as 10% piores escolas.

Tabela 8: Resultados da Estimação por Efeito Fixo (EF) para proficiência em Matemática

Variável	OLS	EA	EF	RQ 0.1	RQ 0.25	RQ 0.50	RQ 0.75	RQ 0.9
LOCALIZACAO	1.837397	3.151684	1.963702	3.64659	4.41196	3.43920	-0.55537	4.87216
DISTORCAO	-0.033645	0.102227*	0.121314*	0.13678*	0.08023	0.06177	-0.02427	-0.22211*
MATRICULADOS	-0.002521*	-0.005070***	-0.009551***	0.00240	0.00159	0.00265	-0.00047	-0.00115
APROV	-0.231562***	0.014226	0.067943	-0.18170**	-0.24359***	-0.20164***	-0.31492***	-0.29003***
ABN	-1.192148***	-0.525919***	-0.301738***	-0.74535***	-0.89426***	-0.93570***	-1.19665***	-1.27918***
DOCENTES	0.122750***	0.166347***	0.197150***	0.02250	0.04034	-0.01303	0.02613	-0.03647
DOC_SUP	0.043989	-0.008452	-0.035733	-0.04499	-0.02420	-0.04238	0.02088	-0.01106
POP/AREA	0.000432***	0.000119	-0.025112***	0.00037*	0.00041**	0.00042***	0.00049***	0.00043*
DENGUE	-0.000789	-0.001094**	-0.002923***	-0.00188**	-0.00161***	-0.00112**	-0.00154**	0.00006
INFRA	0.438669*	0.113050	-0.107309	1.22531***	0.55586*	0.62232**	0.32479	0.62916*
TRANSPORTE	-0.001579***	0.000571	0.007634***					
TMI	0.019961	-0.022803	-0.057987*					
CONSTANTE	280.4471***	256.4791***	307.7475***	250.04936***	267.69765***	276.34157***	301.79528***	312.32319***
R-quadrado	0.2001	0.1550	0.5970					
Crítério de Schwarz	8.2257		8.6942					
Durbin-Watson	0.9397	1.3669	1.6596					

Fonte: Elaborada pelos autores.

(*), (**) e (***) representam significância estatística aos níveis de 1%, 5% e 10% respectivamente.

5 O motivo da estimação discutida ser apenas a de Efeitos Fixos nos casos das variáveis *Taxa de Mortalidade Infantil* e *Transporte* é de que nos demais modelos não houve convergência.

Comparando o modelo que explica a nota de matemática com o modelo que explica a nota de língua portuguesa, há apenas uma divergência de interpretação dos resultados na análise de Efeitos Fixos (EF). Enquanto a variável INFRA mostrou-se positiva e significativa para explicar a nota de língua portuguesa, a mesma variável não apresenta significância para explicar a nota de matemática das escolas. Vale ressaltar que o crescimento da nota de português entre os anos de 2008 a 2014 foi bem mais acentuado do que o crescimento da nota de matemática (tabela 2), indicando que políticas públicas podem ter sido de fato mais eficazes para os exames de português. Entretanto, esta mesma variável, quando se analisa o recorte quantílico, apresenta significância com sinal positivo para as 50% piores escolas.

Ainda na esteira da análise por quantis, a distorção para a idade certa, que possui significância para todos os quantis no modelo de língua portuguesa, possui significância apenas para o primeiro e o último quantil no modelo de matemática. O efeito tamanho da escola, dado pela variável matriculados, que representa o número de alunos matriculados na escola, não possui significância na análise com recorte quantílico em todos os quantis. Mesmo resultado encontrado para a variável localização, docentes e docentes com nível superior. A variável dengue não apresenta significância para o último quantil 0.9.

4 CONCLUSÃO

Este trabalho investigou a influência de características das escolas e do município na qual a escola está inserida sobre a nota de proficiência em língua portuguesa e matemática, extraídas do SPAECE, para o 3º ano do ensino médio de 441 escolas do Estado do Ceará, entre os anos de 2008 e 2014. O painel balanceado consiste de 3.087 observações e foram estimados modelo de Efeito Fixo (FE) e Efeito Aleatório, e também modelos com regressões quantílicas, segundo proposto por Koenker (2004).

A análise de FE indica que os determinantes da nota em proficiência em Língua Portuguesa são: número de docentes, número de alunos transportados e infraestrutura, com impacto positivo sobre a nota e; a quantidade de alunos da escola, o percentual de docentes com nível superior, números de casos de dengue, taxa de mortalidade infantil e densidade demográfica, com impacto negativo sobre a nota.

Os resultados para proficiência em matemática divergem da nota de português em apenas uma variável. A infraestrutura foi positiva e significativa para explicar a nota em língua portuguesa e estatisticamente insignificante no modelo para a nota de matemática. Um resultado surpreendente foi o sinal negativo da variável docentes com nível superior. Ainda, observa-se que escolas com uma maior heterogeneidade de professores conseguem

melhores resultados nos exames.

As principais conclusões do modelo de regressão quantílica para dados em painel apontam que o sinal negativo e significativo da variável distorção só foi encontrado nas 10% melhores escolas no exame de proficiência de matemática. Estas mesmas escolas também foram as únicas que não sofreram o impacto negativo e significativo da ocorrência de dengue. O número de docentes só foi positivo e significativo naquelas 25% piores escolas no exame de língua portuguesa.

Um resultado marcante encontrado em todos os quantis e para os dois modelos, foi o sinal negativo e significativo das variáveis aprovados e abandono. O sinal negativo da variável aprovados, como discutido ao longo do texto, pode ser resultado do afrouxamento da regras e exigências aos alunos na busca da escola por bons indicadores.

Em termos de políticas públicas, é interessante atentar, então, para o aparente impacto positivo que a infraestrutura causa no desempenho escolar. A evidência indica uma correlação entre a nota em língua portuguesa e a presença de equipamentos apropriados para a frequência e aprendizagem do discente, logo, o mais intuitivo seria incentivar o incremento de tal infraestrutura na escolas do país.

Cabe, numa posterior investigação, aprofundar a análise sobre a ausência de aparente correlação da mesma infraestrutura com o desempenho em matemática, e quais fatores outros poderiam contribuir para uma melhora nos indicadores desta disciplina. Da mesma forma, o resultado inesperado da relação negativa da educação docente e o desempenho precisa ser compreendido plenamente, onde possivelmente dinâmicas locais das escolas não esteja sendo captadas pelo modelo.

Ainda, num trabalho posterior, a inclusão de uma variável ligada à condição socioeconômica, como a renda *per capita*, pode trazer conclusões mais contundentes a respeito dos determinantes de rendimento escolar.

REFERÊNCIAS

- ABREVAYA, J., DAHL, C.M. **The effects of birth inputs on birthweight: evidence from quantile estimation on panel data.** Journal of Business & Economic Statistics. Vol. 26, Issue 4. 2008
- ALBERNAZ, A.; FERREIRA, F. H. G.; FRANCO, C. Qualidade e equidade no ensino fundamental brasileiro. **Pesquisa e planejamento econômico**, Rio de Janeiro, v. 32, n. 3, p. 453-476, dez. 2002.
- ALMEIDA, A.T.C. **Determinantes dos Piores e Melhores Resultados Educacionais dos Alunos da Rede Pública de Ensino Fundamental no Brasil.** Revista Planejamento e Políticas Públicas, PPP, IPEA, 2014.
- ALVES, M. T. G.; SOARES, J. F. **Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos alunos.** Educação em revista, Belo Horizonte, n. 45, p. 25-58, jun. 2007.
- BACHE, S. H.; DAHL, C.M., ; KRISTENSEN, J.T., **Headlights on tobacco road to low birthweight—Evidence from a battery of quantile regression estimators and a heterogeneous panel.** *Empirical Economics*, Volume 44, issue 3, pp 1593-1633. 2013.
- BARROS, R. P., MENDONÇA, R., SANTOS, D.D., QUINTAES, G. **Determinantes do desempenho educacional no Brasil.** Pesquisa e planejamento econômico, Rio de Janeiro, v. 31, n. 1, p. 1-42, abr. 2001.
- BRASIL. Resultados SAEB 2015: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2016
- BRASIL. Ministério da Educação. **PDE: plano de desenvolvimento da educação: Prova Brasil: ensino fundamental: matrizes de referência, tópicos e descritores.** Brasília: MEC, SEB, INEP, 2008.
- BRASIL. Ministério da Educação. **Prova Brasil: avaliação do rendimento escolar, ensino-fundamental, matrizes de referências, tópicos e descritores.** Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2007.
- CAMERON, A.C., TRIVEDI, P.K. *Microeconometrics: Methods and Applications.* Cambridge. 2005.
- CARNEIRO, Diego; IRFFI, Guilherme. **Avaliação comparativa das Leis de Incentivo à Educação no Ceará.** Anais da ANPEC, 2017

CASTELAR, P.U.C., MONTEIRO, V.B., LAVOR, D., **Um estudo sobre as causas de abandono escolar nas escolas públicas de ensino médio no Estado do Ceará.** IPECE. Economia do Ceará em Debate. 2012

CASTRO, M. H. G. de. **Educação para o século XXI:** o desafio da qualidade e da equidade. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 1999.

CASTRO, C. M., FLETCHER, P. R. **A escola que os brasileiros freqüentaram em 1985.** Brasília: Ipea, 1986.

DELMANTO, D. *et al.* **Prova Brasil na escola:** material para professores, coordenadores-pedagógicos e diretores de escolas de ensino fundamental. São Paulo: Centro de Estudos e Pesquisas em Educação, Cultura e Ação Comunitária, 2007.

FELÍCIO, F.; FERNANDES, R. **O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de São Paulo.** In: ENCONTRO NACIONAL DE ECONOMIA DA ANPEC, 33., 2005, Natal, Rio Grande do Norte. Anais... Natal: ANPEC, 2005.

FERNANDES, R. **Índice de desenvolvimento da educação básica.** Brasília: Instituto Nacional de Estudos e Pesquisas Anísio Teixeira, 2007. 26 p. (Série documental. Textos para discussão; 26).

JACOBSEN, Cristina C.; MORI, Nerli N. R. **PROVA BRASIL E DESEMPENHO EM LÍNGUA PORTUGUESA: Um estudo com escolas paranaenses.** Revista HISTEDBR On-line, Campinas, n.38, p. 80-91, jun.2010.

KUENZER, A. Z. **Trabalho pedagógico: da fragmentação à unitariedade possível.** In: AGUIAR, S. A. M.; FERREIRA, C. S. N. (Org.). **Para onde vão a orientação e a supervisão educacional?** Campinas: Papirus, 2002. p. 47-78.

KOENKER, R.; BASSET, G. **Regression quantiles.** Econometrica, v. 46, n. 1, p. 33-49, 1978.

KOENKER, R. **Quantile regression for longitudinal data.** Journal of Multivariate Analysis, v. 91, p. 74-89, 2004.

_____. **Quantile regression.** New York: Cambridge University Press, 2005.

LAMARCHE, C. **Robust Penalized Quantile Regression Estimation for Panel Data.** Journal of Econometrics, 157, 396-408. 2010

LAVOR, Daniel Campos. **Avaliação de impacto e formulação de modelo para política educacional.** Tese de Doutorado CAEN/UFC, 2012.

LEE, V.E., FRANCO, C., ALBERNAZ, A. **Quality and equality in Brazilian secondary schools: a multilevel cross-national school effects study**. Annual Meeting of the American Educational Research Association, San Diego. 2004.

LOUZANO, P. ROCHA, V., MORICONI, G.M., OLIVEIRA, R.P. **Quem Quer Ser Professor? Atratividade, Seleção e Formação Docente no Brasil**. Est. Aval. Educ., São Paulo, v. 21, n. 47, p. 543-568, set./dez. 2010

NERI, Marcelo C.; MOURA, Rodrigo L. **Efetividade do “salário mínimo estadual”: uma análise via regressões quantílicas para dados longitudinais**. Econ. Apl. vol.12 no.2 Ribeirão Preto. 2008

OLIVEIRA, D. A. **Política educacional nos anos 1990. Educação básica e empregabilidade**. In: DOURADO, F.; PARO, V. H. (Org.) **Políticas públicas e educação básica**. São Paulo: Xamã, 2001

SANTOS, M.S., CLARO, D.B., OLIVEIRA FILHO, A.M.S., PAOLI, A.R., BARBOSA, C.D.B., ALMEIDA, E.V.G., CARVALHO, G.D.A, SANTANA, H.J., BULCÃO, L.A., MAROTTA, L.L., SILVA, M.P.U., NOVAES. M.P. **Análise das Infraestruturas do Censo Escolar 2011: Uma Proposta da Disciplina de Tópicos em Banco de Dados**. Anais do SBIE 2014. 2014.

SÁTYRO, N.; SOARES, S. **A infraestrutura das escolas brasileiras de ensino fundamental: um estudo com base nos censos escolares de 1997 a 2005**. Brasília: Ipea, 2007. (Texto para Discussão, n. 1.267).

SOARES, S. **O Valor de Mercado da Educação Pública**. IPEA, Texto para Discussão. 2017.

STOCO, S. **SAEB: uma análise da política**. 160 f. Dissertação (Mestrado em Educação)– Universidade Estadual de Campinas, São Paulo, 2006.

ANEXOS:

Estatísticas descritivas das escolas, por tipo de avaliação e ano.

Obs: Com intuito de balancear o painel, foi extraída uma amostra de 441 escolas dos dados apresentados abaixo.

PORTUGUÊS											MATEMÁTICA											
2008											2008											
VALOR											VALOR											
1º quartil	243,2										1º quartil	254,6										
Mediana	234,0										Mediana	245,8										
3º quartil	225,3	VI Mín.	189,6	VI Máx.	293,2						3º quartil	238	VI Mín.	211,7	VI Máx.	347,5						
	QTDE.	LP	MT	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL		QTDE.	MT	LP	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL	
GRUPO 1	125	252,4	263,3	5,3	22,9	1,4	78,4	8,5	13,1	25,1	GRUPO 1	125	265,8	250,0	5,4	23,0	1,3	78,3	8,7	13,0	25,4	
GRUPO 2	125	238,6	249,5	5,2	23,9	1,5	76,7	7,9	15,4	27,1	GRUPO 2	125	249,9	237,3	5,2	23,2	1,3	75,6	8,1	16,3	27,1	
GRUPO 3	126	229,8	242,9	5,1	24,7	1,4	75,0	7,9	17,2	25,6	GRUPO 3	126	242,0	231,2	5,1	24,8	1,6	75,1	7,6	17,3	25,6	
GRUPO 4	126	217,6	233,4	5,0	24,7	2,1	73,4	7,3	19,3	25,4	GRUPO 4	126	231,4	219,8	5,0	25,2	2,2	74,4	7,1	18,5	25,1	
2009											2009											
VALOR											VALOR											
1º quartil	257,3										1º quartil	265,8										
Mediana	249,9										Mediana	258,1										
3º quartil	241,8	VI Mín.	199	VI Máx.	313,5						3º quartil	250,0	VI Mín.	217,5	VI Máx.	342,5						
	QTDE.	LP	MT	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL		QTDE.	MT	LP	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL	
GRUPO 1	120	266,2	275,1	8,6	27,0	1,4	80,0	7,7	12,3	26,4	GRUPO 1	120	277,0	264,5	8,6	26,6	1,4	80,1	7,7	12,2	27,0	
GRUPO 2	120	253,4	261,0	8,5	28,7	1,4	77,0	7,3	15,7	26,7	GRUPO 2	120	261,5	252,7	8,6	27,7	1,5	77,5	7,4	15,1	25,8	
GRUPO 3	121	246,4	254,6	8,4	27,1	1,6	76,6	7,1	16,3	26,1	GRUPO 3	121	254,0	247,2	8,4	28,3	1,6	75,6	7,6	16,8	26,2	
GRUPO 4	121	234,1	244,2	8,3	28,3	2,4	75,9	6,7	17,4	25,4	GRUPO 4	121	242,3	235,6	8,1	28,4	2,5	76,1	6,1	17,8	25,6	
2010											2010											
VALOR											VALOR											
1º quartil	266,9										1º quartil	265,1										
Mediana	259,0										Mediana	256,8										
3º quartil	250,1	VI Mín.	220,1	VI Máx.	327,9						3º quartil	247,9	VI Mín.	221,4	VI Máx.	366,8						
	QTDE.	LP	MT	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL		QTDE.	MT	LP	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL	
GRUPO 1	127	278,0	277,2	8,7	28,1	0,3	84,8	7,2	8,0	23,4	GRUPO 1	127	276,1	279,9	8,7	27,1	0,4	85,7	6,9	7,4	23,8	
GRUPO 2	127	262,5	260,4	8,7	28,6	0,3	79,2	7,9	12,9	25,8	GRUPO 2	127	262,1	260,7	8,6	28,6	0,3	80,0	7,4	12,6	25,3	
GRUPO 3	128	254,8	253,1	8,4	29,2	0,4	78,4	7,6	14,0	24,0	GRUPO 3	128	255,3	253,2	8,4	28,6	0,4	78,7	8,0	13,3	24,8	
GRUPO 4	128	242,8	244,3	8,0	28,6	0,3	78,6	6,7	14,6	23,0	GRUPO 4	128	244,7	241,2	8,1	30,1	0,3	76,4	7,2	16,4	22,4	
2011											2011											
VALOR											VALOR											
1º quartil	267,5										1º quartil	273,2										
Mediana	257,7										Mediana	260,4										
3º quartil	249,9	VI Mín.	185,4	VI Máx.	326,5						3º quartil	251,4	VI Mín.	207	VI Máx.	359,9						
	QTDE.	LP	MT	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL		QTDE.	MT	LP	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL	
GRUPO 1	135	282,4	290,6	8,7	25,5	0,3	86,4	6,1	7,4	22,4	GRUPO 1	135	281,1	293,7	8,7	24,9	0,4	87,3	6,1	6,7	22,2	
GRUPO 2	135	262,3	265,4	8,6	26,3	0,4	79,7	7,3	13,0	23,4	GRUPO 2	135	261,2	265,4	8,6	25,1	0,4	81,6	6,4	12,0	24,2	
GRUPO 3	136	254,2	257,3	8,4	27,8	0,4	78,0	6,6	15,5	24,0	GRUPO 3	136	254,6	256,0	8,3	28,3	0,3	76,9	7,4	15,7	23,7	
GRUPO 4	136	241,0	245,0	8,3	27,2	0,3	75,5	7,5	17,0	21,9	GRUPO 4	136	242,9	243,3	8,4	28,6	0,3	74,1	7,6	18,3	21,7	
2012											2012											
VALOR											VALOR											
1º quartil	259,4										1º quartil	268,5										
Mediana	248,8										Mediana	256,3										
3º quartil	240,7	VI Mín.	167,8	VI Máx.	321,9						3º quartil	247,0	VI Mín.	214,8	VI Máx.	360,4						
	QTDE.	LP	MT	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL		QTDE.	MT	LP	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL	
GRUPO 1	141	278,8	290,8	6,1	24,6	0,3	89,8	5,2	5,0	19,9	GRUPO 1	141	277,0	292,9	6,0	23,7	0,4	90,0	5,2	4,8	20,2	
GRUPO 2	141	253,3	260,7	5,8	24,1	0,4	80,9	7,4	11,6	22,6	GRUPO 2	141	252,1	261,6	5,9	23,8	0,4	82,3	6,9	10,8	22,2	
GRUPO 3	142	244,9	252,6	5,7	25,8	0,4	78,6	7,9	13,5	22,3	GRUPO 3	142	245,2	252,0	5,7	25,4	0,4	79,2	7,6	13,2	22,0	
GRUPO 4	142	231,6	242,5	5,3	25,7	0,3	78,8	6,2	14,9	20,7	GRUPO 4	142	234,3	240,0	5,4	27,2	0,3	76,5	7,1	16,3	21,0	
2013											2013											
VALOR											VALOR											
1º quartil	267,7										1º quartil	279,2										
Mediana	254,4										Mediana	262,8										
3º quartil	244,8	VI Mín.	211,4	VI Máx.	333,6						3º quartil	251,7	VI Mín.	226,1	VI Máx.	393,8						
	QTDE.	LP	MT	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL		QTDE.	MT	LP	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL	
GRUPO 1	142	286,5	303,9	8,8	23,1	0,4	92,6	4,1	3,3	19,0	GRUPO 1	142	285,5	305,7	8,8	22,8	0,4	92,7	4,3	3,0	19,0	
GRUPO 2	143	260,9	269,2	8,5	22,8	0,4	82,4	8,3	9,4	20,9	GRUPO 2	143	259,2	270,5	8,4	22,8	0,4	82,5	7,9	9,5	20,4	
GRUPO 3	144	249,8	258,5	8,5	23,6	0,4	82,1	7,0	10,9	20,2	GRUPO 3	144	249,5	257,5	8,4	23,1	0,4	82,5	7,0	10,4	20,7	
GRUPO 4	144	236,2	245,6	8,3	25,0	0,3	79,3	6,9	13,8	18,4	GRUPO 4	144	239,1	243,4	8,3	25,8	0,3	78,4	7,2	14,4	18,3	
2014											2014											
VALOR											VALOR											
1º quartil	274,2										1º quartil	277,3										
Mediana	260,5										Mediana	261,1										
3º quartil	250,0	VI Mín.	189,2	VI Máx.	340,2						3º quartil	249,5	VI Mín.	213,6	VI Máx.	385,4						
	QTDE.	LP	MT	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL		QTDE.	MT	LP	INFRA	DIST.	% TR.	%AP	%RP	%AB	DOC/AL	
GRUPO 1	150	293,2	301,1	8,6	21,8	0,4	94,5	2,1	1,0	18,7	GRUPO 1	150	291,2	302,8	8,5	21,3	0,4	96,6	1,7	1,0	18,9	
GRUPO 2	150	266,3	266,4	8,1	22,7	0,4	82,3	7,8	9,0	19,2	GRUPO 2	150	265,7	268,0	8,1	22,4	0,4	81,0	8,0	7,2	19,2	
GRUPO 3	150	255,3	255,9	8,2	23,1	0,4	79,4	9,2	12,5	18,2	GRUPO 3	150	255,4	255,3	8,2	23,1	0,3	81,8	9,5	13,0	18,4	
GRUPO 4	150	241,3	244,7	8,0	23,9	0,3	81,2	8,8	13,9	17,8	GRUPO 4	150	243,7	242,0	8,0	24,6	0,3	79,2	8,9	15,6	17,5	

UMA ANÁLISE ESPACIAL DA POBREZA MULTIDIMENSIONAL NOS BAIRROS DE FORTALEZA - CEARÁ

Ricardo Candéa Sá Barreto*

Eli Izidro dos Santos**

Ícaro Célio Santos de Carvalho***

RESUMO

Este estudo objetivou caracterizar a pobreza nos bairros de Fortaleza – Ceará no período de 2000 e 2010, por meio da utilização de ferramentas de análise espacial, instrumentos que permitem uma visualização das regiões afetadas pela pobreza mais próxima da realidade, o que pode colaborar com a implementação de políticas públicas de alívio desse fenômeno de forma eficaz. Destaca-se que este trabalho se apropria de um enfoque multidimensional da pobreza, ou seja, considera outras variáveis que igualmente privam os cidadãos, além da renda, para realização do cálculo do Índice de Pobreza (IP). Assim, organizar os bairros em *rankings* de pobreza demonstrou que os índices são consistentes para o estudo espacializado da pobreza em Fortaleza. Logo, foram encontrados concentração de alta pobreza em diversas áreas do município, bem como de baixa pobreza, principalmente nas áreas centrais de Fortaleza. Além disso, utilizou-se um modelo econométrico espacial que apontou os determinantes da condição de pobreza, cuja influência é positiva: o índice de desigualdade de Theil, a mortalidade infantil e o nível de analfabetismo de pessoas com 25 anos ou mais.

Palavras-Chave: *Cluster*; Distribuição Espacial; Pobreza.

ABSTRACT

This study aimed to characterize poverty in the neighborhoods of Fortaleza - Ceará in the period of 2000 and 2010, through the use of spatial analysis tools, tools that allow a visualization of regions affected by poverty closer to reality, which may contribute to the Implementation of public policies to alleviate this phenomenon effectively. It is worth

* Doutor em Economia pela Universidade Federal de Viçosa- MG. (85) 99819-80769. ricardocandea@yahoo.com.br.

** Mestre em Economia Regional e Políticas Públicas pela Universidade Estadual de Santa Cruz – UESC – Ilhéus – Bahia. elyizidro@hotmail.com.

*** Doutorando em Administração de Empresas. icarocelio@hotmail.com.

noting that this study appropriates a multidimensional approach to poverty, that is, it considers other variables that equally deprive citizens, in addition to income, for calculating the Poverty Index. Thus, to organize neighborhoods in poverty rankings has shown that the indices are consistent for the spatialized study of poverty in Fortaleza. Therefore, a high concentration of poverty was found in several areas of the city, as well as low poverty, mainly in the central areas of Fortaleza. In addition, a spatial econometric model was used that pointed out the determinants of the poverty condition, whose influence is positive: Theil's inequality index, Infant mortality, and the illiteracy rate of people aged 25 years or over.

Keywords: Cluster; Spatial distribution; Poverty.

1 INTRODUÇÃO

Segundo o Programa das Nações Unidas para o Desenvolvimento - PNUD (2013), Fortaleza possui uma área de 320,12 km² e densidade demográfica de 7.645,29 hab/km². Em 2000, o seu IDHM era de 0,652; em 2010, passou para 0,745, esse aumento demonstra um importante crescimento de 15,64% desse índice. Destacam-se três principais dimensões como as motivadoras desse crescimento: longevidade, renda e educação. Por outro lado, a renda *per capita* média do município cresceu 85,18%, passando de R\$ 610,4, em 2000, para R\$ 846,36, em 2010. Ressalta-se ainda uma redução na desigualdade de renda, medida através do índice de Gini, que passou de 0,64, em 2000, para 0,61, em 2010. Ainda assim, segundo o relatório da PNUD, Fortaleza é um dos cinco municípios mais desiguais do mundo. Para o Instituto de Pesquisa e Estratégia Econômica do Ceará - IPECE (2012), esses aspectos se configuram e se refletem, especialmente nos bairros, na falta de serviços públicos de qualidade como: saúde e educação, no desemprego, na falta de programas de geração de renda, entre outros aspectos que, geralmente, têm potencializado o surgimento de conflitos sociais, manifestados, principalmente, através do aumento da violência.

Nesse universo, a pobreza contribui para o afastamento do indivíduo do convívio social, acentuando, sobretudo, a ausência de outros aspectos além da renda, que igualmente privam as pessoas de suas necessidades básicas. Dessa forma, a pobreza não pode mais ser analisada apenas pela ótica unidimensional da renda. Faz-se necessária a ampliação dos estudos, a fim de entender as necessidades básicas dos indivíduos, como nutrição, saúde, educação, moradia, saneamento básico, entre outros, ou seja, atribuir aos estudos da pobreza um caráter multidimensional (LACERDA; NEDER, 2010).

Entretanto, assumir que a pobreza não se limita à renda não é suficiente, é preciso uma análise mais abrangente para a obtenção de resultados mais explicativos e pertinentes.

É importante conhecer de forma clara o objeto de análise e o comportamento desse, caso contrário, tornam-se extremamente limitadas as tentativas de elaborar e avaliar as políticas públicas voltadas para o alívio da pobreza e, principalmente, de entender a realidade de forma ampla e consistente o suficiente para intervir de forma positiva e duradoura.

Outro fator de destaque, nesse contexto, é a espacialização da pobreza, pois, segundo Ávila (2013), é fundamental conhecer e entender as relações que as múltiplas dimensões da pobreza possuem no espaço, para entre outros aspectos colaborar com a elaboração de políticas públicas mais consistentes. Nesse sentido, a identificação de *clusters* de pobreza e suas dimensões pode permitir o entendimento do melhor tipo de política a ser proposta, com foco no núcleo do problema e, dessa forma, possibilitar uma aplicação de recursos públicos de forma mais eficiente e eficaz.

Portanto, a identificação de *clusters* de pobreza nos bairros de Fortaleza se deu por meio de um Índice de Pobreza (IP), que foi calculado a partir dos dados dos Censos Demográficos de 2000 e 2010 e do Atlas do Desenvolvimento Humano 2013 para os 116 bairros de Fortaleza, cuja análise foi desenvolvida através da utilização de métodos de análise espacial. Nesse contexto, este trabalho tem como objetivo geral desenvolver uma análise espacial da pobreza multidimensional dos bairros de Fortaleza no período de 2000 e 2010. E, mais especificamente, busca-se: (a) mensurar a pobreza nos bairros de Fortaleza para o período de 2000 e 2010, com base no Índice de Pobreza; (b) analisar espacialmente a pobreza em Fortaleza; (c) identificar os bairros do município com maior e menor concentração de pobreza; e, por fim, (d) identificar quais variáveis foram determinantes para a pobreza nos bairros do município.

Logo, a relevância deste trabalho está em chamar a atenção para a urgência de ações que promovam mudanças localizadas, uma vez que o município de Fortaleza, assim como a região Nordeste, apresenta elevados índices de pobreza. Destarte, estudos espaciais sobre o tema auxiliam na focalização de políticas públicas, por exemplo.

Metodologicamente, o objetivo da proposta foi a mensuração do Índice Pobreza (IP) dos bairros de Fortaleza para o ano de 2000 e 2010, utilizando-o, posteriormente, dentro do modelo econométrico espacial *cross section*, como variável representativa da pobreza dos bairros no ano de 2010. A construção dessa variável permitiu verificar os diferenciais no comportamento do IP nos anos de 2000 e 2010 e, por fim, permitiu analisar as desigualdades socioeconômicas entre os bairros, através de uma perspectiva multidimensional.

Para além desta primeira seção introdutória, este artigo encontra-se dividido em mais

quatro seções: na segunda seção, encontra-se a revisão da literatura, que apresenta o referencial teórico utilizado para embasar esta análise acerca da pobreza; na terceira seção, é apresentada a metodologia utilizada; na quarta seção, foram feitas as discussões dos resultados encontrados, bem como a espacialização da pobreza nos bairros de Fortaleza; e, na quinta seção, são destacadas as considerações finais do trabalho, além das referências e anexos I, II e III.

2 REVISÃO DE LITERATURA

Entender como a pobreza se comportou com o passar dos anos nos permite compreender como sua concepção foi se alterando ao longo do tempo. Nesse sentido, a trajetória histórica explica como o contexto social necessitou de abordagens que melhor explicassem a pobreza, assim, é possível perceber a necessidade de expandir os estudos de uma visão unidimensional para estudos que compreendam outras variáveis que melhor expliquem essa problemática. Algumas concepções acerca da pobreza foram desenvolvidas desde o século passado, porém há uma complexa conceituação do termo, devido ao seu caráter subjetivo. Para além disso, o estudo pode ser concentrado de duas maneiras: a primeira sob a ótica econômica, na qual é utilizado como base somente a renda; e a segunda por uma perspectiva que incorpora outras variáveis não econômicas, que igualmente privam o cidadão.

Segundo Crespo e Gurovitz (2002), no último século, foram desenvolvidas três concepções gerais acerca da pobreza, (1) sobrevivência; (2) necessidades básicas e (3) privação relativa. No primeiro caso, o enfoque era mais restritivo e predominou nos séculos XIX e XX, teve origem no trabalho de nutricionistas ingleses, o qual apontava que a renda dos mais pobres não era suficiente para a manutenção do rendimento físico do indivíduo. O segundo caso, sua expansão se deu principalmente partir de 1970, foram incorporadas novas variáveis, como serviços de água potável, saneamento básico, saúde, educação e cultura, para o estudo da pobreza. O terceiro e último caso ganhou destaque a partir de 1980, dando ao conceito um enfoque mais abrangente e rigoroso, buscando formulação científica e comparações entre estudos internacionais, principalmente, enfatizando o aspecto social. Essa ideia foi fortalecida por Amartya Sen, principal teórico dessa nova concepção da pobreza.

Nesta lógica, para Kageyama e Hoffmann (2006), a noção de pobreza refere-se a um tipo de privação, que pode ser apenas material ou incluir elementos de ordem cultural e social, em face dos recursos disponíveis para uma pessoa ou família. A distinção entre a natureza da privação pode ser compreendida por meio das discussões conceituais de

Feres e Villatoro (2013), que aponta a pobreza absoluta como a existência de um núcleo irreduzível, ou seja, um núcleo onde os indivíduos estão numa situação abaixo do nível de rendimento mínimo para sobreviver e que os impossibilita de suprir as suas necessidades básicas.

Já a pobreza relativa contrapõe as concepções absolutas e propõe utilizar uma perspectiva que se refere às condições reais de privação, principalmente, na comparação com outros indivíduos da sociedade. Segundo Townsend (1962), *many people have been uneasily aware of the problems of defining necessities like housing, clothing, or fuel and light*. Ou seja, os indivíduos estão em situação de pobreza quando não têm os recursos para atividades diárias em sua sociedade, estão excluídos do estilo de vida socialmente desejável.

Entretanto, de acordo com Sen (2000), a pobreza tem um caráter tanto absoluto como relativo, que pode variar o tempo todo dentro do convívio social, pois é fortemente influenciado pela posse de bens comuns desse convívio, que, por seu turno, impacta diretamente sobre a percepção de bem-estar, a qual neste sentido pode alternar de acordo com cada sociedade.

Contudo, os estudos acerca da pobreza foram recentemente vistos sob uma ótica diferente da que foi aplicada no século passado. Os estudos realizados nas obras de Amartya Sen (2000) sobre o caráter dinâmico da pobreza revelam um novo horizonte de pesquisas, em que outras variáveis além da renda foram incorporadas às análises da pobreza, os denominados estudos das privações das capacidades. Segundo Sen (2000), as liberdades estreitam a noção de pobreza sob a ótica de renda e ampliam e dinamizam os novos estudos. Nesse contexto, o conceito multidimensional da pobreza é definido como uma ideia antiga com novos arranjos, o que caracteriza o amplo aspecto do termo, em que são envolvidas as dimensões econômicas, sociais e estruturais (POGGI, 2004; CONCONI; HAM, 2007).

Em consonância com Sen (2000), a abordagem das capacitações traz uma definição ligada a uma série de necessidades básicas as quais os indivíduos ficam privados de realizar, como: estar bem nutrido, vestido e abrigado adequadamente, livre de doenças, ser respeitado, entre outras funcionalidades. Nesse contexto, as “capacidades” são as combinações alternativas de funcionamentos de possível realização. Portanto, a capacidade é um tipo de liberdade: a liberdade substantiva de realizar combinações alternativas de funcionamentos ou a liberdade para ter estilos de vida diversos. Por exemplo, uma pessoa abastada que faz jejum por sua livre e espontânea vontade pode ter a mesma realização de funcionamento que uma pessoa pobre forçada a passar fome extrema. Porém, a primeira pessoa possui um “conjunto capacitário” diferente do da segunda. A primeira pode escolher comer bem e ser

bem nutrida, o que seria impossível para a segunda (CRESPO; GUROVITZ 2002; SEN, 2000).

Outro ponto que merece destaque nesse contexto está ligado às linhas de pobreza. De acordo com o IBGE (2011), as discussões sobre os indicadores de pobreza no Brasil ainda precisam de aprofundamento, pois são muito incipientes. O Governo Federal, por exemplo, utiliza-se de vários recortes para a implementação dos programas sociais, como é o caso da política de transferência de renda “Bolsa Família”, que considera pobres as pessoas que auferem renda mensal de até $\frac{1}{4}$ do salário mínimo. Contudo, existem outros indicadores como a Pesquisa dos Orçamentos Familiares – POF, que analisa o consumo, por considerar este menos volátil que a renda e representar o gasto real das famílias em alimentos e outros bens. Porém, essas são definições baseadas apenas na renda das pessoas.

Nesta lógica, cada pesquisador acaba definindo um caminho de acordo com a metodologia utilizada, os dados disponíveis e os objetivos do estudo. Logo, são muitas conceituações, no entanto, não há consenso entre os autores, principalmente, sobre as linhas de pobreza a serem utilizadas nas pesquisas. E, quando se trata de estudos multidimensionais, essa tarefa se torna ainda mais difícil, pois a escolha das dimensões que serão objeto do estudo e quais variáveis serão utilizadas em muito dependem do objetivo do trabalho e do conceito de pobreza utilizado pelo pesquisador (AMARAL; CAMPOS; LIMA, 2016).

Para Lacerda (2009), a grande dificuldade é encontrar um bom indicador, capaz de incorporar o lado multidimensional da pobreza. A autora destaca que, diferente do que acontece com os estudos sobre o prisma unidimensional da renda, não existe, ainda, nas abordagens multidimensionais, um conjunto de medidas estabelecidas e consolidadas. Esse fato ilustra bem a importância da utilização do Índice de Pobreza, utilizado neste trabalho, que engloba no seu arcabouço a renda, a educação, a saúde e a habitação. De acordo com Ávila (2013), apesar de utilizar dados quantitativos, esse índice foca a qualidade de vida dos indivíduos, não restringindo apenas a quantificação monetária da pobreza. O IP é inspirado no Índice de Pobreza Humano (IPH)¹, que foi incorporado ao Relatório de Desenvolvimento Humano da PNUD, a partir de 1997, com o objetivo específico de medir a pobreza, utilizando as mesmas variáveis do IDH, entretanto, com foco nos mais pobres e adotando

1. O IPH foi construído a partir de três indicadores: a) percentual de indivíduos suscetíveis de morrer antes dos 40 anos de idade; b) quantidade de adultos analfabetos; e c) serviços que são ofertados pelo conjunto da economia. Esta última é formada por três variáveis: percentual de pessoas que não têm acesso à água potável, percentual de indivíduos que não têm acesso aos serviços de saúde e percentual de crianças abaixo dos cinco anos de idade que sofrem de desnutrição. Esse índice foi, em 1998, dividido em dois: o IPH-1, para países em desenvolvimento, e o IPH-2, para países desenvolvidos, de forma a refletir melhor as diferentes condições existentes entre os dois grupos de países (ÁVILA, 2013; LACERDA, 2009; PNUD; 1997).

uma perspectiva das privações do indivíduo. Tal índice tem o intuito de medir o tamanho do déficit, utilizando as mesmas dimensões fundamentais consideradas pelo IDH. Neste sentido, ele é mais uma medida de vulnerabilidade que propriamente de pobreza (ÁVILA, 2013; LACERDA, 2009).

Apesar das várias limitações desse tipo de metodologia, ela se mostra analiticamente aceitável, pois, além de medir a pobreza, busca entendê-la, considerando dimensões mais ligadas à qualidade de vida das pessoas. Nestes termos, a interpretação desses índices possibilita a proposição e elaboração de políticas públicas capazes de atender, de fato, às necessidades dos indivíduos e revelam-se bem mais eficazes que os métodos que utilizam apenas a dimensão renda (ÁVILA, 2013; LACERDA, 2009).

3 METODOLOGIA

3.1. Índice de Pobreza – IP

A metodologia deste trabalho foi composta por técnicas de distribuição espacial e baseada no trabalho desenvolvido por Ávila (2013), cuja construção do Índice de Pobreza é estruturada a partir da metodologia de elaboração do Índice de Pobreza Humana – IPH, criado por Anand e Sen (1997). Também faz parte desse arcabouço a Análise Exploratória de Dados Espaciais – AEDE, o Índice de Moran Global e Local, instrumentos os quais possibilitaram espacializar a pobreza e a desigualdade nos bairros de Fortaleza, capital do Ceará, sob uma ótica multidimensional. Também faz parte deste arcabouço a utilização de um modelo econométrico regressivo, com o objetivo de identificar os determinantes da pobreza em Fortaleza, tendo o IP como variável dependente.

Neste sentido, após a coleta dos dados, foi realizado o cálculo do IP, a partir das dimensões escolhidas, dispostas no quadro 1. Inicialmente, elaborou-se uma tabela com os indicadores de privações para cada um dos bairros, para que, na sequência, pudessem ser realizados os cálculos dos referidos índices. Sequencialmente, foi organizado o *ranking* dos bairros, para cada um dos índices calculados. Assim, obteve-se uma análise comparativa dos *rankings*, buscando verificar a consistência dos índices como forma de justificar sua utilização na análise espacial. Para verificar a consistência dos índices, utilizou-se como base o *ranking* do IDHM para os bairros de Fortaleza de 2000 e 2010, divulgados através do Atlas do Desenvolvimento Humano do Brasil/2013 (PNUD, IPEA, FJP, 2013).

Quadro 1 – Dimensões e variáveis que compõem o IP

DIMENSÕES (D)	PRIVAÇÕES (P)
Habitação e saneamento (HS)	5 ou mais moradores por domicílio (IBGE) Sem paredes adequadas (PNUD, IPEA e FJP) Sem coleta de lixo (IBGE) Sem energia elétrica (PNUD, IPEA e FJP) Sem tratamento de esgoto e água potável (PNUD, IPEA e FJP)
Educação (E)	Sem instrução e Analfabetismo com 25 anos ou mais (IBGE)
Saúde (S)	Taxa de Mortalidade Infantil (PNUD, IPEA e FJP)
Renda (R)	Até 1/4 do salário mínimo (SM) ou Sem rendimento (IBGE)

Fonte: Adaptado de Ávila (2013).

Em que: D_i = dimensão a ser calculada; P_{ij} = privação que compõe a variável derivada; i = número que indica a dimensão a ser calculada (); j = número da privação que compõe a dimensão a ser calculada (); e n = quantidade de privações que compõem a dimensão.

Desta forma, aplicando-se a média ponderada às dimensões (D_i) e, reescrevendo-as, tem-se: $HS=D_1, E=D_2, S=D_3, C=R_4$. Logo, o Índice de Pobreza assume a seguinte expressão:

$$IP = \{(1/n) * [D_1^a + D_2^a + D_3^a + D_4^a]\}^{1/a} \quad (2)$$

Ou seja: $D = D_i; i = 1, \dots, n$

Desta forma, na fórmula geral, tem-se:

$$IP = \left[\left(\sum D_i^\alpha \right) / n \right]^{1/\alpha} \quad (3)$$

Em que: n = quantidade de dimensões que compõem o índice; e α = fator de ponderação do peso das dimensões que compõem o índice.

Tem-se, desta forma, os três índices utilizados neste trabalho, que são assim apresentados:

$$IP_1 = \{(1/n) * [HS^a + E^a + S^a + R^a]\}^{1/a} \quad \alpha = n = 4 \quad (4)$$

$$IP_2 = \{(1/n) * [HS^a + E^a + S^a]\}^{1/a} \quad \alpha = n = 3 \quad (5)$$

$$IP_3 = \{(1/n) * [R^a]\}^{1/a} \quad \alpha = n = 1 \quad (6)$$

E finalmente tem-se:

$$D_i = 1/n \left(\sum P_{ij} \right) \quad n= 5 \text{ para HS e } n= 1 \text{ para E, S e R} \quad (7)$$

Assim, a partir do cálculo dos Índices de Pobreza (IP)², foi realizada a análise comparativa dos *rankings* dos bairros, buscando verificar a consistência dos índices calculados em comparação ao IDHM dos bairros de Fortaleza de 2000 e 2010 divulgados pela PNUD (2013), como forma de justificar sua utilização na análise espacial.

Para implementação da análise espacial, definiu-se, primeiramente, a matriz de pesos espaciais (W), que, conforme Almeida *et al.* (2008), é a forma de expressar um determinado arranjo espacial das interações resultantes do fenômeno a ser estudado, como primeiro passo para implementação da Análise Exploratória de Dados Espaciais (AEDE). No entanto, tendo em vista a existência de autocorrelação espacial, foi aplicado à estatística I de Moran Global³, pois, segundo Almeida (2012), é a maneira mais aceitável de identificá-la e testá-la. Contudo, quando se lida com um grande número de dados, há sempre a ocorrência de dependência espacial, assim, foi utilizada a estatística I de Moran Local, que permite a identificação de *clusters* espaciais, em que a comparação é feita não entre bairros, mas entre os indicadores locais e seus vizinhos, verificando, desta forma, se há ou não padrões de concentrações locais.

Segundo Almeida *et al.* (2008), o objetivo central desse método é descrever a distribuição espacial, os padrões de associação espacial, os possíveis *clusters* espaciais, verificar a existência de diferentes regimes espaciais ou outras formas de instabilidade espacial (não estacionariedade) e identificar observações espaciais atípicas, ou seja, os *outliers*. Os autores salientam que para a implementação da AEDE, faz-se necessário definir uma matriz de pesos espaciais (W). Destacam, ainda, que a escolha da matriz de pesos espaciais é muito importante em uma AEDE, pois os resultados da análise são sensíveis a tal seleção. Logo, diante da ideia contida na matriz de vizinhança, há existência de uma maior interação espacial entre os vizinhos mais próximos do que com os mais distantes. Ávila (2013) afirma que o resultado dessa interação esperada é que o índice de pobreza de determinado bairro influencie e seja influenciado pelos índices dos bairros com os quais fazem fronteira e que essa influência vá diminuindo, na medida em que a distância interbairros aumente.

Desta forma, inicialmente, definiram-se a matriz e o nível de contiguidade, para

2. Este é um índice cujo valor varia de 0 a 100. Logo, o valor final do IP indica qual a proporção de pobres no bairro. Dessa forma, quanto mais próximo de 100 for o resultado, maior é o grau de pobreza. Nesse sentido, os bairros que obtiverem índice menor que 15% são considerados de baixa pobreza, os que alcançarem índice de 15% a 29,99% são classificados como de média pobreza, os que galgarem índice de 30% a 49,99% são tidos como de alta pobreza e aqueles que obtiverem índice de 50% ou mais serão classificados com grau de pobreza muito alto. Salienta-se que esse foi um arranjo praticado para este trabalho, na metodologia original, devido à menor escala dos bairros, em relação aos municípios ou estados, por exemplo (ÁVILA, 2013).

3. Consiste em uma medida de autocorrelação espacial que aponta para a existência ou não de agrupamentos espaciais para uma dada variável, isto é, para a presença de índices de pobreza com valores similares entre vizinhos, segundo um determinado indicador de interesse (ALMEIDA *et al.*, 2008).

posteriormente proceder a análise a partir da elaboração dos mapas. Assim, realizou-se o teste de autocorrelação espacial ou I de Moran, que indicou que o uso dos K vizinhos mais próximos seria o mais indicado⁴, pois apresentou o maior nível de significância estatística para ambos os períodos e índices, o que está consonância com os preceitos metodológicos (ANSELIN, 1988; ALMEIDA, 2012).

A partir do cálculo dos Índices de Pobreza (IP), realizou-se a análise comparativa dos *rankings* dos bairros, buscando verificar a consistência dos índices calculados, como forma de justificar sua utilização na análise espacial. Para a verificação da consistência dos índices, utilizou-se o *ranking* do IDHM de 2000 e 2010 para bairros de Fortaleza, divulgado através do Atlas do Desenvolvimento Humano do Brasil (PNUD; IPEA; FJP, 2013).

3.2 Modelo econométrico espacial

A econometria espacial pertence a um campo da econometria tradicional, que leva em consideração a interação espacial nas proposições de possíveis soluções aos problemas econômicos nos quais a dimensão espacial esteja presente. Segundo Pimentel (2004), seria uma coletânea de ferramentas propostas para lidar com as particularidades advindas da influência do espaço sobre a análise estatística de dados espaciais. Por esse motivo, a econometria espacial tem suas raízes em estudos específicos das áreas de ciência regional, geografia econômica e economia urbana.

O modelo clássico de regressão linear nos permite estimar a relação entre uma variável dependente e um conjunto de variáveis explicativas. Assim, quando a variável dependente é contínua, um método que é naturalmente escolhido é o de Mínimos Quadrados Ordinários (MQO). Formalmente, esta relação é expressa como:

$$Y = \beta_0 + \beta x + \varepsilon \quad (8)$$

em que Y é a variável dependente, β_0 é a constante, β é o parâmetro a ser estimado, x é a variável independente e ε é o termo de erro.

Contudo, o modelo clássico de análise de regressão linear não incorpora os componentes espaciais e, muitas vezes, para a explicação de um fato ou comportamento de uma variável em um espaço específico, também, devem-se considerar as interações no interior de tal localidade. De acordo com Anselin (1992), o erro residual de uma regressão pode ser considerado englobando todos os elementos ignorados. Se qualquer um desses elementos mostrar um padrão espacial significativo, deveria ser refletido em um padrão espacial para

4. Para este estudo, adotou-se a configuração dos 10 K vizinhos mais próximos. Contudo, para verificar a robustez dos resultados, foram testadas as configurações com 15 e com 20 vizinhos.

os termos do erro. Assim, utilizaram-se métodos de regressão espacial, que, diferente do MQO, pretendem explorar ou demonstrar a presença de dependência entre as observações no espaço. Em outras palavras, o objetivo é controlar a possibilidade de uma região com altas taxas de pobreza estar influenciando regiões vizinhas, corroborando com o argumento da existência de dependência espacial.

Buscando identificar a melhor especificação do modelo econométrico, Florax, Folmer e Rey (2003) recomendam que seja seguido o seguinte roteiro:

- a) Estimar o modelo clássico de análise de regressão linear por meio de Mínimos Quadrados Ordinários (MQO);
- b) Testar a hipótese de ausência de autocorrelação espacial devido a uma defasagem ou a um erro, por meio das estatísticas Multiplicador de Lagrange (ML), ρ (defasagem espacial) e Multiplicador de Lagrange λ (ML) (erro espacial);
- c) Caso ambos os testes não sejam significantes, a utilização do modelo clássico é mais apropriada. Caso contrário, é necessário seguir o próximo passo;
- d) Se ambos forem significantes, estima-se o modelo apontado como o mais significativo, de acordo com as versões robustas destes testes, ou seja, Multiplicador de Lagrange Robusto (MLR) ρ (defasagem espacial) e Multiplicador de Lagrange Robusto (MLR) λ (erro espacial). Assim, caso $MLR\rho > MLR\lambda$, usa-se o modelo com defasagem espacial como o mais apropriado. Caso contrário, $MLR\rho < MLR\lambda$, usa-se o modelo de erro autorregressivo espacial como o mais apropriado.

Na prática, têm-se dois dos principais modelos aplicados em econometria espacial: defasagem espacial – SAR – e o modelo de erro autorregressivo espacial – SEM – (*Spatial Error Model*). No caso do SAR, ele pode ser expresso por (9).

$$y = Wy + X\beta + \varepsilon \quad (9)$$

Em que y é a variável dependente; ε é o coeficiente autorregressivo espacial; Wy é um vetor $n \times 1$ de defasagens espaciais para a variável dependente; X é a matriz das variáveis explicativas; ε é o termo de erro.

Modelos do tipo (9) indicam que “... uma mudança na variável explicativa numa região afetará não apenas a própria região pelo efeito direto, mas pode afetar o valor da variável dependente em todas as regiões, por meio dos efeitos indiretos” (ALMEIDA, 2012, p.157). Esses efeitos indiretos são interpretados como transbordamentos espaciais.

No caso dos modelos do tipo SEM, tem-se a seguinte especificação: $y = X\beta + \varepsilon$ (10)

Considerando que: $\varepsilon = \lambda W\varepsilon +$ (11)

em que λ é o parâmetro do erro espacial que acompanha a defasagem .

A dependência espacial em (10) se manifesta no termo de erro, destacando que os erros associados com qualquer observação são uma média dos erros nas regiões vizinhas, somados a um componente aleatório. O modelo (11) informa que os efeitos sobre a variável dependente não advêm apenas do choque (representado pelo termo erro) de uma região, mas do transbordamento de choques oriundos de outras regiões vizinhas.

É importante destacar que, no caso do modelo SAR e SEM, os testes focados do tipo Multiplicador de Lagrange conseguem definir qual seria o melhor modelo.

Portanto, verificando toda a base de dados disponível, optou-se, por fim, por estimar um modelo econométrico espacial, com dados *cross section*, utilizando o método da Máxima Verossimilhança, com a intenção de identificar os fatores determinantes da pobreza nos bairros de Fortaleza em 2010.

3.2 Fonte de dados

Para este trabalho, as regiões utilizadas foram os bairros da cidade de Fortaleza e as unidades foram: a quantidade de domicílios e o número de habitantes de cada bairro para composição das dimensões analisadas. Dessa forma, para as dimensões habitação/saneamento e renda, a unidade de análise foi o domicílio e, para as dimensões saúde e educação, a unidade foi o indivíduo. Foram utilizados, para este estudo, os dados do Censo de 2000 e 2010, para os 116⁵ bairros de Fortaleza – CE, obtidos através do Banco de Dados SIDRA do IBGE⁶; bem como o Atlas do Desenvolvimento Humano para o Brasil, elaborado pelo Programa das Nações Unidas para o Desenvolvimento (PNUD, IPEA, FJP, 2013).

Na composição do trabalho, foram utilizados os *softwares* ArcGis10, para integração das bases de dados com os arquivos *shapefile*, o GeoDa8.1, para elaboração dos mapas e análise espacial econométrica, e o Excel 12, para construção dos *rankings* de pobreza.

Para o modelo de regressão espacial, a variável dependente foi o IP1, calculado no estudo, e as variáveis explicativas⁷: renda *per capita*, mortalidade infantil e analfabetismo

5. Saliencia-se que para este estudo foram utilizados apenas 116 bairros, para adequar ao arquivo Shapefile, disponibilizado pela Prefeitura Municipal de Fortaleza, que só possuía essa quantidade de bairros. Entretanto, a Capital cearense é formada, atualmente, por 119 bairros.

6. As informações foram obtidas no banco de dados Sistema IBGE de Recuperação Automática (SIDRA) do IBGE.

7. As variáveis explicativas foram escolhidas com base nos trabalhos de Ávila (2013), Lacerda (2009), Sen (2010) e Crespo e Gurovitz (2002). Contudo, devido a problemas de multicolinearidade dos testes econométricos e à limitação do banco de dados, apenas essas foram mantidas como explicativas.

de pessoas com 25 anos ou mais, extraídas do Censo 2010. Já o índice de Theil foi obtido através do Atlas do Desenvolvimento Humano (PNUD, IPEA, FJP, 2013).

As variáveis explicativas utilizadas tiveram como referencial o trabalho de Giovannetti e Raiher (2015), trabalho empírico realizado para os municípios do Paraná no ano de 2010, com dados do Censo demográfico, em que a Taxa de mortalidade infantil teve uma relação direta proporcional com a pobreza. Já o PIB *per capita* e o percentual de pessoas alfabetizadas maiores de 15 anos apresentaram relação inversamente proporcional com a Pobreza. Os mesmos autores utilizaram o Índice de Gini como variável explicativa em relação à pobreza dos municípios do Paraná e a relação foi diretamente proporcional.

Apesar do foco deste artigo, com a aplicação do modelo de econometria espacial, ser verificar a relação dos determinares para o ano de 2010, bem como verificar quais variáveis são representativas para explicar a pobreza nos bairros de Fortaleza, ainda assim, aplicou-se o teste de causalidade de Granger com os dois pontos no tempo, 2000 e 2010 (ver Anexo III), embora esse teste só tenha sido possível elaborar com uma defasagem temporal. Considerando a limitação do banco de dados disponível, encontramos as seguintes relações de causalidade de Granger⁸ para justificar a escolha das variáveis, além da literatura empírica: o IP1 (pobreza) causa Theil (desigualdade); Theil (desigualdade) não causa IP1 (pobreza); o IP1(pobreza) não causa Renda *per capita*; Renda *per capita* causa IP1 (pobreza); o IP1(pobreza) não causa Mortalidade infantil; Mortalidade infantil não causa IP1(pobreza); o IP1(pobreza) causa Analfabetismo de 25 ou mais; Analfabetismo de 25 ou mais não causa IP1 (pobreza).

4 RESULTADOS E DISCUSSÕES

Este item tem o objetivo de apresentar e discutir os principais resultados deste estudo, obtidos através do cálculo dos índices supracitados e da utilização da Análise Exploratória de Dados Espaciais, bem como com a aplicação do modelo econométrico espacial.

4.1. Ranking dos bairros de Fortaleza em relação à pobreza e ao desenvolvimento

Para análise dos dados, foi utilizada uma amostra dos 05 bairros que se apresentaram em maior situação de pobreza e dos 05 em menor situação, de acordo com os índices de pobreza calculados para este trabalho, em comparação ao IDHM divulgado pela PNUD (2013). Dessa forma, foram criados *rankings* de pobreza (ver anexo I e II), com os bairros

8. O teste de causalidade de Granger procura determinar o sentido causal entre duas variáveis, estipulando que X "Granger-causa" Y se valores passados de X ajudam a prever o valor presente de Y. Para mais detalhes, ver Adams *et al.* (2003) e Devlin e Hansen (2001).

em maior situação de pobreza em primeiro lugar, para todos os índices utilizados.

Nessa lógica, o bairro de Fortaleza que apresentou índice de maior situação de pobreza em 2000 foi Parque Araxá, quando avaliando o IP1, alcançou índice de 40,53% e 42,21% para o IP2, em segundo lugar, está o bairro de Pedras com 35,99% e 37,69%. Em 2010, houve uma redução substancial do índice pobreza, porém, o mesmo bairro continuou em primeiro lugar com 26,83% e 27,86%, respectivamente no IP1 e IP2, porém, em segundo. encontra-se o bairro Curió com 22,60% e 22,30%. Quando se analisa o IP3, que considera apenas a renda como indicador, Curió se apresenta com maior índice de pobreza, com 20,57% em 2000. Em 2010, no entanto, o bairro Parque Presidente Vargas assume o primeiro lugar com 21,52%, na segunda colocação, têm-se o bairro de Siqueira com índice de 20,27%.

Confirmando os resultados encontrados, em comparação ao IDHM, o bairro Parque Araxá apresentou, também, o resultado de menor índice de desenvolvimento humano, com 0,426 em 2000. Entretanto, em 2010 melhorou sua condição, ficando com 0,580, na segunda colocação. Assim, o bairro de Fortaleza de menor índice de desenvolvimento humano, no período, é Curió com 0,567. Esses resultados demonstram que houve uma melhora substancial nos níveis de desenvolvimento humano dos bairros, de uma década para outra, provocadas por uma melhor distribuição da renda no país, mas, sobretudo, uma melhora nos indicadores sociais na região.

Essa comparação dos *rankings* demonstra a consistência dos índices de pobreza utilizados neste trabalho e sugerem que eles podem ser usados para estudar a pobreza em Fortaleza, bem como sua distribuição espacial. Entretanto, percebe-se que a similaridade maior foi entre os bairros que apresentaram os 05 resultados maiores para todos os índices, tanto em 2000 como em 2010. Além de deixar claro que os bairros menos desenvolvidos também são aqueles que apresentaram os maiores índices de pobreza multidimensional, assim como os menores índices são apresentados por aqueles mais desenvolvidos, fato que não é confirmado no IP3, pois os bairros que apresentaram maior pobreza por renda não foram os de menor desenvolvimento humano. Por outro lado, na outra ponta do *ranking*, essa premissa é verdadeira, os menores índices de pobreza por renda são apresentados pelos bairros de maior índice de desenvolvimento humano, ou seja, Meireles e Aldeota, em ambos os períodos em análise.

Quando se classificam os bairros por grau de pobreza em muito alto, alto, médio e baixo, percebe-se que os resultados dos índices corroboram com a análise dos *rankings*, há uma subestimação do número de bairros em situação de pobreza, quando se observa o

problema apenas sob o ponto de vista da renda (Tabela 1). Porém, quando se observam as outras dimensões e se considera como muito alto grau de pobreza, aqueles bairros onde o índice foi de 50% ou mais não são encontrados, não há nenhum bairro nessa condição nos dois períodos em análise. Entretanto, quando se classifica com graduação alta, de 30% a 49,99%, encontram-se 14 bairros, os com médio índice de pobreza, de 15% a 29,99%, percebem-se 83 bairros nessa condição e os com grau de pobreza menor que 15% são apenas 25 bairros. Contudo, quando se considera somente a renda, 99 bairros foram classificados como de baixo índice de pobreza, enquanto que 17 foram considerados de médio índice e nenhum bairro de Fortaleza foi classificado como de alto e muito alto índice de pobreza em 2000.

Tabela 1 – Quantidade de bairros de Fortaleza por grau de pobreza/2000 e 2010

Índice	Quantidade de Bairros/ 2000				Quantidade de Bairros/ 2010			
	Baixo	Médio	Alto	M. Alto	Baixo	Médio	Alto	M. Alto
IP 1	22	83	11	0	85	31	0	0
IP 2	25	77	14	0	89	27	0	0
IP 3	99	17	0	0	101	15	0	0

Fonte: Elaboração própria, com base nos dados do Censo 2000 e 2010.

Em relação a 2010 (Tabela 1), pode se verificar uma estrutura bem diferente dos resultados do período anterior, há uma redução substancial da pobreza multidimensional, nenhum bairro foi classificado como de alta pobreza, encontrou-se 31 bairros com média pobreza e 89 com baixa pobreza. Isso ratifica a consistência do índice utilizado, bem como os resultados das análises anteriores. Porém, quando se analisa somente a renda, percebe-se que a redução foi pequena, apenas 02 bairros mudaram de categoria, sendo 15 com média pobreza e 101 com baixa. Para esse período, foi perceptível uma redução das discrepâncias entre os dois métodos de avaliação da pobreza em Fortaleza. Também foi perceptível uma melhora nas condições de pobreza. Como já salientado, esse fato pode estar ligado a uma evolução nos serviços públicos oferecidos à população. Entretanto, não se pode negar que, também, tal fato foi provocado por um aumento da renda média dos brasileiros registrado no período, bem como das políticas de transferência direta de renda desenvolvidas pelo Governo Federal (ROCHA, 2005; HOFFMAN, 2006).

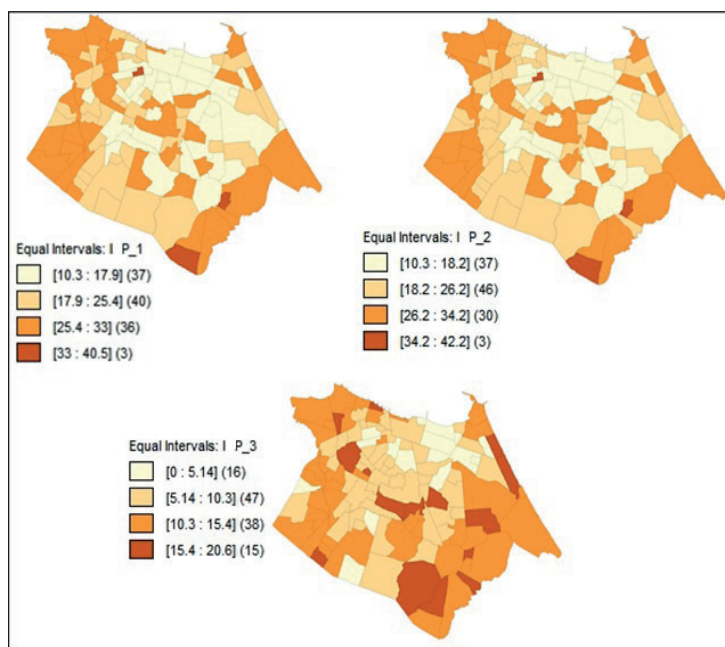
Contudo, cabe destacar que, mesmo com essa melhora na distribuição da renda em Fortaleza, como em várias outras regiões do Brasil, ainda é muito elevada a sua concentração, que somente quando foi espacializada pôde-se perceber sua aglomeração, geralmente, em pequenos grupos de localidades, como foi atestado pelo estudo do IPECE (2012).

4.2 Análise espacial da pobreza em Fortaleza

Este tipo de análise permite visualizar como a pobreza medida pelos índices aqui calculados está distribuída em Fortaleza, possibilitando, com isso, que se proceda mais comparações entre a pobreza pelo prisma multidimensional e aquela baseada apenas na renda, unidimensional.

Os mapas de distribuição da pobreza nos bairros de Fortaleza para o período de 2000 (Figura 1) demonstram a existência de regiões pobres e não pobres, onde o IP1 e IP2 apresentaram resultados muito semelhantes, enquanto que o IP3, quando o foco é a renda, apresentou uma configuração bem mais diferenciada.

Figura 1 – Mapas de distribuição da pobreza em Fortaleza/2000



Fonte: Elaboração própria com auxílio do software GeoDa8.1 e baseado em dados do Censo 2000 e PNUD, IPEA e FJP (2013).

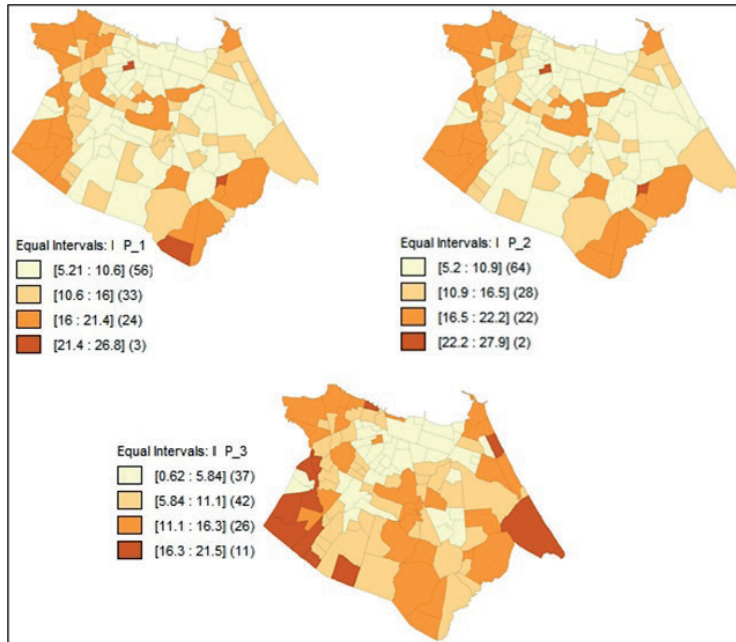
Os mapas ratificam as demonstrações as quais os *rankings* já haviam apresentado, porém com a diferença que aqui pôde-se constatar a espacialização da pobreza, além de apresentar uma grande similaridade⁹ entre os índices em análise. Salienta-se que, apesar de apresentar maior quantidade de bairros na condição de pobreza mais elevada no IP3, os valores correspondentes aos bairros são bem menores que nos índices multidimensionais em todos os intervalos, o que causa uma falsa impressão de que a pobreza é maior nesse

9. O software GeoDa12 utiliza os valores máximo e mínimo de cada *ranking* para determinar os intervalos de análise, por isso, os resultados são diferentes dos utilizados na graduação dos *rankings* anteriores, que variam, conforme já explicitado, de 0 a 100.

índice. Contudo, se observarmos mais atentamente os valores apresentados, percebe-se que no IP3 a pobreza por renda é relativamente menor que a multidimensional.

Em relação a 2010 (Figura 2), os mapas de distribuição espacial da pobreza demonstraram que as disparidades entre os índices de pobreza multidimensionais e o índice unidimensional aumentaram ainda mais em relação a 2000.

Figura 2 – Mapas de distribuição da pobreza em Fortaleza/2010



Fonte: Elaboração própria com auxílio do software GeoDa8.1 e baseado em dados do Censo 2000 e PNUD, IPEA e FJP (2013).

Observa-se nos mapas multidimensionais uma similaridade dos dois índices, ratificando a análise anterior dos *rankings*, que apontou uma redução da pobreza multidimensional nos bairros de Fortaleza. Há claramente uma ampliação dos intervalos dos bairros situados na melhor condição de pobreza. O mesmo acontece com o índice unidimensional, no intervalo dos bairros com melhor índice de pobreza, que apontou 79 bairros, o que demonstra uma melhora substancial dos índices de pobreza por renda nos bairros, em relação a 2000. Quando se consideram os intervalos em que se classificam os bairros em maior condição de pobreza, levando em conta os dois últimos intervalos, 27 e 26 bairros foram elencados nesses intervalos, considerando os índices IP1 e IP2, respectivamente. Esses dados apontam uma redução na quantidade de bairros nessa condição, em torno de 25%, em comparação a 2000, indicando uma melhora nos índices de pobreza multidimensionais. Entretanto, quando se analisa o mapa unidimensional, foram encontrados 37 bairros nessa

mesma condição, representando uma redução ainda mais acentuada, em torno de 30,9%, no mesmo período.

Contudo, as análises espaciais realizadas até o momento têm por base medidas globais e, de acordo com Almeida (2012), deve-se ter em mente que essas medidas podem mascarar a existência de associações locais. Assim, a busca por associações locais lineares, estatisticamente significantes¹⁰, pode ser efetuada com a utilização do Indicador Local de Associação Espacial (LISA).

Assim, a partir do Índice de Moran, foi possível medir a existência de autocorrelação espacial, em que o gráfico de dispersão aponta para a existência ou não de agrupamentos espaciais, para uma dada variável. Nessa lógica, quanto mais próximo de um, maior será o grau de autocorrelação. Esse índice permite verificar se os dados estão ou não correlacionados espacialmente. Os resultados dos gráficos de dispersão para cada índice calculado são apresentados na Tabela 2.

Tabela 2 – Índice de Moran do IP1, IP2 e IP3 para os Bairros de Fortaleza/ 2000 e 2010

	IP1	IP2	IP3
2000	0,199	0,181	0,184
2010	0,161	0,139	0,281

Fonte: Elaboração própria com auxílio do software GeoDa8.1 e baseado em dados do Censo 2000 e PNUD (2013).

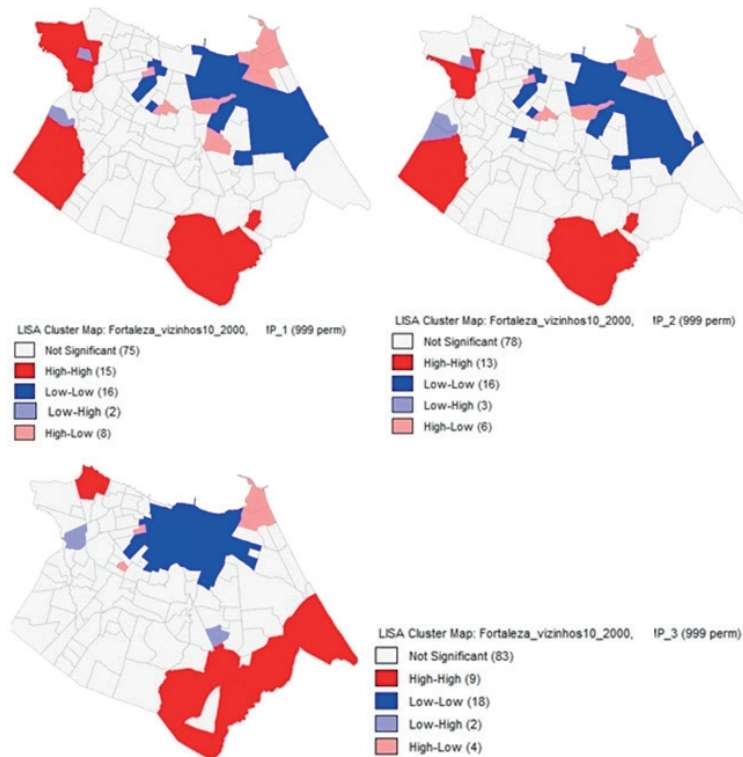
Nos índices multidimensionais, é possível perceber que os valores referentes a 2000 são maiores que os valores encontrados para 2010, no geral, os índices apresentados têm um valor mais próximo de zero, o que aponta para uma baixa autocorrelação espacial positiva da pobreza nos bairros de Fortaleza nesse período. Para o IP3, unidimensional, encontraram-se valores mais relevantes, que apontam uma correlação positiva mais acentuada no período de 2010. No geral, esses resultados sugerem a presença de regimes espaciais semelhantes, ou seja, regiões próximas espacialmente com valores parecidos.

Logo, a aplicação do Índice Local de Moran permite que se identifiquem os *clusters* (ou manchas de pobreza) a partir dos resultados do Índice de Moran Global, nesse caso, comparam-se os indicadores e seus vizinhos, verificando se há concentrações espaciais locais ou não, com resultados semelhantes. Assim, é possível visualizar o comportamento desses índices para os anos estudados através dos mapas de *clusters*, (Figura 3).

10. O software GeoDa8.1 já estipula em 999 permutações, ou seja, o estudo é significativo em 0,001%. Isto é, sua aplicação, obedecendo às mesmas características, tem 99,9% de chance de obter o mesmo resultado. Porém, para este estudo, foi considerada uma significância mínima de até 95%.

A partir dos resultados obtidos para a o período de 2000, é perceptível a presença de *clusters* de bairros classificados como *High-High* (Alto-Alto), que apresentaram altos índices de pobreza rodeados por bairros com alta pobreza, em três pontos distintos do município, no extremo sul, no oeste e norte. Por outro lado, encontraram-se bairros classificados como *Low-Low* (Baixo-Baixo), que apresentam baixos índices de pobreza, cercados por bairros nessa mesma condição, principalmente nas áreas central e leste do município, o que também pode ser denominado como um *cluster* de riqueza. Há, ainda, *clusters Low-High* (Baixo-Alto), indicando que esses *clusters* de riqueza estão rodeados de *clusters* de pobreza. E os *clusters High-Low* (Alto-Baixo) sinalizando alta pobreza cercada por bairros de baixa, em ambos os mapas dos índices multidimensionais (IP1 e IP2), cuja configuração é bem semelhante.

Figura 3 – Mapas de clusters de pobreza nos bairros de Fortaleza/2000



Fonte: Elaboração própria com auxílio do software GeoDa8.1 e baseado em dados do Censo 2000, PNUD, IPEA e FJP (2013).

Com relação ao IP3, esse apresentou um *cluster* com alta pobreza rodeado por alta pobreza no extremo sul de Fortaleza, além de um *outlier* no Norte (Barra do Ceará) que apresentou informação discrepante em relação aos outros bairros. No outro extremo, tem-se um grande *cluster* de baixa pobreza rodeado por bairros de alta pobreza, na área central

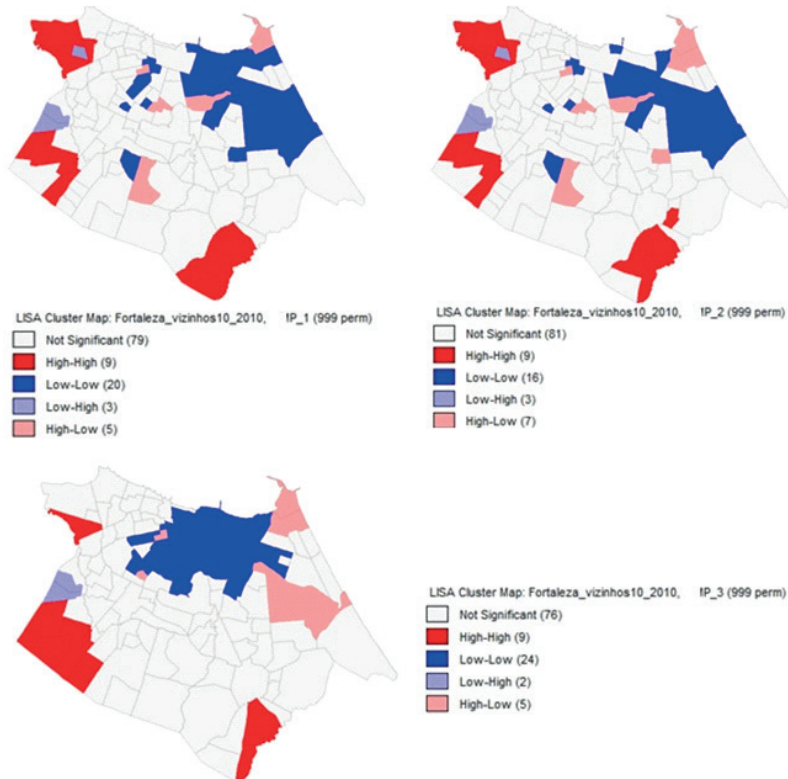
do município, onde estão localizados os bairros Meireles, Praia de Iracema, entre outros mais abastados, conforme o IPECE (2012).

Para 2010 (Figura 4), os resultados obtidos para a o IP1 demonstraram que os bairros que apresentaram *clusters* com alta pobreza rodeada por alta pobreza diminuíram de 15 para 09, por outro lado, nos *clusters* com dispersão negativa (baixa pobreza – rodeada por baixa pobreza) houve um aumento de cerca de 25% em relação a 2000, as regiões de baixa pobreza rodeada por alta pobreza também foram ampliadas, porém as regiões de alta pobreza rodeadas de baixa pobreza foram reduzidas.

Em relação ao IP2, observa-se uma redução dos *clusters* Baixo-Baixo e um aumento dos *clusters* Baixo-Alto e Alto-Baixo.

Acerca do IP3, pode ser observado que os bairros com altos índices de pobreza rodeados por também altos índices permaneceram iguais ao período anterior; já nos bairros de baixo índice de pobreza rodeados por regiões de também baixa pobreza, houve um crescimento em torno de 25% de um período para o outro; para os índices de baixa pobreza rodeados por altos índices de pobreza, não foram registradas alterações; e as regiões com alta pobreza rodeadas por baixa pobreza ampliaram-se de 04 para 05 regiões. Contudo, cabe destacar tanto a redução do *cluster* Alto-Alto, que se localizava no extremo sul do município, no período anterior, quanto o surgimento de outro *cluster* de mesma característica no oeste do município, mais precisamente na região que abrange o bairro de Granja Lisboa. Cabe, ainda, ressaltar a semelhança entre os *clusters* de alta e baixa pobreza nos três índices utilizados neste estudo, todavia o IP3 apresentou um número maior de bairros classificados como Baixo-Baixo. Porém, independente do índice utilizado para análise, é evidente a presença de uma grande concentração de bairros classificados como Alto-Alto na área central de Fortaleza.

Figura 4 - Mapas de clusters de pobreza nos bairros de Fortaleza/2010



Fonte: Elaboração própria com auxílio do software GeoDa8.1 e baseado em dados do Censo 2010 e PNUD, IPEA e FJP (2013).

Esses resultados reforçam as análises realizadas anteriormente com o índice global, ao mesmo tempo em que ratificam a constatação de que o estudo da pobreza, utilizando apenas o indicador renda é insuficiente para explicar e demonstrar os *clusters* de pobreza nos bairros de Fortaleza, necessitando a incorporação de variáveis multidimensionais aos estudos.

4.3. Análise econométrica espacial dos bairros de Fortaleza para dados de 2010

Como apresentado por Florax, Folmer e Rey (2003), foi utilizado primeiramente o método dos mínimos quadrados ordinários, a partir daí gerou-se a regressão apresentada na Tabela 3. Logo, com a utilização deste método, foram testadas todas as variáveis.

Na sequência, realizando uma análise do modelo, constatou-se que o mesmo se mostrou ajustado com o R^2 igual a 0,97. Todas as variáveis apresentaram-se estatisticamente significativas ao nível de 10% e com o sinal esperado compatível com a teoria.

Tabela 3 – Regressão do índice de pobreza contra seus determinantes nos bairros de Fortaleza e seus vizinhos para o ano de 2010 pelo método M.Q.O.

Var. explicativas	Coefficiente	estatística-t	Prob
C	0.604458	1.08637	0.27967
Renda <i>per capita</i>	-0.000190107	-1.64754	0.10228
Mortalidade infantil	0.603567	25.2717	0.00000
Analfabetismo de 25 anos ou mais	0.205497	6.83823	0.00000
Índice de Theil	1.58976	1.64931	0.10191
Diagnósticos da regressão			
Multicolinearidade		20.559490	
Teste de normalidade dos erros		Valor	Prob
<i>Jarke-Bera</i>		2, 24147	0, 3451296
Diagn. de heterocedasticidade			
<i>Teste Breusch-Pagan</i>		18.0256	0.00122
Diagnósticos de dependência espacial			
		Valor	Prob
Moran's I (error - λ)		4.2317	0.00002
Lagrange Multiplier (lag - ρ)		1.7084	0.19119
Robust LM (lag - ρ)		0.5664	0.45168
Lagrange Multiplier (error - λ)		1.3569544	0.00000
Robust LM (error - λ)		1.954789	0.00000
Lagrange Multiplier (SARMA)		1.365472	0.00000
Var. dependente	<i>IPI</i>	N. de observações	116
R	0.976069	G. de liberdade	111
R-ajustado	0.975207		
Crit. de Akaike	270.69	Crit. Schwarz	284.458

Fonte: Elaboração própria com auxílio do software GeoDa8.1 e baseado em dados do Censo 2010 e PNUD (2013).

Por fim, foi estimado um modelo econométrico espacial com dados *cross section*, visando identificar os fatores determinantes da pobreza nos bairros de Fortaleza em 2010. A variável “Renda *per capita*”, que representou a renda média do bairro, mostrou-se correlacionada com a teoria que sugere que em locais em que há maior concentração da pobreza,

há menor nível de renda *per capita*.

A variável “Mortalidade Infantil” representou precárias condições de saúde. Como observado, existe uma correlação positiva entre mortalidade infantil dos bairros e o índice de pobreza. Os bairros com maiores taxas de pobreza são os bairros em que a mortalidade infantil ocorre com mais frequência.

A variável Analfabetismo de 25 anos ou mais expressa o percentual de analfabetismo nos bairros para os maiores de 25 anos, conforme censo do IBGE de 2010, constata-se a relação positiva em que quanto maior o analfabetismo maior a concentração de pobreza nos bairros.

O índice de Theil (concentração de renda) também influencia positivamente o aumento da pobreza dos bairros de Fortaleza, em que um aumento na desigualdade de renda também gera acréscimos na pobreza multidimensional de 1,59%. Assim, o Índice de Theil comprovou que, quanto maior for a concentração de renda, piores são as condições de bem-estar da população. Giovanetti e Raiher (2015) encontraram resultados semelhantes utilizando o Índice de Gini para os municípios do estado do Paraná, que também tiveram uma relação positiva.

O diagnóstico da regressão de dependência espacial revelou que o Modelo de erro Espacial foi o mais indicado, tanto pelo multiplicador de Lagrange, quanto pelo Lagrange robusto. Os erros apareceram indicados como normais, o que viabilizou a utilização do método de Máxima verossimilhança (M.V.S) na geração do modelo. Há presença de homocedasticidade entre os dados, pois o valor do teste *Koenker-Bassett* foi não significativo. Já a multicolinearidade apresentou-se controlada correspondendo ao valor recomendado por Anselin (1992). No caso do teste de *Jarque-Bera*, tem-se como hipótese nula a normalidade, porém, se $p > 0,05$, aceita-se a normalidade.

A regressão com o modelo de erro Espacial para a variável pobreza multidimensional (IP1), conforme Tabela 4, permitiu constatar que as variáveis permaneceram com o sinal esperado, de acordo com a teoria, contudo, algumas variáveis não foram significativas, por exemplo, a variável Renda *per capita* permaneceu com o sinal esperado conforme a teoria, porém deixou de ser significativa a 10%. Já as demais variáveis permaneceram significativas a um nível de 10%.

Tabela 4 - Regressão do modelo de Erro Espacial do índice de pobreza contra seus determinantes nos bairros de Fortaleza e seus vizinhos para o ano de 2010 pelo método M.V.S.

Var. explicativas	Coefficiente	estatística-z	Prob
C	0.677642	1.3839	0.16639
Renda <i>per capita</i>	-0.0001150	-1.00317	0.31578
Mortalidade infantil	0.605495	27.0856	0.00000
Analfabetismo de 15 ou mais	0.198036	6.81036	0.00000
Índice de Theil	1.32525	1.62358	0.10446
λ (Lambda)	0.516759	4.98309	0.00000

Diagnósticos da regressão			
Diagn. de heterocedasticidade		Valor	Prob
<i>Teste Breusch-Pagan</i>		12.8498	0.01203

Diagnósticos de dependência espacial			
		Valor	Prob
<i>Teste de Máximo verossimilhança</i>		15.5136	0.00008

Var. dependente	IP1	N. de observações	116
R	0.980275	G. de liberdade	111
Crit. de Akaike	255.177	Crit. Schwarz	268.945

Fonte: Elaboração própria com auxílio do software GeoDa8.1 e baseado em dados do Censo 2010 e PNUD (2013). Nota: Para esse trabalho chamou-se Máximo verossimilhança de M.V.S.

A variável *lambda* (λ) revela que o peso do impacto da pobreza foi correlacionado espacialmente pelos resíduos como variável explicativa do modelo, ou seja, ocorreu o Erro Autorregressivo Espacial, também conhecido como modelo de Erro Espacial ou SEM – *Spatial Error Model*. O modelo SEM surgiu da incapacidade de se modelar toda a fonte de dependência espacial, oriunda do processo estocástico gerador dos dados espaciais. A parte da dependência não modelada se manifesta no padrão do erro aleatório entre regiões vizinhas, de forma que os erros não sejam autocorrelacionados espacialmente (DARMO-FAL, 2006).

Ressalta-se que a variável Renda *Per capita*, que no modelo de erro espacial deixou de ser significativa, mas continuou com sinal correto, não invalidou os resultados encontrados, pois conforme Araújo, Morais e Cruz (2012), analisar a pobreza multidimensional é utilizar diversas variáveis, fragmentando-as em dimensões (renda, alimentação, saúde, entre outras). Adotar esse método foi uma evolução vista da maneira unidimensional, a qual só considera a renda como indicador. Ainda, segundo os autores, considerar a pobreza multidimensional torna-se mais eficaz, pois, ao tratar a pobreza somente por uma dimensão, pode-se estar negligenciando a real pobreza de uma determinada região.

Com relação ao diagnóstico da regressão, os dados permaneceram homocedásticos e a dependência espacial foi devidamente controlada.

Argui-se que, os resultados obtidos para Renda *per capita* e Índice de desigualdade de Theil, neste estudo, se aproximam de outros trabalhos realizados para o Brasil, ao indicar que a desigualdade contribui de maneira mais representativa para os mais pobres do que o aumento da renda (TAQUES; OLIVEIRA, 2014; PINTO; OLIVEIRA, 2010; NEDER; SILVA, 2004; HOFFMANN, 2005). Segundo Adams (2002), o efeito que o crescimento econômico exerce sobre a pobreza é dado pelo grau de desigualdade, então a desigualdade deve ser utilizada como controle para qualquer função de pobreza. A adição do componente de desigualdade revela que, a princípio, políticas que busquem uma melhora na distribuição da renda são mais efetivas no combate à pobreza, do que as que objetivam unicamente o aumento do nível de renda.

Portanto, este estudo se mostrou relevante para reformulação de políticas públicas voltadas ao alívio e combate da pobreza, pois permitiu um enfoque maior da sensibilidade regional ao englobar os 116 bairros de Fortaleza, e constatou que investimentos em saúde e infraestrutura podem permitir a redução da pobreza multidimensional, além de uma queda expressiva na porcentagem de pobres, com melhoria na oportunidade e qualidade de vida dos residentes fortalezenses, assim como, uma mais robusta transferência de renda. Nesse contexto, destaca-se que, os *policy makers* devem intensificar a alocação dos recursos nos programas de melhoria da saúde e bem-estar, e uma eficiente distribuição de renda, para os mais pobres, para que seja reduzida ainda mais a mortalidade infantil em Fortaleza, e com isso permita reduzir a disparidade entre ricos e pobres.

5 CONSIDERAÇÕES FINAIS

O presente estudo concentrou-se em acrescentar novos dados à discussão da pobreza, utilizando o IP, cuja base metodológica é semelhante à utilizada para composição do IDH. Também utilizou-se a análise de dados espaciais (AEDE), cujo objetivo foi perceber a espacialização da pobreza no município de Fortaleza – CE. Assim, foram utilizados os dados disponibilizados pelo IBGE, através dos Censos de 2000 e 2010, que foram comparados ao IDHM dos mesmos períodos, disponibilizados pela PNUD (2013), por meio do Atlas de Desenvolvimento Humano do Brasil.

Nesta lógica, a organização dos bairros em *rankings* de pobreza demonstrou que os índices são consistentes para o estudo da pobreza espacial em Fortaleza, mesmo não mostrando grandes diferenças quando ocorre a comparação entre o IP1, que inclui a renda no escopo, e o IP2, que a exclui. Entretanto, quando se compara qualquer um dos dois índices multidimensionais com o IP3, que é composto só pela dimensão renda, apresentaram-se

várias modificações na composição dos bairros, tanto para o período de 2000, bem como para 2010, quando as mutações foram mais robustas. Já na comparação com o IDHM, os índices se mostraram como instrumentos relevantes para estudo da pobreza, pois os bairros menos desenvolvidos também foram aqueles classificados como os mais pobres (Ver Anexos I e II), assim como os mais desenvolvidos se encontram na mesma categoria dos menos pobres.

A respeito da análise espacial, os resultados confirmaram a hipótese inicial do estudo, constatando a existência de *clusters* de pobreza em Fortaleza, independente do índice utilizado para fazer a sua medição. Observou-se que o número de *clusters* de baixa pobreza foi bem maior quando se utilizou a análise da pobreza com enfoque apenas na renda, do que quando se utilizaram os índices multidimensionais, apontando uma maior dependência espacial, bem como uma concentração da renda nas localidades mais centrais da cidade. Entretanto, nos mapas multidimensionais, os *clusters*, principalmente de alta pobreza, foram mais numerosos e distribuídos em várias regiões de Fortaleza, indicando que a pobreza com este enfoque está mais distribuída espacialmente.

O diagnóstico de dependência espacial apresentado pelo método M.Q.O. apresentou o modelo de Erro Espacial como o mais indicado. Tal indicador, estimado pelo método M.S.V. revelou que uma política de controle da pobreza tem que abranger outras variáveis, que não só a renda, para uma efetiva política de alívio à pobreza nos bairros de Fortaleza. Portanto, diante dos resultados apresentados, pode-se inferir como determinantes da condição de pobreza em Fortaleza os seguintes fatores, cuja influência foi positiva (ou seja, que tende a elevar a condição de pobreza): o Índice de Theil, Mortalidade infantil e o nível de Analfabetismo de pessoas com 25 anos ou mais. Além disso, o erro espacial se mostrou relevante e, assim, o peso do impacto da pobreza foi correlacionado espacialmente pelos resíduos como variável explicativa do modelo.

Por fim, diante dos resultados encontrados, este estudo evidencia que as políticas públicas de combate à pobreza com foco apenas na renda não são suficientes para combatê-la e, em muitos casos, cria uma minimização do problema, a tempo que aponta uma subestimação da pobreza e do número de pobres. Logo, faz-se necessário os governos inserirem nos seus estudos e programas de combate à pobreza outras dimensões causadoras de privações além da renda e, dessa forma, possibilitarem a criação de políticas públicas capazes de minimizar os efeitos desse fenômeno.

REFERÊNCIAS

- ADAMS, P.; HURD, M. D.; McFADDEN, D.; Merrill, A.; RIBEIRO T. Healthy, Wealthy and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status. *Journal of Econometrics*, v. 112, n. 1, p. 3-56. 2003.
- ALMEIDA, E. S. *Curso de Econometria Espacial Aplicada*. Piracicaba, 2004. Mimeo.
- ALMEIDA, E. S. *Econometria Espacial Aplicada*. Editora Alínea, Campinas, 2012.
- ALMEIDA, E. S.; PEROBELLI, F. S.; FERREIRA, P. G. C. Existe convergência espacial da produtividade agrícola no Brasil? *Revista de Economia e Sociologia Rural*, v. 46, n. 01, p. 031-052, 2008.
- AMARAL, R. F. do; CAMPOS, K. C.; LIMA, P. V. P. S. Distribuição da pobreza no estado do Ceará: uma abordagem multidimensional. *Interações*, v. 16, n. 2, p. 327-337, 2016.
- ANAND, S.; SEN, A. K. *Concepts of Human Development and Poverty: a Multidimensional Perspective*. Human Development Papers, UNDP: New York, 1997.
- ANSELIN, L. *Spacestat tutorial: a workbook for using Spacestat in the analysis of spatial data*. Illinois: Urbana, 1992. Mimeo.
- ANSELIN, L. *Spatial Econometrics: methods and models*. Kluwert Academic: Boston, 1988.
- ARAÚJO, J. A.; MORAES, G. S.; CRUZ, M. S. *Estudo da pobreza multidimensional no Estado do Ceará*. Sobral: Laboratório de Estudos Regionais, 2012. (Texto para discussão)
- ÁVILA, J. F. *Pobreza no Rio Grande do Sul: uma análise exploratória da sua distribuição espacial a partir de indicadores multi e unidimensionais*. 2013. 162f. Dissertação (Mestrado em Economia do Desenvolvimento). Porto Alegre: FACE, PUCRS, 2013.
- BRASIL. *Indicadores Sociais Municipais: uma análise dos resultados do universo do Censo Demográfico 2010*. Rio de Janeiro, RJ: IBGE, 2011.
- CONCONI, A.; HAM, A. *Pobreza Multidimensional Relativa: una aplicación a la Argentina*. CEDLAS. La plata. Agosto, 2007.
- CRESPO, A.; GUROVITZ, E. A pobreza como um fenômeno multidimensional. *Revista ERA*, v. 1, São Paulo, 2002.
- DARMOFAL, D. Spatial econometrics and political science. In: *Annals of the Annual Meeting of the Southern Political Science Association*, Atlanta, GA, 2006.

DEVLIN, N.; HANSEN, P. Health Care Spending and Economic Output: Granger Causality, *Applied Economics Letters*, v. 8, p. 561-64, 2001.

FERES, J. C.; VILLATORO, P. *Cadernos de Estudos Desenvolvimento Social em Debate*. N. 15. Brasília, DF: Ministério do Desenvolvimento Social e Combate à Fome, 2013.

FLORAX, R. J. G. M.; FOLMER, H.; REY, S. J. Specification searches in spatial econometrics: The relevance of Hendry's methodology. *Regional Science and Urban Economics*, v. 33, n. 5, p. 557-79, 2003.

GIOVANETTI, A. E.; RAIHER, A. P. Uma Análise Multidimensional da pobreza nos municípios paraenses em 2010. *Revista de Estudos Sociais*, v. 17, n.33, p. 228-248, 2015.

HOFFMAN, R. Transferência de renda e a redução da desigualdade no Brasil em cinco regiões entre 1997 e 2004. *Econômica*, v. 8, n. 1, p. 55-81, 2006.

_____. . Elasticidade da Pobreza em Relação à Renda Média e à Desigualdade no Brasil e nas Unidades da Federação. *Revista Economia*, 6(2), 255-289, jul. - dez. 2005.

IPECE – Instituto de Pesquisa e Estratégia Econômica do Ceará. Perfil Municipal de Fortaleza. Tema VII: *Distribuição Espacial da Renda Pessoal*. Informe nº 42, Edição Especial. Out., 2012.

KAGEYAMA, A.; HOFFMANN, R. *Pobreza no Brasil: uma perspectiva multidimensional*. São Paulo, Jun., 2006.

LACERDA, F. C. C. *A Pobreza na Bahia Sobre o Prisma Multidimensional: uma análise baseada na abordagem das necessidades básicas e na abordagem das capacitações*. 2009. 210f. Dissertação (Mestrado em Economia), Universidade Federal de Uberlândia. Uberlândia - MG: UFU, 2009.

LACERDA, F. C. C.; NEDER, H. D. Pobreza Multidimensional na Bahia: uma análise fundamentada no indicador multidimensional de pobreza. *Revista Desenhavia*, n. 13, p. 33-70, 2010.

NEDER, H. D., & SILVA, J. L. M. D.. Pobreza e distribuição de renda em áreas rurais: uma abordagem de inferência. *Revista de Economia e Sociologia Rural*, 42(3), 469-486. 2004.

PIMENTEL, E. A. *Desigualdades Regionais no Brasil e sua dinâmica durante a década de 1990: um estudo espacializado*. São Paulo: USP, 2004.

PINTO, M. S.; OLIVEIRA, J. C.. Crescimento pró-pobre: Análise dos estados brasileiros

entre 1995 e 2007. *Revista de Economia Contemporânea*, Rio de Janeiro, 14(2), 327-358, maio/ago. 2010.

PNUD – Programa das Nações Unidas para o Desenvolvimento. *Atlas do Desenvolvimento Humano do Brasil, 2013*. Disponível em: < <http://www.pnud.org.br/atlas/>>. Acesso em: 1 Jun. 2016.

PNUD – Programa das Nações Unidas para o Desenvolvimento. *Poverty in the human development perspectiv: concept end meansurement*. Development. Report, 1997.

POGGI, A. *Social Exclusion in Spain: Measurement Theory and Application*. PhD Thesis, Universitat Autònoma de Barcelona, Spain, 2004.

ROCHA, S. Impacto sobre a pobreza dos novos programas federais de transferência de renda. *Revista de Economia Contemporânea*, v. 9, n. 1, 153-185, 2005.

SEN, A. K. *Desenvolvimento Como Liberdade*. 6ª Ed. São Paulo: Companhia das Letras, 2000.

TAQUES, F. H.; OLIVEIRA, J. C. T.. Existe uma relação entre nível de renda, desigualdade e pobreza? Um estudo a partir de elasticidades. *Economia e Desenvolvimento*, Santa Maria – RS, 26(2), 21 - 42, jul. – dez. 2014.

TOWNSEND, P. The Meaning of Poverty. *The British Journal of Sociology*, v.13., n. 3, p. 210-227, 1962.

ANEXO I - Ranking dos “Piores” Bairros de Fortaleza em Relação à Pobreza e ao Desenvolvimento/2000 e 2010

2000					
Nome do Bairro	IP1	RK	Nome do Bairro	IP2	RK
PARQUE ARAXÁ	40,52	1º	PARQUE ARAXÁ	42,21	1º
PEDRAS	35,98	2º	PEDRAS	37,69	2º
GUAJERU	34,89	3º	GUAJERU	36,22	3º
CURIÓ	32,10	4º	SABIAGUABA	33,59	4º
VILA VELHA	31,95	5º	CURIÓ	32,66	5º
2010					
Nome do Bairro	IP3	RK	Nome do Bairro	IDHM	RK
CURIÓ	20,57	1º	PARQUE ARAXÁ	0,426	116º
ANCURÍ	19,76	2º	GUAJERU	0,461	115º
JARDIM DAS OLIVEIRAS	18,78	3º	PEDRAS	0,466	114º
SAPIRANGA/COITE	18,14	4º	VILA VELHA	0,477	113º
PIRAMBU	18,00	5º	GENIBAÚ	0,481	112º
2010					
Nome do Bairro	IP1	RK	Nome do Bairro	IP2	RK
PARQUE ARAXÁ	26,83	1º	PARQUE ARAXÁ	27,86	1º
CURIÓ	22,60	2º	CURIÓ	22,30	2º
PEDRAS	21,75	3º	PEDRAS	21,68	3º
AUTRAN NUNES	21,22	4º	AUTRAN NUNES	20,63	4º
SIQUEIRA	21,08	5º	SIQUEIRA	20,15	5º
2010					
Nome do Bairro	IP3	RK	Nome do Bairro	IDHM	RK
PARQUE PRESIDENTE VARGAS	21,52	1º	CURIÓ	0,567	116º
SIQUEIRA	20,27	2º	PARQUE ARAXÁ	0,580	115º
GRANJA PORTUGAL	19,40	3º	SIQUEIRA	0,624	114º
GENIBAÚ	19,14	4º	GENIBAÚ	0,624	113º
CANINDEZINHO	18,28	5º	CANINDEZINHO	0,624	112º

Fonte: Elaboração própria com base em dados do Censo 2000, 2010 e PNUD (2013).

Nota: Por se tratar de um estudo sobre a pobreza, a classificação é realizada com os bairros mais pobres em primeiro e os menos pobres em último. No caso do IDHM, os bairros mais desenvolvidos possuem valores mais altos e os bairros menos desenvolvidos, valores menores, e são também apresentados em primeiro.

ANEXO II - Ranking dos “Melhores” Bairros de Fortaleza em Relação à Pobreza e ao Desenvolvimento/2000 e 2010

2000					
Nome do Bairro	IP1	RK	Nome do Bairro	IP2	RK
MEIRELES	10,31	116°	MEIRELES	10,25	116°
ALDEOTA	10,41	115°	ALDEOTA	10,38	115°
COCÓ	10,74	114°	COCÓ	10,78	114°
DIONÍSIO TORRES	11,35	113°	DIONÍSIO TORRES	11,35	113°
DE LURDES	11,37	112°	DE LURDES	11,36	112°
Nome do Bairro	IP3	RK	Nome do Bairro	IDHM	RK
MEIRELES	2,46	116°	MEIRELES	0,894	1°
COCÓ	2,63	115°	ALDEOTA	0,892	2°
DIONÍSIO TORRES	3,27	114°	COCÓ	0,863	3°
PARQUELÂNDIA	3,51	113°	DIONÍSIO TORRES	0,841	4°
PRAIA DE IRACEMA	3,70	112°	FÁTIMA	0,821	5°
2010					
Nome do Bairro	IP1	RK	Nome do Bairro	IP2	RK
MEIRELES	5,21	116°	MEIRELES	5,20	116°
ALDEOTA	5,21	115°	ALDEOTA	5,20	115°
COCÓ	5,57	114°	COCÓ	5,58	114°
PRAIA DE IRACEMA	5,63	113°	PRAIA DE IRACEMA	5,61	113°
GUARARAPES	5,73	112°	GURARAPES	5,81	112°
Nome do Bairro	IP3	RK	Nome do Bairro	IDHM	RK
MEIRELES	0,62	116°	ALDEOTA	0,945	1°
COCÓ	0,87	115°	MEIRELES	0,937	2°
DIONÍSIO TORRES	0,99	114°	COCÓ	0,921	3°
JOSÉ BONIFÁCIO	1,06	113°	LUCIANO CAVALCANTE	0,915	4°
FÁTIMA	1,21	112°	DE LURDES	0,915	5°

Fonte: Elaboração própria com base em dados do Censo 2000, 2010 e do PNUD (2013).

Nota: Por se tratar de um estudo sobre a pobreza, a classificação é realizada com os bairros mais pobres em primeiro e os menos pobres em último. No caso do IDHM, os bairros mais desenvolvidos possuem valores mais altos e os bairros menos desenvolvidos, valores menores, e são também apresentados em primeiro.

Anexo III - Teste de causalidade Granger

Hipótese Nula(H0)	Teste F	Prob.	Rejeição de H0
IP1(pobreza) não causa (no sentido de Granger) Theil (desigualdade)	128446	0,0005	sim
Theil (desigualdade) não causa (no sentido de Granger) IP1(pobreza)	0,95523	0,3305	não
IP1(pobreza) não causa (no sentido de Granger) Renda <i>per capita</i>	0,03893	0,8440	não
Renda <i>per capita</i> não causa (no sentido de Granger) IP1(pobreza)	130850	0,0004	sim
Mortalidade infantil não causa (no sentido de Granger) IP1(pobreza)	0,10303	0,7488	não
IP1(pobreza) não causa (no sentido de Granger) Mortalidade infantil	206886	0,1531	não
Analfabetismo de 25 ou mais não causa (no sentido de Granger) IP1(pobreza)	0,94620	0,3328	não
IP1(pobreza) não causa (no sentido de Granger) Analfabetismo de 15 ou mais	321187	0,00001	sim

Fonte: Elaboração própria com base em dados do Censo 2000, 2010 e PNUD (2013) com auxílio do pacote estatístico Eviews 8.0.

AVALIAÇÃO DE POLÍTICAS EDUCACIONAIS: UM ESTUDO DE CASO DA EDUCAÇÃO EM TEMPO INTEGRAL SOBRE O DESEMPENHO DOS ALUNOS DO ENSINO MÉDIO DO ESTADO DO CEARÁ PARA O ANO DE 2014

Antonio Bruno Fernandes Marcelino*

Wellington Ribeiro Justo**

Matheus Oliveira de Alencar***

RESUMO

A escola pública de educação tradicional vem nos últimos anos atravessando um momento de perda de sua identidade cultural e pedagógica, pois a nova realidade fez com que estas instituições adquirissem funções que vão além da instrução escolar formal. Neste contexto, a educação integral ganhou relevância. O presente trabalho tem como cerne analisar a eficiência da política pública de educação integral no Ceará, comparativamente a eficiência proporcionada pelas escolas de tradição, por meio dos microdados do Exame Nacional do Ensino Médio (ENEM) referentes ao ano de 2014. Para o alcance do objetivo, utilizou-se o *Propensity Score Matching (PSM)* por meio do método *Nearest-Neighbor*. Verificou-se diferença significativa no desempenho escolar favorável ao grupo de tratamento, além de permitir inferir a importância do método para a obtenção de resultados robustos, tendo em vista que sem o pareamento observou-se uma sobrevalorização do efeito médio da política. Apesar disso, a proposta se mostra eficiente, permitindo inferir que a mesma está conseguindo de fato melhorar qualitativamente a ação das escolas sobre os resultados do alunado cearense..

Palavras-Chave: Educação Integral, Desempenho Educacional, PSM.

ABSTRACT

The public school of traditional education has been going through a period of loss of its cultural and pedagogical identity in recent years, as the new reality has made these ins-

* Graduado em Economia pela Universidade Regional do Cariri (URCA). (88)9218-1615. brunnofernandes@outlook.com

** Professor Associado da URCA; Professor do PPGECON. Doutor em Economia pelo PIMES-UFPE. (81)98848-1898. justowr@yahoo.com.br .

*** Mestrando em Economia Rural, Universidade Federal do Ceará (UFC), (88) 981070613, matheusalencar29@gmail.com.

tutions acquire functions that go beyond formal school education. In this context, integral education has gained relevance. The purpose of this study is to analyze the efficiency of the public education policy in Ceará, in comparison to the efficiency provided by the schools of tradition, through the micro data of the National High School Examination (ENEM) for the year 2014. For the Objective, Propensity Score Matching (PSM) was used using the Nearest-Neighbor method. There was a significant difference in the school performance favorable to the treatment group, besides allowing to infer the importance of the method to obtain robust results, considering that without the pairing it was observed an overvaluation of the average effect of the policy. In spite of this, the proposal proves efficient, allowing to infer that it is actually improving qualitatively the action of the schools on the results of the student from Ceará.

Keywords: Integral Education, Educational Performance, PSM.

1 INTRODUÇÃO

É consenso na academia que a educação é uma ferramenta importante para o desenvolvimento de um país. Ciente disso, o Estado foi ao longo dos anos demonstrando grande interesse na melhoria da educação através de medidas importantes como a promulgação da Constituição Federal de 1988, que assegurou o direito universal ao ensino básico, e a reforma da Lei das Diretrizes e Bases da Educação Nacional (LDB), que instituiu as diretrizes que norteiam o sistema educacional do país (ARAÚJO JÚNIOR, *et al.* 2014).

Todavia o ensino no Brasil só foi universalizado no início do século XXI, uma conquista fundamental, mas que não foi o suficiente para melhorar a ação escolar sobre o processo de aprendizagem dos estudantes (CARVALHO, 2006). Para Cavaliere (2002) e Soares (2004), a escola pública de educação tradicional vem nos últimos anos atravessando um momento de perda de sua identidade cultural e pedagógica, pois a nova realidade fez com que essas instituições adquirissem funções, de outras instituições, que vão além da instrução escolar formal, gerando nos pais grandes expectativas quanto ao processo de aprendizagem dos alunos.

Eis que surge como uma possível solução a política de educação integral, cujo debate está em destaque em função da reforma instituída recentemente pelo governo federal no ensino médio brasileiro. Conforme o Centro de Referência em Educação Integral (2014) o termo diz respeito às escolas que ampliam a jornada escolar de seus alunos, adicionando ou não novas disciplinas nos currículos da educação básica, sendo que geralmente as instituições que adotam este modelo programam a extensão do tempo com metade do dia

letivo para as disciplinas obrigatórias e a outra metade para atividades ligadas às artes e aos esportes.

Para o Ministério da Educação, a proposta de tempo integral representa um projeto educativo adaptado à vida, às necessidades, às possibilidades e aos interesses dos estudantes (MEC, 2016a). Foi nesse sentido que em 2008, o governo estadual de Pernambuco tornou a educação integral uma política para o ensino médio e, desde então, seu projeto fundamenta-se na ideia de que investir em mais horas aulas (45 horas aulas semanais), bem como na qualidade do ensino, é indispensável para a formação de jovens mais aptos para a continuidade da vida acadêmica e mais qualificados para o mercado de trabalho (SEE, 2016a). Ademais, Pernambuco é o estado que desde o ano de 2007 mais evoluiu no desempenho educacional e o índice de 3,9 no IDEB de 2015, superior à meta de 3,6, fez com que o estado alcançasse o melhor ensino médio público do país, evoluindo da 4ª para a 1ª posição no Ranking do IDEB, sendo também o estado que possui a menor disparidade em termos de desempenho entre a rede pública e a rede privada (SEE, 2016b).

Neste contexto, a presente pesquisa tem como objetivo analisar a eficiência da política de educação em tempo integral do Estado do Ceará, no recorte temporal de 2014, por meio do desempenho médio dos estudantes das Escolas Estaduais de Educação Profissional (EEEPs), comparativamente ao desempenho dos estudantes das escolas de tradição.

A pesquisa se justifica pela relevância que estudos que envolvem educação e economia possuem como base teórica para a formulação de políticas educacionais, bem como devido ao destaque que a proposta de educação em tempo integral tem ganhado em função do debate acerca da necessidade de mudanças para tornar o ensino mais atrativo e para a melhorar a capacitação dos estudantes para as reais necessidades do século XXI.

Diante do que foi exposto, surge o questionamento: seria a educação na jornada de tempo prolongada uma política pública eficiente, de qualidade e com capacidade para melhorar o desempenho escolar dos estudantes cearenses do ensino médio?

A hipótese a ser testada na pesquisa é que a política de escolas de educação em tempo integral é eficiente para melhorar o desempenho cognitivo dos alunos cearenses frente ao desempenho cognitivo proporcionado pelas escolas tradicionais.

Além desta seção introdutória, o trabalho está estruturado em mais cinco seções: a segunda seção traz um breve referencial teórico; a terceira apresenta os procedimentos metodológicos utilizados; a quarta expõe os resultados obtidos; e a quinta apresenta as considerações finais do estudo.

2 EDUCAÇÃO INTEGRAL

No Brasil a origem da educação integral remonta ao movimento de renovação do ensino denominado Escola Nova, que ganhou notoriedade com a divulgação do Manifesto dos Pioneiros da Educação Nova do ano de 1932, que defendia, dentre suas ideias, a universalização, a laicidade e a gratuidade da escola pública (DUTRA, 2014). O autor destacou que em meados dos anos 1950, na cidade de Salvador (BA), Anísio Teixeira, um dos mentores intelectuais do Manifesto dos Pioneiros da Educação Nova, colocou suas ideias em prática e criou a Escola Parque, considerada a primeira experiência de educação integral destinada a crianças nas séries iniciais da educação básica, que incluía no processo de escolarização diversas ações de socialização e práticas de aprendizagem não formais, como atividades culturais e de lazer.

Para Cavaliere (2002), a Escola Nova ou Corrente Pedagógica Escolanovista, teve como objetivo associar as atividades escolares com as experiências de vida, ou seja, a corrente tinha como base o entendimento de educação como vida, e não como um processo de preparação para a vida. Com o passar do tempo, mais precisamente a partir da abertura política da década de 1980, algumas experiências surgiram na direção da jornada de tempo integral, sendo os Centros Integrados de Educação Pública (CIEPs), no Rio de Janeiro, os de maior duração e de maior repercussão e, posteriormente, a partir do século XXI, com base no artigo 34 da Lei das Diretrizes e Bases da Educação Nacional (LDB) de 1996 – que trouxe como agenda o ensino fundamental em regime de tempo integral – intensificou-se o surgimento de propostas envolvendo o aumento do tempo diário de permanência dos jovens nas escolas públicas estaduais e municipais (CAVALIERE, 2007).

Nesse sentido, para constituir a jornada exigida na LDB, vários municípios já estão desenvolvendo ou buscando aplicar nas redes de ensino uma política que junte educação, cultura, esportes, assistência social e reforço escolar para constituir a jornada de tempo integral em detrimento da jornada escolar convencional (CARVALHO, 2006). Ademais, por ser também uma das metas (6) do Programa Nacional de Educação (PNE) do governo federal, a proposta figura-se como uma das políticas públicas mais difundidas pelas autoridades governamentais brasileiras.

A política, segundo Carvalho (2006), ganhou relevância, visto que o ensino no Brasil só foi universalizado no início do século XXI, mas que esta conquista, embora fundamental, não foi o suficiente para a melhoria da aprendizagem dos estudantes. No entanto, de acordo com a autora, a educação integral é compreendida de diferentes formas no país, sendo vista por vezes apenas como instituição com jornada escolar ampliada, outras vezes

como uma reivindicação em virtude do baixo desempenho escolar dos estudantes – na expectativa de que a ampliação do tempo de escola aumente o aprendizado – ou ainda como um complemento socioeducativo à escola tradicional.

Cavaliere (2007) apontou que existem pelo menos quatro concepções de escolas de tempo integral no Brasil. Para a autora, existe a visão assistencialista (1), que considera a escola de tempo integral como uma instituição voltada para suprir as deficiências de aprendizagem de pessoas desprivilegiadas; existe a visão autoritária (2), que trata a escola de jornada prolongada como uma instituição que tira crianças e jovens das ruas para preveni-los da violência e de crimes; há ainda a visão democrática (3) que considera a escola de tempo integral como uma instituição emancipatória que prepara os alunos de forma crítica para o mundo; e por fim, surgiu mais recentemente a concepção de escola de tempo integral numa visão multissetorial (4), que independe do tempo de escola.

Para Gonçalves (2006), na visão multissetorial, que atualmente é a visão mais difundida pelos educadores, a educação integral é definida como aquela que vai além da concepção cognitiva, considerando os indivíduos também na sua concepção corpórea, isto é, na compreensão de que o sujeito está em um contexto de relações sociais e afetivas. Segundo o autor, deve-se pensar nos estudantes como sujeitos inteiros, em uma dimensão biopsicosocial, sendo importante considerar os seus interesses, não no sentido de trabalhar apenas com o que desejam, mas no sentido de que o conteúdo da grade curricular seja ofertado considerando a vivência dos sujeitos. Ainda para o autor, o diálogo entre educadores e educandos é fundamental, pois só faz sentido pensar na educação integral se a perspectiva da expansão da jornada escolar for de encontro à ampliação das oportunidades que promovem a melhoria do aprendizado.

Conforme Carvalho (2006) e Cavaliere (2007), esta proposta não pode ser centrada exclusivamente na escola, ou seja, ela pode e deve ser feita fora da instituição. Carvalho (2006) destacou que a rede escolar atual prioriza a demanda por vagas operando em vários turnos e isso inviabiliza a proposta de educação em tempo integral, pois a política exige uma articulação entre escolas públicas e programas socioeducativos na comunidade local, de modo a proporcionar o debate sobre a educação integral e não apenas sobre a ampliação da jornada escolar.

Cavaliere (2002) apontou que as instituições de ensino tradicional enfrentam um momento de perda de sua identidade cultural e pedagógica. Para a autora, a escola pública no país destinava-se efetivamente a poucos, mas seguiu na segunda metade do século XX o processo de escolarização das grandes massas, que se desenvolveu tendo como foco a

transmissão formal dos conhecimentos, se distanciando das demais responsabilidades sociais e culturais. A mesma destacou que esse processo se deu, entre outros fatores, pelas precárias instalações escolares, pela redução da jornada de horas aulas, pela multiplicação dos turnos e pela baixa qualidade da formação dos professores.

Mas, acontecimentos recentes como a ampliação dos currículos escolares; a influência de políticas públicas sociais, como o Bolsa Família, que exige a permanência dos jovens nas escolas; a retirada da classe média urbana da rede pública para a rede privada; e o processo de privatização do ensino são alguns fatores que fizeram surgir a necessidade de debater sobre uma nova identidade para a escola pública, pautada na integração efetiva dos jovens à vida escolar, fazendo com que as escolas atuais procurem cada vez mais assumir novos compromissos pedagógicos, consolidando a ideia de baixa qualidade da escola pública de tradição (CAVALIERE, 2002).

Cavaliere (2007) explicou que a ampliação da jornada escolar diária pode ser entendida e justificada de diferentes formas, entre elas: (1) para melhorar a ação escolar sobre os resultados do alunado; (2) também como forma de se adequar as novas condições familiares da vida urbana, particularmente às novas condições da mulher no século XXI; (3) ou ainda em decorrência da mudança na concepção de educação escolar, ou seja, do papel da escola na vida e na formação dos estudantes.

O governo federal possui dois programas de ensino em tempo integral: o Mais Educação para o nível fundamental e o Ensino Médio Inovador (ProEMI) para o nível médio, ambos possuem em sua essência o propósito de tornar o ensino brasileiro mais dinâmico através de ações como o acompanhamento pedagógico às áreas do conhecimento, bem como o estímulo e o desenvolvimento de atividades artísticas, culturais, esportivas, de lazer, de comunicação, de cultura digital, entre outras, que desenvolvem as múltiplas dimensões humanas (cognitiva, psicomotora, socioafetiva, estética e ética), visando atender as expectativas dos estudantes e as demandas da sociedade contemporânea (MEC, 2016b; 2016c).

Com outro ponto de vista sobre o tema, Arroyo (1988) fez uma crítica sobre o rumo que a educação integral tomou. Para o autor, a história mostrou que o tempo integral, ora na rede pública ora na rede privada, foi de fato se tornando uma prática de ensino em diversos países, mas a lógica que tem conduzido este processo não caminhou em direção à vinculação da ampliação do tempo de escola à proposta de educação integral onde crianças e adolescentes seriam internados em escolas dirigidas pelo estado, que ofereceria a educação necessária. Para o mesmo, a proposta tem seguido a lógica mercantil do mercado de

trabalho, isto é, como a maioria da população é obrigada a vender sua força de trabalho, a escola passou, portanto, a ser mais demandada em função da ideologia de que a condição para obter uma posição mais elevada na hierárquica organização do trabalho depende do tempo formal de escolarização.

2.1 Educação profissionalizante no Brasil e no Estado do Ceará

A educação técnica ou profissionalizante no Brasil teve sua origem no século XIX, mais precisamente no ano de 1809 com a criação do Colégio das Fábricas por D. João VI para o ensino de ofícios aos jovens em situação de miserabilidade que não conseguiam se empregar no mercado, possuindo, portanto, caráter assistencialista (SILVÉRIO e ALBUQUERQUE, 2012). Analogamente, Barros (2013) destacou que o objetivo era transmitir ofícios aos órfãos que vieram na frota com a família real portuguesa para o Brasil, mas posteriormente foi acrescentado à proposta o ensino primário completo.

Para Silvério e Albuquerque (2012), no século XX a política muda de assistencialista e passa a focar na formação de operários para o exercício profissional, empobrecendo a formação dos estudantes, principalmente a partir da década de 1970 com o regime militar, que estabeleceu através da Lei Nº 5.682/71 a obrigatoriedade da educação profissional para os brasileiros, que serviriam de mão de obra qualificada para os projetos de expansão econômica do período que ficou conhecido como milagre econômico.

Barros (2013) destacou que foi a partir da Constituição Federal de 1937 que a educação profissionalizante deixou sua função compensatória para adotar os princípios técnico-científicos do fordismo e do taylorismo, mas destacou também que a proposta desde o seu primórdio sempre esteve indiretamente atrelada a formação de mão de obra, já que era destinada efetivamente às camadas sociais mais populares, revelando um caráter dualista na educação. Ainda de acordo com a autora, na década de 1940 foram criados pelo empresariado industrial e comercial o Serviço Nacional de Aprendizagem Industrial (SENAI) e o Serviço Nacional de Aprendizagem Comercial (SENAC) para o provimento de cursos profissionalizantes de curto prazo para atender as demandas da divisão internacional do trabalho.

Durante a década de 1980 a educação profissionalizante foi desaparecendo em função da Lei Nº 7.044/82 que flexibilizou sua adoção e foi somente a partir da década de 1990 quando se criou a atual Lei das Diretrizes e Bases da Educação Nacional (LDB) – Lei Nº 9.394/96 – que o ensino profissionalizante passou a ser integrado ao ensino médio regular, surgindo também as primeiras instituições especializadas, tais como as Escolas Técnicas

Federais e as Escolas Agrotécnicas Federais (SILVÉRIO e ALBUQUERQUE, 2012).

Ainda segundo Silvério e Albuquerque (2012), no final dos anos 1990, em virtude da conjuntura econômica e da ascensão das políticas neoliberais, o Estado reduz o custeio com a educação profissionalizante e a desagrega da educação regular, até que nos anos 2000, mais precisamente no governo Lula, por meio da Portaria Nº 3.621/2003 foi criado o Fórum Nacional de Educação Profissional e Tecnológica e em 2004 a educação profissionalizante voltou a integrar o ensino médio regular com o estado retomando o investimento nessa mobilidade de ensino.

No Estado do Ceará, a política de educação integral na modalidade profissionalizante teve início no ano de 2008 quando o governo estadual decidiu adaptar 25 escolas de ensino regular em 20 municípios para receber 4 mil alunos que desejavam uma formação técnica no ensino médio. Desde então, o Estado tem se destacado no cenário nacional pela ampliação e pelo desenvolvimento da mesma (CENTRO DE REFERÊNCIA EM EDUCAÇÃO INTEGRAL, 2015; SEDUC, 2017).

Conforme Silvério e Albuquerque (2012), as Escolas Estaduais de Educação Profissional (EEEPs) do Ceará, criadas pela Lei Nº 14.273/2008, são pautadas na formação geral, cultural e científica dos jovens e foram pensadas a partir do Programa de Desenvolvimento dos Centros de Ensino Experimental (PROCENTRO) do Estado de Pernambuco, que tinham como princípios: protagonismo juvenil, formação continuada, tecnologia empresarial e níveis de resultados.

Segundo a SEDUC (2017), no início a oferta de cursos técnicos era restrita (Enfermagem, Finanças, Segurança do Trabalho e Guia de Turismo), mas atualmente já são 53 cursos nas mais diversas áreas do conhecimento distribuídos em 116 escolas profissionalizantes por todo o Estado a depender das características socioeconômicos dos municípios em que as instituições estão instaladas. Ainda de acordo com a Secretaria de Educação, a política contava no ano de 2015 com 12% dos estudantes do ensino médio matriculados em escolas profissionalizantes, o equivalente a 43.280 alunos, e até esse mesmo ano cerca de 42.816 estudantes concluíram a educação técnica em alguma destas escolas, sendo enfermagem, informática e redes os cursos com mais profissionais formados (SEDUC, 2017).

A depender do curso profissionalizante, as turmas dedicam entre 800 e 1200 horas aulas para formação profissional com conteúdo técnico e para a conclusão do curso os alunos devem ter obrigatoriamente alguma experiência de estágio, sendo a remuneração a cargo do governo do Estado do Ceará (CENTRO DE REFERÊNCIA EM EDUCAÇÃO INTEGRAL, 2015).

Ainda segundo o Centro de Referência em Educação Integral (2015), para acompanhar o desempenho dos estudantes da educação técnica foi criado nas ETECs o diretor de turma, projeto que estabelece um professor para cada sala de aula com a função que acompanhar os estudantes por meio de reuniões e desenvolver projetos interdisciplinares e projetos de vida com os jovens.

O Diretor de Turma é um plano desenvolvido nas escolas públicas de Portugal e chegou ao Brasil quando a Secretaria de Educação do Ceará tomou conhecimento do projeto no XXI Encontro da Associação Nacional de Política e Administração da Educação (ANPAE) sediado no próprio estado no ano de 2007 (SILVÉRIO e ALBUQUERQUE, 2012). Os autores destacaram que no início o projeto foi implantado apenas nas ETECs, mas no ano de 2010 se estendeu para as demais escolas de ensino médio regular. Ainda conforme os mesmos, o professor Diretor de Turma é responsável pela elaboração de um dossiê completo com dados estatísticos sobre todos os alunos, o que permite conhecer as especificidades de cada discente, além de facilitar o trabalho e o empenho do docente em lograr sucesso com a turma.

3 METODOLOGIA

3.1 Natureza dos dados

A presente pesquisa utilizou dados de natureza secundária do Exame Nacional do Ensino Médio (ENEM), fornecidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), autarquia do Ministério da Educação (MEC).

Para este trabalho foram utilizados os microdados do ano de 2014 pelo fato dos de serem os mais atuais até então disponibilizados pelo INEP no momento do início da análise, bem como pelo fato de que é necessário um período de tempo considerável para conseguir verificar os efeitos de uma determinada política pública.

Com base no que a literatura julga importante para a medição de índices de proficiência, foram selecionadas para a análise as seguintes variáveis do banco de dados do ENEM: gênero, idade, raça, estado civil, atuação em atividade remunerada, tamanho da família, renda familiar, escolaridade do pai, escolaridade da mãe, localização da residência, tipo de escola, tipo de ensino, tipo de ensino fundamental, tipo de ensino médio, tempo para concluir o ensino fundamental, tempo para concluir o ensino médio, dependência administrativa, localização da escola, acesso à internet e status da redação, esta última apenas para verificar a performance dos participantes quanto a prova subjetiva do exame.

4.2 Método de análise

Inicialmente foi realizada uma análise descritiva a fim de se conhecer as principais características e atributos socioeconômicos, familiares e escolares dos estudantes que realizaram o Exame Nacional do Ensino Médio (ENEM) no ano de 2014. Após a realização da análise descritiva, foi aplicado o método de pareamento baseado em escore de propensão para avaliar o efeito da política cearense de escolas de tempo integral no desempenho educacional dos estudantes, ou seja, para avaliar o impacto (tamanho da discrepância) entre o desempenho dos alunos de escolas de tempo integral sobre o desempenho cognitivo dos alunos de escolas de tempo convencional.

Para satisfazer o objetivo principal da pesquisa adaptou-se o modelo desenvolvido por Becker e Ichino (2002) e também aplicado por Pereira, Justo e Lima (2015) e dar-se o exemplo de uma escola i de jornada diária tradicional, que possui desempenho educacional igual a Y_i . Considerando essa mesma escola i , agora de tempo integral, o resultado observado para a variável de interesse será dado pela Equação 1:

$$Y_i = DY_i + (1-D)Y_i \quad (1)$$

é uma variável *dummy* que assume valor 1 para as escolas de tempo integral e valor 0 para as escolas de tempo convencional.

Para verificar o tamanho da diferença do efeito proporcionado pelas escolas de tempo integral no desempenho cognitivo dos alunos é necessário medir a diferença total do desempenho dessa escola i quando ela passa a ser de tempo integral dado pela Equação 2:

$$\Delta Y_i = Y_i \quad D=1 - Y_i \quad D=0 \quad (2)$$

Ampliando a expressão para a população total n introduz-se o impacto médio (δY_i) das escolas de tempo integral sobre as escolas de tempo convencional dado pela Equação 3:

$$\delta Y_i = E(Y_i \mid D=1 - Y_i \mid D=0 \mid D=1) \quad i = 1, 2, \dots, n \quad (3)$$

Adaptado de Duarte, Sampaio e Sampaio (2009), $E(Y_i \mid D=1 - Y_i \mid D=0 \mid D=1)$ é o valor esperado condicional da diferença do desempenho de uma escola i , dado que esta escola passou a ser de tempo integral. Com base nos mesmos autores, como não é possível observar as escolas em ambas as situações no mesmo instante de tempo, utiliza-se para a avaliação dois grupos: um composto pelas escolas que são de tempo integral (grupo de tratamento), que será comparado a outro, composto pelas escolas que não são de tempo integral (grupo de controle), conforme expresso pela Equação 4:

$$\begin{aligned}
 & E(Y_i | D=1 | D=1) - E(Y_i | D=0 | D=0) = \\
 & E(Y_i | D=1 | D=1) - E(Y_i | D=0 | D=1) + E(Y_i | D=0 | D=1) - E(Y_i | D=0 | D=0) = \quad (4) \\
 & = \delta^* Y_i + E(Y_i | D=0 | D=1) - E(Y_i | D=0 | D=0)
 \end{aligned}$$

A Equação (4) demonstra o viés causado pela diferenciação das escolas pelo simples fato destas serem ou não de tempo integral, pois o efeito decorrente da simples comparação entre as escolas pode refletir características socioeconômicas anteriores à adoção da política. Para que esse problema não ocorra, segundo Pereira, Justo e Lima (2015) deve-se adotar uma medida de análise que minimize o viés e a medida adotada foi o *Propensity Score Matching* (PSM).

Então, para avaliar a hipótese formulada foi utilizado o PSM, criado por Rosenbaum e Rubin (1983) para reduzir o viés do cálculo do efeito médio do impacto de uma variável de interesse em um grupo de tratamento sobre um grupo de controle. O *propensity score*, chamado de $p(X)$, é então definido como a probabilidade condicional de participar de um grupo de tratamento dado um vetor de características observáveis X , dado pela Expressão 5:

$$p(X_i) = Pr(D_i=1 | X_i) = E(D_i | X) \quad (5)$$

Após a estimação de $p(X)$, torna-se possível o cálculo do impacto médio sobre uma população de participar ou não do grupo de tratamento. Assim, surge o conceito de *Average Effect of Treatment on the Treated* (ATT) ou efeito médio do tratamento sobre os tratados (τ), apresentado na Equação 6, adaptada de Becker e Ichino (2002):

$$\begin{aligned}
 \tau &= E(Y_i | D=1) - E(Y_i | D=0 | D_i=1) = \\
 & E[E\{Y_i | D=1 - Y_i | D=0 | D_i=1, p(X_i)\}] = \\
 & = E[E\{Y_i | D=1 | D_i=1, p(X_i)\} - E\{Y_i | D=0 | D_i=0, p(X_i)\} | D_i=1]
 \end{aligned} \quad (6)$$

Na literatura existem diversas metodologias de pesos para a estimação do ATT utilizando o propensity score matching, tais como: *nearest-neighbor*, *radius*, *stratification*, *kernel* e *local linear matching* (DANTAS e TANNURI-PIANTO, 2013). Para o cálculo do ATT desta pesquisa foi considerado o método de *Nearest Neighbor*¹ (NN) ou vizinho mais próximo. Neste método, calcula-se a média da diferença da variável de interesse entre uma observação do grupo de tratamento com uma equivalente do grupo de controle levando em consideração o $p(X_i)$ mais próximo, expresso pela Equação 7, adaptada de Santos *et al.* (2014) e Pereira, Justo e Lima (2015):

1. Esse procedimento permite um maior número de observações pareadas.

$$V(i) = \min_j \|p_i - p_j\|, i \in I \quad (7)$$

Em que: $V(i)$ representa o conjunto de observações do grupo de controle a serem relacionadas com uma escola i do grupo de tratamento; $p_i - p_j$ representam as probabilidades de ser uma escola de tempo integral; e I representa o grupo de escolas de tempo integral.

Por fim, para a estimação do PSM foi utilizado o modelo *Probit*. As variáveis explicativas do modelo estarão incluídas no vetor de características X_i e a variável dependente será uma *dummy* (escola de tempo integral = descint), que assumirá valor 1 se a escola for de tempo integral e valor 0, caso contrário, adaptado de Pereira, Justo e Lima (2015).

As variáveis explicativas empregadas na regressão são as seguintes: gênero (dsexo), idade, raça (draça), estado civil (destadocivil), atuação em atividade remunerada (dtrabalha), tamanho da família (residentes), renda familiar (renda), escolaridade da mãe (escolar-mãe), localização da residência (dlocalresid), tipo de ensino (dtipoensino), tipo de ensino fundamental (tipoescolanoensfund), tipo de ensino médio (tipoescolanoensmedio), tempo para concluir o ensino fundamental (tempofundamental), tempo para concluir o ensino médio (tempoensinomedio), dependência administrativa (dependência), localização da escola (dlocalização) e acesso à internet (internet).

Apresenta-se no Quadro 01 a descrição das variáveis explicativas binárias (*dummies*) empregadas na regressão com modelo *Probit*.

Quadro 01: Descrição das variáveis dummies da regressão com modelo *probit*

Variável	Descrição
Dsexo	Atribui-se valor 1 se masculino e 0 caso contrário
Draça	Atribui-se valor 1 se negro e 0 caso contrário
Destadocivil	Atribui-se valor 1 se solteiro e 0 caso contrário
Dtrabalha	Atribui-se valor 1 se trabalha e 0 caso contrário
Dlocalresid	Atribui-se valor 1 se na zona urbana e 0 caso contrário
Dtipoensino	Atribui-se valor 1 se ensino regular e 0 caso contrário
Dtipoescolanoensfund	Atribui-se valor 1 se público e 0 caso contrário
Dtipoescolanoensmedio	Atribui-se valor 1 se público e 0 caso contrário
Dependência	Atribui-se valor 1 se pública e 0 caso contrário
Dlocalização	Atribui-se valor 1 se na zona urbana e 0 caso contrário

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

As variáveis escolaridade do pai e tipo de escola (dummy, que recebeu valor 1 para escolas públicas e 0 caso contrário) precisaram ser retiradas da regressão em virtude do problema de multicolinearidade, ou seja, da relação linear perfeita ou exata.

4 RESULTADOS E DISCUSSÕES

Na primeira subseção deste capítulo são apresentados os resultados da análise descritiva, ao passo que na segunda subseção são apresentados os resultados com a ferramenta *Propensity Score Matching* através das regressões múltiplas com modelo *Probit*.

4.1 Análise descritiva

Com o uso de estatísticas descritivas, apresenta-se a seguir tabelas com distribuição de frequências (frequência absoluta e frequência relativa) e medidas de dispersão (média e desvio-padrão) para algumas das variáveis utilizadas na pesquisa e apresentadas anteriormente na seção de metodologia².

Na Tabela 01 apresentam-se os resultados das notas de acordo com o sexo e observou-se que a maioria dos cearenses que participou do Exame Nacional do Ensino Médio (ENEM) no ano de 2014 era do gênero feminino (54,71%). Apesar da maior participação feminina, as médias apontam que os participantes do gênero masculino perfizeram o melhor desempenho médio em quase todas as áreas do conhecimento avaliadas: Ciências da Natureza (CN), Ciências Humanas (CH) e Matemática (MT).

Tabela 01: Distribuição de frequências³ e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com o gênero

Sexo	Provas									
	Fi	%	CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Feminino	164.770	54,71	464,36	66,97	523,87	73,36	496,57*	70,79	443,50	90,46
Masculino	136.389	45,29	479,54*	76,07	535,03*	80,28	493,13	75,41	479,05*	109,68
Total	301.159	100	471,19	71,61	528,89	76,75	495,03	72,92	459,43	101,09

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Nota: Foram feitos testes de igualdade de médias para cada prova por gênero. * as diferenças são estatisticamente significantes a 1%,

Sabe-se que no passado as mulheres possuíam papel secundário na sociedade, papel esse que vem mudando gradativamente. Segundo Tiba (1996), os papéis tradicionais desempenhados por homens (pais/provedores) e mulheres (mães/rainhas do lar) sofreram grandes alterações, pois as mulheres agora pleiteiam as mesmas faculdades e ocupam os mesmos espaços que os homens, e cada vez maiores, sendo agora além de donas de casa e mães, sócias dos homens nas tarefas de arcar com as despesas familiares.

Quanto a elevada participação feminina no exame, observa-se uma ascensão social

2. Por falta de espaço não são apresentadas para todas as variáveis. Mas podem ser disponibilizadas pelos autores.

3. O número de observações varia em função da quantidade de pessoas que responderam a cada pergunta do questionário do exame. O mesmo vale para o número de observações nas demais tabelas que serão apresentadas.

do sexo feminino na busca pela realização profissional, dentre outras maneiras, através da educação e o ENEM representa um dos principais meios de acesso à educação de nível superior em qualquer instituição de ensino do País, pública ou privada.

Quanto a performance no exame, identificou-se que há diferença no desempenho médio por gênero. Os resultados corroboram com os de Silva (2015), que através de um modelo *Probit* Ordenado vislumbrou, em cursos de graduação da Universidade Federal da Paraíba (UFPB), que estudantes do sexo feminino possuíam rendimento escolar na ordem de 41% a menos que estudantes do sexo masculino, evidenciando também que apesar de a maioria dos estudantes concludentes fosse do sexo feminino, os do sexo masculino concluía com coeficientes de rendimento escolar mais elevado.

Para Saavedra (2001) existe um estereótipo ou bipolarização do conhecimento nas disciplinas e nos cursos acadêmicos. Ainda de acordo com a autora, isso reflete o dualismo de gênero, que está profundamente enraizado na cultura, em que atributos geralmente ligados ao gênero feminino (emocionalidade, subjetividade, intuição e sensibilidade) estão associados às Línguas, as Ciências Humanas e as Artes, enquanto os traços geralmente ligados ao gênero masculino (racionalidade, objetividade, impessoalidade e impassibilidade) estão associados às Ciências Físicas, Tecnológicas e à Matemática.

Na Tabela 02 encontram-se os resultados por raça. Verificou-se que a maioria dos estudantes que realizaram o exame se declarou pardo (68,23%). Quanto à performance, observou-se que os indivíduos que se declararam brancos perfizeram o melhor desempenho médio em todas as áreas do conhecimento.

Tabela 02: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com a raça

Raça	Provas									
	Fi	%	CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Branca	61.978	20,97	488,13	79,91	546,79	78,38	511,30	74,07	483,27	116,20
Preta	22.796	7,71	471,17	69,24	533,58	76,46	496,78	71,93	459,39	96,70
Parda	201.678	68,23	465,57	68,01	522,32	75,29	489,38	71,80	451,61	94,78
Amarela	6.865	2,32	477,23	68,32	539,69	71,13	506,01	68,52	466,80	99,64
Indígena	2.281	0,77	459,76	62,30	519,74	69,25	483,34	66,80	443,75	84,62
Total	295.598	100	471,09	71,40	528,84	76,59	495,03	72,76	459,34	100,80

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

O Censo Demográfico de 2010 mostrou que quanto a composição racial brasileira, a população negra (pardos e pretos) figura-se como maioria (VOLPE, *et al.* 2012). Soares (2008) analisou que isso se deve a ampliação de indivíduos que se reconhecem como negros, pois se fosse levado em consideração apenas o impacto da diferença das taxas de

fecundidade, os negros só seriam maioria no ano de 2020.

O Presidente do INEP, Francisco Soares, atribui a elevação da participação de negros no ENEM à Lei Nº 12.711/2012 (Lei de Cotas para o Ensino Superior), uma oportunidade que antes não estava aberta, mas que agora trouxe os alunos negros para o exame que é tido como a porta de entrada para a educação superior (MEC, 2014).

Quanto ao desempenho estudantil por raça, especificamente para a etnia negra, os resultados corroboram com os de Barbosa (2014) e Benevides e Soares (2016). Para Cordeiro (2007), os negros após longo período de escravidão não conseguiram ascensão econômica ou social, ocupando, de modo geral, nichos menos privilegiados na sociedade como um todo, sendo poucas as exceções. Já para os indígenas, a autora destacou que foi criado um estereótipo de hostilidade e ociosidade, sendo a passagem para a condição de índio civilizado um resultado de pressões de toda ordem que os obrigaram a modificar o modo de vida e a aderir as práticas capitalistas que possibilitaram condições mínimas de sobrevivência.

A seguir, apresenta-se o rendimento familiar declarado pelos participantes e notou-se que a maioria das famílias sobrevive com rendimento mensal baixo. Como pode ser visualizado na Tabela 03, o maior grupo, 56,61% da amostra, representa aqueles que declararam renda mensal de até 1 salário mínimo (até R\$ 724,00). No que diz respeito a performance dos participantes do exame, notou-se, de modo geral, que quanto mais elevado for o rendimento mensal da família, maior será o desempenho médio do estudante.

Percebe-se que 83,65% das famílias possuem rendimento mensal de até 2 salários mínimos. O resultado corrobora com o de Barbosa (2014), que vislumbrou que parcela majoritária (70%) das famílias de estudantes que realizaram o ENEM 2011, possuía renda familiar de até 2 salários mínimos. Para o autor, um cenário de baixa renda limita o investimento familiar em capital humano, o que possivelmente compromete a formação e a distribuição espacial deste tipo de capital, e consequentemente compromete o desempenho estudantil, principalmente nas regiões onde a pobreza se faz mais presente. Ainda conforme o autor, o Nordeste concentra mais de 50% das famílias que não possuem nenhuma renda ou possuem renda menor ou igual a 1 salário mínimo, ao passo que a Região Sudeste concentra mais de 50% das famílias que possuem renda a partir de 2 salários mínimos, sendo que mais da metade dos que declararam renda superior a trinta salários mínimos se encontra nesta mesma região.

Tabela 03: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com a renda familiar

Renda Familiar	Provas									
	Fi	%	CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Sem Renda	8.685	2,88	445,45	57,22	493,38	69,24	463,46	67,44	421,07	74,01
Até 1 salário mínimo	170.495	56,61	454,09	59,39	508,66	69,85	477,48	68,07	434,19	78,82
De 1 até 1,5	56.849	18,88	473,85	66,87	538,90	71,83	504,55	68,35	463,77	93,91
De 1,5 até 2	24.579	8,16	485,46	70,94	551,11	72,60	513,95	69,05	480,71	102,53
De 2 até 2,5	10.308	3,42	501,07	76,53	566,84	71,66	526,68	68,90	502,30	110,56
De 2,5 até 3	7.969	2,65	510,13	80,86	574,43	72,60	534,02	68,43	515,10	117,58
De 3 até 4	5.921	1,97	523,49	81,17	586,39	70,85	541,66	67,99	533,79	123,52
De 4 até 5	4.656	1,55	537,87	85,62	596,47	70,73	551,87	69,34	558,11	130,59
De 5 até 6	2.889	0,96	549,30	84,64	601,83	69,75	557,75	65,45	566,83	133,43
De 6 até 7	1.853	0,62	560,88	87,98	609,57	71,51	565,88	63,61	591,75	137,95
De 7 até 8	1.065	0,35	573,20	84,48	615,68	70,37	570,98	63,80	603,82	135,89
De 8 até 9	926	0,31	563,81	83,20	611,73	68,74	564,00	68,39	591,80	132,87
De 9 até 10	889	0,30	582,28	84,30	622,83	66,24	576,31	61,38	625,92	135,72
De 10 até 12	925	0,31	584,38	84,99	626,22	65,08	579,05	64,91	629,80	133,24
De 12 até 15	1.026	0,34	590,22	84,48	628,36	64,11	582,00	62,14	637,52	138,88
De 15 até 20	848	0,28	592,43	83,55	632,39	66,12	585,91	63,34	645,47	135,71
Mais de 20 salários mínimos	1.276	0,42	596,09	87,93	631,37	67,04	585,12	65,29	654,93	139,93
Total	301.159	100	471,19	71,61	528,89	76,75	495,03	72,92	459,43	101,09

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

* O salário mínimo do ano era de R\$ 724,00

Na Tabela 04 estão expostos os resultados referentes à escolaridade da mãe dos participantes do exame. A maioria das mães de participantes, mais precisamente 29,61% da amostra, só cursou até o antigo primário (da 1ª a 4ª série do ensino fundamental). Analogamente à influência proporcionada pela escolaridade do pai, a escolaridade da mãe também influencia no desempenho do aluno. Logo, quanto mais elevado for o grau de instrução da mãe, mais elevado será o desempenho do aluno no exame.

Os resultados corroboram com a literatura que afirma que o grau de instrução dos pais influencia no rendimento dos estudantes. Ramos e Reis (2008) vislumbraram claramente, ao comparar indivíduos com o mesmo grau de escolaridade, que o nível educacional é maior no caso de indivíduos com pais mais escolarizados e as médias de rendimento do trabalho aumentam de acordo com o nível educacional dos pais, concluindo, que as diferenças de desempenho educacional tendem a ser transmitidas de geração para geração contribuindo para a persistência dos níveis de rendimento familiar. Ademais, os autores perceberam também que o impacto da escolaridade da mãe é ligeiramente maior que o impacto da escolaridade do pai.

Os números também corroboram com os resultados de Benevides e Soares (2016). Os autores vislumbraram que o impacto sobre o desempenho do aluno é positivo e crescen-

te à medida que a mãe completa os ciclos educacionais do ensino primário, fundamental, médio e superior, respectivamente. Logo, isso reflete que o nível de instrução dos pais é de fato importante para o melhor desempenho escolar dos alunos, principalmente o grau de instrução da mãe, que possui papel fundamental no processo de aprendizagem, visto que a mesma é tida, geralmente, como a figura central na educação dos filhos em função do maior acompanhamento e da maior cobrança.

Tabela 04: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com a escolaridade da mãe

Escolaridade da Mãe	Provas									
	Fi	%	CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Da 1ª a 4ª série do Ensino Fundamental	83.920	29,61	456,95	60,91	513,74	72,67	480,12	69,56	436,58	82,29
Da 5ª a 8ª série do Ensino Fundamental	64.678	22,82	461,87	63,36	520,87	72,36	488,09	69,19	447,22	87,32
Ensino Médio Incompleto	17.688	6,24	474,26	69,46	536,07	73,57	503,87	69,04	466,24	97,47
Ensino Médio Completo	53.682	18,94	487,55	73,82	550,13	73,32	515,40	68,63	483,23	105,73
Ensino Superior Incompleto	6.423	2,27	506,72	82,59	566,29	76,33	529,44	71,84	510,00	122,62
Ensino Superior Completo	15.534	5,48	519,57	88,62	573,68	78,04	535,63	73,20	530,13	133,38
Pós-Graduação	9.501	3,35	533,24	89,56	585,31	76,91	546,52	71,75	550,75	135,08
Não sabe	31.987	11,29	449,48	62,18	499,54	71,80	468,91	69,90	430,80	83,67
Total	283.413	100	472,28	72,07	530,06	76,85	496,35	72,90	461,33	102,01

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Na Tabela 5 constam os resultados para o tipo de escola de ensino médio dos participantes cearenses do ENEM 2014. Os números indicam que 85,72% dos candidatos do exame cursavam/cursaram o ensino médio em uma escola da rede pública, mas foram as escolas da iniciativa privada que registraram o melhor desempenho médio dos alunos em todas as provas.

Tabela 5: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com a tipo de escola do ensino médio

Tipo de escola	Provas									
	Fi	%	CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Pública	163.704	85,72	468,61	64,21	531,07	72,72	495,69	69,60	453,80	90,81
Privada	27.268	14,28	549,17	87,64	602,55	70,48	558,77	66,29	564,83	136,13
Total	190.972	100	480,91	74,20	541,98	76,81	505,38	72,75	470,86	106,91

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Na Tabela 06, expõem-se de os resultados para a forma como os participantes

cursaram o ensino fundamental, por tipo de escola. Nota-se que a maioria dos participantes que prestou o exame cursou os anos iniciais da educação básica somente em escola pública (76,05%). Em termos de desempenho médio, estes ficaram atrás de todos os demais estratos.

Tabela 06: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com o tipo de ensino fundamental

Tipo de Ensino Fundamental	Fi	%	Provas							
			CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Somente em Escola Pública	100,21	76,05	460,13	61,86	523,83	72,56	485,44	70,34	441,98	84,36
Maior Parte em Escola Pública	12,70	9,64	482,82	67,05	556,42	70,10	516,07	66,84	476,90	98,05
Somente em Escola Particular	12,21	9,26	517,75	79,09	585,62	72,04	543,45	68,91	527,30	121,03
Maior Parte em escola Particular	6,65	5,05	493,61	70,08	566,10	71,29	526,64	66,94	491,81	103,66
Total	131,77	100	469,81	67,32	535,36	75,20	496,37	72,45	456,48	95,16

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Em relação ao Ensino Médio, na Tabela 07 notou-se que a maioria dos participantes que prestou o exame também cursou esta etapa da educação somente na rede pública (87,98%). Quanto a performance no exame, esse estrato também registrou o menor desempenho médio.

Tabela 07: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com o tipo de ensino médio

Tipo de Ensino Médio	Fi	%	Provas							
			CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Somente em Escola Pública	116.666	87,98	463,63	63,19	528,67	73,06	490,33	70,75	448,09	88,05
Maior Parte em escola Pública	3.885	2,93	482,71	70,12	554,53	75,23	511,42	70,60	470,40	100,51
Somente em Escola Particular	9.807	7,40	526,83	79,81	594,07	70,47	549,92	68,29	534,19	123,83
Maior Parte em Escola Particular	2.241	1,69	501,28	73,68	573,70	72,67	528,81	69,07	499,88	109,10
Total	132.599	100	469,69	67,32	535,23	75,22	496,20	72,48	456,28	95,10

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Em relação a maior frequência de estudantes oriundos da educação pública no ENEM, tal resultado pode ser explicado pelo próprio objetivo do exame, que é o de democratizar o acesso à educação superior para todos os brasileiros. Além do mais, a Lei N° 12.711/2012 (Lei das Cotas para o Ensino Superior) pode ter influência sobre os resultados, visto que além de garantir vagas por condições socioeconômicas, também garante um determinado número de vagas para alunos oriundos da educação pública.

Quanto às diferenças de desempenho, conforme Oliveira, Silva e Siqueira (2008), as disparidades entre o ensino público e o privado são visíveis e demonstram uma condição real da educação brasileira, onde estudantes da educação privada possuem uma maior preparação para vestibulares do que estudantes da educação pública. Ao avaliar os determinantes do desempenho de estudantes no processo seletivo da Universidade Federal da Paraíba (UFPB) em 2006, os autores vislumbraram que o tipo de escola onde os alunos estudaram influenciou no desempenho e, conseqüentemente, na aprovação no processo seletivo, pois os que estudaram na rede privada tiveram maior aproveitamento no vestibular em cerca de 56%.

Ainda conforme Oliveira, Silva e Siqueira (2008), o resultado pode ser explicado pela melhor qualidade do ensino das escolas privadas em comparação ao ensino oferecido pelas escolas públicas e quiçá maior quantidade de horas aulas no ensino médio. O projeto do ministério da educação, que prevê uma série de mudanças no ensino médio, contempla a elevação de 800 para 1000 horas anuais nas escolas públicas de ensino médio enquanto atualmente as melhores escolas particulares já oferecem 1600 horas.

Na Tabelas 08 apresentam-se os resultados por tempo para a conclusão do ensino médio, respectivamente. Notou-se que a maior parcela da amostra corresponde aos participantes que concluíram esta etapa da educação básica em 3 anos (43,52%), vindo logo em seguida a parcela de estudantes que declararam não ter concluído, os denominados treineiros que não concorrem as vagas nas instituições de ensino superior, realizando o exame apenas para efeito de conhecimento. No que diz respeito ao desempenho médio no exame, os indivíduos que concluíram o ensino médio no período exato de 3 anos apresentam resultados melhores que aqueles que terminaram a educação básica em menos tempo e que aqueles que terminaram em mais tempo. O estrato composto pelos treineiros foi o que registrou a menor performance, exceto na prova de matemática, embora o resultado tenha sido próximo à performance dos que registraram o menor desempenho nesta prova (grupo dos participantes que concluíram o ensino médio em 6 anos ou mais).

Tabela 08: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com o tempo para concluir o ensino médio

Tempo para concluir o Ensino Médio	Provas									
	Fi	%	CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Menos de 3 anos	6.832	2,28	472,96	69,93	537,86	77,23	495,32	75,43	453,21	101,46
3 anos	130.231	43,52	486,17	76,20	549,15	76,36	512,01	72,21	479,02	110,47
4 anos	9.946	3,32	475,29	68,66	542,34	75,85	501,09	71,55	462,24	98,17
5 anos	2.377	0,79	473,19	65,04	544,49	72,82	497,90	72,04	451,71	89,38
6 anos ou mais	2.610	0,87	466,64	63,25	537,34	74,35	491,87	69,23	443,21	84,95
Não concluiu	147.223	49,20	458,72	65,47	510,93	72,68	480,86	70,48	443,90	90,21
Total	299.219	100	471,28	71,62	529,00	76,76	495,13	72,92	459,56	101,15

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Em linhas gerais, o baixo desempenho estudantil dos alunos que terminaram tanto o ensino fundamental como o ensino médio em mais tempo pode estar associado a deficiências no processo de aprendizagem que resultou em repetência de séries, principalmente em disciplinas mais sistematizadas como é o caso da matemática e das ciências da natureza (química, física e biologia), atrasando o término das etapas no período regular. Quanto aos treineiros, o baixo desempenho em relação aos demais estudantes pode ser explicado pelo fato destes não terem ainda concluído os anos finais da educação básica, estando, portanto, com menos bagagem de conteúdo.

A Tabela 09 apresenta os resultados por dependência administrativa. O sistema educacional brasileiro é composto, dentre outras características, por um conjunto de instituições de ensino que são distribuídas entre quatro dependências, são elas: a federal, a estadual, a municipal e a privada. De acordo com os dados abaixo, percebe-se que a maioria dos estudantes são oriundos da dependência administrativa estadual, mas foram os estudantes da rede federal que obtiveram o melhor desempenho médio no exame.

Os resultados corroboram com os de Barbosa (2014), que vislumbrou que parcela majoritária dos estudantes do ENEM 2011 (75,82%) declarou ser oriunda de escolas da rede estadual. O autor atribuiu a esse resultado à missão dos estados de priorizar o ensino médio para todos os estudantes que o demandam, tal como consta na Lei das Diretrizes e Bases da Educação Nacional (LDB). Ademais, o mesmo também identificou que a dependência administrativa da escola tem repercussão sobre a performance do aluno e concluiu que o melhor desempenho médio também ficou a cargo da rede pública federal.

Tabela 09: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com a dependência administrativa

Dependência	Provas									
	Fi	%	CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Federal	348	0,42	563,70	79,85	615,49	58,41	572,15	57,55	603,22	124,87
Estadual	73.028	88,01	456,19	59,62	507,82	70,37	477,54	68,21	436,97	81,00
Municipal	162	0,20	449,49	53,08	513,44	66,79	458,49	70,18	420,86	58,55
Privada	9.439	11,38	544,88	87,66	591,03	74,61	550,80	67,24	558,77	137,81
Total	82.977	100	468,27	71,00	519,20	76,51	487,66	72,66	453,88	100,29

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Na Tabela 10 são apresentados os resultados de acordo com o local da escola dos participantes, respectivamente.

Quanto aos resultados para a localização da escola dos participantes, a maior parcela das escolas situa-se na zona urbana (97,32%), enquanto a zona rural possui apenas 2,68% do total de escolas. Quanto à performance no exame, são os participantes oriundos de instituições da zona urbana que possuem o melhor desempenho médio.

Tabela 10: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com a localização da escola

Local da Escola	Provas									
	Fi	%	CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Zona Urbana	80.752	97,32	468,81	71,27	519,87	76,66	488,35	72,70	454,54	100,79
Zona Rural	2.225	2,68	448,08	56,50	493,95	65,69	461,53	66,10	428,61	74,84
	82.977	100	468,27	71,00	519,20	76,51	487,66	72,66	453,88	100,29

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

O baixo desempenho estudantil logrado pelos estudantes oriundos da zona rural, quando comparados aos alunos da zona urbana, pode ser explicado, entre outros motivos, pela herança histórica resultante do passado escravocrata do Brasil, sendo possível também que haja uma menor qualificação dos docentes nessas escolas.

Até meados da década de 1920 existia a defesa da diferenciação das escolas rurais das urbanas, que se associava a ideia de que a educação destinada para as crianças do campo não precisava possuir o mesmo nível de elaboração da educação oferecida para as crianças da zona urbana, já que no meio rural a chegada da época das colheitas prejudicava a frequência escolar (VICENTINI e GALLEG0, 2006). Em conformidade, Aliardi e Westermann (2012), afirmaram que até 1934 a educação no campo foi negligenciada. Ademais para os mesmos a infraestrutura precária, a ausência de profissionais qualificados

em virtude dos baixos salários e das condições de trabalho desfavoráveis, as escolas mal distribuídas no espaço e os currículos baseados na cultura descontextualizados compõem o cenário da educação no campo.

Os resultados mencionados anteriormente corroboram com os encontrados por Barbosa (2014) que vislumbrou que a maioria dos estudantes do ENEM no ano de 2011 residia no meio urbano (87,23%) e que a maioria de escolas (97,48%) se situavam na zona urbana. O mesmo identificou também que o efeito de residir na zona urbana proporciona um melhor desempenho aos estudantes e que as escolas localizadas na zona rural refletiam negativamente no desempenho estudantil.

Na Tabela 11, encontram-se descritos os resultados por tempo de escola. Observou-se que os estudantes das escolas de tempo convencional representam parcela majoritária da amostra, com 97,22% do total, ao passo que as escolas de tempo integral correspondem a apenas 2,78% dos estudantes que realizaram o exame em 2014 no Estado do Ceará. Apesar disso, estes estudantes das instituições de jornada escolar prolongada perfizeram melhor desempenho médio que aqueles das demais instituições de tradição. O resultado aponta para a eficácia desta política no ensino médio, mas é necessário realizar uma análise mais profunda, que será feita adiante com a metodologia PSM.

Tabela 11: Distribuição de frequências e medidas de dispersão dos participantes Cearenses do ENEM em 2014 de acordo com o tempo de escola

Tempo de Escola	Fi	%	Provas							
			CN		CH		LC		MT	
			Média	DP	Média	DP	Média	DP	Média	DP
Tradicional	292.777	97,22	470,43	71,76	527,98	77,01	494,19	73,21	458,48	101,13
Tempo Integral	8.382	2,78	491,86	63,79	553,59	64,56	517,29	60,63	484,91	96,51
Total	301.159	100	471,19	71,61	528,89	76,75	495,03	72,92	459,43	101,09

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Os resultados a cima podem indicar resquícios dos investimentos realizados na política de educação integral via educação técnica, foco do Governo do Estado do Ceará, e coincidem com o que foi identificado pelo IPECE (2013), que analisou a participação das escolas públicas estaduais cearenses no ENEM de 2009 a 2011. O Instituto, observou que o desempenho destas instituições de ensino teve uma leve melhora, evoluindo de 448,8 pontos para 449,5 pontos, devido à participação dos alunos das EEEPs, principalmente no ano de 2011, pois se considerado apenas a média destas instituições de ensino no exame (480,8) o Ceará ascenderia para a 8ª posição na classificação geral do ENEM 2011, ao pas-

so que se fosse considerado apenas a média das escolas não profissionalizantes (442,5) o Ceará seria rebaixado para a penúltima posição, à frente apenas do Maranhão. Mas, ainda segundo o órgão, o aumento substancial de participantes de escolas de ensino médio não profissionalizantes (de 4,72% em 2009 para 7,60% em 2011) pode ter puxado a média destas escolas para baixo, já que o ENEM pode incentivar tanto a participação dos estudantes de alto rendimento como dos estudantes de baixo rendimento.

4.2 Propensity Score Matching (PSM)

Para fazer a análise do impacto da política de educação em tempo integral nas escolas profissionalizantes de ensino médio do Estado do Ceará, foi realizado em um primeiro momento a estimação de regressões logísticas *Probit* para cada prova objetiva do Exame Nacional do Ensino Médio (ENEM) do ano de 2014 a fim de se obter a probabilidade de cada estudante fazer parte da política estadual.

Com a determinação do escore de propensão do indivíduo fazer parte, ou não, da política, seguiu-se para a formação dos pares entre estudantes de escolas profissionalizantes (tempo integral) e estudantes de escolas não-profissionalizantes (educação tradicional) com as características mais semelhantes possíveis. Com esse procedimento foi possível analisar o impacto da política sobre o desempenho dos estudantes no ENEM de 2014. Além do mais, é importante destacar que todos os testes estatísticos foram realizados considerando um nível de significância de 5%.

4.2.1 Resultados do modelo *probit*

Na Tabela 12 e apresenta-se a regressão estimada com as variáveis utilizadas no pareamento entre estudantes de escolas de tempo integral e estudantes de escolas de tempo convencional. Foi realizada a análise da significância dos coeficientes e a análise dos sinais dos coeficientes que são significantes estatisticamente para determinar se as variáveis da regressão apresentam influência positiva ou negativa na probabilidade de o indivíduo ser um aluno de escola de tempo integral. Logo, o modelo estimado é empregado apenas para a determinação dos vizinhos para o pareamento e estimação dos ATT's.

Tabela 14: Regressão probit para as provas de CN, CH, LC e de MT

Descint	Coefficiente	Erro padrão	Z	P> z
Idade	-0,2139719	0,0242249	-8,83	0,000
Dlocalização	0,2865483	0,2811422	1,02	0,308
Dependência	-0,1087709	0,1082894	-1,00	0,315
Dsexo	0,2132954	0,0564298	3,78	0,000
Dtrabalha	-0,8009526	0,0711228	-11,26	0,000
Destadocivil	0,1266475	0,2240149	0,57	0,572
Dtipoensino	0,1633828	0,0845077	1,93	0,053
Draça	0,0840526	0,0911172	0,92	0,356
Escolarmãe	-0,0018958	0,0111268	-0,17	0,865
Dlocalresid	-0,088297	0,0762545	-1,16	0,247
Internet	-0,1256033	0,0704969	-1,78	0,075
Tempofundamental	0,0418783	0,0249579	1,68	0,093
Tempoensinomedio	-0,0348342	0,0143856	-2,42	0,015
Dtipoescolanoensmedio	-0,03557634	0,1054342	-3,37	0,001
Dtipoescolanoensfund	0,1216239	0,0295616	4,11	0,000
Renda	-0,0135607	0,0187066	-0,72	0,469
Residentes	-0,021703	0,0190739	-1,14	0,255
Constante	2,655644	0,6239897	4,26	0,000

Número de Observações = 6644
Log likelihood = -1231,1927
PseudoR² = 0,1292

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Os coeficientes das variáveis quantitativas escolaridade da mãe, renda familiar e tamanho da família não apresentaram valores $P > |z|$ dentro do limite aceitável e, desta forma, nada se pode dizer do efeito dessas variáveis sobre a variável dependente. Por outro lado, o coeficiente da variável idade mostrou-se significante estatisticamente a 1% com sinal negativo e, deste modo, significa dizer que quanto mais elevada for a idade do indivíduo, menor será a probabilidade de ser estudante de escola de tempo integral.

Ademais, o coeficiente da variável tempo para conclusão do ensino médio mostrou-se significante estatisticamente a 5% e os coeficientes das variáveis acesso à internet e tempo para conclusão do ensino fundamental mostraram-se significante a 10%. A variável tempo para conclusão do ensino fundamental apresentou coeficiente com sinal positivo e, conseqüentemente, possui efeito positivo sobre a probabilidade de o indivíduo pertencer a uma escola de tempo integral. Já quanto às variáveis tempo para a conclusão do ensino médio e acesso à internet, observou-se o oposto. Desta forma, indivíduos que levaram mais tempo para concluir o ensino fundamental possuem maior probabilidade de serem atendidos pela política, ao passo que indivíduos que levaram mais tempo para concluir o ensino médio e indivíduos que possuem mais meios de acesso à internet, possuem menor probabilidade de serem estudantes de escolas de tempo integral.

Em relação as variáveis binárias, os coeficientes para a localização da escola, para a dependência administrativa, para o estado civil, para a raça e para o local de residência

não apresentaram valores $P > |z|$ dentro do limite aceitável. Logo, implica dizer que não há diferença na probabilidade de pertencer à escola de tempo integral entre essas categorias e as respectivas categorias de referência.

Os coeficientes das *dummies* de gênero, de atuação em atividade remunerada, do tipo de escola de ensino fundamental e do tipo de escola de ensino médio mostraram-se significantes estatisticamente a 1%, ao passo que o coeficiente da *dummy* para o tipo de ensino mostrou-se significativo estatisticamente a 10%, indicando que há diferença na probabilidade de pertencer a uma escola de tempo integral a depender das categorias. Deste modo, significa dizer que homens, quem estudou maior quantidade de anos do ensino fundamental em escola pública e quem possui ensino ofertado pela modalidade de educação regular possuem maiores chances, ao passo que quem trabalha e quem estudou a maior quantidade de anos do ensino médio em escola pública possuem menores chances.

Acerca dos indicadores de ajuste, tem-se que os Pseudo R^2 apresentam valores satisfatórios para esses tipos de modelo e similar aos valores encontrados por Dantas e Tanuri-Pianto (2013). As estatísticas LR atesta a validade do modelo a 1% de significância. Ou seja, todas as variáveis explicativas em conjunto são importantes para a explicação da variável dependente.

4.2.2 Resultados do PSM

A tabela 13 apresenta o valor dos ATTs, que representa o efeito médio da política de educação em tempo integral sobre o desempenho dos alunos em cada área de conhecimento com base na probabilidade de serem ou não alunos da política, sem o pareamento e considerando o método de *Nearest Neighbor* (vizinho mais próximo).

Como pode ser observado, sem o método de pareamento há uma diferença média de 33,88 pontos na prova de Ciências da Natureza, de 38,73 pontos na prova de Ciências Humanas, de 31,84 pontos na prova de Linguagens e Códigos e de 53,73 pontos na prova de Matemática entre o grupo de tratamento (escolas de tempo integral) e o grupo de controle (escolas de tradição). Pelo método de pareamento, observa-se que a diferença média cai para 25,87 pontos na prova de Ciências da Natureza, para 28,60 pontos na prova de Ciências Humanas, para 22,24 pontos na prova de Linguagens e Códigos e para 40,30 pontos para a prova de Matemática.

Tabela 13: Efeito do grupo de tratamento

Método	Tratado	Controle	Diferença	Erro padrão	Estatística t
Sem pareamento para CN	489,99	456,112821	33,8789598	3,26031246	10,39
<i>Nearest Neighb para CN</i>	489,99	464,123288	25,868492	5,11528974	5,06
Sem pareamento para CH	552,14	513,411865	38,7319698	3,72077136	10,41
<i>Nearest Neighb para CH</i>	552,14	523,541369	28,6024657	5,56433841	5,14
Sem pareamento para LC	512,549046	480,707245	31,8418007	3,58650745	8,88
<i>Nearest Neighb para LC</i>	512,549046	490,30545	22,243596	4,93367474	4,51
Sem pareamento para MT	492,463215	438,7299	53,7333147	4,40569842	12,20
<i>Nearest Neighb para MT</i>	492,463215	452,167847	40,2953677	6,57445839	6,13

Fonte: Elaborado pelos autores com base nos microdados do ENEM (2014).

Diante disso fica evidente a necessidade do método para analisar com precisão a eficácia da política, tendo em vista que sem o pareamento há uma distorção significativa de 8 pontos na prova de CN, de aproximadamente 10 pontos nas provas de CH e de LC e de aproximadamente 13 pontos na prova de MT. Apesar disso, pôde-se perceber o êxito da política na melhoria do rendimento escolar dos estudantes do ensino médio da rede pública cearense, tendo em vista os ganhos acima da média para os estudantes das escolas de tempo integral em todas as provas avaliadas⁴.

CONSIDERAÇÕES FINAIS

Diante da importância atribuída pela literatura aos gastos com educação para o desenvolvimento das economias no longo prazo, vive-se um momento em que políticas educacionais ganharam papel de destaque por possuírem o poder de influenciar tanto a renda das gerações futuras, como o bem-estar da população. No Brasil, a política de educação integral apresenta-se como uma proposta que visa tanto a construção de uma nova identidade para a escola, ao rever o seu papel na vida escolar dos estudantes, como melhorar a ação das instituições sobre os resultados do alunado. No Estado do Ceará o governo investiu na educação integral via educação profissionalizante e, desde então, vem se destacando com reconhecimento nacional pelo bom desempenho da política.

Neste contexto, o presente trabalho, através da utilização do método de pareamento baseado em escore de propensão, analisou a eficiência da política de educação integral adotada pelo Ceará, por meio das Escolas Estaduais de Educação Profissional (EEEPs), utilizando os microdados do Exame Nacional do Ensino Médio (ENEM) referentes ao ano de 2014.

Acerca do perfil dos estudantes, pôde-se perceber que a maioria dos cearenses que

4. Foram realizadas estimativas do efeito de tratamento com bootstrap. Os resultados apontam que o viés é maior na prova de Ciências Humanas e menor em matemática. A análise da distribuição de Kernel para cada uma das provas mostrou que a diferença favorável ao tratamento ocorre em toda a distribuição e não somente na média.

se enquadrava nas seguintes estratos/características: gênero feminino, idade entre 16 e 20 anos, cor de pele parda, estado civil solteiro, nunca exerceu atividade remunerada, núcleo familiar composto por 4 ou 5 pessoas, renda de até 1 salário mínimo (R\$ 724,00), pais com baixa escolaridade (da 1ª a 4ª série do ensino fundamental), ensino médio em escola pública, conclusão do ensino fundamental e do ensino médio somente escola pública, conclusão do ensino fundamental e do ensino médio no tempo regular, ensino de modalidade regular, escola de dependência administrativa estadual, residência localizada na zona urbana, escola localizada na zona urbana, escola de educação tradicional e um meio de acesso à internet.

Ainda que que em fase de expansão, a política de educação em tempo integral atendia no ano de 2014 um universo de 8.282 alunos. Quanto ao retorno da política, notou-se um indicativo de que a mesma vinha se mostrando eficiente, pois percebeu-se através de estatística descritiva que o desempenho médio no exame dos estudantes provenientes das Escolas Estaduais de Ensino Profissionalizante (EEEPs) era superior ao desempenho médio dos estudantes das escolas de ensino tradicional.

Para a obtenção de dados mais robustos foi utilizado o PSM. Através desse método foi possível visualizar a diferença de desempenho entre os resultados com pareamento e sem pareamento, notando-se que sem o pareamento o efeito da política era sobrevalorizado. Feito o pareamento, constatou-se que o desempenho dos estudantes das escolas de tempo integral, ainda assim, é superior ao desempenho das demais escolas em todas as áreas do conhecimento, tendo em vista os ganhos acima da média em todas as provas avaliadas, com destaque para o ganho na performance da prova de Matemática.

Portanto, tendo em vista os resultados expostos pode-se aceitar a hipótese deste trabalho, ou seja, de que a educação em tempo integral é uma política eficiente para melhorar a ação escolar sobre o desempenho dos estudantes. Neste sentido, o Governo do Estado do Ceará pode considerar os investimentos nessa política como ações bem sucedidas tanto por melhorar o desempenho escolar dos alunos, como também por contribuir para a formação de capital humano local, uma vez que além de direcionar esforços para a proposta de educação integral, que defende o ensino de qualidade e o desenvolvimento dos jovens em todas as suas dimensões, acrescenta a educação técnica que oferece capacitação profissional a depender da área de interesse dos jovens que visam ingressar no mercado de trabalho a médio ou a longo prazo com alguma experiência técnica no currículo.

Ademais, tendo em vista que atualmente o mundo é regido pela revolução técnico-científica e por um mercado de trabalho cada vez mais desafiante e exigente, a política de educação integral apresenta-se como uma ação importante por preparar os jovens da

educação pública para o mundo adulto e para competir em pé de igualdade com os jovens preparados pela rede de ensino particular. Apesar do sucesso da política, essa ainda não abrange uma grande população alvo, sendo interessante a ampliação da mesma no sentido de melhorar cada vez mais os resultados educacionais e o processo de desenvolvimento do Estado do Ceará via educação.

REFERÊNCIAS

ALIARDI, R. T.; WESTERMANN, L. D.. Educação no Campo: integração entre escola e comunidade. **Revista Modelos - FACOS**, v. 02, p. 53-72, 2012.

ARAÚJO JÚNIOR, J. N.; JUSTO, W. R.; ROCHA, R. M.; GOMES, S. M. F. P. O. Eficiência Técnica das Escolas Públicas dos Estados do Nordeste: Uma Abordagem em dois Estágios. In: XXII Fórum Banco do Nordeste de Desenvolvimento e do XXI Encontro Regional de Economia. **Anais...** Fortaleza, CE, 2016.

ARROYO, M. G.. O direito ao tempo de escola. **Cadernos de Pesquisa**. São Paulo, v. 65, p. 3-20, Mai. 1988.

BARBOSA, W. F. **Educação no Brasil é para quem quer ou para quem pode?** Uma avaliação dos fatores determinantes do desempenho educacional dos estudantes brasileiros. 2014. Monografia (Graduação em Economia) – Universidade Regional do Cariri - URCA. Crato, CE, 2014.

BECKER, S. O.; ICHINO, A. **Estimation of average treatment effects based on propensity scores**. The Stata Journal (2002). 2, Number 4, pp. 358–377.

BENEVIDES, A. A.; SOARES, R. B.. Diferencial de desempenho das escolas militares: bons alunos ou boa escola?. In: **Anais do XXI Encontro Regional de Economia**. Fortaleza, 2016.

BRINGMANN, G.; JUNG, M. M.; WELTER, E.; PRATES, F. L.; HENZ, L. F.; CRUZ, L. R.. A população acadêmica da terceira idade na Universidade de Santa Cruz do Sul. **Revista Brasileira de Ciências do Envelhecimento Humano**, v. 8, p. 148-157, 2011.

BRITO, A. E. C. A.; QUIRINO, D. L. C.; PORTO, L. C. M.. Educação Especial e Inclusiva no Ensino Superior. **Revista de Educação**. São Paulo, v. 16, n. 20/21, p. 14-20, 2013.

CARVALHO, M. C. B. O lugar da educação integral na política social. **Cadernos CENPEC**. São Paulo, v. 5.000, p. 7 - 11, 01, Out. 2006.

CAVALIERE, A. M. Educação Integral: uma nova identidade para a escola brasileira?.

Educação e Sociedade. São Paulo, v. 23, n. 81. Dez. 2002.

_____. Tempo de escola e qualidade na educação pública. **Educação e Sociedade**. São Paulo, v. 28, p. 1015-1035. Out. 2007.

CENTRO DE REFERÊNCIA EM EDUCAÇÃO INTEGRAL. **Ceará foca em programa de Educação Integral profissionalizante**. 2015.

_____. **Educação em Tempo Integral**. 2014.

CETIC - Centro de Estudos sobre as Tecnologias da Informação e da Comunicação. O uso da Internet por alunos brasileiros de Ensino Fundamental e Médio. **Panorama setorial da Internet**, v.5, n° 2. Ago. 2013.

DANTAS, R. S.; TANNURI-PIANTO, M. E. Avaliação de Impacto de Reconhecimento de Direito de Propriedade de Facto: uma Análise de Propensity Score Matching. In: Encontro Nacional de Economia. **Anais...** Foz do Iguaçu, PR, 2013.

DUARTE, G. B.; SAMPAIO, B.; SAMPAIO, Y. Programa Bolsa Família: impacto das transferências sobre os gastos com alimentos em famílias rurais. **Revista de Economia e Sociologia Rural**. Piracicaba, vol. 47, n° 04, p. 903-918, Out/Dez 2009.

DUTRA, P. F. V. Educação Integral no Estado de Pernambuco: Uma realidade no ensino médio. In: IV Congresso Ibero-Americano de Política e Administração da Educação/VII Congresso Luso Brasileiro de Política e Administração da Educação. **Anais...** Porto, Portugal, 2014.

GONÇALVES, A. S.. Reflexões sobre Educação Integral e Escola de Tempo Integral. **Cadernos CENPEC**. São Paulo, v. 1, p. 129-135, 2006.

IPECE - Instituto de Pesquisa e Estratégia Econômica do Ceará. **Análise da Participação das Escolas Públicas Estaduais Cearenses no Exame Nacional do Ensino Médio (ENEM): 2009-2011**. Informe nº 54 - Fevereiro. Fortaleza, CE, 2013.

LIMA, M. C. B. A Educação como Direito Fundamental. **Revista da Escola da Magistratura do Estado do Rio de Janeiro**. Rio de Janeiro, v. 04, n.13, p. 212-233, 2001.

MEC - Ministério da Educação. **Educação Integral**. Acesso: 23 Out 2016a.

_____. **Ensino Fundamental**. Disponível: Acesso: 23 Out 2016b.

_____. **Ensino Médio**. Acesso: 23 Out 2016c.

_____. **Maior parte dos candidatos é formada por mulheres e negros**. Outubro, 2014.

_____. **Resumo Técnico - Censo Escolar 2010**. Brasília, 2010.

OLIVEIRA, I. S. V.; SILVA, M. V. B.; SIQUEIRA, L. B. O.. Determinantes do desempenho dos estudantes no vestibular da Universidade Federal da Paraíba. **Economia e Desenvolvimento** (Recife), v. 07, p. 102-120, 2008.

PEREIRA, A. F. C.; JUSTO, W. R.; LIMA, J. R. F. Impactos das Rendas Não-Agrícolas sobre os Indicadores de Pobreza Foster-Greer-Thorbecke (Fgt) para as Famílias Rurais do estado do Ceará. In: BARRETO, F. A. F. D.; MENEZES, A. S. B.; ASSIS, D. N. C.; SOUSA, F. J. (Org.). **Economia do Ceará em Debate 2015**. Fortaleza, v-1, p. 325- 348.

RAMOS, L.; Reis, M. C.. A escolaridade dos pais e os retornos à educação no mercado de trabalho. IPEA, Nota Técnica, Rio de Janeiro, 2008.

ROSENBAUM, P. R; RUBIN, D. B. **The Central Role of the Propensity Score in Observational Studies for Causal Effects**. Biometrika, Vol. 70, No. 1. 1983. p. 41-55.

SAAVEDRA, L.. **Sucesso/insucesso escolar**: A importância do nível socio-económico e género. Psicologia, 2001, p. 67-92.

SEDUC – Secretaria de Educação. **Educação Profissional - Criação das ETECs**. Fortaleza, 2017.

SEE - Secretaria de Educação do Estado do Pernambuco. **Educação Integral?**. Disponível em: <<http://www.educacao.pe.gov.br/porta1/?pag=1&men=70>>. Acesso em: 23 Out 2016a.

_____. **Pernambuco alcança melhor ensino público do Brasil**. Pernambuco, 2016.

SILVA, H. G. Fatores Determinantes do Desempenho Acadêmico no Ensino Superior: Uma Abordagem por Meio do Estado da Arte. In: **Anais do VIII Encontro de Pesquisa em Educação e II Conferência Internacional de Trabalho Docente e Processos Educativos**. Minas Gerais, 2015.

SILVÉRIO, R. A.; ALBUQUERQUE, F. C. A. A gestão na escola de educação profissional integrada no Ceará: análise do modelo gerencial de resultados. In: IV Fórum Internacional de Pedagogia. **Anais...** Parnaíba, PI, 2012. Disponível:

SOARES, J. F. O efeito da escola no desempenho cognitivo de seus alunos. **Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**. Madrid, Espanha, vol. 2, núm. 2, p. 83-104, 2004.

SOARES, S.. A demografia da cor: a composição da população brasileira de 1890 a 2007. In: THEODORO, M.; JACCOUD, L.; OSÓRIO, R.; SOARES, S. (Org). **As políticas públicas e a desigualdade racial no Brasil**: 120 anos após a abolição. IPEA. Brasília, P. 97-118, 2008.

VICENTINI, P. P.; GALLEGO, R. C.. Escolas primárias urbanas e rurais: um estudo dos debates acerca de suas especificidades na organização do sistema de ensino paulista (1890 - 1945). In: IV Congresso Brasileiro de História da Educação, 2006, Goiânia. **Segundo Caderno de Resumos do IV CBHE**. Goiânia: Ed. da UCG/Ed. Vieira, 2006. p. 94-95.

VOLPE, A. P. S.; GOES, F. L.; LOBO, M. S. S. H.; SILVA, T. D.. Igualdade Racial. In: IPEA. **Políticas Sociais: acompanhamento e análise**. IPEA, v. 1, p. 313-368, 2012.

DILEMA EQUIDADE-EFICIÊNCIA PROPOSTA METODOLÓGICA E EVIDÊNCIAS EMPÍRICAS PARA O CEARÁ

Lyana Maria França da Costa Ribeiro*

Christiano Modesto Penna**

RESUMO

Criamos uma conexão entre os testes empíricos das hipóteses de convergência e da curva de Kuznets, e mostramos que esse aparato nos permite encontrar uma medida de “desigualdade de renda de estado estacionário”. Mostramos que essa medida, juntamente com a renda *per capita* associada à mesma, estabelecem prontamente o que se conhece na literatura por “*trade-off equidade-eficiência*”. Mostramos em que condições o referido *trade-off* passa a valer, assim como explicamos a possibilidade de ocorrência de diferentes tipos de relações entre crescimento e desigualdade. Mais especificamente, mostramos que estas ocorrências guardam relação com a renda per capita de estado estacionário e a renda *per capita* associada ao ponto de máximo da curva de Kuznets. Posteriormente, conduzimos uma análise com enfoque nos municípios cearenses.

A eficiência indica que a sociedade está obtendo o máximo que pode de seus recursos escassos, enquanto equidade significa que os benefícios advindos desses recursos estão sendo distribuídos com justiça entre os membros da sociedade. em outras palavras: a eficiência refere-se ao tamanho do bolo, já a equidade, à maneira como o bolo é dividido [mankiw (2004)].

1 INTRODUÇÃO

Em 1975, Arthur Okun publicou o livro *Equality and Efficiency: The Big Tradeoff*. A expressão inglesa “*trade-off*” define um dilema, ou seja, uma situação em que há conflito de escolhas. Este conflito surge do fato de que, em alguns casos específicos, a resolução de um problema acaba acarretando em outro.

A ideia de *trade-off* sempre é apresentada nos livros de introdução a economia, desde o livro introdutório de Samuelson e Nordhaus, até o *best-seller* de Gregory Mankiw. No caso específico do *trade-off* equidade-eficiência, o que se pontua é que as sociedades

* TV Diário – Sistema Verdes Mares. (85) 99854-5545. email: lyanaribeiro30@gmail.com.

** Professor do MAER e do CAEN/UFC, (85) 99736-6006. email: cmp@caen.ufc.br.

devem escolher entre uma economia eficiente ou uma sociedade igualitária.

A explicação para esse dilema é relativamente simples, e está calcada no Primeiro Teorema do Bem-Estar: em uma economia que se baseia na iniciativa privada, os esforços públicos para promover a igualdade representam uma interferência deliberada nos resultados gerados pelo livre mercado; assim, se as políticas pró-igualdade distorcem a alocação (eficiente) de mercado, o desempenho econômico só pode ser melhorado à custa de uma distribuição de renda menos equitativa.

A existência do dilema equidade-eficiência finda numa relação negativa entre crescimento econômico e igualdade de renda: ou o país cresce (devido ao uso eficiente de seus recursos), ou há uma melhor distribuição de renda.

No final dos anos 90, diversos estudos adentraram nessa questão. Do lado teórico, o *trade-off* tentava ser explicado através da interferência, por parte do governo, na maneira como a renda era distribuída. Os trabalhos que buscaram explicar esse processo eram basicamente calcados em modelos de economia política com enfoque no repasse da arrecadação de tributos [Alesina e Rodrick (1994), Persson e Tabellini (1994), Benabou (1996), Benhabib (2003)]. Houve, também, modelos que adentraram em questões como comércio [Martin (1998)] e tecnologia [Aghion, Caroli e García-Peñalosa (1999)].

Do ponto de vista empírico, os primeiros estudos tendem a apoiar prontamente o *trade-off* [ver Benabou (1996) para uma revisão desses estudos], no entanto, como sustenta Voitchovsky (2005), essas análises, com base em dados de corte transversal, eram muito sensíveis à inclusão de *dummies* regionais e à seleção amostral [Perotti (1996) e Partridge (1997)].

Com base nos dados de Deininger e Squire (1996),¹ alguns trabalhos passaram a relatar um efeito positivo da desigualdade de renda sobre o crescimento econômico, utilizando uma amostra diversificada de países desenvolvidos e em desenvolvimento [ver Li e Zou (1998) e Forbes (2000)].

Explorando melhor essa possibilidade, Barro (2000) constatou que a desigualdade parece encorajar o crescimento apenas nos países ricos, e tende a ter um efeito negativo sobre os países mais pobres. Banerjee e Duflo (2003) permitem uma não-linearidade na relação crescimento-desigualdade e encontram que uma mudança na desigualdade - em qualquer direção - poderia prejudicar o crescimento.

1. Essa base de dados tornou possível reduzir o erro de medição nas estatísticas de desigualdade, controlar a heterogeneidade invariante não observada no tempo entre países e usar técnicas de painel para mitigar as preocupações de endogeneidade.

Enquanto os estudos *cross-country* retornavam uma relação negativa entre crescimento e desigualdade, estudos com dados em painel apontavam que essa relação era costumeiramente positiva. Com base em um painel para os estados dos EUA, Panizza (2002) usa estimadores de efeito fixo e GMM e encontra evidência de uma relação negativa entre desigualdade e crescimento. Entretanto, o trabalho mostra que essa relação não é robusta e que pequenas diferenças no método utilizado para medir a desigualdade podem resultar em grandes diferenças na relação estimada entre desigualdade e crescimento.

Ehrhart (2009) faz um levantamento detalhado da literatura teórica e empírica que analisa efeito da desigualdade no crescimento e pontua, assim como Bourguignon (1998), que seria um exagero considerar que a redistribuição de renda e / ou de ativos produtivos seriam uma panaceia para o crescimento econômico; além disso, não haveria razão para considerar que a redistribuição é sistematicamente ineficiente.

O que se observa dessa discussão é que parece não haver um consenso sobre a real existência do *trade-off* apontado por Okum (1975). Apesar dessa ausência de consenso, pode-se afirmar que os modelos teóricos que buscaram lançar luz sobre o *trade-off* eram, todos eles, modelos ditos micro-fundamentados, ou seja, eles se concentravam nas árvores, mas não na floresta.²

O presente trabalho sugere um outro rumo para a pesquisa que busca analisar a relação desigualdade-crescimento. Aqui, nos embasamos em macro-fundamentos que permitem explicar, estabelecer e testar o referido *trade-off*. Para tanto, criamos uma conexão entre duas das mais antigas hipóteses nas áreas de crescimento e desenvolvimento econômico: a hipótese da convergência e a hipótese da curva de Kuznets.

Nosso estudo está em linha com a análise inicial de Schin (2008), e é calcado nos testes empíricos destas duas hipóteses: a hipótese da convergência prescreve uma renda *per capita* de estado estacionário e sustenta que, se as economias transpuserem esse *steady-state*, elas passam a fazer uso ineficiente de seus recursos. Já a hipótese de Kuznets sugere uma relação entre desigualdade e desenvolvimento em forma de U invertido; essa relação sugere uma renda *per capita* associada à um ponto de desigualdade máxima, e sustenta que o sinal da correlação entre desigualdade e crescimento estaria condicionado à essa renda.

Com base nessas duas hipóteses mostramos que é possível se identificar um par ordenado (renda *per capita*, distribuição de renda) de estado estacionário, e descrevemos de

2. Há vários canais fundamentados na microeconomia através dos quais a desigualdade pode se relacionar com o crescimento. Por exemplo, o canal das imperfeições do mercado de capitais, da abordagem da fertilidade endógena, do argumento relativo ao tamanho do mercado interno, do enfoque da política fiscal endógena, e o canal de instabilidade política.

que forma o dilema equidade-eficiência pode ser prontamente estabelecido a partir deste par ordenado.

Além disso, demonstramos que a relação crescimento-desigualdade é condicionada à distância entre a renda *per capita* de estado estacionário e a renda *per capita* associada ao ponto de desigualdade máxima da curva de Kuznets, e que é essa distância que explica os diferentes sinais da correlação crescimento-desigualdade apresentados pela literatura.

Após expor a ideia central do trabalho, conduzimos uma análise empírica para os municípios do Ceará. Uma inovação empírica ainda é apresentada, frente aos estudos nacionais, pois conduzimos o teste proposto por Lind e Mehlum (2010) para identificar um intervalo de confiança apropriado para o ponto de máximo da curva de Kuznets.

O trabalho está estruturado da seguinte forma: após essa introdução apresentamos a estrutura teórica que embasa as hipóteses da convergência e da curva de Kuznets e, posteriormente, conectamos essas duas teorias para elucidar de que forma o dilema equidade-eficiência se forma. Na terceira seção, apresentamos a metodologia empírica levando em consideração os municípios do Ceará. Na quarta seção, com base nos resultados, traçamos nossos comentários finais. A quinta seção, como de praxe, traz as referências bibliográficas utilizadas nesse estudo.

2 ESTRUTURA TEÓRICA

A hipótese da Convergência é um resultado direto do modelo de crescimento introduzido por Robert Solow, em 1956, enquanto a curva com o formato de “U” relacionando desigualdade e renda foi observada por Simon Kuznets, em 1955.

Embora haja alguns trabalhos anteriores, a análise econométrica formal da hipótese de Kuznets parece ter tido início com o trabalho de Ahluwalia (1976), já os testes da hipótese da Convergência só tiveram início 15 anos depois, a partir do trabalho seminal de Barro (1991).

Ambas as hipóteses reúnem uma vasta gama de trabalhos empíricos, entretanto, em se tratando de refinamentos metodológicos, a revisão destas duas literaturas sugere que os testes da hipótese de Kuznets se concentraram mais na direção causal da relação crescimento-desigualdade e na qualidade e disponibilidade de dados sobre desigualdade;³ já a hipótese da convergência parece ter avançado mais em termos de técnicas econométricas

3. Ver Deininger e Squire (1998) para uma discussão relacionada às diferentes bases de dados utilizadas na literatura.

que buscassem contribuir para o processo de formação de clubes de convergência.^{4,5}

Com o intuito de explicar os efeitos diferenciados da desigualdade no crescimento econômico sugeridos em Barro (2000, 2008), Shin (2008) conectou simultaneamente a teoria da convergência e a teoria de Kuznets. A análise proposta por Shin (2008) traz inter-relações importantes para o entendimento dos processos de crescimento e desenvolvimento econômico, entretanto, pode-se explorar em mais detalhes as potencialidades da conexão entre estas duas teorias. Isso é o que se tenta fazer a seguir.

2.1 Convergência

Inicialmente, é interessante adentrar na hipótese da convergência em maiores detalhes. Considere o modelo de crescimento de Solow expresso em termos de capital por trabalhador efetivo \tilde{k} , com função de produção Cobb-Douglas $\tilde{y} = \tilde{k}^\alpha$, taxa de poupança s , taxa de depreciação δ , taxa de crescimento populacional n , e taxa de progresso tecnológico g .

Nestes termos, a equação fundamental de movimento do modelo é:

$$\dot{\tilde{k}} = s\tilde{k}^\alpha - (\delta + g + n)\tilde{k}. \quad (1)$$

O modelo de crescimento de Solow prevê um nível de capital de estado estacionário único para cada taxa de poupança possível, s . A taxa de poupança da regra de ouro, s_g , é a taxa que maximiza o consumo *per capita* no longo prazo. O consumo no estado estacionário, por trabalhador efetivo pode ser escrito como:

$$\dot{\tilde{k}} = s\tilde{k}^\alpha - (\delta + g + n)\tilde{k}. \quad (2)$$

onde usamos o fato de que $\dot{\tilde{k}} = 0$ implica $sf(k^*) = (n+g+\delta)k^*$. Tomando a derivada parcial com relação a s e igualando a mesma a zero, temos:

$$(3)$$

Isto implica que a taxa de poupança da regra de ouro deve satisfazer:

$$f'(k^*) = (n + g + \delta) \quad (4)$$

4. Quah (1996) utiliza métodos não paramétricos e matrizes de transição de Markov; Durlauf e Johnson (1995) empregaram técnicas de *Classification and Regression Tree Analysis* (CART); Islam (1995) propõe uma análise de painel dinâmico; Hansen (2000) utiliza um modelo econométrico não-linear com efeito *threshold*; Canova (2004) propõe uma análise com base na densidade preditiva dos dados; Phillips e Sul (2007) fazem uso do teste $\log t$; Cheng e Lin (2009) fazem um mix dos *Common Correlated Effects Pooled* (CCEP), proposto por Pesaran (2006), e do efeito *threshold* de Hansen (2000).

5. Revisões destas literaturas e uma comparação criteriosa entre as mesmas fogem completamente do escopo deste trabalho, entretanto, uma revisão de literatura sobre os trabalhos empíricos relacionados à hipótese de Kuznets é feita por Kimhi (2004) e Chong (2001) e, em relação ao processo de convergência, o trabalho de Durlauf e Johnson (2005) parece ser o mais completo.

Há dois casos onde a economia não está caracterizada por sua regra de ouro:

i) A economia pode estar economizando muito ($s > s_g$) e, neste caso, ela é dita dinamicamente ineficiente, pois, em todos os pontos no tempo, há uma trajetória de consumo viável que seria maior do que a corrente.

ii) A economia pode estar economizando muito pouco ($s < s_g$) e, neste caso, ela também é dita ineficiente, pois, embora o consumo atual seja maior que o preconizado pela regra de ouro, o consumo *per capita* de longo prazo não está sendo maximizado.

A dinâmica de transição para o estado estacionário é determinada pela lei do movimento do capital: $\dot{k}(t) = sf(k(t)) - (n + g + \delta)k(t)$. (5)

Sabemos que o sistema é estável, ou seja, que ele retorna ao estado estacionário quando k se desvia de k^* . Uma questão mais interessante é: quão rápido chegamos lá? Para responder a esta pergunta, podemos escrever a equação acima, mais sucintamente, como,

$$\dot{k} = k'(k) \quad (6)$$

Uma aproximação de Taylor-Series de (5) em torno de k^* gera:

$$\dot{k}(t) \cong \left[\frac{\partial \dot{k}(k)}{\partial k(t)} \Big|_{k=k^*} \right] (k(t) - k^*), \quad (7)$$

que pode ser reescrito como:

$$\dot{k}(t) = \lambda(k(t) - k^*). \quad (8)$$

Usando (4), podemos calcular a velocidade da convergência através de

$$\lambda = - \left. \frac{\partial \dot{k}(k)}{\partial k(t)} \right|_{k=k^*} = -[sf'(k^*) - (n + g + \delta)]. \quad (9)$$

Para uma função de produção Cobb-Douglas (com $f(k) = k^\alpha$), temos

$$\lambda = -(n + g + \delta) \left[\frac{f'(k^*)k^*}{f(k^*)} \right] = (1 - \alpha)(n + g + \delta). \quad (10)$$

Como o modelo Solow é dinamicamente estável, se o sistema estiver fora do estado estacionário, ele voltará a convergir. Este princípio também pode ser aplicado para um conjunto de economias. Se as economias A e B são idênticas em todos os sentidos, exceto que a economia A tem menos capital, o modelo de Solow prediz que a economia A crescerá mais rápido e alcançará a economia B. Isso é o que caracteriza a hipótese de convergência absoluta, ou seja, a hipótese de que economias idênticas devem convergir para um estado estacionário comum. Isso implica que economias pobres crescerão mais rápido que economias ricas até que todas as taxas de crescimento cessem e elas atinjam padrões de vida iguais.

A equação (8) nos ajuda a testar empiricamente a hipótese de convergência. Sabemos que $k(t)$ converge para k^* à taxa λ . Isso também implica que $y(t)$ converge para y^* , também à taxa λ . Portanto, podemos escrever,

$$y(t) - y^* = e^{-\lambda t} [y(0) - y^*]. \quad (11)$$

Dividindo ambos os lados por y^* gera:

$$\frac{y(t) - y^*}{y^*} = e^{-\lambda t} \left[\frac{y(0) - y^*}{y^*} \right]. \quad (12)$$

Podemos aproximar (12) como

$$\ln \left[\frac{y(t)}{y^*} \right] = e^{-\lambda t} \ln \left[\frac{y(0)}{y^*} \right], \quad (13)$$

que pode ser reescrita como:

$$\ln(y(t)) - \ln(y^*) = e^{-\lambda t} \ln(y(0)) - e^{-\lambda t} \ln(y^*). \quad (14)$$

Subtraindo de ambos os lados e reorganizando, temos:

$$\ln(y(t)) - \ln(y(0)) = (1 - e^{-\lambda t}) \ln(y^*) - (1 - e^{-\lambda t}) \ln(y(0)). \quad (15)$$

Isso sugere uma regressão para um conjunto de economias com a forma:

$$\left(\frac{1}{t} \right) \ln(y_i(t)/y_i(0)) = \gamma + \beta \ln(y_i(0)) + \epsilon_i, \quad (16)$$

onde $i=1, \dots, N$. A hipótese de convergência é validada quando não se rejeita $H_0: \beta_1 < 0$.

Também é usual incluir variáveis de controle nessa regressão. Assim fazendo, e trabalhando com uma notação mais simplificada, (16) dá lugar à:

$$\Delta y_i = \gamma + \beta y_i + \kappa Z_i + \epsilon_i, \quad (17)$$

onde, $\Delta y_i = \left(\frac{1}{t} \right) \ln \left(\frac{y_i(t)}{y_i(0)} \right)$, $y_i = \ln(y_i(0))$, e Z é uma matriz com variáveis de controle.

A equação (17) prescreve que o processo de convergência pode ser diretamente testado com base numa regressão onde se tem, como variável dependente, as taxas anuais médias de crescimento para um conjunto de economias contra uma série de variáveis de controle mais o logaritmo neperiano da renda *per capita* inicial de cada economia.

Graficamente, caso o processo de convergência seja validado, o que se teria é algo próximo da figura a seguir:

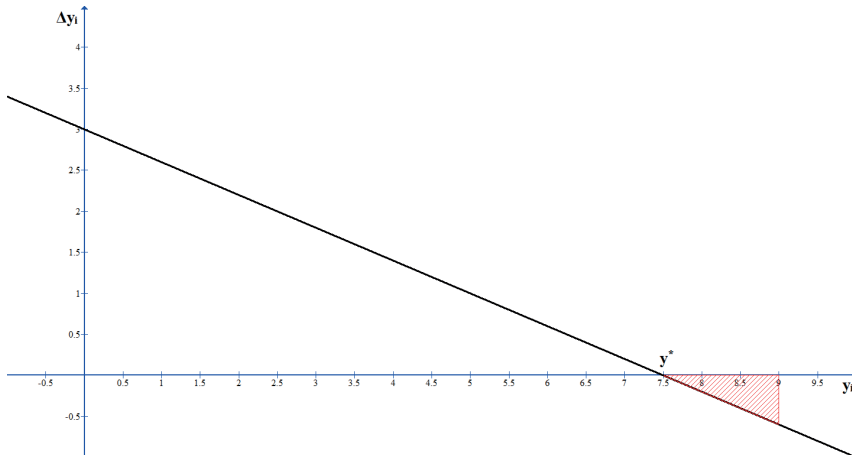


Figura 1 – Processo de Convergência

Em estado estacionário, a taxa de crescimento cessaria e teríamos $\ln(y^*) = 7.5$. O que é essencial dessa análise, é que o modelo de Solow prescreve estabilidade, ou seja, a área hachurada – que seria condizente com uma renda *per capita* superior a $\ln(y^*) = 7.5$ – representa rendas *per capita*s dinamicamente ineficientes.⁶

Grosso modo, essa ineficiência dinâmica caracteriza-se por uma superacumulação de capital.⁷ Apesar de se poder estender o modelo básico de Solow em diversas direções, de um modo ou de outro, a regressão acima permanece indicando ineficiência para as economias que ultrapassam suas posições de estado estacionário. Isso implica que, caso as economias atuem acima de suas posições de *steady-state*, então elas entrariam numa região dinamicamente ineficiente e, automaticamente, passariam a observar crescimento negativo, o que faria com que elas retornassem para seu equilíbrio estacionário.

$$\frac{\partial y}{\partial s} = [f'(k^*) - (n + g + \delta)] \frac{\partial k}{\partial s} = 0.$$

2.2 Curva de Kuznets

A análise inicial de Kuznets (1955) da relação entre crescimento e desigualdade foi baseada em pouquíssimos dados. Como o próprio autor sustenta na conclusão de seu trabalho, 5% das informações trazidas no artigo seriam evidências empíricas e 95% seriam especulação.

Com uma maior disponibilidade de dados, e buscando dar uma resposta à especulação de Kuznets, Ahluwalia (1976) investigou mais formalmente a relação entre crescimento econômico e desigualdade. Este autor utilizou dados de corte transversal para uma amostra de 60 países onde se especificava a apropriação da renda (dos 20% mais ricos; dos

6. Para traçarmos a figura acima plotamos a seguinte função: $(\frac{1}{3}) \ln(\frac{y_t(0)}{y_t(0)}) = 3 - 0.4 \cdot \ln(y_t(0))$.

7. O Professor Fernando Holanda Barbosa tem uma *apostila* onde ele discute uma série de variações do modelo de Solow, assim como condições específicas que seriam capazes de gerar essa ineficiência dinâmica.

40% intermediários e; dos 60%, 40% e 20% mais pobres) como função quadrática da renda *per capita*, da participação da agricultura no PIB e da taxa de urbanização, além de algumas variáveis de controle.

Haja vista que a taxa de urbanização e a participação da agricultura no PIB se davam de acordo com o que foi concebido por Kuznets, e que a apropriação da renda se elevava e, posteriormente, se reduzia, conforme a renda *per capita* da sociedade crescia (exceto para os 20% mais ricos), o autor conclui que a hipótese de Kuznets não podia ser rejeitada.

O estudo de Ahluwalia (1976) deu início a uma série de trabalhos que incorporavam novos elementos a análise, tais como suposições econômicas mais sustentáveis, modelos teóricos mais enorpados, e técnicas econométricas mais sofisticadas.⁸ De um modo geral, o teste da hipótese de Kuznets costuma se basear na seguinte regressão:

$$g_i = \alpha + \beta y_i + \delta f(y_i) + \psi Z_i + \varepsilon_i, \quad (18)$$

onde g é, usualmente, o índice de desigualdade de Gini ou Theil; y é o logaritmo da renda *per capita*, X é um vetor com variáveis de controle, α, β, δ e ψ são parâmetros, e ε é um termo de erro. $f(\cdot)$ é escolhida *ad hoc* e dá a curvatura de (18) e, dependendo dos parâmetros β e δ , a reta de regressão pode ter formato de “U”, de “U invertido”, ou ser uma função monótona.⁹

Estudos anteriores se concentravam em testar os sinais e a significância estatística dos parâmetros β e δ . Caso se verificasse $\beta > 0$ e $\delta < 0$, então seria côncava em y e a hipótese de Kuznets não deveria ser rejeitada. A partir desta motivação, diversas técnicas, estruturas de dados, períodos e coberturas regionais foram exploradas. De um modo geral, o que se observa é que a hipótese de Kuznets parece ter alguma aderência aos dados.¹⁰

Note que, se a hipótese de Kuznets for validada, então ela prescreve um “ponto de máximo estimado”. Supondo uma função quadrática, esse ponto seria dado por $y^* = -(\beta / 2\delta)$. Entretanto, esta estimativa tem como base a razão entre dois estimadores supostamente independentes e normalmente distribuídos. Lind e Mehlum (2010) observam que a distribuição do teste para o ponto máximo estimado segue uma Cauchy, cuja média não é definida e, conseqüentemente, não se torna possível estabelecer um desvio padrão.

8. Um sumário dos estudos empíricos nacionais mais relevantes pode ser visto em Penna et al (2015) e, para a literatura internacional, ver Gallup (2012).

9. As duas especificações mais comuns são a quadrática e a especificação inversa. No caso da primeira, que é a mais recorrente na literatura, temos a especificação conforme a descrita em (2.1). Anand e Kambur (1993) sugerem que para que a hipótese de Kuznets seja válida é necessário que se adotem diferentes especificações para diferentes índices de desigualdade (Theil, Gini, etc). O trabalho a ser possivelmente desenvolvido terá como base o índice de Gini com especificação quadrática, ficando as demais possibilidades deixadas para estudos futuros.

10. Como discutido na introdução, a hipótese de Kuznets tem uma série de possíveis explicações teóricas, entretanto, aqui não estamos prontamente interessados nestas teorias. Nosso ponto de partida é que alguma dessas motivações teóricas seja capaz de explicar a relação mecanicista do U invertido preconizado por Kuznets.

Isso dificulta sobremaneira o cômputo de intervalos estatisticamente confiáveis para y^{\sim} . Com efeito, políticas públicas balizadas por tal parâmetro podem equivocarse caso este não seja apropriadamente estimado.

Trabalhando em cima desta problemática, Lind e Mehlum (2010) fizeram uso do Teorema de Fieller (1954) para estabelecer intervalos de confiança em pequenas amostras condizentes com o ponto crítico estimado e propõem um teste estatisticamente mais robusto, o qual fornece condições necessárias e suficientes exatas para o teste da forma funcional de “U invertido” em amostras finitas. Este teste é formalizado na Seção 3.1.

2.3 Reunindo Hipóteses

Note, inicialmente, que o mesmo regressor – o log da renda *per capita* – está presente em ambas as especificações para o teste empírico das hipóteses:

$$\Delta y_i = \gamma + \beta y_i + \kappa Z_i + \epsilon_i \quad (19)$$

$$g_i = \alpha + \beta y_i + \delta f(y_i) + \psi Z_i + \epsilon_i \quad (20)$$

A conexão destas duas (possíveis) regularidades empíricas, juntamente com a ideia de renda *per capita* de estado estacionário, y^* , e de renda *per capita* associada ao ponto de máximo da curva de Kuznets, y^{\sim} , finda em três situações possíveis:

- i) $y^* < y^{\sim}$, o *steady-state* é menor do que o ponto de máximo;
- ii) $y^* = y^{\sim}$, o *steady-state* é igual ao ponto de máximo;
- iii) $y^* > y^{\sim}$, o *steady-state* é maior do que o ponto de máximo;

Note-se que, independentemente da situação em que a análise se enquadre, é possível descrever um nível de desigualdade de estado estacionário. Essa ideia é condizente com o trabalho de Benabou (1996), que sugere que o acréscimo de choques idiossincráticos em modelos neoclássicos de crescimento, quando acompanhados de formações de clubes de convergência de renda *per capita*, seria capaz de gerar convergência em distribuição. Em outras palavras, economias com os mesmos fundamentos macroeconômicos tenderiam para um mesmo nível de distribuição de renda *per capita*.

Os gráficos a seguir ajudam a ilustrar o raciocínio.

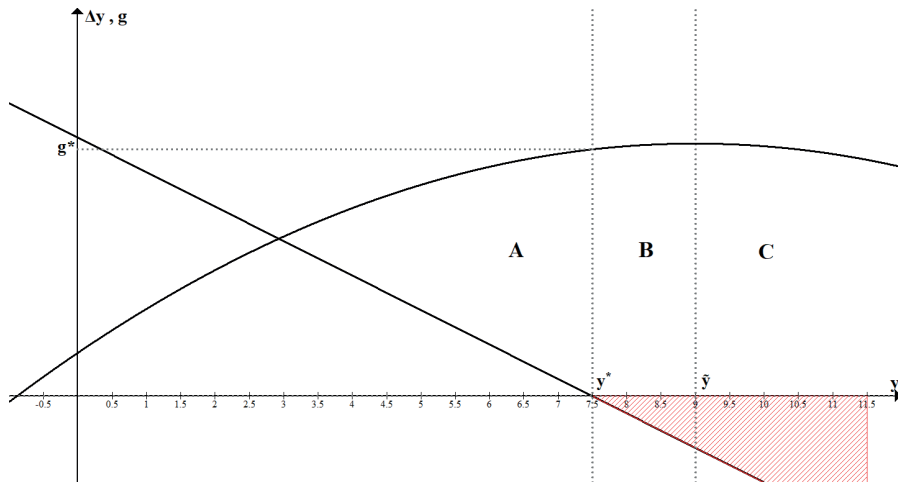


Figura 2 – Kuznets e Convergência com

A Figura 2, acima, relata o caso com a renda *per capita* de estado estacionário inferior à renda *per capita* associada ao máximo da curva de Kuznets, isto é, $y^* < \tilde{y}$. Neste caso, conforme as economias forem convergindo para y^* , elas também estariam se deslocando ao longo da parte ascendente da curva de Kuznets. Isso é retratado pela área da figura, que sugere pontos onde o crescimento econômico ocorreria concomitantemente com o aumento da desigualdade de renda. Essa situação iria prevalecer até que se chegue numa situação de estado estacionário (y^*, g^*) .

Se o padrão estrutural das economias que convergiram para não se alterar, políticas públicas que promovam o crescimento para além deste ponto teriam um efeito extremamente deletério às economias, pois, ao longo da área, observa-se que haveria crescimento econômico negativo (ineficiência) juntamente com desigualdade de renda crescente (pois ainda se está à esquerda do ponto de máximo da curva de Kuznets).

A partir do ponto de máximo da curva de Kuznets - indicado via \tilde{y} - estaria se entrando na área C do gráfico, e passaria a valer o dilema equidade-eficiência: é possível reduzir a desigualdade de renda, mas ao custo de se colocar a economia em uma situação ineficiente. Note-se que, neste caso, o grau de ineficiência é bem mais elevado.

Vejamos, agora, o que ocorre quando a renda *per capita* de estado estacionário coincide com a renda *per capita* associada ao máximo da curva de Kuznets, ou seja, $y^* = \tilde{y}$. Este caso é apresentado na Figura 3, abaixo.

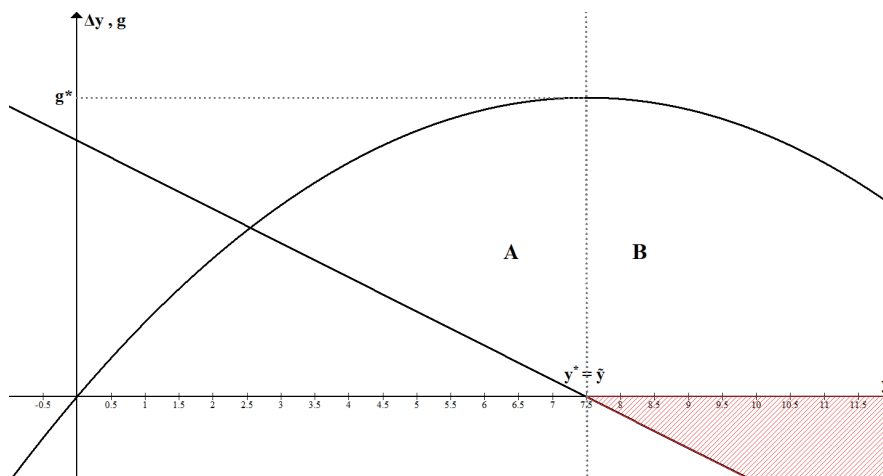


Figura 3 – Kuznets e Convergência com

Neste caso, conforme as economias convergem para seu *steady-state*, a renda *per capita* e a desigualdade tendem a se ampliar. Isso é descrito na parte da figura. Atingindo sua posição de estado estacionário, e não havendo alterações estruturais, então passa a valer o Dilema equidade-eficiência, e a redução da desigualdade só poderá ocorrer ao custo de perda de eficiência. Essa situação é descrita na área da figura. Aqui, o dilema equidade-eficiência é dito “fortemente caracterizado”, pois a estabilidade do estado estacionário tende a colocar a relação (y^*, g^*) numa situação estática onde, qualquer perturbação gerada com o intuito de se reduzir a desigualdade irá gerar ineficiência econômica.

Devido à natureza do equilíbrio de estado estacionário, alterações em e serão apenas transitórias, ou temporárias. Note-se também que essa situação é extremamente insalubre, do ponto de vista distributivo, pois g^* coincide com a desigualdade máxima preconizada pela curva de Kuznets. Situações como esta requerem bastante atenção dos formuladores de políticas, e requerem alterações estruturais para que se tenham melhorias no binômio equidade-eficiência.

Por fim, temos ainda a possibilidade de observar casos onde a renda *per capita* de estado estacionário é superior à renda *per capita* sugerida pela desigualdade máxima da curva de Kuznets. Isso é exposto na Figura 4, a seguir.

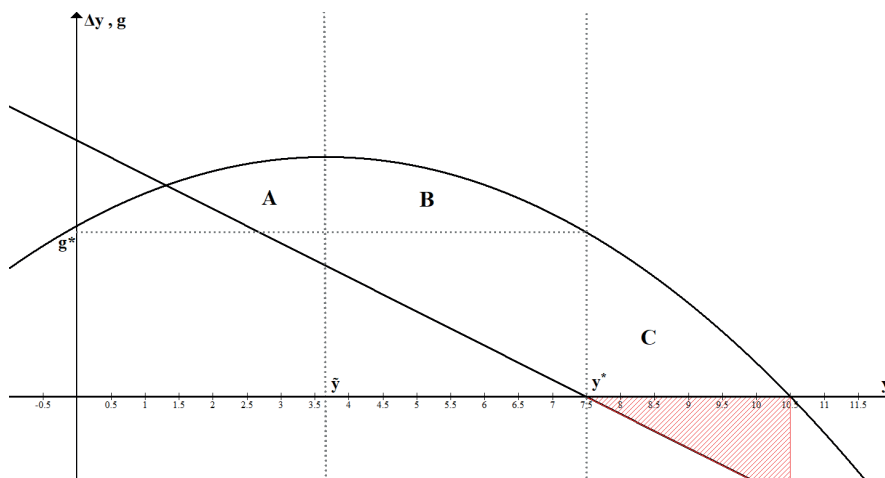


Figura 4 – Kuznets e Convergência com

Aqui temos o que deve ocorrer no caso em que se tem $y^* > \tilde{y}$. Esta é a situação ideal, pois se tem uma área onde há uma ruptura do *trade-off* equidade-eficiência, e quanto mais a direita y^* estiver de \tilde{y} , mais se explora a parte ascendente da curva de Kuznets.

Na parte *A* da Figura 4, o processo de convergência faz com que a renda e a desigualdade cresçam conjuntamente. Isso equivale à relação negativa entre equidade e eficiência, ou seja, tem-se evidências do *trade-off*. A renda cresce até atingir o ponto de máximo da curva de Kuznets e, a partir daí, entra-se na área .

A área *B* da Figura 4 sugere a área onde o *trade-off* equidade-eficiência é rompido. O processo de convergência para o estado estacionário ocorre simultaneamente ao deslocamento na parte descendente da curva de Kuznets, ou seja, há crescimento da renda e redução das desigualdades.

Note-se que, ao atingir o *steady-state*, o dilema equidade-eficiência volta a valer (área *C*), entretanto, para um dado \tilde{y} , quanto mais à direita y^* estiver, mais a curva de Kuznets pode ser explorada.

Em termos práticos, é de se esperar que sociedades avançadas, como os países nórdicos, por exemplo, tenham posições de estado estacionário muito próximas da maior raiz da função quadrática que representa a curva de Kuznets, o que caracterizaria elevado nível de renda e baixa desigualdade. Já sociedades comunistas tendem a ser vistas na outra extremidade.

O que se tira do exposto acima é que, constatadas as regularidades empíricas da hipótese de Kuznets e da hipótese da convergência, e supondo que não haja mudança estrutural para um conjunto de economias que passam por estes processos, então, quanto maior

for a renda de estado estacionário com relação à renda associada ao ponto de máximo da curva de Kuznets, melhor será o aproveitamento da relação crescimento-equidade observada por este grupo de economias, havendo, inclusive, possibilidade da ruptura do *trade-off* equidade-eficiência.

De outro modo, e supondo estabilidade estrutural, quanto mais aquém a renda *per capita* de estado estacionário estiver da renda *per capita* associada ao ponto de máximo da curva de Kuznets, maiores serão os desgastes, em termos de equidade e eficiência, gerados por políticas que promovam rendas *per capitas* superiores à de estado estacionário.

Apresentada a teoria, nosso próximo passo é buscar evidências empíricas para essa discussão analisando o conjunto de municípios do estado do Ceará. Para tanto, a metodologia empírica utilizada com esse intuito é exposta a seguir.

3 METODOLOGIA EMPÍRICA

O primeiro passo da investigação empírica é calcado nas equações de regressões (17) e (18). Aqui fizemos uso dos seguintes dados, para todos os 184 municípios cearenses:

Variável	Descrição
y_i	Renda <i>per capita</i> média - Razão entre o somatório da renda de todos os indivíduos residentes em domicílios particulares permanentes e o número total desses indivíduos. Anos: 1991, 2000 e 2010. Valores em reais de 01/agosto de 2010.
Δy_i	Taxa média de crescimento: $\Delta y_i = \left(\frac{1}{T}\right) \ln \left(\frac{y_{i,t}}{y_{i,t-1}}\right)$. Anos: 2000 e 2010.
g_i	Índice de Gini - Mede o grau de desigualdade existente na distribuição de indivíduos segundo a renda domiciliar <i>per capita</i> . Seu valor varia de 0, quando não há desigualdade (a renda domiciliar <i>per capita</i> de todos os indivíduos tem o mesmo valor), a 1, quando a desigualdade é máxima (apenas um indivíduo detém toda a renda). O universo de indivíduos é limitado àqueles que vivem em domicílios particulares permanentes. Anos: 1991, 2000 e 2010.
Z_i	Índice de Desenvolvimento Humano Municipal - Média geométrica dos índices das dimensões Renda, Educação e Longevidade, com pesos iguais. Anos: 1991, 2000 e 2010.

Alguns pontos precisam ser enfatizados:

- i) A regressão de convergência requer que se utilize como regressor a renda *per capita* inicial; buscando ampliar o *span* de dados, usamos as taxas de crescimento com relação a 1991-2000 e 2000-2010; isso dá uma maior dinâmica à análise, pois incorpora-se maior informação acerca da trajetória para o estado estacionário.¹¹
- ii) Com o intuito de se manter um modelo simples, optou-se por utilizar apenas o IDHM como variável de controle, tanto na regressão de convergência, quanto na regressão da curva de Kuznets; ao final do trabalho discutimos os prós e contras dessa

11. Há respaldo para isso na literatura. Ver Islam (1995), apud Penna e Linhares (2013).

estratégia.

iii) Devido a exiguidade amostral, ambas as regressões devem ser estimadas por mínimos quadrados ordinários, pois, como o spam temporal é muito pequeno ($T=2$, no caso da convergência e $T=3$, no caso da curva de Kuznets), não é aconselhável a se trabalhar com dados em painel.

iv) Como observado por Lind e Mehlum (2010), é preciso cautela ao se fazer inferências sobre a renda *per capita* associada ao ponto de máximo da curva de Kuznets. Os autores também propuseram intervalos de confiança adequados para esse ponto de máximo. Embora esse não seja o propósito do artigo, mas como nossa revisão de literatura não retornou nenhuma aplicação nacional deste teste, em seguida discorremos um pouco mais sobre o mesmo.

3.1 O Teste de Lind e Mehlum (2010)

Lind e Mehlum (2010) salientam que, ao assumir que (18) tem formato de um “U” invertido com um único ponto extremo, então é necessário que a inclinação da curva de Kuznets seja inicialmente positiva e, posteriormente, negativa para um intervalo razoavelmente escolhido $[y_l, y_h] = [\min(y), \max(y)]$; ou seja, é necessário que $f'(\cdot)$ seja monótona neste intervalo. O formato de “U” invertido requer, portanto,

$$\beta + 2\delta y_l > 0 > \beta + 2\delta y_h \quad (19)$$

Para se testar quando as condições descritas acima são suportadas pelos dados é preciso se testar quando as hipóteses nulas combinadas,

$$H_0: \quad \beta + 2\delta y_l \geq 0 \quad e/o \quad \beta + 2\delta y_h \leq 0, \quad (20)$$

podem ser rejeitadas em favor das hipóteses alternativas combinadas,

$$H_1: \quad \beta + 2\delta y_l < 0 \quad e/o \quad \beta + 2\delta y_h > 0 \quad (21)$$

Como (19) é linear em β e δ , o teste de (20) contra (21) é um teste de restrições lineares em β e δ , entretanto, Lind e Mehlum (2010) ressaltam que tal teste envolve um conjunto de restrições de desigualdade, isto é, que o conjunto de (β, δ) que satisfaz H_1 é um setor no \mathbf{R}^2 contido entre as duas linhas, $\beta + 2\delta y_l = 0$ e $\beta + 2\delta y_h = 0$.

Assumindo que $\varepsilon_{it} \sim NID(0, \sigma^2)$, Sasabuchi (1980) demonstra que o teste da hipótese nula em (20), pelo Princípio da Razão da Máxima Verossimilhança, equivale a: rejeitar H_0 com $\alpha\%$ de confiança somente se, H_0^l , ou H_0^h , ou ambas, puderem ser rejeitadas ao nível α de significância, onde H_0^l e H_0^h referem-se às hipóteses nulas nos dois testes unilaterais:

$$\begin{aligned} H_0^l: \beta + 2\delta y_l \geq 0 & \quad vs \quad H_1^l: \beta + 2\delta y_l < 0 \\ H_0^h: \beta + 2\delta y_h \leq 0 & \quad vs \quad H_1^h: \beta + 2\delta y_h > 0 \end{aligned} \quad (22)$$

Ou seja,

$$\begin{aligned}
 H_0^l: y_l &\geq -\left(\frac{\beta}{2\delta}\right) & vs & & H_1^l: y_l < -\left(\frac{\beta}{2\delta}\right) & (23) \\
 H_0^h: y_h &\leq -\left(\frac{\beta}{2\delta}\right) & vs & & H_1^h: y_h < -\left(\frac{\beta}{2\delta}\right)
 \end{aligned}$$

Para sanar tal problema existem três alternativas disponíveis: o método delta, *bootstrapping* e o Teorema de Fieller (1954). Para pequenas amostras o método delta pode gerar resultados extremamente viesados se comparado aos outros dois [Hirschberg e Lye (2005)]; o *bootstrapping* é confiável, mas sua implementação torna-se relativamente custosa ao incorporar a metodologia descrita a seguir; o Teorema de Fieller, entretanto, é de fácil tratamento e garante que a região de rejeição deste teste descrito em (23) é o cone convexo¹²

$$R_\alpha = \left\{ (\beta, \delta): \frac{\beta + 2\delta y_l}{\sqrt{s_{11} + 4(y_l)s_{12} + (2y_l)^2 s_{22}}} < -t_\alpha \cap \frac{\beta + 2\delta y_h}{\sqrt{s_{11} + 4(y_h)s_{12} + (2y_h)^2 s_{22}}} < -t_\alpha \right\} \quad (24)$$

onde t_α é o valor crítico da estatística t para um nível de significância com os graus de liberdade apropriados e s_{11} , s_{12} e s_{22} são, respectivamente, as variâncias estimadas de β e δ e a covariância entre elas.

A área de rejeição pode ser manipulada de modo que se expresse a mesma em função de y_l e y_h ; assim, utilizando $(\hat{\beta}, \hat{\delta})$ em R_α tem-se:

$$\begin{aligned}
 \mu_l < \theta_l &\equiv \frac{s_{12}t_\alpha^2 - \hat{\beta}\hat{\delta} - t_\alpha\sqrt{(s_{12}^2 - s_{22}s_{11})t_\alpha^2 + \hat{\delta}^2s_{11} + \hat{\beta}^2s_{22} - 2s_{12}\hat{\beta}\hat{\delta}}}{2(\hat{\delta}^2 - s_{22}t_\alpha^2)} & (25) \\
 \mu_h > \theta_h &\equiv \frac{s_{12}t_\alpha^2 - \hat{\beta}\hat{\delta} - t_\alpha\sqrt{(s_{12}^2 - s_{22}s_{11})t_\alpha^2 + \hat{\delta}^2s_{11} + \hat{\beta}^2s_{22} - 2s_{12}\hat{\beta}\hat{\delta}}}{2(\hat{\delta}^2 - s_{22}t_\alpha^2)}
 \end{aligned}$$

Note-se que o ponto extremo estimado, $y \hat{=} -(\beta/2\delta)$, referente à renda *per capita* associada ao ponto de máximo da curva de Kuznets, requer $y_l \hat{=} y \hat{=} y_h$, portanto, um intervalo de confiança de $(1-2\alpha)\%$ para $-(\beta/2\delta)$ será dado por $[\hat{\theta}_l, \hat{\theta}_h]$.

Com efeito, o teste proposto por Lind e Mehlum (2010) requer que se verifique, além dos sinais estimados de β e de δ , quando o intervalo para y vai estar dentro do range dos dados, $[\hat{\theta}_l, \hat{\theta}_h] \subset [y_l, y_h]$, ou seja, o teste descrito é estatisticamente mais severo que os testes anteriores.

12. Este teste é conhecido na literatura por “Teste Interseção-União”. O método de construção de testes deste tipo pode ser útil quando a hipótese nula é convenientemente expressa como uma interseção, ou seja, $H_0: \theta \in \bigcap_{\lambda \in L} \Theta_\lambda$, onde Θ_λ é um conjunto de índices arbitrário que pode ser finito ou infinito, dependendo do problema. Supondo que há testes disponíveis para se testar cada uma das hipótese $H_{0\lambda}: \theta \in \Theta_\lambda$; $\theta \in \bigcap_{\lambda \in L} \Theta_\lambda$ e, supondo que a região de rejeição para o teste de $H_{0\lambda}$ é $\{x: T_\lambda(x) \in R_\lambda\}$ é, para cada $\lambda \in L$, então, a região de rejeição do teste é dada por $\bigcup_{\lambda \in L} \{x: T_\lambda(x) \in R_\lambda\}$, daí o nome do teste.

4 COMENTÁRIOS FINAIS

O presente trabalho faz uma conexão entre os testes empíricos das hipóteses de convergência e da curva de Kuznets. A partir desta conexão, se estabelece uma distância entre a renda *per capita* de estado estacionário, y^* , e a renda *per capita* associada à máxima desigualdade da curva de Kuznets, \tilde{y} , além de se estabelecer uma distribuição de renda de estado estacionário, .

Mostrou-se que a condição $y^* > \tilde{y}$ é desejável à condição $y^* \leq \tilde{y}$, pois torna-se possível crescer explorando a parte descendente da curva de Kuznets, ou seja, seria possível crescer, pelo menos até certo ponto, sem incorrer no *trade-off* equidade-eficiência.

Em termos de causalidade, nossa análise pressupõe que convergência causa distribuição. Essa é uma premissa embasada na ideia de que a condição de *steady-state* é um ponto de equilíbrio, e que este ponto de equilíbrio é inexistente na teoria de Kuznets. Assim sendo, ao conglomerar as duas teorias, a desigualdade na curva de Kuznets passa a ficar condicionada à renda *per capita*, que deveria necessariamente convergir para o equilíbrio do modelo de Solow.

No gráfico a seguir são expostas as retas de regressão estimadas para as hipóteses de convergência e da Kuznets. Elas foram traçadas supondo a média do IDHM.

A renda *per capita* associada ao máximo da curva de Kuznets foi estimada em . O intervalo de confiança de Fieller para essa estimativa equivale à [5.1246 ,5.9849]. Já a renda *per capita* de estado estacionário foi estimada em . Todas as estatísticas são significantes a um nível de significância de 5%. Já a desigualdade de estado estacionário é condizente com o valor de $g^*=0.5485$.

Nossos resultados sugerem que as hipóteses de Kuznets e de Convergência são válidas para os municípios do Ceará. Com relação às estimativas, encontramos que y^* e \tilde{y} e equivalem às rendas *per capita*s médias (em Reais de 2010) de R\$ 418.86 e R\$ 237,34, respectivamente.

É interessante realizar um exercício para se mensurar qual seria a taxa de crescimento da renda *per capita* dos municípios de 2010 até a renda *per capita* de estado estacionário. Isso pode ser feito computando-se a média da seguinte métrica: $\frac{y^* - y_{i,2010}}{y_{i,2010}}$. Assim fazendo, encontramos o valor de 0.4765, que sugere que ainda há espaço para os municípios crescerem, em média, 47.65% até chegarem a renda *per capita* de estado estacionário.

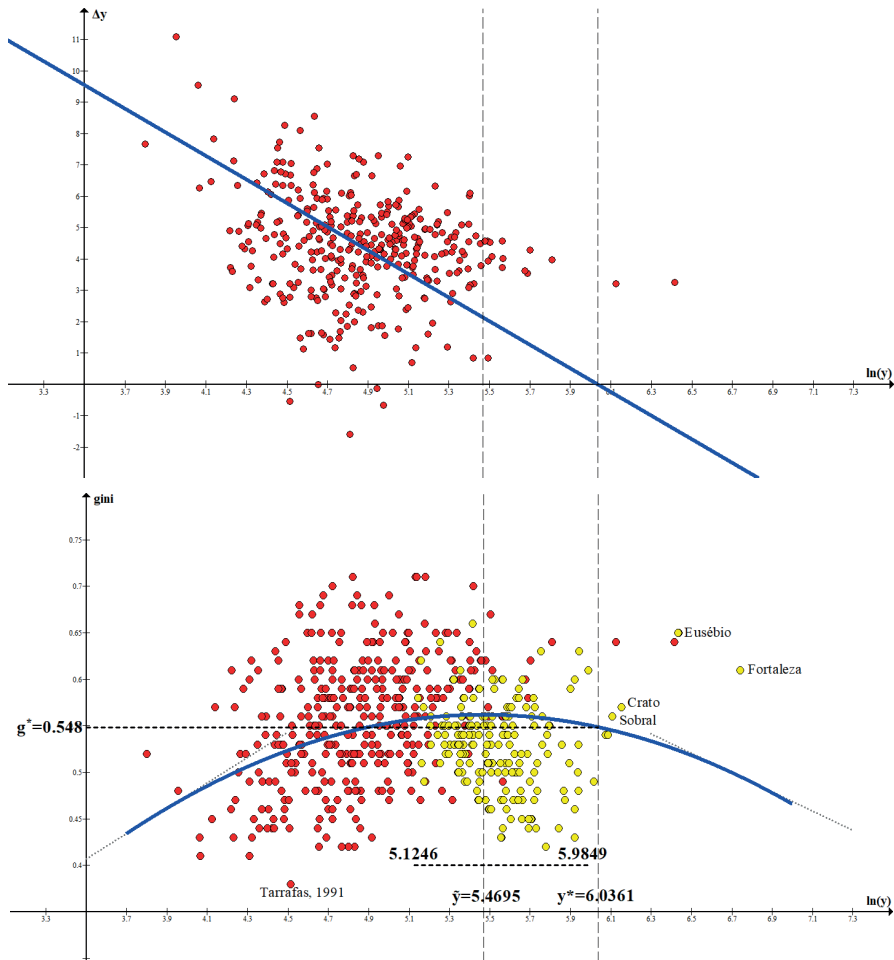


Gráfico 1 – Convergência e Kuznets no Ceará – Elaboração Própria.

Pontos Amarelos – Dados de 2010

O mesmo exercício pode ser feito levando-se em consideração o índice de Gini. Neste caso, ainda há espaço para o índice de Gini crescer, em média, 0.0472, ou seja, em média, a desigualdade municipal ainda pode aumentar aproximadamente 4.7% até chegar em sua posição de *steady-state*.

À primeira vista, o cômputo dessas métricas parece ser algo contraditório, pois $y^* > \tilde{y}$ indica uma posição onde o aumento da renda deveria ocorrer simultaneamente à queda na desigualdade.

Com relação a esse ponto, são necessários alguns adendos: no gráfico exposto, fizemos questão de ressaltar os dados de 2010 na cor amarela. Uma análise desses dados revela que ainda há uma série de municípios que estariam numa situação anterior ao ponto de máximo, da curva de Kuznets, o que sugere que eles deveriam crescer ampliando de desigualdade.

Note-se, também, que muitos deles detêm níveis de renda *per capita* que pertencem ao intervalo de confiança de Fieller. Assim, se o ponto de máximo da curva de Kuznets for condizente com o ponto , por exemplo, então boa parte do processo de crescimento dos municípios ainda estaria ocorrendo na parte crescente da curva de Kuznets, o que indica que ainda se teria crescimento econômico com elevação da desigualdade. Com efeito, essa imprecisão acerca do ponto de máximo abre espaço para esse crescimento de renda coexistente com o aumento da desigualdade sugerido pelos dados.

No gráfico, também fizemos questão de ressaltar que os municípios de Crato, Fortaleza, Euzébio e Sobral já observaram renda *per capita* e desigualdade superiores aos prescritos pelo estado estacionário. A princípio, isso estaria sugerindo que o *trade-off* equidade-eficiência estaria passando a valer para esses municípios. Entretanto, há uma fragilidade nesse raciocínio, pois há uma hipótese muito importante que foi negligenciada na análise de convergência.

Nosso arcabouço empírico pressupõe convergência absoluta, desconsiderando a possibilidade de clubes de convergência. A existência de clubes de convergência estaria indicando mais de uma reta de regressão de convergência e, por conseguinte, posições de *steady-state* distintas.

Assim, se ficar evidenciado que os municípios de Crato, Fortaleza, Euzébio e Sobral formam um clube que vem convergindo para um nível de renda *per capita* de estado estacionário mais elevado, então isso amenizaria a situação atual destes municípios com relação aos seus níveis de desigualdade, além de explicar as elevadas rendas *per capita* observadas.

Também não sabemos ao certo se a formação de clubes de convergência é condizente com a existência de curvas de Kuznets distintas para cada clube de economias. Isso precisa ser estudado com mais cautela, e testado, mas, infelizmente, para o caso dos municípios cearenses, a exiguidade de dados não nos permite adentrar nessa questão, pois temos uma dimensão temporal muito restrita para embasar a análise.

A pesquisa apresentada ainda está em fase de germinação, mas ela parece já dar indícios importantes aos pesquisadores que buscam uma melhor compreensão sobre a relação crescimento-equidade no Ceará. As questões e fragilidades apresentadas podem ser exploradas em maiores detalhes em estudos futuros, que se façam valer de bases de dados mais robustas. Contudo, sugere-se que estes estudos realizem análises que findem em evidências condizentes com as hipóteses da curva de Kuznets e do processo de convergência.

Outra questão crucial é a possibilidade de mudanças estruturais no processo de crescimento. Como é sabido, alterações estruturais podem levar um conjunto de economias para uma posição de estado-estacionário com renda *per capita* mais elevada. Se essa mudança não gerar alterações no padrão da curva de Kuznets, então torna-se possível explorar melhor a parte descendente desta curva, ou seja, torna-se possível crescer reduzindo-se desigualdade, ou ainda, o Estado pode se desvencilhar temporariamente do *trade-off* equidade-eficiência. Neste sentido, as atuais apostas do Governo do Estado na área de educação parecem estar no caminho certo.

5 BIBLIOGRAFIA

AGHION, Philippe; CAROLI, Eve; GARCIA-PENALOSA, Cecilia. Inequality and economic growth: the perspective of the new growth theories. **Journal of Economic literature**, v. 37, n. 4, p. 1615-1660, 1999.

AHLUWALIA, Montek S. Inequality, poverty and development. **Journal of development economics**, v. 3, n. 4, p. 307-342, 1976.

ALESINA, Alberto; RODRIK, Dani. Distributive politics and economic growth. **The quarterly journal of economics**, v. 109, n. 2, p. 465-490, 1994.

ANAND, Sudhir; KANBUR, SM Ravi. The Kuznets process and the inequality—development relationship. **Journal of development economics**, v. 40, n. 1, p. 25-52, 1993.

BANERJEE, Abhijit V.; DUFLO, Esther. Inequality and growth: What can the data say?. **Journal of economic growth**, v. 8, n. 3, p. 267-299, 2003.

BARRO, Robert J. Economic growth in a cross section of countries. **The quarterly journal of economics**, v. 106, n. 2, p. 407-443, 1991.

BARRO, Robert J. Inequality and Growth in a Panel of Countries. **Journal of economic growth**, v. 5, n. 1, p. 5-32, 2000.

BARRO, Robert J. **Inequality and growth revisited**. ADB Working paper series on regional economic integration, 2008.

BENABOU, Roland. Inequality and growth. **NBER macroeconomics annual**, v. 11, p. 11-74, 1996.

BENHABIB, Jess et al. The tradeoff between inequality and growth. **Annals of Economics**

and Finance, v. 4, p. 491-507, 2003.

BOURGUIGNON, François et al. Distribution, Redistribution and Development: where do we stand?. **REVISTA DESARROLLO Y SOCIEDAD**, 1998.

CANOVA, Fabio. Testing for convergence clubs in income per capita: a predictive density approach. **International Economic Review**, v. 45, n. 1, p. 49-77, 2004.

CHENG, Jen-Chieh; LIN, Chang-Ching; WANG, Chien-Ho. Estimation of growth convergence using common correlated effects approaches. **Working Paper**, 2009.

CHONG, Alberto. **Inequality, Democracy and Redistribution: Is there a Political Kuznets curve**. IADB Working Papers 445, 2001.

DEININGER, Klaus; SQUIRE, Lyn. A new data set measuring income inequality. **The World Bank Economic Review**, v. 10, n. 3, p. 565-591, 1996.

DEININGER, Klaus; SQUIRE, Lyn. New ways of looking at old issues: inequality and growth. **Journal of development economics**, v. 57, n. 2, p. 259-287, 1998.

DURLAUF, Steven N.; JOHNSON, Paul A. Multiple regimes and cross-country growth behaviour. **Journal of applied econometrics**, v. 10, n. 4, p. 365-384, 1995.

EHRHART, Christophe et al. The effects of inequality on growth: a survey of the theoretical and empirical literature. **ECINEQ WP**, v. 107, 2009.

FIELDER, Edgar C. Some problems in interval estimation. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 175-185, 1954.

FORBES, Kristin J. A Reassessment of the Relationship between Inequality and Growth. **American economic review**, p. 869-887, 2000.

GALLUP, John Luke. Is there a Kuznets curve. **Portland State University**, 2012.

HANSEN, Bruce E. Sample splitting and threshold estimation. **Econometrica**, v. 68, n. 3, p. 575-603, 2000.

HIRSCHBERG, Joe; LYE, Jenny N. Inferences for the extremum of quadratic regression models. 2005.

ISLAM, Nazrul. Growth empirics: a panel data approach. **The Quarterly Journal of Economics**, v. 110, n. 4, p. 1127-1170, 1995.

- KIMHI, Ayal. Growth, inequality and labor markets in LDCs: a survey. 2004.
- KUZNETS, Simon. Economic growth and income inequality. **The American economic review**, v. 45, n. 1, p. 1-28, 1955.
- LI, Hongyi; ZOU, Heng-fu. Income inequality is not harmful for growth: theory and evidence. **Review of development economics**, v. 2, n. 3, p. 318-334, 1998.
- LIND, Jo Thori; MEHLUM, Halvor. With or without U? the Appropriate Test for a U-Shaped Relationship. **Oxford bulletin of economics and statistics**, v. 72, n. 1, p. 109-118, 2010.
- MANKIW, N. Gregory; À ECONOMIA, Introdução. Princípios de micro e macroeconomia. **Editora Campus–1999**.
- MARTIN, Philippe. Public policies, regional inequalities and growth. **Journal of public economics**, v. 73, n. 1, p. 85-105, 1999.
- OKUN, Arthur M. Equality and Efficiency: The Big Tradeoff (Washington, DC: Brookings Institution, 1975).
- PANIZZA, Ugo. Income inequality and economic growth: evidence from American data. **Journal of Economic Growth**, v. 7, n. 1, p. 25-41, 2002.
- PARTRIDGE, Mark D. Is inequality harmful for growth? Comment. **The American Economic Review**, v. 87, n. 5, p. 1019-1032, 1997.
- PENNA, Christiano Modesto et al. Trabalho, transferências e desigualdade: a curva de kuznets para o Nordeste. **Revista Brasileira de Estudos Regionais e Urbanos**, v. 7, n. 2, p. 34-51, 2015.
- PEROTTI, Roberto. Growth, income distribution, and democracy: What the data say. **Journal of Economic growth**, v. 1, n. 2, p. 149-187, 1996.
- PERSSON, Torsten; TABELLINI, Guido. Is inequality harmful for growth?. **The American Economic Review**, p. 600-621, 1994.
- PESARAN, M. Hashem. Estimation and inference in large heterogeneous panels with a multifactor error structure. **Econometrica**, v. 74, n. 4, p. 967-1012, 2006.
- PHILLIPS, Peter CB; SUL, Donggyu. Transition modeling and econometric convergence tests. **Econometrica**, v. 75, n. 6, p. 1771-1855, 2007.

QUAH, Danny T. Empirics for economic growth and convergence. **European economic review**, v. 40, n. 6, p. 1353-1375, 1996.

SAMUELSON, Paul; NORDHAUS, William. Principles of economics. **McCraw-Hill, New York, any edition**, 1985.

SASABUCHI, Syoichi. A test of a multivariate normal mean with composite hypotheses determined by linear inequalities. **Biometrika**, v. 67, n. 2, p. 429-439, 1980.

SHIN, Inyong. Income inequality and economic growth. 2008.

SOLOW, Robert M. A contribution to the theory of economic growth. **The quarterly journal of economics**, v. 70, n. 1, p. 65-94, 1956.

VOITCHOVSKY, Sarah. Does the profile of income inequality matter for economic growth?. **Journal of Economic Growth**, v. 10, n. 3, p. 273-296, 2005.

EDUCAÇÃO NO ESTADO DO CEARÁ: ANÁLISES UTILIZANDO O MÉTODO DO CONTROLE SINTÉTICO

Francisco Germano Carvalho Lúcio*

Natanael Soares Leite**

RESUMO

Este trabalho visa avaliar o impacto da gestão governamental do Estado do Ceará nos indicadores educacionais referentes ao ensino médio, no período de 2007 a 2014. As análises serão baseadas em indicadores tanto de cunho quantitativo, proporção de indivíduos entre 15 e 17 anos matriculados no ensino médio em escolas públicas, quanto qualitativo, desempenho do ensino público estadual no ENEM. Para tanto serão utilizadas análises descritivas e o método do controle sintético para ambas abordagens. A análise descritiva mostra que o estado do Ceará apresentou uma evolução contínua, permanecendo acima da média da região Nordeste. Em relação à média nacional, o estado do Ceará a partir de 2009 ultrapassa a nacional e permanece acima até o fim da série considerada, 2014. Na abordagem qualitativa, destaca-se o rápido crescimento de 2011 a 2014 de tal forma que o estado do Ceará, anteriormente abaixo, acompanha o desempenho nacional em 2014. Utilizando o método do controle sintético, os resultados de ambas abordagens mostram que o estado do Ceará apresentou bom desempenho. Em relação especificamente ao resultado no ENEM destaca-se a trajetória crescente desde 2011 e o crescimento acentuado de 2013 para 2014. A partir dos resultados pode-se concluir que a intervenção analisada, a gestão do Governo do Estado do Ceará entre os anos de 2007 e 2014, apresentou bom rendimento na educação, tanto em resultados quantitativos quanto em qualidade da educação.

Palavras-Chave: Educação. Controle sintético. Ceará.

ABSTRACT

This study aims to evaluate the impact of the government's management of the State of Ceará on educational indicators for secondary education, from 2007 to 2014. The analysis will be based on quantitative and qualitative indicators. We use the proportion of individuals between 15 and 17 years enrolled in public high schools and state public education

* Doutorando em Economia - CAEN/UFC. germanocarvalho@caen.ufc.br Telefone: 85 996266258.

** Doutorando em Economia - CAEN/UFC. natanaelsoaresleite@gmail.com.

performance in the ENEM, respectively. Both approaches used descriptive analysis and the synthetic control method. The descriptive analysis shows that the state of Ceará presented a continuous evolution, always remaining above the average of the Northeast region. Relation to the national average, the state of Ceará as of 2009 surpasses the national level and remains above until the end of the period, 2014. The qualitative approach highlights the rapid growth from 2011 to 2014, so that the state of Ceará catch up the national performance in 2014. On the other hand, using the synthetic control method the results of both approaches show that the state of Ceará achieve good results. Regarding specifically about the result on ENEM stands out the increasing trajectory since 2011 and the sharp growth from 2013 to 2014. From the results we can conclude that the intervention analyzed, the management of the Government of the State of Ceará between the years 2007 and 2014, has achieved good results in education, both on quantitative and on quality of education.

Keywords: Education. Synthetic control. Ceará.

1 INTRODUÇÃO

Um fato bastante aceito na literatura econômica é a relação positiva entre educação e indicadores econômicos e sociais. Frequentes análises associam, por exemplo, educação e capital humano¹. Este, por sua vez, tem fortes relações com crescimento econômico.

Uma forma de evidenciar a atenção destinada a importância da educação é ter ciência de esta ser um dos oito objetivos/metasp do milênio, definidos na Declaração do Milênio², um compromisso estabelecido entre 191 países-membros da Organização das Nações Unidas – ONU, no ano 2000.

Considerando o segundo “objetivo do milênio”, a oferta de educação básica de qualidade para todos, determinou-se a meta de garantir que as crianças de todos os países, de ambos os sexos, terminem um ciclo completo de ensino, até 2015. No caso específico do Brasil essa meta foi adaptada de forma a garantir o cumprimento em todas as regiões do país.

Segundo Cunha, Heckman e Schennach (2010), as habilidades adquiridas em uma etapa da vida persistem nos períodos seguintes e podem vir a facilitar a aquisição de habilidades posteriores. De forma similar, a interferência cedo o suficiente pode afetar posi-

1. Tendo em vista as externalidades positivas geradas pela educação, Grossman (2006) afirma que a educação tem impacto sobre todas as atividades do indivíduo influenciando, por exemplo, na saúde. Esta que, segundo ele, é uma forma específica de capital humano.

2. Disponível em: <https://www.unric.org/html/portuguese/uninfo/DecdoMil.pdf>

vamente o desenvolvimento das habilidades dos indivíduos. Institucionalizando, de certa forma, o fomento das habilidades supracitadas, Curi e Menezes-Filho (2006) apontam que a pré-escola tem impacto positivo e significativo na conclusão dos ciclos escolares.

O governo do Estado do Ceará implantou em 2007 o Programa Alfabetização na Idade Certa – PAIC³. Tal política converge em interesse com a literatura exemplificada no parágrafo anterior.

Seguindo o Programa Brasil Profissionalizado, instituído por meio do Decreto nº 6.302, de 12 de dezembro de 2007⁴, o governo do Estado do Ceará, por meio da Lei Estadual nº 14.273, de 19 de dezembro de 2008⁵, criou as Escolas Estaduais de Educação Profissional (EEEP). Nas EEEP, além do ensino regular com as disciplinas básicas do currículo do ensino médio, compunham a carga horária dos alunos cursos profissionalizantes voltados para áreas escolhidas de forma a atender a realidade e necessidades locais do município e/ou região sede da implantação da instituição. Esse arranjo escolar/institucional desempenha um duplo benefício, uma vez que possibilita ao aluno cursar o ensino médio capacitando-o a ingressar no ensino superior simultaneamente que se profissionaliza para o mercado de trabalho. Implicando, assim, em maiores possibilidades de escolha ao indivíduo.

Decorrida uma década da elaboração da Declaração do Milênio o Instituto de Pesquisa e Estratégia Econômica do Ceará – IPECE elaborou um relatório⁶ acerca do desempenho do Estado do Ceará no que diz respeito as metas definidas dentro dos objetivos do milênio. Das análises referentes a educação, quase totalidade referem-se a análises com abordagem quantitativa, como taxas de escolarização e de alfabetização dentre outras, sendo apenas uma de cunho qualitativo, a saber, o resultado do SAEB em Língua Portuguesa na 4ª e 8ª série do Ensino Fundamental.

Muitos estudos que combinam economia e educação utilizam variáveis de caráter quantitativo para representação e/ou análises em educação. No entanto, ultimamente há uma busca por maior refino e acurácia de resultados utilizando variáveis que medem a qualidade do ensino. Hanushek e Kimko (2000) analisam qualidade da educação e crescimento agregado verificando que a variável de qualidade da educação é estatisticamente significativa, concluindo que a variável que expressa tão somente quantidade perde poder explicativo.

3. O objetivo do programa é que todos os alunos do Estado cheguem ao 5º ano do Ensino Fundamental na idade certa e com domínio das prerrogativas de sucesso na leitura, escrita e cálculos condizentes ao seu nível de escolarização.

4. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Decreto/D6302.htm

5. Disponível em: http://educacaoprofissional.seduc.ce.gov.br/index.php?option=com_content&view=article&id=12&Itemid=128

6. Disponível em: http://www.ipece.ce.gov.br/metas_milenio/ODM_Ceara_IPECE_2010.pdf

A crescente demanda da sociedade por serviços públicos, dentre eles a educação, bem como a atual e cada vez maior exigência de eficiência nos gastos públicos traz consigo a necessidade de avaliar as políticas implementadas nessa vertente e mensurar seus impactos. Dentro desse contexto evolui, dentro da economia, uma área que concentra esforços em mecanismos de avaliação de impacto de políticas públicas e de outras intervenções.

Para fins de exemplificação consideremos estudos de avaliação de políticas públicas e intervenções referentes a educação e casos do Estado do Ceará, como Lavor e Arraes (2014) e Carneiro e Irffi (2014, 2017).

Lavor e Arraes (2014), por exemplo, avaliam a relação entre o desempenho dos alunos do Ceará na avaliação de língua portuguesa do 5º ano do ensino fundamental e a implantação do Programa Alfabetização na Idade Certa – PAIC. Sob o argumento de contornar a dificuldade de encontrar um estado ou grupos de estados adequados para representar o grupo de controle utilizam o MCS. No entanto, os autores utilizam apenas dois anos como período pré-tratamento e Abadie, Diamond e Hainmueller (2010) não recomendam o uso dessa metodologia quando o período de pré-intervenção é curto.

No intuito de fornecer resultados mais gerais em relação ao desempenho do estado do Ceará, no que diz respeito a indicadores educacionais, este trabalho visa avaliar o impacto da gestão governamental no referido estado no período de 2007 a 2014. O argumento para utilização deste recorte temporal, além da atualidade dos dados, é que permaneceu sob o comando do mesmo gestor. Some-se a isso o fato de ser um recorte temporal razoável para análises e de situar-se integralmente dentro do período estipulado para atingir os objetivos do milênio⁷. Uma vez que a intervenção analisada, a gestão do governo do Estado do Ceará, caracteriza-se como universal, utilizar-se-á o método do controle sintético. Além disso, frisa-se que as análises basear-se-ão em indicadores tanto de cunho quantitativo quanto qualitativo.

Este trabalho conta com mais 4 seções além desta introdução. A próxima seção elenca alguns trabalhos significativos na literatura acerca do método do controle sintético. A seção 3 descreve minimamente esse método e os dados utilizados. A seção 4 versa sobre os resultados obtidos e, por fim, uma seção é destinada para concluir o trabalho.

7. Uma vez que o foco do estudo, Educação, seja relativo a um dos objetivos do milênio.

2 MÉTODO DO CONTROLE SINTÉTICO

O Método do Controle Sintético, doravante MCS, é um método proposto inicialmente por Abadie e Gardeazabal (2003) e refinado, posteriormente, por Abadie, Diamond e Hainmueller (2010).

Segundo Athey e Imbens (2016) o método do controle sintético é, provavelmente, a mais importante inovação em avaliação de políticas dos quinze últimos anos. Os autores destacam que o MCS é construído sobre a estimativa de diferenças em diferenças e apontam escolhas e restrições alternativas que podem ser impostas para melhorar o desempenho do método em relação ao método de estimação de diferenças em diferenças puro.

Com dados de 1970 a 2000, Abadie, Diamond e Hainmueller (2010) avaliaram os efeitos da Proposição 99, um programa de controle de fumo de larga escala implementado na Califórnia em 1988. Os resultados mostraram que os efeitos desse programa de controle do tabagismo foram muito maiores do que as estimativas desenvolvidas por outros métodos. Isso mostra, portanto, maior acurácia na captação de efeitos de intervenções.

Com resultados ainda preliminares, Ando e Sävje (2013) discutem a importância da hipótese de Distribuição Idêntica e Independente para o processo de inferência e propõem duas novas estatísticas de teste que possuem tamanhos adequados e maior poder quando aplicadas aos testes de hipóteses vis-à-vis os propostos por Abadie, Diamond e Hainmueller (2010, 2015). Abadie, Diamond e Hainmueller (2010) detêm a primazia na proposição de um procedimento de inferência que consiste em estimar os P-valores através de testes de permutação.

Por outro lado, Firpo e Possebom (2016), propõem uma nova maneira de estimar Conjuntos de Confiança para o MCS e descobrem que as estatísticas de teste que o utilizam superam as estatísticas de teste comumente utilizadas na literatura de avaliação.

Em relação a imparcialidade assintótica do MCS discutido em Abadie, Diamond e Hainmueller (2010), Kaul et al. (2015) demonstram, teórica e empiricamente, que o uso de todos os resultados pré-intervenção como preditores econômicos pode provocar viés, forçando o estimador de controle sintético a ignorar todas as outras covariadas predictoras, válido independentemente da importância dessas covariadas.

Kaul et al. (2015) exemplificam que os resultados da estimativa podem mudar consideravelmente quando a armadilha supracitada é evitada. Uma alternativa seria incluir, além das covariadas, apenas a média dos resultados da pré-intervenção tal como em Abadie e Gardeazabal (2003) ou Kleven, Landais e Saez (2013). Pode-se, também, ao invés da mé-

dia dos resultados optar por incluir o último valor de pré-tratamento, além do conjunto de covariadas. Em suma, se os dois esquemas supracitados diferirem significativamente, talvez seja melhor refazer o modelo econômico e/ou suas covariadas. Contudo, se eles forem aproximados, pode-se escolher uma dessas alternativas.

Mais recentemente, Carling e Li (2017), ao examinarem as propriedades estatísticas do estimador de controle sintético, apontam que o MCS é poderoso com um número limitado de controles no grupo de doadores e um período de pré-intervenção bastante curto. Isso é válido desde que o parâmetro de intervenção seja uma especificação paramétrica do efeito de intervenção, a duração do período pós-intervenção seja razoavelmente longa e o ajuste da unidade de controle sintético para a unidade exposta no período pré-intervenção seja bom, ou seja, tenha havido perfeita mimetização dos grupos tratado e controle antes do tratamento.

Como características críticas do MCS, Carling e Li (2017) ranqueiam a discrepância na variável de resultado entre a unidade tratada e a unidade de controle sintético no período pré-tratamento, o período pós-tratamento ou de duração da política/intervenção e o tamanho do grupo de doadores. Outras características, como discrepância nas covariadas, correlação de unidades de controle, erros de medição, tempo do efeito de intervenção, duração do período pré-tratamento e ajuste do modelo entre resultados e covariadas são menos críticas.

Direcionando o discurso para as aplicações temos que, apesar de relativamente recente, o MCS espalhou-se rapidamente para diversas áreas e abordagens, como por exemplo: recursos e desastres naturais (Barone e Mocetti (2014) e Smith (2015)), finanças internacionais (Jinjarak, Noy e Zheng (2013)), liberalização econômica (Billmeier e Nannicini (2013)), reformas políticas (Billmeier e Nannicini (2009) e Ribeiro, Stein e Kang (2013)) e desenvolvimento local (Ando (2015) e Gobillon e Magnac (2016)).

Consideremos alguns exemplos da literatura nacional: Carrasco, De Mello e Duarte (2014) analisam a gestão do governo federal de um partido político específico; Ribeiro et al. (2014) analisam o impacto/custo econômico das chuvas de 2008 em Santa Catarina; Signor e Petterini (2017) analisam o impacto de uma inovação no mercado de motocicletas.

No que tange a avaliações especificamente de políticas públicas, consideremos como exemplo Silveira-Neto et al. (2013) que analisam o Programa Pacto pela Vida⁸ e Souza (2014) que analisa os efeitos sobre taxaço e inflação de um programa de nota fiscal no estado de São Paulo. Além dos já citados Lavor e Arraes (2014), dentre outros.

8. Programa de segurança pública implementado no estado do Pernambuco.

3 ESTRATÉGIA EMPÍRICA

A estratégia empírica para avaliar o impacto da gestão do governo do estado do Ceará em um período específico consiste em utilizar as informações do estado do Ceará como tratado, e os demais estados brasileiros como possíveis candidatos a grupo de controle. Para testar se a gestão governamental alterou os indicadores educacionais no Ceará empregar-se-á o método do Controle Sintético.

Esse método é utilizado em avaliações de impacto, principalmente em casos nos quais a aplicação do tratamento é universal. Nesses casos, há a necessidade de um método estatístico objetivo que simule o comportamento da unidade tratada em situação hipotética na qual não houvesse sido exposta ao tratamento.

Assim, no método de controle sintético o contrafactual é formado com base nas características das potenciais unidades de comparação, uma vez que tal método geralmente reproduz melhor a unidade de interesse vis-à-vis a utilização de apenas uma unidade comparativa. Em outras palavras, o controle é construído de forma a ser o mais assemelhado possível ao tratado no período anterior ao tratamento.

Basicamente, a ideia é que se o desempenho do tratado e sintético é similar no período anterior à intervenção, possíveis diferenças em desempenho após o tratamento representam o efeito resultante da intervenção.

Isto posto, tem-se grupos de comparação que são definidos não a partir de uma escolha do pesquisador, mas a partir de informações sobre a variável de interesse e sobre uma série de variáveis econômicas que exercem influência sobre a variável de interesse. Com base nessas informações, o método define a contribuição (pesos) de cada unidade considerada como doador na formação do grupo sintético, assim como as similaridades entre tratado e controle em termos de desempenho das variáveis consideradas.

Uma vez que o objetivo deste estudo é verificar o desempenho da gestão do governo do Estado do Ceará no que diz respeito a educação no ensino médio, o controle sintético considerado neste estudo foi concebido a partir de uma combinação dos outros estados brasileiros e do Distrito Federal. O governo do Estado do Ceará nos anos de 2007 a 2014 representa a intervenção de interesse deste estudo. A partir dessa definição tem-se que os anos anteriores a 2007⁹, estão inseridos no período conhecido como pré-tratamento.

3.1 Modelagem do Controle Sintético

A seguir apresenta-se formalmente a metodologia de estimação do controle

9. Mais precisamente de 2001 a 2007.

sintético com base em Abadie, Diamond e Hainmueller (2010). Considerando-a, desde já, para o caso em análise neste estudo.

Considera-se que existam $J + 1$ estados (unidades) no Brasil e que apenas o Ceará é submetido a intervenção, a gestão do governo do Estado do Ceará, em determinada data, restando J unidades como potenciais controles. Definem-se Y_{it}^N como o resultado observado da educação em i estados, no tempo t na ausência de intervenção e, Y_{it}^I o resultado observado para unidade i no tempo t se a unidade i for submetida à intervenção nos períodos entre $T_0 + 1$ e T . A hipótese usual é de que não há efeitos da intervenção no período pré-tratamento, isto é, para $t \in \{1, \dots, T_0\}$ e para $i \in \{1, \dots, J + 1\}$ temos $Y_{it}^I = Y_{it}^N$.

Logo, considerando $\alpha_{it} = Y_{it}^I - Y_{it}^N$ o efeito da intervenção para o estado i no período t , tem-se que:

$$Y_{it}^I = Y_{it}^N + \alpha_{it}$$

Seja D_{it} um indicador que recebe valor um se a unidade i sofre intervenção no tempo t e valor zero caso contrário. Pode-se inferir que o resultado observado para a unidade i é $Y_{it}^I = Y_{it}^N + \alpha_{it}D_{it}$. Como apenas o primeiro modelo sofre a intervenção e somente após o período T_0 tem-se que:

$$D_{it} = \begin{cases} 1 & \text{se } i = 1 \text{ e } t > T_0 \\ 0 & \text{c. c.} \end{cases}$$

O objetivo consiste em estimar o efeito da intervenção sobre o estado de interesse, Ceará, no período posterior ao tratamento, ou seja, estimar-se-á o vetor $(\alpha_{1T_0+1}, \dots, \alpha_{1T})$ para $t > T_0$:

$$\alpha_{1t} = Y_{1t}^I - Y_{1t}^N = Y_{1t} - Y_{1t}^N$$

Dado que Y_{1t}^I é observado, para encontrar o efeito de interesse é necessário apenas estimar Y_{1t}^N . Seja Y_{it}^N dado por um modelo de fatores tal como:

$$Y_{it}^N = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \varepsilon_{it} \quad (1)$$

Temos que δ_t denota um fator comum a todos os estados, Z_i representa um vetor $(r \times 1)$ de características observadas, θ_t um vetor $(1 \times r)$ de parâmetros desconhecidos, λ_t é um vetor $(1 \times F)$ de fatores comuns desconhecidos, μ_i um vetor $(F \times 1)$ de cargas fatoriais e ε_{it} é um choque aleatório com média zero.

Considere um vetor $J \times 1$ de pesos $W = (w_2, \dots, w_{J+1})'$ tal que $w_j \geq 0$ para $j = 2, \dots, J + 1$ e $w_2 + \dots + w_{J+1} = 1$. Cada valor particular do vetor W representa um potencial controle sintético. Ou seja, cada vetor W representa uma combinação particular

de pesos associados aos estados em comparação. O valor da variável resultado para cada controle sintético representado por W é dado por:

$$\sum_{j=2}^{J+1} w_j Y_{jt} = \delta_t + \theta_t \sum_{j=2}^{J+1} w_j Z_j + \lambda_t \sum_{j=2}^{J+1} w_j \mu_j + \sum_{j=2}^{J+1} w_j \varepsilon_{jt} \quad (2)$$

Definido o vetor $K = (k_1, \dots, k_{T_0})'$ com tamanho $T_0 \times 1$ como uma combinação linear de resultados pré-intervenção, tem-se: $\bar{Y}_i^K = \sum_{s=1}^{T_0} k_s Y_{is}$. Considerando, para simplificar, o caso $k_1 = k_2 = \dots = k_{T_0} = 1/T_0$, o qual resulta em $\bar{Y}_i^K = T_0^{-1} \sum_{s=1}^{T_0} Y_{is}$, teria-se uma simples média das variáveis de resultado para os períodos anteriores à intervenção.

Suponha que se possa escolher $(w_2^*, \dots, w_{J+1}^*)'$ tal que:

$$\sum_{j=2}^{J+1} w_j^* \bar{Y}_{j1}^K = \bar{Y}_1^K \quad \text{e} \quad \sum_{j=2}^{J+1} w_j^* Z_j = Z_1 \quad (3)$$

Dessa forma, se $\sum_{s=1}^{T_0} \lambda_s / T_0 \neq 0$, então:

$$Y_{1t}^N - \sum_{j=2}^{J+1} w_j^* Y_{jt} = \frac{\lambda_t}{\sum_{s=1}^{T_0} \lambda_s / T_0} \sum_{j=2}^{J+1} w_j^* \frac{1}{T_0} \sum_{s=1}^{T_0} (\varepsilon_{js} - \varepsilon_{1s}) - \sum_{j=2}^{J+1} w_j^* (\varepsilon_{jt} - \varepsilon_{1t}) \quad (4)$$

É possível mostrar também que, sob condições usuais, a média do lado direito dessa equação se aproxima de zero. Dessa forma, para $t \in \{T_0 + 1, \dots, T\}$ um estimador para α_{1t} pode ser representado por:

$$\hat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \quad (5)$$

Uma vez que a equação (3) vale somente com igualdade se (\bar{Y}_1^K, Z_1) pertencer ao conjunto $\{(\bar{Y}_2^K, Z_2), \dots, (\bar{Y}_{J+1}^K, Z_{J+1})\}$, o que na prática dificilmente ocorre. Nesse caso, um conjunto de pesos é escolhido de forma que a equação (4) seja válida aproximadamente.

No modelo, $X_1 = (Z_1', \bar{Y}_1^{K_1}, \dots, \bar{Y}_1^{K_M})'$ é considerado como um vetor $(k \times 1)$, com $k = r + M$, de características pré-intervenção para o estado considerado tratado. De maneira similar, X_0 é uma matriz $(k \times J)$ que contém as mesmas características para os estados que não sofreram a intervenção. O vetor W^* é selecionado de forma a minimizar a distância, $\|X_1 - X_0 W^*\|$, entre X_1 e $X_0 W^*$, sujeito a $w_2 \geq 0, \dots, w_{J+1} \geq 0, w_2 + \dots + w_{J+1} = 1$. Em particular, considera-se $\|X_1 - X_0 W^*\|_V = \sqrt{(X_1 - X_0 W^*)' V (X_1 - X_0 W^*)}$, em que V é uma matriz simétrica semi-definida $k \times k$, que atribui pesos às covariadas de acordo com seu poder preditivo sobre o resultado.

Posto isso, é possível estimar o resultado para a unidade tratada, o Estado do Ceará, e para o controle sintético e, a partir disso, inferir sobre o efeito da intervenção.

3.2 Dados

Este estudo visa fornecer uma análise tanto de cunho quantitativo quanto qualitativo. Uma vez que a unidade analisada é o governo do Estado do Ceará e que suas responsabilidades, no que diz respeito a educação, recaiam principalmente sobre o ensino médio, buscar-se-á transcender a análise do fornecimento do serviço de educação, avaliando também a qualidade da educação no ensino médio do estado.

Dessa forma, para a análise quantitativa será utilizado o número de matrículas¹⁰ no ensino médio. Por outro lado, como indicador qualitativo analisar-se-á o resultado médio obtido pelo estado no Exame Nacional do Ensino Médio – ENEM no período referente aos anos de 2001 a 2014.

Os dados referentes às características da população e renda foram obtidos na Pesquisa Nacional por Amostra e Domicílio – PNAD¹¹, já os dados relacionados ao desempenho dos estados no ENEM foram obtidos no site do INEP¹². Os dados adquiridos na PNAD foram compatibilizados e deflacionados utilizando o *datazoom*¹³.

Os dados sobre o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério – FUNDEF e Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação – FUNDEB, foram obtidos no site do Tesouro Nacional¹⁴, em frequência mensal e agregados anualmente de forma a compatibilizar com os dados obtidos na PNAD. Permite-se, com isso, deflacionar os dados utilizando o deflator gerado pelo *datazoom*, o mesmo processo foi utilizado para deflacionar os dados referentes ao PIB estadual *per capita*.

Devido as diferenças metodológicas entre PNAD e CENSO os dados referentes a 2010 foram obtidos via interpolação. Tal mecanismo foi utilizado considerando a afirmação de Signor e Petterini (2017) que para o cálculo do controle sintético a série em análise não pode apresentar dados faltantes do início ao fim do período analisado. Por outro lado, no que diz respeito aos preditores, estes podem apresentar *missing*. No entanto, é necessário que não exista nenhum ano sem pelo menos uma informação durante o período estudado.

De forma a facilitar a compreensão do estudo e entendimento das variáveis

10. Devido fato de se buscar indivíduos entre 15 e 17 anos matriculados no ensino médio regular e público os dados foram obtidos com base na PNAD, uma vez que os dados apresentados no site do INEP não permitem tal procedimento.

11. Disponível em: <https://ww2.ibge.gov.br/home/default.php>

12. Disponível em: <http://portal.inep.gov.br/web/guest/microdados>

13. O *Datazoom* deflaciona variáveis derivadas de rendimento para setembro de 2012, utilizando um deflator para rendimentos baseado no INPC do IBGE, construído a partir da média geométrica do índice de preços entre agosto e setembro de cada ano, de modo a centrá-lo no início de setembro, quando, em geral, são pagos os salários.

14. Disponível em: [http://sisweb.tesouro.gov.br/apex/f?p=2600:1:::~:](http://sisweb.tesouro.gov.br/apex/f?p=2600:1:::)

utilizadas, o quadro abaixo apresenta um compilado das principais informações sobre as variáveis.

Quadro 1 – Descrição das Variáveis.

Variável	Descrição
Variáveis de Interesse	
<i>Ind. Matriculados</i>	Número de indivíduos entre 15 e 17 anos matriculados no ensino médio regular e público, obtido com base na PNAD.
<i>ENEM</i>	Resultado no ENEM por candidato obtido no site do IPEA, normalizados* e tomado a média por estado.
Preditores	
<i>Tx. Desemprego</i>	Proporção de indivíduos desocupados dos indivíduos em idade economicamente ativa no estado. (PNAD)
<i>Log Transf.</i>	Transferências do FUNDEB e FUNDEF para os estados, obtido no site do tesouro nacional, tomado em logaritmo.
<i>Log Renda</i>	Rendimento mensal familiar médio por estado, obtido na PNAD, e tomado em logaritmo.
<i>Log PIB</i>	PIB estadual per capita, obtido no site do <i>Ipeadata</i> , e tomado em logaritmo.
<i>Sexo</i>	Proporção de indivíduos do sexo feminino no estado entre 15 e 17 anos, obtido na PNAD.
<i>Região</i>	<i>Dummy</i> de região.

Fonte: Dados da Pesquisa. Elaboração Própria.

Obs.: Sobre todos os dados obtidos na PNAD utilizou-se o peso amostral.

Nota: *A normalização deve-se à mudança de metodologia de atribuição de resultados em 2009 para a Teoria de Resposta ao Item – TRI.

4 RESULTADOS E DISCUSSÃO

Para fins de melhor explanação, a análise dos resultados foi dividida em duas subseções. Primeiramente, trata-se de análises gráficas de cunho descritivo. Em seguida, realiza-se uma análise utilizando o método do controle sintético, parte principal deste trabalho. Ambas análises abordam a educação no estado do Ceará tanto de forma quantitativa quanto qualitativa, a lembrar, utilizando a taxa de matrícula e o desempenho no ENEM.

Em relação a abordagem quantitativa, com o intuito de transcender a análise acerca de indivíduos matriculados e devido ao perfil de interesse sobre estes, optou-se pela utilização de dados da PNAD, pesquisa realizada no mês de setembro. Acredita-se que a partir desta abordagem específica pode-se supor que o indivíduo que frequenta a escola em setembro permanecerá até o final do período letivo. Ou seja, supõe-se que o indi-

víduo possui menor propensão a abandonar a escola. Isso pode indicar uma espécie de frequência/permanência ao invés de tão somente uma taxa de matrícula. Grosso modo, tal procedimento permite captar se o indivíduo permaneceu na escola.

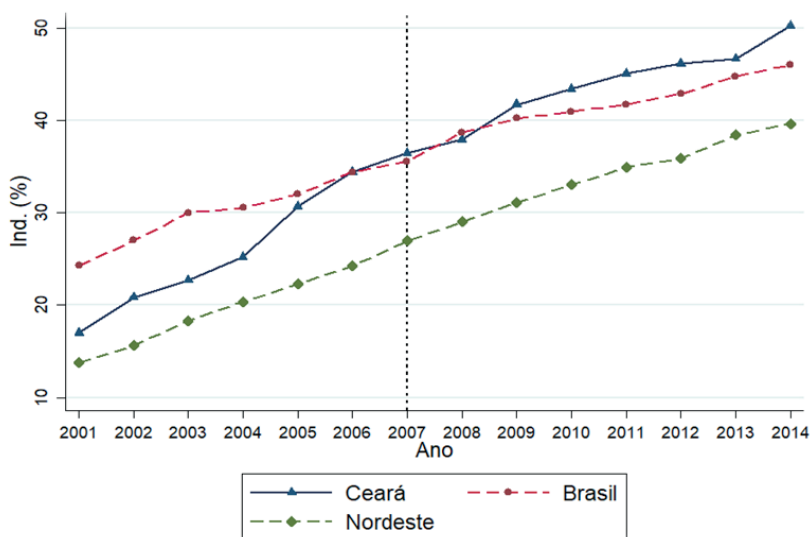
4.1 Análise Descritiva

Primeiramente, adotou-se uma abordagem de teor quantitativo. A análise, então, utiliza a evolução da proporção de indivíduos entre 15 e 17 anos¹⁵ matriculados no ensino médio regular do ensino público durante o período de 2001 a 2014.

Em relação ao discutido no parágrafo anterior, o gráfico 1, abaixo, mostra que o estado do Ceará apresentou uma evolução crescente e contínua ao longo do período analisado. A média do estado do Ceará permaneceu sempre acima da média da região Nordeste. Ainda que a média da região Nordeste tenha apresentado um aumento praticamente constante ao longo do tempo, o estado do Ceará conseguiu obter uma maior distância entre as médias vis-à-vis o início do período analisado.

Já em relação à média nacional, o estado do Ceará apresentou resultados abaixo dessa média entre os anos de 2001 a 2006, passando a oscilar em torno da média nacional entre 2006 e 2008. Contudo, a partir de 2009 ultrapassa a nacional e permanece acima, até o fim da série considerada, 2014. Isso mostra um bom desempenho do estado do Ceará nesse quesito, ou ainda da intervenção em análise, a lembrar, o governo do estado do Ceará, de 2007 a 2014.

Gráfico 1 – Proporção de Indivíduos entre 15 e 17 anos no ensino médio público



Fonte: Elaboração própria.

15. Faixa etária tomada como parâmetro para entrada e conclusão do ensino médio regular.

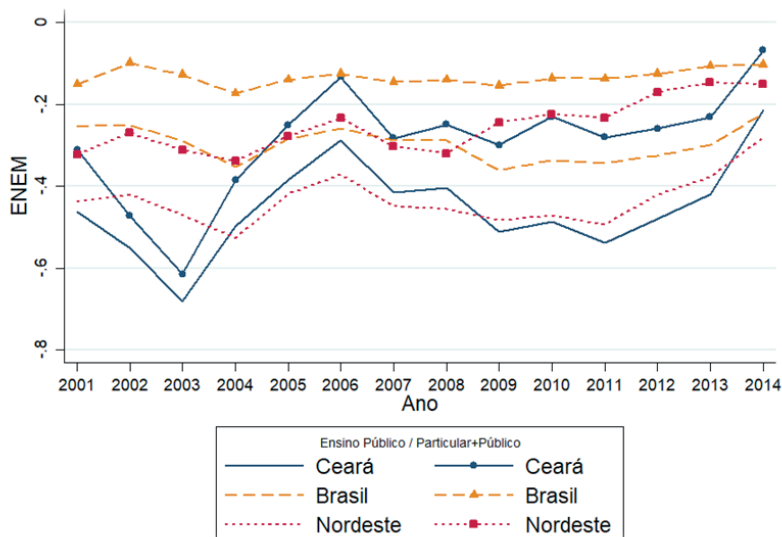
Por outro lado, uma abordagem descritiva que contemple aspectos da qualidade da educação no ensino público do estado do Ceará pode ser visualizada no gráfico 2. Este mostra, no mesmo recorte temporal abordado neste estudo, a evolução do desempenho no ENEM do ensino público do estado comparado a região Nordeste e ao Brasil.

Além disso, agregou-se à descrição os resultados sob os mesmos aspectos citados no parágrafo anterior agregando também os resultados das escolas privadas. Isto fora feito, supondo que o governo do estado fomenta alguma espécie de sinergia entre as instituições privadas e/ou que as mesmas dependam de alguma forma de políticas adotadas de cunho universal a estudantes e/ou às próprias instituições enquanto empresas sediadas no estado.

Neste caso, em relação a tão somente as escolas públicas, os resultados mostram que em nenhum momento da série histórica o rendimento do Ceará é maior que a média nacional. Porém, destaca-se o rápido crescimento de 2011 a 2014 de tal forma que o estado do Ceará acompanha o desempenho nacional nesse último ano considerado neste estudo. Já em relação ao desempenho da região Nordeste não existe uma tendência definida, pois as linhas circundam entre si durante toda a trajetória.

A análise da evolução do desempenho agregado (público e particular) é análoga a apresentada até agora, logo não carece de descrições. Fica, portanto, a cargo do leitor.

Gráfico 2 – Desempenho do ensino público do estado do Ceará no ENEM



Fonte: Elaboração própria.

4.2 Análise via Método do Controle Sintético

Esta parte do trabalho destina-se a análises utilizando o método do controle sintético.

Seguindo a estrutura proposta no início desta seção desenvolveu-se, primeiramente, uma análise de teor quantitativo.

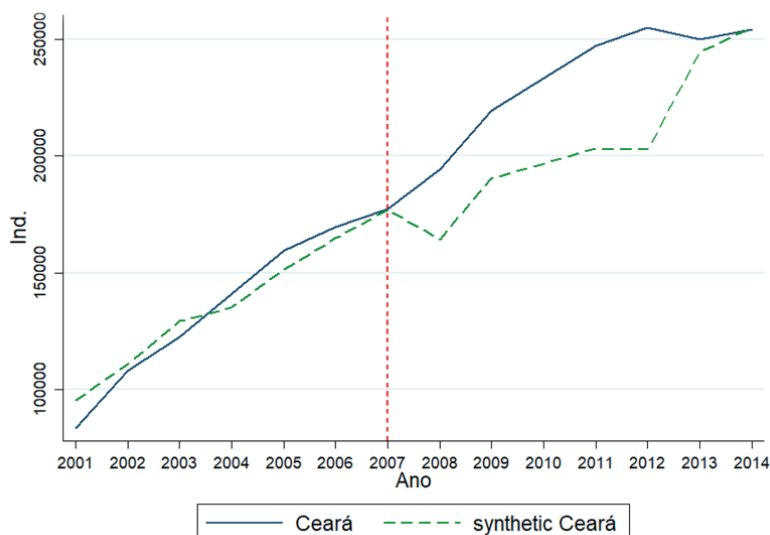
Observando o gráfico 3, percebe-se uma ótima mimetização entre o tratado, Estado do Ceará, e o controle sintético no período pré-tratamento. Compuseram o controle presente no gráfico 3, abaixo, os estados do Maranhão, Bahia, São Paulo e o Distrito Federal¹⁶.

Os resultados apontam que o estado do Ceará apresenta melhor desempenho em relação ao quesito analisado, matrícula e manutenção de indivíduos no ensino médio, durante todo o período da intervenção considerada. Tal desempenho ocorreu, inclusive, desde o primeiro período da intervenção. Isso se deve ao fato de ser um indicador quantitativo que, em geral e especialmente neste caso, responde rapidamente¹⁷ ao estímulo.

Tal resultado é fundamental para o bom desempenho dos resultados qualitativos, uma vez que o ingresso e permanência de mais indivíduos no ensino médio implica em maior utilização do capital humano disponível na sociedade.

Além disso, uma vez que vagas criadas e não preenchidas e/ou abandonadas (ociosas de alguma forma) implicam em ineficiência/desperdício de recursos públicos, pode-se interpretar esse resultado como um indicativo de melhor utilização da capacidade instalada para tal segmento da educação básica, pelo governo do Estado.

Gráfico 3 – Controle Sintético: Indivíduos entre 15 e 17 anos no ensino médio público.



Fonte: Elaboração própria.

16. A tabela completa com os pesos de cada estado na concepção do Ceará Sintético está disposta na tabela 2, no Apêndice.

17. Pode-se dizer que responde “imediatamente” no sentido de referir-se ao mesmo ano, uma vez que os dados são anuais.

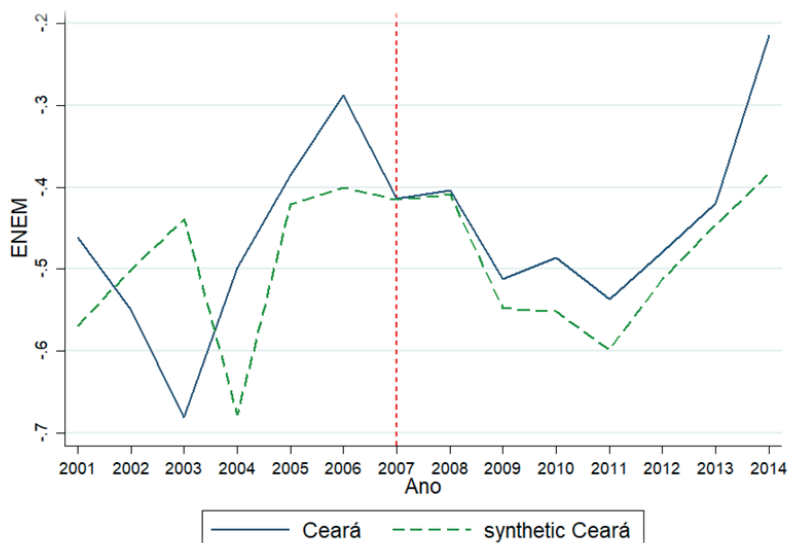
Agora, na análise qualitativa utilizando o controle sintético tem-se o desempenho do ensino médio das escolas públicas¹⁸ no ENEM. Muito embora não se tenha conseguido uma mimetização perfeita entre as trajetórias do tratado e do sintético no período pré-tratamento, percebe-se que há uma perseguição de um pelo outro seguindo a mesma trajetória com uma espécie de defasagem. Isso é comumente observado quando não se tem uma tendência bem definida na trajetória durante esse período.

Um ponto importante a se destacar é a igualdade no ponto de divisão dos períodos pré e pós-tratamento, implicando numa situação de igualdade no início da intervenção. Pode-se, inclusive, apontar essa igualdade como um requerimento mínimo¹⁹ para a análise/avaliação de qualquer intervenção via método do controle sintético. Alguns estudos, ou abordagens componentes dos mesmos, não conseguem essa igualdade.

Exemplos de casos nos quais tanto não há uma boa mimetização quanto não há igualdade no ponto de início da intervenção podem ser vistos em Carrasco, De Mello e Duarte (2014), Grier e Maynard (2016) e Signor e Petterini (2017).

Para esta análise, compuseram o controle presente no gráfico 4, abaixo, os estados do Amazonas, Maranhão e Paraná²⁰.

Gráfico 4 – Controle Sintético: Desempenho do ensino público do estado do Ceará no ENEM.



Fonte: Elaboração própria.

18. Inclui-se os Institutos Federais (IF's). Porém, por serem uma pequena parcela não influencia nos resultados.

19. Condição necessária, não suficiente, para uma boa análise. Pois o mais indicado seria uma mimetização ao longo do período pré-tratamento, tal como no gráfico 3.

20. A tabela completa com os pesos de cada estado na concepção do Ceará Sintético está disposta na tabela 4, no Apêndice.

O resultado mostra que o estado do Ceará obteve um bom desempenho também nesse quesito de análise. Uma vez que o gráfico que representa o tratado está acima do controle durante todo o período de intervenção pode-se concluir que a intervenção foi benéfica para o desempenho no ENEM.

Destaca-se alguns momentos na trajetória obtida. Primeiramente a queda no ano de 2009. Fato coincidente com a mudança na metodologia de avaliação do ENEM, ano em que passou vigorar a Teoria de Resposta ao Item – TRI. Essa metodologia ganha em acurácia de análise de desempenho uma vez que pune/bonifica os candidatos ponderando questões por níveis de dificuldade. Depois disso, no resultado imediatamente posterior, 2010, há uma melhora no desempenho que pode ser vista como uma resposta positiva a nova metodologia de atribuição das notas, via preparação dos alunos especificamente e/ou direcionamento por meio de divulgação da existência/mudança da metodologia.

Entretanto, em 2011 há uma nova queda. Coincidente a esse fato ocorre a mudança no ingresso nas universidades federais, tomados desde então pelo Sistema de Seleção Unificada - SiSU, que utiliza o ENEM como prova de seleção. Essa sinalização de novas possibilidades de cursar o ensino superior em uma instituição federal, e como tal totalmente gratuita, pode ter aumentado a quantidade de participantes egressos do ensino médio²¹. Uma vez que é naturalmente esperado que os egressos apresentem menor rendimento, dada sua não preparação para a prova vis-à-vis aqueles alunos matriculados, pode-se residir nesse fato uma explicação, pelo menos parcial, dessa oscilação/redução.

Depois desse período cíclico, 2007 - 2011, o desempenho no ENEM apresenta trajetória crescente até o fim da análise, 2014. Destaca-se, ainda o crescimento acentuado de 2013 para 2014. O período de crescimento do desempenho coincide com as primeiras aplicações no ENEM das escolas estaduais de ensino profissionalizante de tempo integral. Afinal, como obtido por Mariano, Arraes e Barbosa (2016), as escolas de ensino profissionalizante do estado do Ceará demonstram maior competência em potencializar habilidades do conhecimento cognitivo vis-à-vis as escolas regulares de ensino.

Tendo em vista a necessidade natural de um maior período de tempo para obtenção de resultados referentes a aspectos qualitativos, sobretudo de forma sustentada, os resultados obtidos nesse período possuem grande parcela de esforços desempenhados no passado, seja desde o início do período da intervenção ou até mesmo anterior.

Em suma, tais resultados podem ser vistos como uma confluência de esforços e

21. Somente em dados mais recentes (desde 2014) pode-se saber se o inscrito na prova é egresso. Os dados utilizados contêm apenas a informação do tipo de escola que o inscrito cursou o ensino Médio, uma divisão binária pública e privada.

sinergia por parte do governo do estado do Ceará e demais *stakeholders* no que diz respeito a gestão da pasta da educação.

5 CONSIDERAÇÕES FINAIS

Considerada um dos principais fatores para gerar crescimento e desenvolvimento econômico, a educação mantém-se como prioridade na agenda de políticas públicas. Este trabalho avaliou o impacto da gestão governamental do Estado do Ceará nos indicadores educacionais referentes ao ensino médio, no período de 2007 a 2014. Para tanto utilizou-se todos os demais estados e o Distrito Federal como doadores para gerar o controle por meio do método do controle sintético.

As análises foram baseadas em indicadores educacionais tanto de cunho quantitativo quanto qualitativo. Para a análise quantitativa utilizou-se a proporção de indivíduos entre 15 e 17 anos matriculados no ensino médio em escolas públicas. Já para a qualitativa, utilizou-se o desempenho do ensino público do estado no ENEM.

Os resultados dividiram-se em análises descritivas e utilizando o método do controle sintético. A primeira considera apenas análises gráficas da evolução dos resultados/variáveis objetivo para ambas abordagens, quantitativa e qualitativa. Já a segunda utiliza todo o arcabouço da metodologia adotada e um conjunto de dados como informações candidatas a explicar o efeito resultante da intervenção.

A análise descritiva da abordagem quantitativa mostra que o estado do Ceará apresentou uma evolução crescente e contínua ao longo do período analisado, não somente permanecendo sempre acima da média da região Nordeste como também obtendo uma maior distância entre as médias analisadas. Já em relação à média nacional, o estado do Ceará a partir de 2009 ultrapassa a nacional e permanece acima até o fim da série considerada, 2014.

Por outro lado, em relação ao desempenho no ENEM, os resultados mostram que em nenhum momento da série histórica o rendimento do Ceará é maior que a média nacional. Porém, destaca-se o rápido crescimento de 2011 a 2014 de tal forma que o estado do Ceará acompanha o desempenho nacional nesse último ano considerado neste estudo.

Compuseram o controle sintético para a análise quantitativa os estados do Maranhão, Bahia, São Paulo e o Distrito Federal. Os resultados apontam que o estado do Ceará apresenta melhor desempenho em relação ao quesito analisado, matrícula e manutenção de indivíduos na escola no ensino médio durante todo o período da intervenção.

Já em relação a análise qualitativa, compuseram o controle sintético os estados do Amazonas, Maranhão e Paraná. O resultado também mostra que o estado do Ceará obteve um bom desempenho no ENEM durante o período da intervenção. Destaca-se, a trajetória crescente desde 2011 e, ainda, o crescimento acentuado de 2013 para 2014, período coincidente com os primeiros resultados das escolas estaduais de ensino profissionalizante (EEEP), de tempo integral.

Dos resultados obtidos pode-se inferir que o ingresso e permanência de mais indivíduos no ensino médio implica em maior utilização do capital humano disponível na sociedade. Tal fato indica melhor utilização da capacidade instalada para a Educação pelo governo do Estado, desencadeando, provavelmente, parcela dos bons resultados do desempenho no ENEM.

Em suma, pode-se concluir que a intervenção analisada, a gestão do governo do Estado do Ceará no período de 2007 a 2014, obteve bons resultados na Educação tanto no que diz respeito a resultados quantitativos, matrícula e permanência, quanto qualitativo, desempenho no ENEM.

REFERÊNCIAS

ABADIE, Alberto; GARDEAZABAL, Javier. The Economic Costs of Conflict: A Case Study of the Basque Country, **American Economic Review**, Vol. 93. 2003.

ABADIE, Alberto; DIAMOND, Alexis; HAINMUELLER, Jens. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. **Journal of the American Statistical Association**, June 2010, Vol. 105, No. 490, Applications and Case Studies. 2010.

ABADIE, Alberto; DIAMOND, Alexis; HAINMUELLER, Jens. Comparative Politics and the Synthetic Control Method. **American Journal of Political Science**, 59(2):495–510, 2015.

ANDO, Michihito. Dreams of Urbanization: Quantitative Case Studies on the Local Impacts of Nuclear Power Facilities using the Synthetic Control Method. **Journal of Urban Economics**, 85:68–85, 2015.

ANDO, Michihito; SÄVJE, Fredrik. Hypothesis Testing with the Synthetic Control Method. Working Paper. 2013. Disponível em: https://www.dropbox.com/sh/8emim3q5nid05pp/AABt07GpLnzMPe54Xe-BohmTa?dl=0&preview=Ando_Savje_synth_130605.pdf

Acesso em: 04 de outubro de 2017.

BARONE, Guglielmo. MOCETTI, Sauro. Natural Disasters, Growth and Institutions: a Tale of Two Earthquakes. **Journal of Urban Economics**, p. 52–66, 2014.

BILLMEIER, Andreas. NANNICINI, Tommaso. Assessing Economic Liberalization Episodes: A Synthetic Control Approach. **The Review of Economics and Statistics**, 95(3):983–1001, 2013.

CARLING, Kenneth; LI, Yujiao. The Power of the Synthetic Control Method. Working **papers in transport, tourism, information technology and microdata analysis N°2016:10**. Dalarna University, Sweden. 2017. Disponível em: <https://www.diva-portal.org/smash/get/diva2:1061375/FULLTEXT01.pdf> Acesso em: 25 de maio de 2017.

CARNEIRO, D.; IRFFI, Guilherme. Política de Incentivos a Escola melhora a proficiência no Ensino Fundamental? Uma avaliação do Prêmio Escola Nota Dez. In: **Anais do X Encontro Economia do Ceará em Debate**. Fortaleza. 2014.

CARNEIRO, D.; IRFFI, Guilherme. Avaliação comparativa das leis de incentivo à educação no Ceará. In: **Anais do XII Encontro Regional de Economia da ANPEC**. Fortaleza. 2017.

CARRASO, Vinicius. DE MELLO, João M. P. DUARTE, Isabela. A década perdida: 2003 – 2012. **Texto para Discussão N°626**. Departamento de Economia. PUC-RIO. 2014. Disponível em: <http://www.econ.puc-rio.br/uploads/adm/trabalhos/files/td626.pdf>. Acesso em: 25 de maio de 2017.

CUNHA, Flávio; HECKMAN, James; SCHENNACH, Susanne. Estimating the Technology of Cognitive and Noncognitive Skill Formation. **Econometrica**, Vol. 78, N° 3, p. 883–931. 2010.

CURI, Andréa Z.; MENEZES-FILHO, Naércio A. Os efeitos da pré-escola sobre os salários, a escolaridade e a proficiência escolar. **Anais do XXXIV Encontro Nacional de Economia**. Associação Nacional dos Centros de Pós-Graduação em Economia - ANPEC, 2006.

FIRPO, Sérgio; POSSEBOM, Vítor. Synthetic control estimator: a generalized inference procedure and confidence sets. **Annals of the 2016 Latin American Workshop in Econometrics (Econometric Society)**. Montevideo, Uruguay. 2016.

GRIER, Kevin. MAYNARD, Norman. The economic consequences of Hugo Chavez: a synthetic control analysis. **Journal of Economic Behavior & Organization**, v. 125, p.

1-21. 2016.

GROSSMAN, Michael. Education and nonmarket outcomes. In: HANUSHEK, E.A.; WELCH, F. **Handbook of the economics of education**, North-Holland, 2006.

HANUSHEK, Eric; KIMKO, Dennis. **Schooling, Labor Force Quality, and the Growth of Nations**. Unpublished, University of Rochester, forthcoming in *American Economic Review*. 2000.

HINRICHS, P. The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. **The Review of Economics and Statistics**, 94(3):712–722. 2012.

JINJARAK, Yothin. NOY, Ilan. ZHENG, Huanhuan. Capital Controls in Brazil — Stemming a Tide with a Signal? **Journal of Banking & Finance**, 37:2938–2952, 2013.

KAUL, Ashok; KLÖBNER, Stefan; PFEIFER, Gregor; SCHIELER, Manuel. Synthetic Control Methods: Never Use All Pre-Intervention Outcomes as Economic Predictors. **Working Paper**. 2015.

KLEVEN, H. J., LANDAIS, C., SAEZ, E. Taxation and international migration of superstars: Evidence from the European football market. **The American Economic Review**, 103(5):1892–1924. 2013.

LAVOR, Daniel C.; ARRAES, Ronaldo de A. Qualidade da educação básica e uma avaliação de política educacional para o Ceará. In: **Anais do X Encontro Economia do Ceará em Debate**. Fortaleza. 2014.

MARIANO, F. Z. ARRAES, R. A. BARBOSA, R. B. Análise Longitudinal para Avaliação do Ensino Profissionalizante. In: **Anais do XII Encontro Economia do Ceará em Debate**. 2016. Disponível em: <http://www2.ipece.ce.gov.br/encontro/2016/index.htm>. Acesso em: 04 de outubro de 2017.

RIBEIRO, Felipe. STEIN, Guilherme. KANG, Thomas. The Cuban Experiment: Measuring the Role of the 1959 Revolution on Economic Performance using Synthetic Control. 2013. Disponível em: <http://economics.ca/2013/papers/SG0030-1.pdf>. Acesso em: 04 de outubro de 2017.

RIBEIRO, F. G. STEIN, Guilherme. CARRARO, André. RAMOS, Pedro Lutz. O impacto econômico dos desastres naturais: o caso das chuvas de 2008 em Santa Catarina. **Planejamento e Políticas Públicas**. nº 43. Julho/dezembro, 2014.

SEVERNINI, Edson R. The Power of Hydroelectric Dams: Agglomeration Spillovers. **IZA**

Discussion Paper, No. 8082, 2014. Disponível em: <http://ftp.iza.org/dp8082.pdf>. Acesso em: 04 de outubro de 2017.

SIGNOR, Diogo. PETERINI, Francis Carlo. O impacto do motor *flex* no mercado brasileiro de motocicletas na liderança da Honda. **Revista Brasileira Inovação**, Campinas – SP, p. 351-380, julho/dezembro 2017.

SILVEIRA-NETO, R. da Mota. RATTON, J. L. MENEZES, T. A. de. MONTEIRO, Circe. Avaliação de Política Pública para Redução da Violência: o Caso do Programa Pacto pela Vida do Estado de Pernambuco. In: **Anais do XLI Encontro Nacional de Economia**. 2013. Disponível em: <https://econpapers.repec.org/paper/anpen2013/227.htm>. Acesso em: 04 de outubro de 2017.

SMITH, Brock. The Resource Curse Exorcised: Evidence from a Panel of Countries. **Journal of Development Economics**, 116:57–73, 2015.

SOUZA, Fernando Friaça Asmar. **Tax Evasion and Inflation: Evidence from the Nota Fiscal Paulista Program**. Dissertação de Mestrado. Pontifícia Universidade Católica. Rio de Janeiro, 2014.

APÊNDICE

Tabela 1 – Balanço dos preditores do Ceará

	Tratado	Sintético
Tx. Desemprego	0.0766667	0.0744081
Ind. (2001(1)2006)	130733.3	131186.6
Log Transf.	19.40622	19.55687
Log Renda	7.109409	7.194913
Log PIB	9.368105	9.368062
Sexo	259501.3	259903.8

Fonte: Elaboração própria.

Tabela 2 – Pesos dos Estados para construção do Ceará Sintético: Indivíduos entre 15 e 17 anos no 2º grau regular do ensino público

Estado	Peso da Und.
Rondônia	0.000
Acre	0.000
Amazonas	0.000
Roraima	0.000
Pará	0.000
Amapá	0.000
Tocantins	0.000
Maranhão	0.792
Piauí	0.000
Rio grande do Norte	0.000
Paraíba	0.000
Pernambuco	0.000
Alagoas	0.000
Sergipe	0.000
Bahia	0.078
Minas Gerais	0.000
Espirito Santo	0.000
Rio de Janeiro	0.000
São Paulo	0.040
Paraná	0.000
Santa Catarina	0.000
Rio Grande do Sul	0.000
Mato Grosso do Sul	0.000
Mato Grosso	0.000
Goiás	0.000
Distrito Federal	0.090

Fonte: Elaboração própria.

Tabela 3 – Balanço dos preditores do Ceará

	Tratado	Sintético
Tx. Desemprego	0.0766667	0.0756819
ENEM (2001(1)2004)	- 0.54814	- 0.5466372
Log Transf.	19.40622	19.58741
Log Renda	7.109409	7.12797
Log PIB	9.368105	9.319085
Região	2	1.76

Fonte: Elaboração própria.

Tabela 4 – Pesos dos Estados para construção do Ceará sintético: Desempenho no ENEM do ensino público do estado do CE

Estado	Peso da Unidade
Rondônia	0.000
Acre	0.000
Amazonas	0.255
Roraima	0.000
Pará	0.000
Amapá	0.000
Tocantins	0.000
Maranhão	0.740
Piauí	0.000
Rio grande do Norte	0.000
Paraíba	0.000
Pernambuco	0.000
Alagoas	0.000
Sergipe	0.000
Bahia	0.000
Minas Gerais	0.000
Espirito Santo	0.000
Rio de Janeiro	0.000
São Paulo	0.000
Paraná	0.005
Santa Catarina	0.000
Rio Grande do Sul	0.000
Mato Grosso do Sul	0.000
Mato Grosso	0.000
Goiás	0.000
Distrito Federal	0.000

Fonte: Elaboração própria.

UMA ANÁLISE DOS DETERMINANTES SOCIOECONÔMICOS DA (IN)SEGURANÇA ALIMENTAR EM DOMICÍLIOS CEARENSES

Maria Hosana P. Moreira*

Vitor Hugo Miro**

RESUMO

O presente trabalho aborda o problema da insegurança alimentar. O objetivo central é analisar a situação de insegurança alimentar, mensurada pela Escala Brasileira de Insegurança Alimentar – EBIA, e avaliar a significância de um conjunto de características socioeconômicas associado a ocorrência desta. A análise emprega dados da Pesquisa Nacional por Amostra de Domicílios e avalia a incidência de insegurança alimentar no estado do Ceará. Aplicando um modelo *logit* ordenado, estima-se as probabilidades de domicílios cearenses estarem em um dos quatro níveis de segurança alimentar da EBIA. Assim como a literatura aponta, a análise permitiu verificar que a renda domiciliar per capita é o principal determinante da situação de segurança alimentar nos domicílios cearenses. A análise para os anos de 2004, 2009 e 2014 mostra a redução da insegurança alimentar no Ceará ao longo da década considerada.

Palavras-Chave: Insegurança Alimentar, Ceará, *logit* ordenado.

ABSTRACT

This paper approaches the problem of food insecurity. The central objective is to analyse the food insecurity situation, measured by the Brazilian Food Insecurity Scale (Escala Brasileira de Insegurança Alimentar – EBIA), and to evaluate a significance of a set of socioeconomic characteristics associated to its occurrence. The data used in this analysis are from the National Household Sample Survey and the evaluation of the incidence of food insecurity is realized to the state of Ceará. Applying a logistic model, it is estimated that households in Ceará are likely to be in one of the four food security levels of EBIA. Like as the literature presents, the results show that per capita household income is the main determinant of the food security situation in households of Ceará. An analysis

* Mestranda em Economia Rural - MAER/UFC. Email: moreira.h.maria@gmail.com, (85) 99604-9385.

** Professor do Departamento de Economia Agrícola da UFC e do Mestrado em Economia Rural - MAER/UFC. Pesquisador do Laboratório de Estudos da Pobreza - LEP/CAEN.

for the years 2004, 2009 and 2014 shows a reduction of food insecurity in Ceará over the decade considered.

Keywords: Food Security, Ceará, ordered logit.

1 INTRODUÇÃO

O presente trabalho aborda o problema da insegurança alimentar. O objetivo central é estudar a situação de insegurança alimentar, mensurada no âmbito da Escala Brasileira de Insegurança Alimentar – EBIA, e avaliar a significância de um conjunto de características socioeconômicas associado a ocorrência desta. A análise emprega dados da Pesquisa Nacional por Amostra de Domicílios, a PNAD, levada a campo pelo Instituto Brasileiro de Geografia e Estatística (IBGE), e avalia a incidência de insegurança alimentar no estado do Ceará.

Estudos sobre a insegurança alimentar no Brasil ganharam um novo impulso com o desenvolvimento e validação da EBIA. Incorporada à Pesquisa Nacional por Amostra de Domicílios (PNAD) em 2004, 2009 e 2013, a EBIA permitiu não só a construção de um indicador de (in)segurança alimentar, mas também a realização de análises que relacionassem o indicador com um amplo conjunto de informações socioeconômicas de pessoas e domicílios.

A análise aplicada ao estado do Ceará justifica-se pela necessidade de subsidiar pesquisas e estratégias de combate às desigualdades sociais que ainda persistem no Brasil e de forma mais intensa na região Nordeste. Estudos em âmbito subnacional também são importantes pela possibilidade de considerar especificidades regionais, uma vez que as condições de oferta de alimentos e de condições para a aquisição e consumo não são distribuídas de forma homogênea em todo o território nacional.

No caso específico para o Ceará, a questão da segurança alimentar vem sendo abordada no âmbito de uma política estadual de segurança alimentar e nutricional, orientada pela Célula de Segurança Alimentar e Nutricional sediada na Secretaria do Trabalho e Desenvolvimento Social (STDS). As principais diretrizes da política recente foram estabelecidas no Plano de Segurança Alimentar e Nutricional do Estado do Ceará (PLANSAN/CE) para os anos de 2012-2015, que faz parte do próprio Plano Plurianual do Governo do Estado do Ceará.

De modo geral, o diagnóstico apresentado pelo PLANSAN apontava a insuficiência de renda das famílias mais pobres como o principal determinante para a situação de

insegurança alimentar. Além da condição de pobreza, são destacados grupos vulneráveis em relação ao acesso à uma alimentação adequada, como o caso de crianças. Também se destacam deficiências na oferta de alimentos, em função da estrutura produtiva e de fatores climáticos próprios do semiárido nordestino. Com base nesse diagnóstico, diversos objetivos e diretrizes para alcançá-los foram estabelecidos, compreendendo ações no âmbito das políticas de combate à pobreza, de atendimento de grupos vulneráveis e de estímulos à produção agropecuária (Governo do Estado do Ceará, 2012).

Neste trabalho a (in)segurança alimentar é dividida conforme a orientação da EBIA, construída a partir de um conjunto de 14 questões, inseridas no questionário da PNAD. Tais questões investigam sobre o acesso dos moradores dos domicílios, menores e maiores de 18 anos, à comida e se esta esteve presente de maneira suficiente. Desta forma, a EBIA não tem a intenção de avaliar o grau de nutrição das famílias, uma vez que se baseia em resposta dos moradores de domicílios sem que haja uma avaliação do grau biológico e/ou químico da qualidade dos alimentos ingeridos por essas pessoas. Assim, a escala mensura a percepção das famílias em relação ao acesso à alimentos.

Pela EBIA um domicílio é classificado em situação de segurança alimentar ou insegurança alimentar, qualificada nos níveis leve, moderada ou grave. O domicílio é considerado em situação de segurança alimentar quando seus moradores conseguem ter uma alimentação em quantidades suficientes e de forma continuada, sem que haja a preocupação de restrição alimentar no futuro próximo. Já a situação de insegurança alimentar é presente nos domicílios onde os moradores são assolados pela falta de comida em quantidade suficiente ou pela preocupação de que esta acabe antes que se possa comprar mais.

Ao tratar deste tema, busca-se uma melhor compreensão dos determinantes socioeconômicos relacionados à condição de insegurança alimentar, sejam as características do domicílio, como posses de bens duráveis utilizados no preparo e conservação de alimentos, acesso aos serviços básicos de saneamento, número de moradores, rendimento domiciliar; sejam as características das pessoas responsáveis por eles, como o sexo, cor/raça, escolaridade e a situação de ocupação no mercado de trabalho. Adota-se a premissa de que quanto maior o conhecimento dos determinantes da situação de insegurança alimentar, melhores condições serão reunidas para enfrentá-la, seja na elaboração de políticas públicas ou no aprimoramento das ações existentes.

O estudo realiza uma análise econométrica para medir e avaliar o tamanho e a significância estatística do efeito de um conjunto de características socioeconômicas dos domicílios e dos seus responsáveis sobre a distribuição de probabilidade dos indicadores de

insegurança alimentar. Mais especificamente, estima-se um modelo *logit* ordenado em que a classificação de insegurança alimentar, dada pela EBIA e implementada na PNAD, foi definida como variável explicada pelo modelo. Dessa forma pretende-se contribuir para a literatura sobre segurança alimentar, ainda bastante escassa na literatura brasileira na área de economia social.

Com esta proposta, o trabalho é constituído por essa introdução e mais sete seções. Na seção 2, apresenta-se uma breve discussão sobre a definição e as formas de mensuração da segurança alimentar, situando o contexto em que se desenvolveu a Escala Brasileira de Insegurança Alimentar. Na terceira seção faz-se uma breve revisão da literatura nacional com foco nos trabalhos que analisaram o tema adotando uma abordagem econômica e econométrica. Na seção 4 expõe-se o referencial teórico, a metodologia utilizada e detalhes a respeito da base de dados e variáveis empregadas na análise empírica. A seção 5 se dedica a fazer uma análise exploratória com base no indicador de segurança alimentar, mostrando o contexto do Ceará, a evolução da segurança alimentar no estado em relação à região Nordeste e ao Brasil, além de explorar a associação da segurança alimentar como características socioeconômicas dos domicílios e seus responsáveis. Na sexta seção são apresentados os resultados obtidos com o exercício de estimação do modelo *logit* ordenado e, por fim, as considerações finais do trabalho.

2 SEGURANÇA ALIMENTAR: DEFINIÇÃO E MEDIDA

Definir segurança alimentar é uma tarefa complexa, pois existe uma série de possibilidades ao definir a disposição de alimentos para as famílias. De modo simples, segurança alimentar significa, sobretudo, acesso em quantidade e qualidade suficiente aos alimentos sem que haja, no futuro próximo, a preocupação com a escassez destes.

Segundo Maxwell (1995) uma definição completa de segurança alimentar (*food security*) inclui conceitos de acesso, suficiência, vulnerabilidade e sustentabilidade. Elementos desta definição podem ser encontrados na literatura, em que a própria definição de segurança alimentar é posta em análise como em Maxwell e Frankenberger (1992) e Maxwell (1995).

Hoffman (2008) destaca que o conceito de segurança alimentar se apresenta de forma vaga, mas mesmo assim, não perde utilidade e importância ao se analisar essa questão. Este autor cita o projeto de Lei Orgânica de Segurança Alimentar e Nutricional (LOSAN), que viria a ser aprovado como Lei nº 11.346/2006, e que define Segurança Alimentar e Nutricional como:

a realização do direito de todos ao acesso regular e permanente a alimentos de qualidade, em quantidade suficiente, sem comprometer o acesso a outras necessidades essenciais, tendo como base práticas alimentares promotoras de saúde, que respeitem a diversidade cultural e que sejam social, econômica e ambientalmente sustentáveis.

Após a definição de segurança alimentar, um próximo passo lógico é estabelecer uma forma de mensurar e classificar indivíduos e domicílios/ famílias segundo sua situação de segurança alimentar. Medidas bastante intuitivas podem ser construídas com base na observação e mensuração do gasto com itens alimentício, medidas de consumo de alimentos, frequência e ingestão de alimentos ao longo de um período (24 horas, por exemplo) ou mesmo da qualidade nutricional destes. Os indicadores antropométricos também são tradicionalmente empregados para estimar a magnitude da insegurança alimentar pela sua consequência nutricional (MAXWELL e FRANKENBERGER, 1992; MAXWELL, 1995).

Ao longo da segunda metade do século XX, preocupações com a questão de segurança alimentar conduziram para um esforço de definição e mensuração. Logo após a 2ª guerra, a percepção de que a segurança alimentar consistia na limitação de disponibilidade de alimentos, ocasionada por insuficiência na oferta destes, levou o Fundo das Nações Unidas para a Alimentação (FAO), a propor um indicador de medida padronizado a partir da disponibilidade calórica *per capita*. Progressivamente e à medida que o conceito de segurança alimentar ganhava complexidade, outros indicadores foram sendo introduzidos visando captar e mensurar as novas dimensões da segurança alimentar incorporadas (SEGALL-CORRÊA; MARIN-LEON, 2009).

Na década de 1990, novas investigações realizadas por dois projetos nos Estados Unidos, permitiram a elaboração de uma escala psicométrica para medir de forma direta a magnitude da segurança ou insegurança alimentar da população, com base na percepção das pessoas quanto a suficiência de alimentos, captada por questões que investigam desde a falta de alimentos para as refeições até a preocupação sobre a escassez de alimentos e as condições para aquisição em um futuro próximo.

Esse desenvolvimento de um instrumento de medida se concretizou com o *Household Food Security Survey Module* – HFSSM, aplicado pelo Departamento de Agricultura dos Estados Unidos (*United States Department of Agriculture* – USDA). Com base em 18 itens o instrumento de medida do HFSSM compõe uma escala capaz, não apenas de mensurar a magnitude do problema da insegurança alimentar na população, mas, também, de identificar diferentes graus de acesso aos alimentos, desde a completa satisfação das

necessidades alimentares (segurança alimentar), até as experiências em níveis de progressiva gravidade de restrição de alimentos: insegurança alimentar leve, moderada e grave (SEGALL-CORRÊA; MARIN-LEON, 2009).

No Brasil, foi possível adaptar o HFSSM a partir dos esforços de um grupo acadêmico coordenado na Universidade Estadual de Campinas (UNICAMP) em conjunto com pesquisadores de outras cinco instituições¹. Como resultado, foi elaborada e validada a Escala Brasileira de Insegurança Alimentar (EBIA), que também constitui uma escala psicométrica que se baseia na percepção de indivíduos entrevistados sobre questões pertinentes às condições da alimentação nos domicílios em que residem (SEGALL-CORRÊA; MARIN-LEON, 2009).

O desenvolvimento e validação da EBIA promoveu um grande avanço na mensuração da segurança alimentar para a realidade brasileira. Em 2004, com incentivos do Ministério do Desenvolvimento Social e Combate à Fome, a EBIA foi incorporada à PNAD na forma de uma pesquisa suplementar. Surgiu assim, o primeiro diagnóstico, no Brasil, de segurança e insegurança alimentar domiciliar, com abrangência e representatividade nacionais. Até o momento, o suplemento de insegurança alimentar foi replicado na PNAD de 2009 e 2013, com alguns aprimoramentos, fato que indica a intenção de um levantamento periódico de dados permitindo a avaliação de indicadores de insegurança alimentar no Brasil².

O método estabelecido define pontos de corte fixados de acordo com o registro de respostas positivas ou negativas em um conjunto de 14 perguntas, que foram incorporadas no questionário da pesquisa suplementar da PNAD. As questões são vinculadas em seu enunciado à falta de dinheiro para comprar comida: a) a preocupação de que o alimento acabe; b) a redução da quantidade de alimentos; c) deixar de fazer alguma refeição; d) comer menos; e) sentir fome e não comer; f) perder peso; g) quantidade de peso perdida; h) fazer apenas uma refeição ou ficar o dia todo sem comer; i) deixar de ter alimentação saudável e variada; j) não comer quantidade suficiente de comida; e k) diminuir a quantidade de alimentos das refeições (GOMES; PESSANHA; MITCHELL, 2010). O Quadro 1 a seguir, apresenta este conjunto de perguntas.

1. Universidade Estadual de Campinas-UNICAMP (coordenação), Universidade de Brasília-UNB, Universidade Federal da Paraíba-UFPB, Universidade Federal do Mato Grosso-UFMT e Instituto Nacional de Pesquisa da Amazônia-INPA, e da Universidade de Connecticut-UCONN-USA (SEGALL-CORRÊA; MARIN-LEON, 2009).

2. Além da PNAD, questões necessárias para a aplicação da EBIA foram incorporadas na Pesquisa Nacional de Demografia e Saúde de 2006 (PNDS 2006) do Ministério da Saúde, entre outras pesquisas específicas.

Quadro 1. Perguntas da Escala Brasileira de Insegurança Alimentar.

Nº PERGUNTA	ENUNCIADO
1	Nos últimos três meses, os moradores desse domicílio tiveram a preocupação de que os alimentos acabassem antes de poderem comprar mais comida?
2	Nos últimos três meses, os alimentos acabaram antes que os moradores desse domicílio tivessem dinheiro para comprar mais comida?
3	Nos últimos três meses, os moradores desse domicílio ficaram sem dinheiro para ter uma alimentação saudável e variada?
4	Nos últimos três meses os moradores desse domicílio comeram apenas alguns alimentos que ainda tinham porque o dinheiro acabou?
5	Nos últimos três meses, algum morador de 18 anos ou mais de idade deixou de fazer alguma refeição porque não havia dinheiro para comprar comida?
6	Nos últimos três meses, algum morador de 18 anos ou mais de idade, alguma vez, comeu menos do que achou que devia porque não havia dinheiro para comprar comida?
7	Nos últimos três meses, algum morador de 18 anos ou mais de idade, alguma vez, sentiu fome, mas não comeu porque não havia dinheiro para comprar comida?
8	Nos últimos três meses, algum morador de 18 anos ou mais de idade, alguma vez, fez apenas uma refeição ao dia ou ficou um dia inteiro sem comer porque não tinha dinheiro para comprar comida?
9	Nos últimos três meses, algum morador com menos de 18 anos de idade, alguma vez, deixou de ter uma alimentação saudável e variada porque não havia dinheiro para comprar comida?
10	Nos últimos três meses, algum morador com menos de 18 anos de idade, alguma vez, não comeu quantidade suficiente de comida porque não havia dinheiro para comprar comida?
11	Nos últimos três meses, alguma vez, foi diminuída a quantidade de alimentos das refeições de algum morador com menos de 18 anos de idade, porque não havia dinheiro para comprar comida?
12	Nos últimos três meses, alguma vez, algum morador com menos de 18 anos de idade deixou de fazer alguma refeição porque não havia dinheiro para comprar a comida?
13	Nos últimos três meses, alguma vez, algum morador com menos de 18 anos de idade sentiu fome, mas não comeu porque não havia dinheiro para comprar comida?
14	Nos últimos três meses, alguma vez, algum morador com menos de 18 anos de idade fez apenas uma refeição ao dia ou ficou sem comer por um dia inteiro porque não havia dinheiro para comprar comida?

Fonte: IBGE.

A classificação ocorre conforme o somatório do número de respostas “sim”, constituindo scores de pontuação, diferenciados para domicílios com maiores de 18 anos de idade e sem menores de 18 anos. De acordo com o score de pontuação os domicílios são classificados em quatro categorias: segurança alimentar (SA), insegurança alimentar leve (IA leve), insegurança alimentar moderada (IA moderada) e insegurança alimentar grave

(IA grave). O quadro 2 apresenta a pontuação para a classificação dos domicílios nas quatro categorias.

Quadro 2. Pontuação para classificação dos domicílios

Classificação	Pontos de corte para domicílios	
	Com menores de 18 anos	Sem menores de 18 anos
Segurança Alimentar	0	0
Insegurança Alimentar Leve	1-5	1-3
Insegurança Alimentar Moderada	6-9	4-5
Insegurança Alimentar Grave	10-14	6-8

Fonte: IBGE.

Nesse sentido o desenvolvimento e validação da EBIA, que mede a experiência e a percepção de insegurança alimentar no nível familiar, tornou disponível um indicador direto das condições de desigualdade alimentar e nutricional (PANIGASSI, G. *et al.* 2008). Isso representou um grande avanço para a análise da questão de segurança alimentar e constitui um fomento importante para que a literatura acadêmica sobre o tema possa se desenvolver no país.

3 UMA BREVE REVISÃO DE LITERATURA NACIONAL

Estudos sobre a insegurança alimentar no Brasil ganharam um novo impulso com o desenvolvimento e validação da EBIA, e sua incorporação ao questionário da PNAD. Apesar de terem sido produzidos trabalhos e pesquisas nas mais diversas áreas; nesta seção destaca-se alguns trabalhos que utilizaram uma abordagem econômica e econométrica para estudar os determinantes da (in)segurança alimentar. O objetivo aqui não é realizar uma revisão exaustiva sobre a literatura na área, mas estabelecer alguns pontos importantes que podem nortear a análise empírica que o presente estudo se propõe a realizar.

Hoffmann (1995) utiliza dados da Pesquisa Nacional sobre Saúde e Nutrição (PNSN) e análise regressões simples para inferir sobre a prevalência de desnutrição crônica e de retardo no crescimento entre crianças menores de cinco anos de idade. O estudo mostra que o fator primordial para garantir às pessoas uma alimentação adequada é o nível de renda, mas também demonstra que, nas regiões Norte e Nordeste, mesmo depois de descontado o efeito do rendimento domiciliar, a prevalência de desnutrição está associada ao menor

acesso da população dessas regiões a serviços como saneamento, fornecimento de água tratada, serviços de saúde e educação.

Resultados semelhantes são encontrados em Hoffmann (2008, 2013 e 2014). Nestes trabalhos foram utilizados os dados levantados pelas PNADs realizadas em 2004, 2009 e 2013 respectivamente. O autor aplica a modelagem *logit* para avaliar como características socioeconômicas dos domicílios e dos chefes de domicílios afetam a probabilidade de insegurança alimentar nos domicílios brasileiros. A análise mostra que o principal determinante da situação de segurança alimentar é a renda domiciliar *per capita*. O autor ressalta que, mesmo depois de controlado o efeito da renda, se mostraram importantes as características dos chefes de domicílio, que afetam não apenas o nível de renda, mas também a estabilidades desta, como a escolaridade e o tipo de colocação no mercado de trabalho. Também são destacados o acesso a serviços públicos, características regionais e a localização dos domicílios entre áreas urbanas e rurais.

O estudo de Camelo, Tavares e Saiani (2009) avaliou o impacto do Programa Bolsa Família (PBF) sobre a situação de segurança alimentar dos domicílios beneficiários. Para este objetivo os autores empregaram os dados da Pesquisa Nacional de Demografia e Saúde de 2006 (PNDS 2006) do Ministério da Saúde e a EBIA como medida de segurança alimentar dos domicílios. Foram considerados também indicadores de saúde infantil, baseados em medidas antropométricas, e de mortalidade infantil ao nível domiciliar.

Dentre os resultados desta avaliação, destaca-se a defesa que os autores fazem do PBF como instrumento para melhorar a condição de segurança alimentar de domicílios, uma vez que a maior parcela das transferências do PBF é gasta com alimentos, no entanto, afirmam que o Bolsa Família poderia ser mais eficaz em melhorar a condição de segurança alimentar de domicílios mais próximos da linha de pobreza. Pois nestes casos um pequeno aumento da renda pode ser o suficiente para que a família em situação de vulnerabilidade alimentar “salte” para uma classe melhor de segurança alimentar. Em contrapartida esse acréscimo de renda pode não ser suficiente para melhorar a situação alimentar de famílias extremamente pobres (mais longe da linha de pobreza) que apresentam um grau de severidade alimentar maior.

O estudo de Lima *et al.* (2016) se propôs a analisar os determinantes da segurança (e insegurança) alimentar nos domicílios dos municípios não autorrepresentativos do meio rural da Região Nordeste. Para isso, os autores utilizam os dados da PNAD de 2009. O procedimento metodológico adotado se baseia na estimação das probabilidades de um domicílio se enquadrar nos níveis de segurança e insegurança alimentar da EBIA, a partir de

um modelo *logit* ordenado. Analisando os efeitos marginais, os resultados mostram a forte associação da insegurança alimentar com a renda, e os autores destacam que mulheres e as pessoas da raça negra têm maior probabilidade de residir em domicílios com insegurança alimentar, assim como os que se dedicam às atividades agrícolas e têm baixo nível de escolaridade.

O trabalho de Miro e Sales (2012) analisa os determinantes da situação de insegurança alimentar para o Ceará com dados da PNAD de 2009. Segundo estes autores, aproximadamente 4,5 milhões de pessoas residiam em domicílios que reportaram insegurança alimentar em algum nível, compreendendo um contingente de quase 50% dos domicílios do estado. Com exceção da região metropolitana, domicílios cearenses localizados em áreas urbanas possuem uma probabilidade maior de estar em situação de insegurança alimentar frente aos domicílios localizados na área rural. Isso se deve ao fato das famílias que se encontram na zona rural poderem contar com a produção própria ou doações de alimentos para garantir sua disponibilidade e acesso.

Sales e Miro (2012) também destacam que a renda dos domicílio possui um efeito direto sobre a segurança alimentar no momento em que possibilita a aquisição de alimentos suficientes. Para estes autores, admite-se também um efeito indireto, uma vez que a renda também afeta a capacidade de transporte e armazenagem de alimentos.

Como pode ser observado, a segurança alimentar está intimamente associada ao conceito de pobreza monetária, determinada diretamente pela renda ou poder aquisitivo das famílias e dos responsáveis pelos domicílios. A literatura nacional confirma esta hipótese ao estabelecer a renda domiciliar *per capita* (RDPC) como o principal determinante da situação de insegurança alimentar. Nesse contexto, as transferências condicionais do PBF se mostraram um dos instrumentos mais efetivos da estratégia de redução da insegurança alimentar no Brasil.

Evidencia-se também que, mesmo controlando os efeitos da RDPC, há diferenças regionais consideráveis e efeitos importantes de outros fatores associados a características da pessoa de referência do domicílio. Além da escolaridade, destacam-se características que afetam a inserção e o nível de rendimentos no mercado de trabalho, como cor/raça, setor de atuação, qualidade do emprego e posição.

Cabe ressaltar que há situações especiais nas quais o acesso a alimentos não depende exclusivamente da renda monetária, como é o caso da agricultura de subsistência, onde o acesso à terra e aos meios de produção são particularmente mais importantes do que a ren-

da propriamente dita. Isso foi confirmado nos estudos de Hoffmann (2008, 2013 e 2014) e Sales e Miro (2012), por um efeito positivo da localização de domicílios em áreas rurais.

A existência de água encanada e acesso a serviços de saneamento, como esgotamento sanitário e coleta de lixo, além do acesso à energia elétrica mostram a relevância da atuação do Estado na oferta de serviços básicos e os potenciais impactos sobre a condição de segurança alimentar.

4 ARCABOUÇO TEÓRICO E METODOLÓGICO

Esta seção apresenta o arcabouço teórico utilizado para determinação de insegurança alimentar de uma família. O arcabouço teórico oferece algumas fundamentações importantes que devem ser consideradas na análise empírica que, por sua vez, estima a contribuição de diversas características de domicílios (e respectivos responsáveis) sobre as chances deste se encontrar em determinado grau de insegurança alimentar.

O arcabouço teórico adotado no presente estudo tem como referência o modelo microeconômico apresentado por Rose *et al.* (1998)³. Já a metodologia aplicada se baseia em um modelo *logit* ordenado, que será apresentado aqui com tendo como referência os textos de Long (1997) e Long e Freese (2014).

4.1 Arcabouço Teórico

A metodologia aplicada neste trabalho possui amparo no arcabouço teórico que apresenta um conjunto de relações que determinam o consumo de nutrientes pelas pessoas de um domicílio. Tais relações levam em conta que uma família produz refeições combinando alimentos adquiridos no mercado, o tempo gasto na compra e no preparo, bem como o uso de bens duráveis como refrigeradores e fogões.

Assim, uma família maximiza a utilidade derivada da satisfação das necessidades alimentares e dos efeitos sobre a saúde proporcionados pelo consumo de nutrientes. A utilidade é maximizada de acordo com a função:

$$U = U(s, n, x, l) \tag{1}$$

onde s é um vetor de gostos e preferências alimentares, n é um vetor de alimentos ou nutrientes, x é um vetor de outros bens e l representa o tempo alocado em atividades de lazer.

3. Rose *et al.* (1998) realiza uma análise dos determinantes socioeconômicos da insegurança alimentar nos Estados Unidos. Na análise empírica, os autores estimam um modelo *logit* multivariado que avalia os efeitos de características socioeconômicas, incluindo o status de pobreza e benefícios do *Food Stamp Program*, (um auxílio para pessoas de baixa renda ou que não tenham renda para comprar alimentos) na propensão de um domicílio norte-americano estar em situação de insegurança alimentar.

Assume-se, ainda, que a família maximiza essa função utilidade sujeita a uma função de produção de refeições e a restrições de renda e tempo. A função de produção de alimentos é dada por:

$$A = f(x_a, l_a, k, D) \tag{2}$$

onde x_a representa um vetor de bens alimentícios adquiridos, l_a é o tempo despedido na aquisição e no preparo de refeições, k é um vetor de bens de capital e capital humano, e D é um vetor de características demográficas do domicílio, tais como tamanho e composição da família.

O problema de maximização do domicílio está ainda sujeito à uma restrição orçamentária dada por:

$$p_A x_A + p_0 x_0 = r + w(T - l_A - l) \tag{3}$$

onde p_A é um vetor de preços dos bens de alimentação, p_0 é um vetor de preços dos outros bens, w se refere a rendimentos do trabalho (basicamente salários), r representa rendimentos que não são provenientes do trabalho, T é o tempo total disponível para os membros da família.

Na forma reduzida a equação por demanda de refeições é dada por:

$$A = f(p_A, p_0, r, w, k, D) \tag{4}$$

Para definir a condição de insegurança alimentar define-se uma cesta de bens alimentícios com determinada qualidade mínima, seja na diversidade, na quantidade de itens ou de nutrientes necessários para garantir uma condição saudável às pessoas da família. Definida como β_{min} , pode-se definir um indicador de insegurança alimentar para uma família i , I_i , construindo uma *dummy* onde:

$$I_i = 1 \text{ se } B_i < B_{min} \tag{5}$$

$$I_i = 0 \text{ se } B_i > B_{min}$$

Assim, um domicílio apresentaria um indicador de insegurança alimentar se o seu consumo de alimentos estiver abaixo de β_{min} , esse indicador de insegurança alimentar é baseado em uma cesta de bens alimentícios de determinada qualidade, que por sua vez, é dada pelas condições apresentadas na função de demanda por refeições.

4.2 Metodologia econométrica

O objetivo da análise empírica realizada no presente trabalho é mensurar a magnitude e a significância estatística do efeito de um conjunto de características socioeconômicas

dos domicílios e dos seus responsáveis sobre a distribuição de probabilidade de um indicador de insegurança alimentar. O indicador utilizado é baseado na EBIA, implementada a partir dos dados da Pesquisa Suplementar de Segurança Alimentar levada a campo pelo IBGE em conjunto com a PNAD nos anos de 2004, 2009 e 2013.

Conforme apresentado na seção 2, esta escala estabelece pontos de corte fixados de acordo com o registro de respostas positivas ou negativas em um conjunto de 14 questões apresentadas no Quadro 1. O somatório das respostas constitui scores de pontuação, diferenciados para domicílios com maiores de 18 anos de idade e sem menores de 18 anos. De acordo com essa escala os domicílios são classificados em quatro categorias: segurança alimentar (SA), insegurança alimentar leve (IL), insegurança alimentar moderada (IM) e insegurança alimentar grave (IG).

Como é possível perceber, a partir do indicador de (in) segurança alimentar baseado na EBIA tem-se uma variável qualitativa e ordenada. Nesse sentido, a análise empírica aqui realizada adota um modelo *logit* ordenado, em que o indicador de (in) segurança alimentar é a variável explicada. A utilização do *logit* ordenado em vez da versão multinomial, foi feita tendo em vista a clara definição de estratos ordinais relacionados à variável dependente.

O modelo *logit* é dado por uma equação na qual uma variável latente y_i^* definida em um intervalo $-\infty \leq y_i^* \leq +\infty$, é dada por:

$$y^* = x' \beta + \varepsilon \quad (6)$$

A variável latente é estimada por um indicador observado y , dado da seguinte forma:

$$y = j \text{ se } \alpha_{j-1} < y^* < \alpha_j \text{ para } j = 1, 2, \dots, J \quad (7)$$

Onde cada termo α_j , com $\alpha_1 < \alpha_2 < \dots < \alpha_j$, representa um “valor de corte” ou “limiar”, que delimita as categorias ordenadas da variável dependente. Adotando o ordenamento da escala de insegurança alimentar com quatro categorias, segurança alimentar (SA) e insegurança alimentar leve (IL), moderada (IM) e grave (IG), pode-se definir a relação da variável observada com a variável latente de acordo com a seguinte regra:

$$\begin{aligned} y &= 0, \text{ se } -\infty \leq y^* \leq \alpha_1 \\ y &= 1, \text{ se } \alpha_1 \leq y^* \leq \alpha_2 \\ y &= 2, \text{ se } \alpha_2 \leq y^* \leq \alpha_3 \\ y &= 3, \text{ se } \alpha_3 \leq y^* \leq +\infty \end{aligned} \quad (8)$$

A probabilidade de um domicílio se encontrar em determinada situação de segurança alimentar, dado o vetor de características x , é obtida com o seguinte desenvolvimento:

$$\begin{aligned}
 P[y = j | x] &= P[\alpha_{j-1} < y^* \leq \alpha_j | x] = P[\alpha_{j-1} < x'\beta + u \leq \alpha_j | x] \\
 P[y = j | x] &= P[\alpha_{j-1} - x'\beta < u \leq \alpha_j - x'\beta | x] \\
 P[y = j | x] &= F(\alpha_j - x'\beta) - F(\alpha_{j-1} - x'\beta)
 \end{aligned} \tag{9}$$

Onde F é a função de distribuição acumulada de u , que aqui será na forma logística. A última expressão representa a probabilidade de um resultado particular $y = j$. Assim, aplicando tal modelo na análise empírica aqui proposta tem-se:

$$\begin{aligned}
 P[y = 1 | x] &= F(\alpha_1 - x'\beta) \\
 P[y = 2 | x] &= F(\alpha_2 - x'\beta) - F(\alpha_1 - x'\beta) \\
 P[y = 3 | x] &= F(\alpha_3 - x'\beta) - F(\alpha_2 - x'\beta) \\
 P[y = 4 | x] &= 1 - F(\alpha_3 - x'\beta)
 \end{aligned} \tag{10}$$

A estimação dos parâmetros do modelo é realizada pelo método de máxima verossimilhança. O logaritmo desse modelo é dado por:

$$\ln L(\beta, \alpha | y, x) = \sum_{i=1}^N \sum_{j=0}^{j=3} y_{ij} \ln p_{ij} \tag{11}$$

, onde a p_{ij} é obtida na Equação (10). Esta equação é maximizada a partir de um método numérico para estimar os vetores α e β , na qual as estimativas são consistentes, assintoticamente normais e eficientes. E as probabilidades p_{ij} , são obtidas através da equação anterior (10).

Uma abordagem de apresentação e interpretação dos coeficientes dos modelos da classe *logit* é por meio da razão de chance ou *odds ratios* para a probabilidade acumulada (Long, 1997). O primeiro passo é definir a chance de que um resultado, condicionado ao conjunto de variáveis explicativas x , seja menor ou igual a j em detrimento de resultados superiores a j . Essa função de chance é dada pela razão de probabilidades:

$$\Omega_{\leq j | > j}(x) = \frac{P[y_i \leq j | x_i]}{P[y_i > j | x_i]} \quad \text{para } j = 1, J - 1 \tag{12}$$

Com isso, pode-se calcular, inclusive, o logaritmo das chances $\ln \Omega_{\leq j | > j}(x) = \alpha_j - x'\beta$. Desta equação, tem-se:

$$\Omega_{\leq j | > j}(x) = \exp(\alpha_j - x'\beta)$$

Em que o vetor de coeficiente β é o mesmo para todas os resultados j , assumindo a suposição de egressões paralelas). Para o exemplo que a variável de qualidade $\Omega_{\leq 2 | > 2}(x) = \exp(\alpha_2 - x'\beta)$ é ordenada, tem-se:

$$\Omega_{\leq 3 | > 3}(x) = \exp(\alpha_3 - x'\beta)$$

Para determinar o efeito de uma mudança de δ em x_k , calcula-se a razão de chance de x_k da seguinte forma:

$$\frac{\Omega_{\leq j | > j}(x, x_k + \delta)}{\Omega_{\leq j | > j}(x, x_k)} = \exp(\delta \times \beta_k) \quad (13)$$

$$\frac{\Omega_{\leq j | > j}(x, x_k + 1)}{\Omega_{\leq j | > j}(x, x_k)} = \exp(\beta_k) \quad (14)$$

A razão de chance também pode ser interpretada em termos percentuais, obtendo $[(\exp(\beta) - 1) \times 100]$. Vale ressaltar a razão de chance para uma mudança em x_k é derivada mantendo todas as demais variáveis explicativas constantes. Consequentemente, não é aplicável o cálculo das razões de chance para variáveis na forma polinomiais ou em interação com outras variáveis.

4.3 Base de dados

Para viabilizar a análise empírica aqui proposta, foram utilizados os microdados da Pesquisa Nacional por Amostra Domiciliares (PNAD) realizada pelo Instituto Brasileiro de Estatística e Geografia (IBGE). Em função da periodicidade que a pesquisa suplementar foi levada a campo, analisou-se os anos de 2004, 2009 e 2013.

Para esse estudo utilizam-se as informações apenas para o estado do Ceará nos casos de entrevista realizada e domicílios particulares e permanentes. A amostra foi delimitada pela restrição de que todas as informações incorporadas nas variáveis dependente e independentes estivessem presentes em todos os domicílios considerados.

As variáveis explicativas empregadas na estimação desse modelo são inspiradas no arcabouço teórico, na literatura existente e nas possibilidades que a base de dados permite trabalhar. Estima-se a propensão de um domicílio apresentar um indicador de segurança alimentar em função de um conjunto de características do domicílio tais como rendimentos, capital humano, posse de bens e serviços, características demográficas entre outras.

5 SEGURANÇA ALIMENTAR NO CEARÁ

Nesta seção serão apresentadas informações que permitem verificar a situação de

segurança alimentar no Ceará. Alguns dados mostram a evolução do estado do Ceará tendo como referência a evolução observada na região Nordeste e no Brasil. Outros irão relacionar os índices de segurança alimentar com características dos domicílios e de seus moradores, com o objetivo específico de desenhar um perfil socioeconômico básico dos domicílios e da prevalência de insegurança alimentar no Ceará.

Os resultados das pesquisas realizadas em 2004, 2009 e 2013 apontam para uma redução da prevalência de insegurança alimentar em todas as dimensões geográficas. O gráfico apresentado na Figura 1 a seguir permite visualizar a redução do percentual de domicílios que reportaram algum nível de insegurança alimentar entre os anos de 2004, 2009 e 2013. São apresentadas as informações para o Brasil, a região Nordeste e o Estado do Ceará.

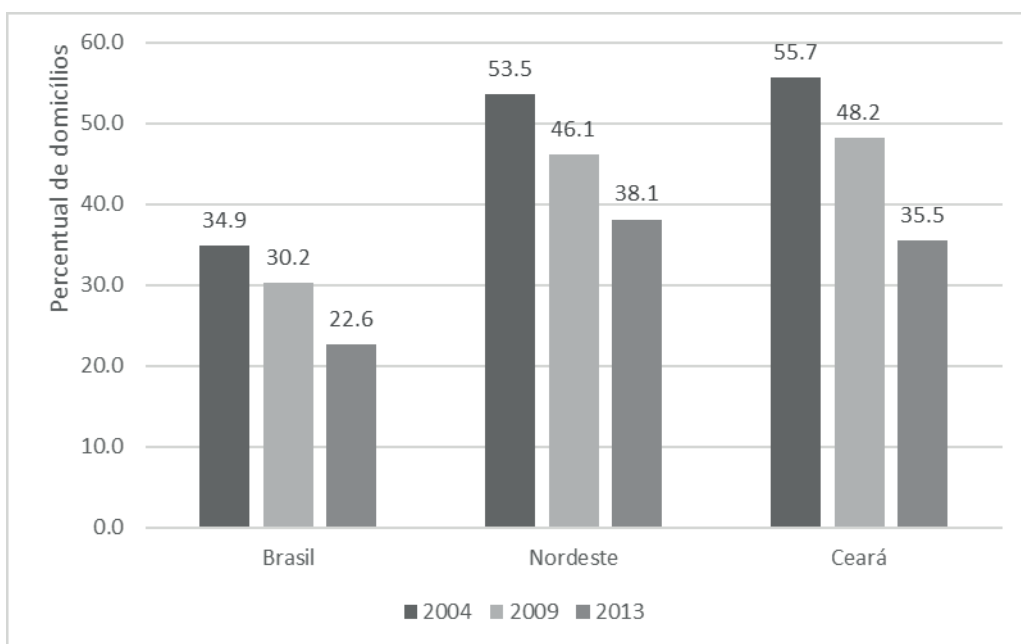


Figura 1. Prevalência de insegurança alimentar em algum nível. Brasil, Nordeste e Ceará (2004, 2009 e 2013)

Fonte: Elaboração própria com base nos dados do IBGE/ PNAD.

No gráfico apresentado na Figura 1 tem-se que, em 2004, aproximadamente 35% dos domicílios brasileiros reportaram algum grau de insegurança alimentar. No Nordeste este percentual era superior a 53% e no Ceará era maior que 55%. Cinco anos depois, os dados da pesquisa de 2009 já mostrava uma queda no percentual de domicílios em insegurança alimentar e 2013 tem-se a confirmação da tendência de redução. Em 2013, a prevalência de insegurança alimentar no Brasil era de 22,6% dos domicílios, de 38,1% no Nordeste e 35,5% no Ceará. De forma ainda mais detalhada, a Tabela 1 apresenta o percentual de

domicílios conforme os quatro níveis de (in)segurança alimentar estabelecidos na EBIA.

Em termos relativos, o Ceará apresentou uma queda ainda acentuada que a redução de insegurança alimentar no Brasil e no Nordeste ao longo do período analisado, principalmente na redução dos níveis de insegurança moderada e grave. Na última coluna da Tabela 1 esse resultado é bastante explícito, com o aumento de 45,5% dos domicílios em situação de segurança alimentar e reduções de 57% e 64,3% de domicílios em insegurança moderada e grave, respectivamente. Esse resultado demonstra que esforços para a redução da vulnerabilidade alimentar e nutricional no estado vem ao longo do tempo melhorando a situação alimentar das famílias cearenses.

Tabela 1. Percentual de domicílios conforme situação de segurança alimentar. Brasil, Nordeste e Ceará (2004, 2009 e 2013)

		2004	2009	2013	Var. % (2004-2013)
Brasil	Segurança Alimentar	65.1	69.8	77.4	18.9
	Inseg. Alim. leve	18.0	18.7	14.8	-18.1
	Inseg. Alim moderada	9.9	6.5	4.6	-53.9
	Inseg. Alim grave	6.9	5.0	3.2	-53.7
Nordeste	Segurança Alimentar	46.5	53.9	61.9	33.3
	Inseg. Alim. leve	22.9	24.8	23.6	3.3
	Inseg. Alim moderada	17.5	12.0	8.9	-49.1
	Inseg. Alim grave	13.2	9.3	5.5	-58.0
Ceará	Segurança Alimentar	44.33	51.82	64.49	45.5
	Inseg. Alim. leve	24.28	24.38	23.06	-5.0
	Inseg. Alim moderada	17.13	13.49	7.37	-57.0
	Inseg. Alim grave	14.26	10.31	5.09	-64.3

Fonte: Elaboração própria com base nos dados do IBGE/ PNAD.

Levando em consideração a importância da renda como determinante da situação de segurança alimentar, a Tabela 2 a seguir apresenta o percentual de domicílios conforme a situação de segurança alimentar ao longo dos intervalos dos decis da distribuição da renda domiciliar *per capita* (RDPC). Para estabelecer uma referência adicional e tornar a apresentação destes dados ainda mais didática, são mostrados os valores da RDPC média em cada intervalo.

Nota-se que a prevalência de insegurança alimentar declina de forma expressiva conforme avançamos para os estratos de renda mais elevada, principalmente nos níveis moderado e grave. Essa evidência reforça a importância dos programas e políticas de transferência de renda e inclusão produtiva. De modo mais geral, pode-se afirmar sem qualquer erro que, crescimento econômico, redução das desigualdades e inclusão social são os gran-

des meios para erradicar a vulnerabilidade alimentar e a fome.

Um fato que pode chamar a atenção na Tabela 2 é a presença de uma proporção não nula de domicílios com algum grau de insegurança na parte superior da distribuição da renda. Para estes casos, vale reforçar o caráter parcialmente subjetivo da EBIA, que depende diretamente da forma como as questões são reportadas pelos moradores entrevistados. Outra questão que poderia ser abordada é a significância estatística de tais valores, mas essa análise é, no presente estudo, deixada para valorações posteriores.

Tabela 2. Percentual de domicílios conforme situação de segurança alimentar e segundo intervalos dos decis da distribuição da renda domiciliar *per capita* (RDPC). Ceará (2004, 2009 e 2013).

Decis da distribuição da RDPC - Ceará (2004)										
	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°
RDPC média (em R\$)	45.4	102.9	158.8	218.8	278.5	386.0	502.1	686.5	1017.3	2619.6
Segurança Alimentar	16.4	29.4	43.7	52.6	69.2	76.6	82.6	87.7	87.7	95.3
Inseg. Alim. leve	26.8	31.5	29.3	27.5	16.9	13.8	11.4	10.4	8.8	3.2
Inseg. Alim moderada	28.4	22.8	16.9	11.0	7.6	4.4	3.2	1.9	1.6	1.5
Inseg. Alim grave	28.4	16.3	10.1	8.9	6.3	5.3	2.9	0.0	2.0	0.0
Decis da distribuição da RDPC - Ceará (2009)										
	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°
RDPC média (em R\$)	42.9	104.8	159.5	224.3	295.8	381.7	491.6	680.2	1007.5	2804.3
Segurança Alimentar	23.1	28.7	38.2	45.2	51.3	56.3	67.4	81.2	83.8	92.5
Inseg. Alim. leve	29.0	29.1	29.5	29.9	29.4	25.7	18.7	13.8	12.8	6.8
Inseg. Alim moderada	23.5	21.5	19.1	15.3	13.5	10.7	8.4	3.9	1.6	0.3
Inseg. Alim grave	24.5	20.7	13.2	9.6	5.9	7.3	5.5	1.1	1.8	0.3
Decis da distribuição da RDPC - Ceará (2013)										
	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°
RDPC média (em R\$)	29.8	104.1	159.9	222.5	290.8	377.2	496.3	684.6	1011.5	2810.9
Segurança Alimentar	41.8	34.8	52.3	51.4	54.5	62.7	69.8	75.5	82.3	92.9
Inseg. Alim. leve	27.7	40.6	31.0	31.3	31.7	24.6	22.1	17.0	12.9	5.8
Inseg. Alim moderada	12.5	15.5	10.6	10.1	8.9	8.0	5.4	5.2	2.9	0.9
Inseg. Alim grave	18.0	9.1	6.1	7.2	4.9	4.7	2.7	2.3	2.0	0.4

Fonte: Elaboração própria com base nos dados do IBGE/ PNAD.

A Tabela 3 a seguir apresenta a distribuição (%) dos domicílios em áreas urbanas e rurais conforme a situação de segurança alimentar. O primeiro ponto a se observar é que, em qualquer dos anos, a prevalência de insegurança alimentar é maior entre os domicílios localizados em áreas rurais. Além disso, a prevalência de insegurança alimentar diminuiu em maior proporção nas áreas urbanas do que nas áreas rurais.

Tabela 3. Percentual de domicílios conforme situação de segurança alimentar em áreas urbana ou rural. Ceará (2004, 2009 e 2013).

	Domicílios urbanos			Domicílios rurais		
	2004	2009	2013	2004	2009	2013
Segurança Alimentar	46.7	55.5	69.1	35.6	37.8	51.3
Inseg. Alim. leve	22.6	22.3	19.8	30.5	32.2	32.4
Inseg. Alim moderada	17.1	12.3	6.4	17.2	17.9	10.1
Inseg. Alim grave	13.6	9.8	4.7	16.7	12.1	6.3

Fonte: Elaboração própria com base nos dados do IBGE/ PNAD.

Em 2004, a proporção domicílios urbanos classificados com algum nível de insegurança alimentar era de aproximadamente 53% enquanto que em áreas rurais essa proporção era superior a 64%. Em áreas urbanas, esta proporção diminuiu para 44,5% em 2009 e para algo em torno de 31% em 2013. Em áreas rurais, o número de domicílios classificados com algum nível de insegurança alimentar também diminuiu, mas ainda se mantem em um nível relativamente elevado, de quase 50% em 2013. Essa evidência mostra que particularidades presentes em áreas rurais podem ser determinantes sobre a situação de segurança alimentar e nutricional.

Considera-se que nas áreas rurais, o acesso a alimentos não é realizado, em grande medida, por meio da aquisição, como se espera que aconteça nas áreas urbanas. Nos domicílios rurais há a possibilidade de cultivo e as condições para que isso se realize dependem de uma infinidade de fatores, dentre os quais se destacam o acesso à terra e a água e as condições ambientais e climáticas. Na região do semiárido nordestino, onde se localiza 93% do território cearense⁴, a produção de alimentos está associada a um elevado grau de risco e vulnerabilidade, em função da dependência de condições climáticas.

Assim, o acesso à terra e aos meios de produção como água e tecnologia, constituem uma importante ferramenta de redução da insegurança alimentar, não apenas no Ceará, mas em outros estados e regiões do país. E isso não é válido apenas para as áreas rurais, considerado que a produção de alimentos que abastece os domicílios urbanos também tem as áreas rurais como origem. Tais condições atribuem um papel importante para as políticas de convivência com a seca no semiárido, de acesso à terra, ao crédito rural e de assistência técnica.

Apesar da renda ou poder aquisitivo surgir como sendo o principal determinante da situação de segurança alimentar, diversos fatores também são apontados como determi-

4. Instituto de Pesquisa e Estratégia Econômica do Ceará. Ceará em Mapas. Disponível em <http://www2.ipece.ce.gov.br/atlas/capitulo1/12.htm> . Consultado em 06 de outubro de 2017.

nantes da capacidade de uma família exercer o direito de ter uma alimentação adequada. Algumas características dos moradores como escolaridade e inserção no mercado de trabalho também são importantes, pois afetam a capacidade de geração de renda no domicílio, a alocação de tempo entre trabalho, lazer e atividades domésticas, além de determinarem a forma como os indivíduos e famílias processam informações a respeito de uma alimentação adequada. Características demográficas dos domicílios, o acesso a serviços de saneamento e energia elétrica, a posse de bens duráveis como geladeiras e fogões, também possuem forte associação com a situação alimentar dos domicílios, como aponta a literatura. A Tabela 4 sumariza estatísticas descritivas (médias e proporções) relativas à estas características para os domicílios cearenses nos anos considerados.

6 RESULTADOS DA ESTIMAÇÃO ECONÔMETRICA

Esta seção apresenta e discute os resultados obtidos na estimação do modelo *logit* ordenado conforme a análise empírica proposta e os objetivos do trabalho. Os coeficientes estimados (Coef) e respectivas estimativas de erro padrão (EP) são apresentados na Tabela 5 a seguir. Conforme a proposta do trabalho, os resultados são reportados para os anos de 2004, 2009 e 2013.

Tabela 5. Modelo Logit Ordenado - coeficientes estimados – Ceará (2004, 2009 e 2013)

	2004		2009		2013	
	Coef	(EP)	Coef	(EP)	Coef	(EP)
RDPC (25° - 50° percentil)	-1,094	0,005	-0,635	0,004	-0,297	0,005
RDPC (50° - 75° percentil)	-1,884	0,008	-1,138	0,005	-0,687	0,005
RDPC (75° - 100° percentil)	-2,509	0,013	-2,045	0,008	-1,264	0,007
Nº de moradores	0,051	0,002	0,066	0,002	0,022	0,002
Nº de menores de 18 anos	0,049	0,002	0,073	0,002	0,068	0,002
Água canalizada	-0,481	0,005	-0,260	0,005	-0,219	0,005
Esgot. sanitário adequado	0,130	0,005	-0,240	0,004	-0,585	0,004
Coleta de lixo	0,285	0,005	-0,089	0,005	-0,026	0,006
Energia elétrica	0,673	0,007	0,172	0,014	-0,025	0,026
Geladeira	-0,420	0,004	-0,504	0,005	-0,537	0,007
Fogão	-1,132	0,008	-0,677	0,011	-0,379	0,023
Urbano	0,132	0,006	0,359	0,005	0,160	0,006
Metropolitano	0,010	0,004	-0,221	0,004	-0,135	0,004
Mulher	0,392	0,004	0,391	0,004	0,211	0,004
Branco	-0,181	0,004	-0,214	0,004	-0,092	0,004
Idade	0,009	0,000	0,012	0,000	-0,005	0,000
5 a 8 anos de estudo	-0,114	0,005	-0,221	0,004	-0,275	0,004
9 a 12 anos de estudo	-0,466	0,006	-0,384	0,005	-0,450	0,005
13 anos de estudo ou mais	-0,735	0,011	-0,670	0,010	-0,745	0,009
Setor Agrícola	0,080	0,005	0,252	0,005	0,160	0,005
Emprego Formal	0,440	0,011	0,314	0,008	-0,131	0,008
Emprego informal	0,823	0,010	0,439	0,007	0,222	0,006
Func. Público e militares	0,465	0,013	0,186	0,011	0,009	0,010
Conta própria	0,399	0,010	0,163	0,007	-0,142	0,006
Empregador	-0,238	0,013	-0,365	0,011	-0,648	0,012
Aposentado	-0,417	0,007	-0,160	0,007	-0,038	0,007
Corte1	-1,122	0,013	-1,089	0,019	-1,405	0,033
Corte2	0,256	0,013	0,298	0,019	0,204	0,033
Corte3	1,452	0,013	1,462	0,019	1,256	0,033
Número de observações (ponderadas)	4.401 (1.481.780)		4.946 (1.667.053)		3.748 (1.737.586)	
LR chi2 (26)	490.416		441.039		285.852	
Prob > chi2	0,0000		0,0000		0,0000	
Pseudo R2	0,1286		0,1109		0,0845	
Log likelihood	-1.661.881		-1.767.928		-1.547.871	

Fonte: Elaboração própria com base nos dados do IBGE/ PNAD.

*

Os resultados do modelo estimado são condizentes com algumas expectativas formuladas *a priori*, baseadas na literatura pré-existente. Os valores dos coeficientes na Tabela 5 não podem ser interpretados como uma influência quantitativa sobre a probabilidade da situação de segurança alimentar, uma vez que sua relação não é linear. Para realizar inferências das variáveis explicativas sobre a dependente, calcula-se a razão de chance (*odds ratio*) para o *logit* por meio do antilogaritmo.

A Tabela 6, a seguir, apresentada as razões de chance (RC) e as variações percentuais (Δ %) na razão de chance para o modelo estimado. Assim como os coeficientes estimados, as razões de chance são as mesmas para as diferentes categorias de (in)segurança alimentar (segurança, inseguranças leve, moderada e grave), devido à suposição de chances proporcionais.

Tabela 6. Modelo Logit Ordenado: Razões de chance (RC) e variações percentuais da RC estimadas – Ceará (2004, 2009 e 2013)

	2004		2009		2013	
	RC	%	RC	%	RC	%
RDPC (25° - 50° percentil)	0,335	-66,5	0,530	-47,0	0,743	-25,7
RDPC (50° - 75° percentil)	0,152	-84,8	0,320	-68,0	0,503	-49,7
RDPC (75° - 100° percentil)	0,081	-91,9	0,129	-87,1	0,282	-71,8
Nº de moradores	1,053	5,3	1,068	6,8	1,022	2,2
Nº de menores de 18 anos	1,050	5,0	1,076	7,6	1,071	7,1
Água canalizada	0,618	-38,2	0,771	-22,9	0,803	-19,7
Esgot. sanitário adequado	1,139	13,9	0,787	-21,3	0,557	-44,3
Coleta de lixo	1,329	32,9	0,915	-8,5	0,974	-2,6
Energia elétrica	1,961	96,1	1,188	18,8	0,975	-2,5
Geladeira	0,657	-34,3	0,604	-39,6	0,585	-41,5
Fogão	0,322	-67,8	0,508	-49,2	0,684	-31,6
Urbano	1,141	14,1	1,432	43,2	1,173	17,3
Metropolitano	1,010	1,0	0,802	-19,8	0,874	-12,6
Mulher	1,480	48,0	1,478	47,8	1,235	23,5
Branco	0,834	-16,6	0,807	-19,3	0,912	-8,8
Idade	1,009	0,9	1,012	1,2	0,995	-0,5
5 a 8 anos de estudo	0,892	-10,8	0,802	-19,8	0,759	-24,1
9 a 12 anos de estudo	0,627	-37,3	0,681	-31,9	0,637	-36,3
13 anos de estudo ou mais	0,479	-52,1	0,512	-48,8	0,475	-52,5
Setor Agrícola	1,083	8,3	1,287	28,7	1,174	17,4
Emprego Formal	1,553	55,3	1,369	36,9	0,877	-12,3
Emprego informal	2,278	127,8	1,551	55,1	1,248	24,8
Func. Público e militares	1,593	59,3	1,204	20,4	1,009	0,9
Conta própria	1,490	49,0	1,178	17,8	0,867	-13,3
Empregador	0,789	-21,1	0,695	-30,5	0,523	-47,7
Aposentado	0,659	-34,1	0,852	-14,8	0,962	-3,8

Fonte: Elaboração própria com base nos dados do IBGE/ PNAD.

*

Considerando as características dos domicílios, destaca-se a renda domiciliar *per capita* (RDPC) que, como era de se esperar, constitui o principal determinante da situação de segurança alimentar dos domicílios. Os coeficientes estimados e as *odds ratios* calculadas mostram que a propensão de uma família estar em uma situação de insegurança alimentar grave reduz sensivelmente com aumentos no nível de renda. Mantidas as demais variáveis constantes, em 2004 um domicílio com RDPC no 2º quartil da distribuição, possui 66,5% a menos de chance do que um domicílio no 1º quartil, de se encontrar em situação de insegurança alimentar mais rigorosa. Essa diferença se reduziu em anos posteriores. Em 2013, um domicílio com RDPC no 2º quartil da distribuição, possui 25,7% a menos de chance do que um domicílio no 1º quartil, de se encontrar em situação de insegurança alimentar grave.

Conforme se avalia as chances de domicílios com RDPC nos quartis mais elevados, reduzem as chances de vulnerabilidade alimentar. Em 2004, para um domicílio com RDPC no 4º quartil da distribuição, a chance de insegurança alimentar diminui em mais de 90% quando comparado a um domicílio no estrato mais baixo de renda. Em 2013, essa chance se reduzia em quase 72%.

O efeito do número de moradores no domicílio (membros na família), assim como a presença de crianças e jovens até 18 anos, elevam as chances de um domicílio estar em situação mais precária de insegurança alimentar.

A posse de alguns bens relacionados ao preparo e conservação de alimentos também se mostrou importante, e estatisticamente relevantes. As chances de um domicílio estar em situação de insegurança alimentar grave são menores quando o domicílio possui itens como fogão e geladeira.

A presença de água encanada em pelo menos um cômodo do domicílio também apresentou um efeito redutor das chances de insegurança alimentar. A presença de saneamento básico, como esgotamento sanitário adequado e coleta de lixo, também se mostrou significativa no âmbito de redução da insegurança alimentar nos anos de 2009 e 2013. A presença de energia elétrica não apresentou uma evidência muito clara nos anos considerados, apresentando um efeito inesperado.

Com relação à localização dos domicílios, destaca-se que estar situado na zona urbana (não-metropolitana) elevava as chances da situação mais grave de insegurança alimentar grave em aproximadamente 14%, em 2004, e pouco mais de 17% em 2013, quando comparados com domicílios na zona rural. Este resultado já era esperado como mencionado anteriormente, pela possibilidade que há na zona rural de se produzir o próprio alimento

e assim reduzir o risco de insegurança alimentar. Por sua vez, domicílios localizados na Região Metropolitana de Fortaleza, possuíam chances de insegurança alimentar quase 20% menor em 2009 e quase 13% menor do que os demais, em 2013.

Com relação às características dos chefes de domicílio como cor e gênero podem ser importantes, uma vez que tais características estão associadas à inserção destes indivíduos no mercado de trabalho e na capacidade de obtenção de renda. Ter declarado cor/raça branca reduz as chances de estar em insegurança alimentar, em relação aos que se declararam negros ou pardos. Por sua vez, um chefe de domicílio do sexo feminino eleva a chance de insegurança alimentar, quando comparado a um domicílio chefiado por um homem. Nos anos 2004 e 2009, as chances de um domicílio chefiado por mulheres estar em situação de insegurança alimentar mais severa chegaram a ser quase 50% maiores do que nos domicílios chefiados por homens.

Como era de se esperar, uma maior escolaridade do chefe de domicílio também reduz, de forma crescente, as chances do domicílio ser classificado em um nível de insegurança alimentar mais severo. Tal efeito foi bastante significativo nos anos analisados. Nos três anos, domicílios em que o chefe tinha mais de 13 anos de estudo, a chance de o domicílio estar na pior situação de vulnerabilidade alimentar é mais ou menos a metade das chances de um domicílio chefiado por um indivíduo com o menor nível de escolaridade (que têm até 4 anos de estudo).

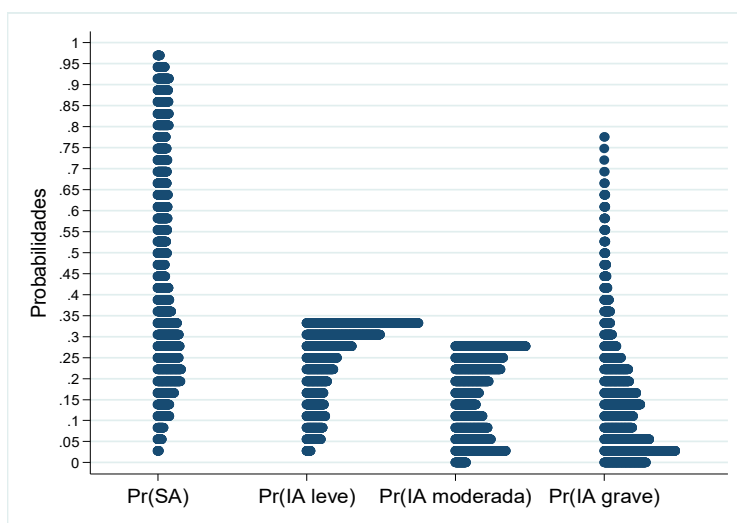
A inserção do chefe de domicílio no mercado de trabalho também se mostrou importante. No modelo estimado a categoria de referência foram os ocupados em atividades voltadas para o próprio consumo e sem remuneração. Com relação a essa categoria, em 2013, as chances dos domicílios se encontrarem em um nível mais elevado de insegurança alimentar foram menores para empregados formais, trabalhadores por conta própria e empregadores.

Comparado a um chefe de domicílio empregado informalmente, ter um emprego formal reduz as chances de insegurança alimentar do domicílio. Em 2013 essa redução é foi de quase 25%, enquanto que o emprego por conta própria reduz em 27,3%, e o fato de ser empregador em aproximadamente 57%. O setor de atividade também se mostra significativo, mas os resultados variaram entre os anos. Em 2009 e 2013, chefes de domicílio empregados em atividades agrícolas elevaram a chance de insegurança alimentar mais severa quando comparados aos empregados em outros setores (indústria, comércio, serviços, administração pública).

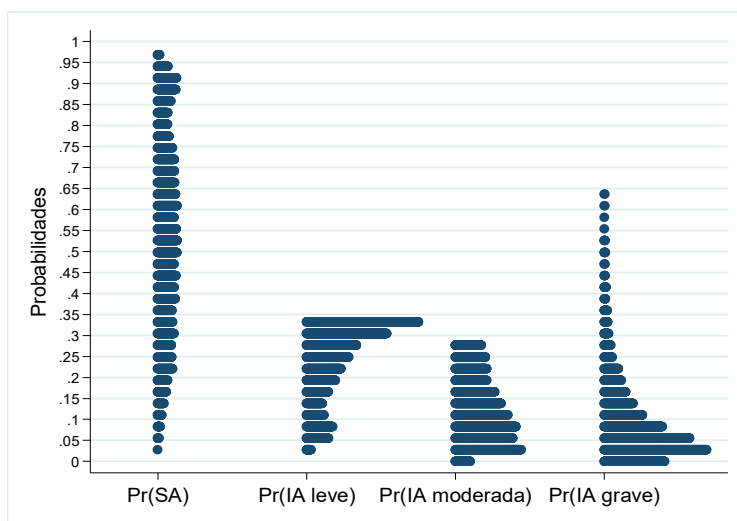
A Figura 2 a seguir, apresenta gráficos com as probabilidades previstas para as qua-

tro categorias de insegurança alimentar, de acordo com o modelo *logit* ordenado estimado para cada ano (2004, 2009 e 2013). Com a comparação entre os três gráficos é notável a redução das probabilidades de insegurança alimentar, principalmente da probabilidade de insegurança alimentar grave.

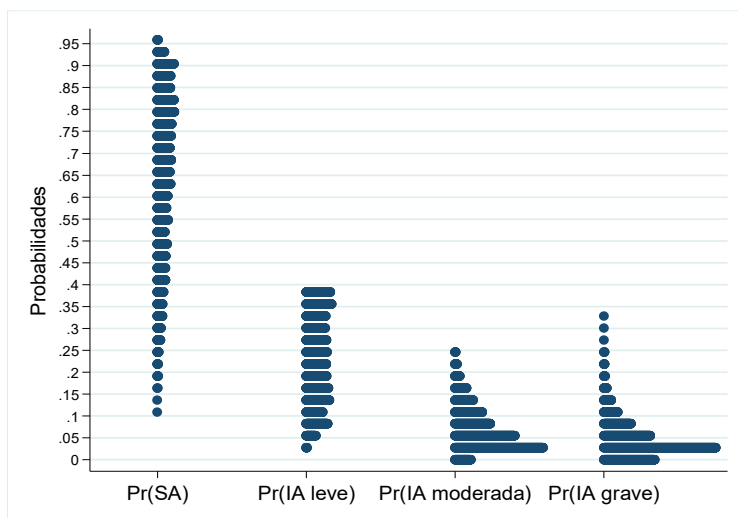
A redução observada nas probabilidades de insegurança alimentar, em qualquer um dos três níveis, aponta para evidências de que as políticas de combate à pobreza e a redução da desigualdade de renda observada no período analisado (2004 a 2013) permitiram um melhor acesso à alimentos, e a redução da vulnerabilidade alimentar e nutricional. Em particular, para o recorte geográfico do presente estudos, pode-se inferir que o estado do Ceará apresentou importantes avanços nesse sentido.



2004



2009



2013

Figura 2. Probabilidades estimadas para as categorias de segurança alimentar – Ceará (2004, 2009 e 2013).

Fonte: Elaboração própria com base nos dados do IBGE/ PNAD.

7 CONSIDERAÇÕES FINAIS

Com a proposta de analisar o problema da insegurança alimentar no Ceará, o presente trabalho estimou um modelo *logit* ordenado para verificar a significância estatística do efeito de características socioeconômicas sobre as chances de um domicílio estar em situação de insegurança alimentar. Foram utilizados dados do suplemento de segurança alimentar da PNAD, aplicada nos anos de 2004, 2009 e 2013.

A estimação auxilia na identificação de um perfil de domicílios que possui uma maior propensão a se encontrar em situação de insegurança alimentar. Informação que pode ser bastante útil, não apenas como subsídio para pesquisas na área, mas também na formulação de estratégias e políticas de combate à insegurança alimentar e a fome.

De forma isolada, o principal determinante da propensão de um domicílio se encontrar em situação de insegurança alimentar é a renda domiciliar *per capita*. Nesse sentido, os programas de transferência de renda, como o Programa Bolsa Família, poderiam exercer efeitos significativos sobre as chances de um domicílio pobre se encontrar em situação de insegurança alimentar.

A presente análise também revelou que alguns indicadores, que dependem diretamente da ação governamental, foram importantes e significantes para situação de segurança alimentar dos domicílios, como o acesso à água tratada e esgoto, escolaridade dos chefes

do domicílio, entre outros.

A posse de alguns bens utilizados no armazenamento e preparação de alimentos, assim com características socioeconômicas dos moradores e dos chefes de domicílio que também foram estatisticamente significativas, demonstrando que a falta de refrigerador e fogão no domicílio assim como a presença do chefe da família do sexo feminino, de cor/raça negra ou parda, com baixa escolaridade, sem ocupação ou ocupados no setor agrícola, aumentam sensivelmente as chances do domicílio apresentarem algum grau de insegurança alimentar.

De forma geral, a análise comparativa entre os anos considerados aponta para uma redução da incidência de insegurança alimentar no Ceará. No entanto, esse fato não elimina a importância e necessidade de permanente monitoramento dos determinantes da condição de insegurança alimentar; muito menos a necessidade da manutenção de políticas voltadas para a questão alimentar.

REFERÊNCIAS BIBLIOGRÁFICAS

BRASIL. Congresso Nacional. Lei 11.346, de 15 de setembro de 2006. Lei Orgânica de Segurança Alimentar e Nutricional (LOSAN). Brasília, DF, 2006

CAMELO, R. S.; TAVARES, P. A.; SAIANI, C. C. S.. Alimentação, nutrição e saúde em programas de transferência de renda: evidências para o Programa Bolsa Família. *Economia*, Brasília, v.10, n.4, p.685–713, dez./ 2009.

GOVERNO DO ESTADO DO CEARÁ. Plano Estadual de Segurança Alimentar e Nutricional 2012/2015. Fortaleza, 2015.

HOFFMANN, R. Pobreza, insegurança alimentar e desnutrição no Brasil. *Estudos Avançados*. São Paulo, vol. 9, n. 24, 1995.

HOFFMANN, R. Determinantes da Insegurança Alimentar no Brasil: Análise dos Dados da PNAD de 2004. *Revista Segurança Alimentar e Nutricional*, Campinas, v.15 n. 1, p. 49-61, 2008.

HOFFMANN, R. Determinantes da Insegurança Alimentar no Brasil em 2004 e 2009: Análise dos Dados da PNAD de 2004. *Revista Segurança Alimentar e Nutricional*, Campinas, v.20 n. 2, p. 219-235-61, 2013.

HOFFMANN, R. Brasil, 2013: mais segurança alimentar. *Revista Segurança Alimentar e Nutricional*, Campinas, v. 21 n. 2, p.422-436, 2014.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA– IBGE. **Pesquisa Nacional por Amostra de Domicílios: Suplemento de Segurança Alimentar 2004**. Rio de Janeiro, 2005.

_____. **Pesquisa Nacional por Amostra de Domicílios. Segurança Alimentar 2004/2009**. Rio de Janeiro, 2010.

_____. **Pesquisa Nacional por Amostra de Domicílios: Suplemento de Segurança Alimentar 2013**. Rio de Janeiro, 2014.

LIMA, J. R. F.; DA PAIXÃO, A. N.; DE SALES SILVA, J.. Determinantes da (in) segurança alimentar nos domicílios rurais dos municípios não autorrepresentativos da Região Nordeste. *Ensaio FEE*, v. 36, n. 4, p. 921-944, 2016.

LONG, J. S.. *Regression Models for Categorical and Limited Dependent Variables: Advanced Quantitative Techniques in the Social Sciences*. Sage Publications. 1997.

LONG, J. S.; FREESE, J. *Regression Models for Categorical Dependent Variables Using Stata*. College Station, Tex: Stata Press; 2014.

MAXWELL, D.G. Measuring food insecurity: the frequency and severity of coping strategies. International Food Policy Research Institute. FCND *discussion paper* n.8. Washington, 1995.

MAXWELL, S.; FRANKENBERGER, T.R. Household food security: concepts, indicators, measurements. UNICEF. *Technical Report*, 1992.

PANIGASSI, G., SEGALL-CORRÊA, A. M., MARIN-LEÓN, L., PÉREZ-ESCAMILLA, R., SAMPAIO, M. D. F. A., MARANHA, L. K.. Insegurança alimentar como indicador de iniquidade: análise de inquérito populacional. *Cadernos de Saúde Pública*, Rio de Janeiro, v. 24, n. 10, 2008.

ROSE, D.; GUNDERSEN, C.; OLIVEIRA, V. Socio-Economic Determinants of Food Insecurity in the United States: Evidence from the SIPP and CSFII Datasets. *Technical Bulletin 1869*. Economic Research Service, U.S. Department of Agriculture, Washington, DC (in press), 1998.

SALES, R. S. ; MIRO, V. H. . Segurança Alimentar no Ceará: uma análise com os dados da PNAD 2009. In: *X Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos*, 2012, Recife. X Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos, 2012.

SEGALL-CORRÊA, A. M.; MARIN-LEON, L. A Segurança Alimentar no Brasil: Proposição e Usos da Escala Brasileira de Medida da Insegurança Alimentar (EBIA) de 2003 a 2009. *Revista Segurança Alimentar e Nutricional*, Campinas, v.16 n. 2, p. 1-19, 2009.

ANÁLISE DA SIMULTANEIDADE ENTRE DESENVOLVIMENTO ECONÔMICO, TRANSPARÊNCIA MUNICIPAL E DESIGUALDADE DE RENDA PARA O CEARÁ

Márcio Aurélio Frota Pereira*

Ronaldo de Albuquerque e Arraes**

RESUMO

Este estudo investiga a relação simultânea entre desenvolvimento econômico, qualidade institucional e desigualdade de renda, tomando os municípios do estado do Ceará como referência. A conexão entre estes fenômenos endógenos é feita a partir de um amplo número de variáveis exógenas, as quais podem ser divididas em grupos, contendo características socioeconômicas, institucionais, financeiras e políticas. Utilizando IDHM para desenvolvimento, índice de Gini para a desigualdade e índice de transparência municipal para a qualidade institucional, a base metodológica se baseia em um modelo econométrico de equações simultâneas aplicados a dados *cross-section* dos municípios cearenses. As evidências apontam para uma relação de causalidade entre desigualdade e desenvolvimento e delas sobre a transparência. Por sua vez, a qualidade institucional tem relevância explicativa sobre o desenvolvimento, todavia, esta relevância não foi constatado sobre a desigualdade. Neste sentido, se o município apresenta elevado grau de transparência, sua chance de ser classificado entre os mais desenvolvidos decresce 14,1%. A maior parte das variáveis exógenas mostrou-se significativa, com destaque para o índice de gestão fiscal que impacta positivamente na transparência e negativamente sobre a desigualdade de renda. Municípios com residentes que apresentem maior média de idade e menor taxa de desemprego, acabam por reduzir a desigualdade. No âmbito das variáveis políticas, municípios de prefeitos reeleitos ou que possuem ideologia de centro ou de esquerda tendem a reduzir sua transparência.

Palavras-chave: Transparência Municipal. Desigualdade de Renda. Desenvolvimento Econômico. Equações Simultâneas. Ceará.

* Doutorando em economia da UFC (CAEN/UFC), (85) 9914-25309, marcio.frota10@hotmail.com.

** PhD, Professor CAEN/UFC, ronald@ufc.br.

ABSTRACT

This study investigates the simultaneous relationship between economic development, institutional quality and income inequality, taking the municipalities of the state of Ceará as a reference. The connection between these endogenous phenomena is made from a large number of exogenous variables, which can be divided into groups, containing socio-economic, institutional, financial and political characteristics. Using the GHI for development, the Gini index for inequality and the index of municipal transparency for institutional quality, the methodological basis is based on an econometric model of simultaneous equations applied to cross-section data from the municipalities of Ceará. The evidence points to a causal relationship between inequality and development and of them about transparency. In turn, institutional quality has explanatory relevance on development, however, this relevance has not been verified on inequality. In this sense, if the municipality has a high degree of transparency, its chances of being classified among the most developed grow 14,1%. Most of the exogenous variables were significant, with emphasis on the fiscal management index that positively impacts on transparency and negatively on income inequality. Municipalities with residents with higher average age and lower unemployment rate, end up reducing inequality. In the context of political variables, municipalities of re-elected mayors or those with a center or left ideology tend to reduce their transparency.

Keywords: Municipal Transparency. Income Inequality. Economic Development. Simultaneous Equations. Ceará.

1. INTRODUÇÃO

Um ponto marcante da economia do Brasil é a falta de transparência municipal e estadual. Na tentativa de reduzir a ocorrência deste problema, foi criado pela CGU (Controladoria Geral da União) o Índice de Transparência Municipal, que classifica os municípios brasileiros de acordo com o grau de disponibilização de informações financeiras. Com isso, segundo o Ranking Nacional da Transparência¹ o nível dos estados e municípios, no período de seis meses (tendo início em outubro de 2015 e finalizado em maio de 2016), cresceu 33 pontos percentuais, a partir da atuação coordenada do ministério público. O índice de transparência aumentou de 3,92, em 2015, para 5,21, em 2016. Entretanto, alguns estados apresentam nível muito aquém dos demais, como o estado de Manaus, que apresentou o menor grau de transparência com 2,84 para o âmbito municipal, por outro lado, Santa Ca-

1. Publicado pelo MPF (Ministério Público Federal).

tarina obteve grau de 8,24, cerca de três vezes maior.

Outra característica muito relevante da economia brasileira é a forte desigualdade de renda. Apesar da queda desse indicador entre os anos de 2001 e 2009, a saber, o índice de Gini da distribuição de renda² foi de 0,594 para 0,539 (redução de 9,2%), o Brasil ainda está entre os países que apresentam maior intensidade da desigualdade de renda. Esta redução da disparidade foi observada em todas as regiões do país, todavia, a dinâmica de redução de desigualdade no período entre 2001 e 2010 foi diferenciada entre as regiões brasileiras. De acordo com os censos de 2000 e 2010³, as que apresentam maior redução em seus níveis de desigualdade foram, respectivamente, Sul (14,9%), Centro-Oeste (11,9%) e Sudeste (10,1%). As regiões Nordeste (9,3%) e Norte (9,2%) apresentaram reduções relativamente menores.

Um importante efeito da desigualdade é aquele exercido sobre o crescimento econômico. No entanto, a literatura não chegou a um consenso sobre a direção exata dessa relação, determinando a inexistência de conclusão a seu respeito. O início desta discussão nos remete ao trabalho de Kuznets (1995), no qual, encontrou uma relação de “U” invertido entre as duas medidas. Barro (2000) encontrou resultados parecidos, onde, corroborou com a hipótese e indicou que a simultaneidade entre as variáveis pode diferenciar de acordo com a dinâmica e os estágios de desenvolvimento.

Todavia, por diversas vezes, consideram-se crescimento e desenvolvimento como fenômenos equivalentes ou sinônimos. Crescimento econômico é o aumento do Produto Interno Bruto (PIB), ou seja, uma elevação da produção da região estudada. O PIB é calculado através da soma de todos os produtos e serviços finais de uma região para um determinado período. Já o conceito de desenvolvimento econômico está relacionado a melhoria do bem estar da população, sendo um processo mais amplo, no qual, engloba educação, saúde, igualdade de gênero. Atualmente o Índice de Desenvolvimento Humano - IDH é o critério mais utilizado para comparar o desenvolvimento de diferentes economias. Entretanto, Chang (2009) faz uma crítica à ampliação do conceito de desenvolvimento, que outrora se referia exclusivamente ao progresso das economias.

Na literatura está bem estabelecido que não podemos obter uma visão abrangente do desenvolvimento humano, considerando apenas uma única dimensão, como renda (Seth e Villar, 2015). Isto é assim devido ao fato de que nem todas as variáveis que afetam esta

2. O índice de Gini varia de 0 a 1 e mede o grau de desigualdade na distribuição de renda. Quanto mais próximo de 1 maior será a desigualdade, todavia, quanto mais próximo de 0 melhor a equidade.

3. Publicado periodicamente a cada dez anos pelo IBGE (Instituto brasileiro de Geografia e Estatística) – acesso <<http://www.ibge.com.br>>.

variável evoluem de forma similar. Existindo defasagens entre desempenho de variáveis monetárias e não monetárias. Portanto, a mensuração do desempenho tem que ser estabelecida através de uma variável multidimensional, no qual, consiga captar o maior grau possível de variação da qualidade de vida (Seth & Villar, 2015). Com isso, a variável com maior cumprimento desse quesito será IDHM.

Desse modo, cabe a seguinte contestação, será que desenvolvimento, desigualdade e seus determinantes estão conectados? Pode-se incluir outra medida que esteja perpetuando sobre os canais de transmissão das mesmas? Nesse sentido, a literatura traz um possível candidato para preencher esta lacuna, no qual, seria a qualidade institucional. Acemoglu, Johnson e Robinson (2004) em conformidade com os estudos de North (1991) e Hodgson (1998), propõem que o nível de desenvolvimento econômico é determinado pelo nível das instituições econômicas. Kotschy e Sunde (2017) afirmam que na determinação do comportamento econômico, a desigualdade é um dos principais influentes da qualidade institucional.

Um forte candidato a qualidade institucional é a transparência, no qual, intuitivamente reduz a desigualdade e aumenta o desenvolvimento, isto é, com a maior disponibilização de informações financeiras seria uma forma de combater a corrupção em vários níveis governamentais, repassando os recursos sociais de forma mais limpa e igualitária.

A divulgação de informações fiscais é um tema amplo e muito pesquisado internacionalmente. Diversos estudos divulgados recentemente voltaram-se para a análise da transparência e seus determinantes em nível local ou regional, como os dados espanhóis que foram analisados por Esteller-Moré e Polo Otero (2012) e Sol (2013), além disso, dados da Nova Zelândia foram levados em conta por Laswad, Fisher e Oyelere (2005). Por outro lado, as análises estudadas ao nível dos países, como: De Renzio e Angemi (2012); Ríos et al. (2014); Andreula e Chong (2015) são menos divulgadas. Ao nível nacional são escassos estudos voltados ao tema. Pode-se citar Zuccolotto e Teixeira (2014) que analisaram a transparência dos estados brasileiros.

O presente estudo tem como objetivo averiguar como a relação de desenvolvimento, qualidade institucional e desigualdade, e seus fatores determinantes, estão conectados e podem explicar diversos problemas no território cearense, levando em consideração algumas características específicas, a partir do conceito de IDHM (Índice de Desenvolvimento Humano Municipal) para o desenvolvimento, índice de Gini para a desigualdade e índice de transparência municipal para a qualidade institucional. Desse modo, uma análise *cross-section* é feita a partir de um modelo de equações simultâneas relacionando desenvolvi-

mento, transparência e desigualdade.

Além desta introdução, este estudo contém mais quatro seções. A seção 2 traz o referencial teórico, no qual, é dividido em três partes, a saber, uma voltada para desenvolvimento e a desigualdade, outra para a qualidade institucional e sua causalidade com desigualdade e desenvolvimento e a última para a transparência. Na seção 3, é exibida a estratégia empírica empregada. A mesma traz as bases de dados utilizadas e sua descrição, além disso, apresenta o modelo probit, a técnica *Limited-Information Maximum Likelihood* (LIML) e especifica como serão modeladas as variáveis de interesse. Na seção 4, são apresentados os resultados obtidos através da estimação das equações simultâneas. Por fim, a seção 5 traz as considerações finais.

2. REFERENCIAL TEÓRICO

2.1. Desenvolvimento e Desigualdade

A relação entre desigualdade e desenvolvimento está longe de ser um assunto já consolidado na teoria econômica. No qual, não possui um sentido único, uma vez que exista uma relação simultânea onde os canais que perpetuam o desenvolvimento atuam no intuito de promover a melhora na qualidade de vida. Algumas teorias propõem que o desenvolvimento afeta a desigualdade, enquanto outras argumentam que a desigualdade tem impacto sobre o desenvolvimento. Existem, também, teorias em que a correlação entre as duas variáveis é positiva, enquanto em outras esta correlação é negativa. Nesse intuito, segundo Mizrahi (2010) a redução da desigualdade é um fator imprescindível para que se assegure um desenvolvimento sustentável.

Estudos, a partir da metade do século XX, vêm apresentando preocupação com as origens e as implicações da desigualdade. Nesse sentido, surgiu a hipótese de Kuznets na tentativa de explicar a relação entre crescimento econômico e desigualdade, no qual, pode ser rastreada até Simon Kuznets em 1955, com artigo seminal “Crescimento Econômico e Desigualdade de Renda”. Com isso, a hipótese supõe que as variáveis apresentam uma relação de “U” invertido, onde, quando o crescimento ocorre, a desigualdade inicialmente sobe e ao atingir certo grau de crescimento tende a cair.

Em relação a causalidade entre crescimento e desigualdade, o senso comum aponta para uma correlação positiva, baseando-se em duas linhas de argumentação, a saber, poupança e incentivo. O argumento voltado à poupança origina-se da ideia de que ricos tendem a economizar mais que os pobres, levando à conclusão de que a concentração de renda im-

plica maiores níveis de poupança. Nesse sentido, Banerjee e Duflo (2011) resumem alguns estudos empíricos e teóricos sobre o comportamento dos pobres, no qual, indicam que os menos desprovidos, de fato, tendem a poupar menos que os ricos.

Em relação ao incentivo é mais fácil visualizar, pois o indivíduo racional fará um esforço maior para alcançar um objetivo somente se ele receber um incentivo proporcional aos seus esforços. Desse modo, maior será a produtividade e o nível de produção quando as recompensas são pagas desigualmente. Segundo Okun (2015) para altos níveis de equidade econômica, os agentes econômicos não serão incentivados a aumentar seus resultados, pois tais esforços não serão recompensados com aumento da renda.

Entretanto, as imperfeições do mercado de crédito impactam negativamente sobre o crescimento, pois excluirá os pobres, já que problemas de assimetria informacional e limitações impostas por instituições fariam o mercado destinar recursos aos investimentos mais seguros ao invés dos mais rentáveis (Kim, 2016).

No entanto, por muitas vezes, consideram-se crescimento e desenvolvimento como fenômenos equivalentes ou sinônimos. Entretanto, o desenvolvimento é um processo mais amplo, no qual, engloba educação, saúde, igualdade de gênero, entre outros, e por si só o crescimento econômico não é capaz de explicá-lo. Todavia, Chang (2009) faz uma crítica à ampliação do conceito de desenvolvimento, que outrora se referia exclusivamente ao progresso das economias.

O PIB (produto interno bruto) *per capita* era utilizado como indicador de qualidade de vida até o surgimento do IDH (Índice de Desenvolvimento Humano). Nesse sentido, Scarpin e Slomski (2007) defendem que PIB *per capita* não é uma medida suficiente para avaliar a qualidade de vida das pessoas. Com isso, o Índice de Desenvolvimento Humano foi criado tendo como um de seus objetivos, oferecer um contraponto ao PIB *per capita*, o qual considera apenas uma única dimensão econômica do desenvolvimento. Desse modo, O IDH populariza o conceito de desenvolvimento centrado nos indivíduos, e não a visão de que desenvolvimento se limita a crescimento econômico.

Segundo o PNUD⁴ (Programa das Nações Unidas de Desenvolvimento) o Índice de Desenvolvimento Humano (IDH) é uma medida resumida em três dimensões básicas do desenvolvimento humano: renda, educação e saúde. No qual, pretende ser uma medida geral e sintética que, apesar de ampliar a perspectiva sobre o desenvolvimento humano, não abrange nem esgota todos os aspectos de desenvolvimento.

4. Ver <<http://www.br.undp.org>>.

Desse modo, foi desenvolvido pelo economista paquistanês Mahbud Ul Haq e é utilizado pelo PNUD desde o ano de 1993. O índice varia de 0 a 1, sendo considerados de baixo desenvolvimento os países que atingem menos de 0,499 pontos, de médio desenvolvimento os que possuem notas de 0,500 até 0,799, e de alto desenvolvimento os países que atingem pontuação superior a 0,800.

Segundo Seth e Villar (2015) está bem estabelecido que não podemos obter uma visão abrangente do desenvolvimento humano, considerando apenas uma única dimensão, como renda. Isto é assim devido ao fato de que nem todas as variáveis que afetam esta variável evoluem de forma similar. Existindo defasagens entre desempenho de variáveis monetárias e não monetárias.

Portanto, a mensuração do desempenho tem que ser estabelecida através de uma variável multidimensional, no qual, consiga captar o maior grau possível de variação da qualidade de vida (Seth & Villar, 2015). Desse modo, a variável com maior cumprimento desse quesito será IDHM.

2.2 O papel das Instituições na Desigualdade e no Desenvolvimento

Há uma vasta literatura que busca dimensionar a relação de causalidade entre qualidade institucional e desempenho econômica (HALL, JONES, 1999; ACEMOGLU, JOHNSON e ROBINSON, 2001, 2002, 2004; ENGERMAN e SOKOLOFF, 2002; EASTERLY e LEVINE, 2002; CHONG e GRADSTEIN, 2004; BLUHM e SZIRMAI, 2011; KOTSCHY e SUNDE, 2017). Entender como desenvolvimento ou desigualdade se relaciona com as instituições econômicas é fundamental como suporte nas decisões de política econômica. Identificar as causas das desigualdades regionais é condição necessária para reduzi-las. Daí a importância de estudos empíricos para tentar entender tais relações e a direção de causalidade.

Acemoglu, Johnson e Robinson (2004) em conformidade com os estudos de North (1991) e Hodgson (1998), propõem que o nível de desenvolvimento econômico é determinado pelo nível das instituições econômicas, pois influenciam investimentos em capital físico, em capital humano, em tecnologia, além disso, moldam os incentivos dados aos agentes na sociedade.

Os autores relatam que, apesar de que aspectos culturais e geográficos tenham um papel importante na determinação do desenvolvimento econômico, o grau da qualidade institucional é a grande fonte das disparidades entre países no nível de desenvolvimento

e renda. Essas instituições não somente determinam o desempenho econômico e o nível de desenvolvimento da economia, mas também a distribuição de recursos no futuro. Ou seja, elas induzem não somente sobre o grau de atividade econômico, mas também a forma como será distribuída entre diferentes indivíduos e grupos na sociedade.

Kotschy e Sunde (2017) na tentativa de analisar a interação significativa entre instituições políticas e desigualdade na determinação da qualidade de Instituições econômicas. Utilizaram como medida de qualidade institucional um índice de Liberdade Econômica, no qual, é um de índice composto que se baseia nos componentes principais dos índices de Liberdade Econômica e Liberdades Civas. Além disso, utilizaram o índice de Gini para medir a igualdade econômica.

Segundo os autores, na determinação do comportamento econômico, a desigualdade é um dos principais influentes da qualidade institucional. Por outro lado, Chong e Gradsstein (2004) indicam que, instituições de baixa qualidade causam desigualdade, bem como a desigualdade aumenta as chances da existência de instituições de baixa qualidade. Por fim, indicam que a causalidade da desigualdade sobre as instituições parece dominar a causalidade reversa.

Continuando nesta linha de pesquisa, Chong e Gradstein (2007a) mostram que uma maior desigualdade pode prejudicar a qualidade institucional. Além disso, Chong e Gradstein (2007b) fornecem provas de uma dupla causalidade entre a qualidade das instituições e a distribuição de renda.

A compreensão da relação entre desigualdade, desenvolvimento e qualidade institucional é de suma relevância, embora esta causalidade e os canais em que se propagam começassem a ser pesquisada e desenvolvida apenas recentemente. Este link é importante na avaliação do desempenho econômico, relatando como a economia está organizada e como produção e distribuição serão executadas. Isso, por sua vez, pode aumentar o entendimento do por que os países podem ser ricos ou pobres. Desse modo, segundo Savoia et al. (2010), o entendimento atual sugere que sociedades com baixo desempenho econômico desenvolvem instituições exploradoras e ineficazes.

Além disso, na literatura empírica há uma gama de estudos que buscam dimensionar o papel das instituições sobre o desempenho econômico através de análises de países (ACEMOGLU; JOHNSON; ROBINSON, 2001, 2002, 2004; ENGERMAN; SOKOLOFF, 2002; HALL; JONES, 1999; EASTERLY; LEVINE, 2002). Segundo Naritomi (2007) restringir a análise para um único país permite investigar dimensões específicas de institui-

ções locais. O caso brasileiro, em particular, permite que tais instituições capturem efeitos do ambiente institucional *de facto*, uma vez que as instituições *de jure* são centralizadas e determinadas a nível nacional.

Entretanto, essa análise entre qualidade institucional, desigualdade e desenvolvimento exige cautela, pois é possível que haja uma causalidade reversa entre instituições e desempenho econômico, ou seja, regiões mais desenvolvidas economicamente são mais aptas a sustentar arcabouços institucionais melhores. Sendo esse fenômeno conhecido como endogeneidade das variáveis explicativas.

A endogeneidade torna necessária a busca por fontes exógenas de variação para a realização de análises empíricas. Desse modo, grande parte da literatura internacional para tratar a endogeneidade institucional recorreu à história e a geografia. Por exemplo, Acemoglu, Johnson e Robinson (2001 e 2002) e Engerman e Sokoloff (2002) consideram a característica exógena dos países terem sido ex-colônias europeias como variável instrumental para instituições. Além disso, seguindo esta linha, Naritomi (2007) utilizou variáveis históricas, a saber, o ciclo do ouro e o ciclo do açúcar, como determinantes da qualidade de quatro dimensões institucionais específicas: desigualdade de distribuição de terras, concentração política, capacidade gerencial e acesso à justiça.

O argumento que dá suporte a utilização das instituições do passado como instrumentos, no qual, são claramente exógenas em relação à renda corrente dos países, como variáveis representantes das instituições atuais, é a ideia de inércia das instituições, segundo a qual as instituições que se desenvolvem, sejam elas negativas ou positivas ao processo de desenvolvimento econômico, tendem a persistir ao longo dos anos.

2.3 Índice de Transparência

2.3.1. *Estudos sobre Transparência Municipal*

Pesquisadores de vários campos do conhecimento têm voltado seu interesse para a relevância do tema transparência e do seu potencial poder explicativo e que, por causa da diversidade de interesses envolvendo o fluxo de informações, o conceito transparência não foi monopolizado por nenhuma área de estudo, podendo ser encontrado em pesquisas sobre comércio, corrupção, gestão pública, conflitos internacionais, organizações internacionais, ambiente político, política monetária, etc.

A literatura existente em relação aos fatores que levam à divulgação de informações fiscais pode ser dividida em duas partes. Por um lado, os documentos centrados na transpa-

rência fiscal do país e seus determinantes, com base em dados regionais ou locais (observando estudos recentemente divulgados, os dados espanhóis foram analisados por: Gandía e Archidona (2008); García e García-García (2010); Guillamón, Bastida e Benito (2011); Esteller-Moré e Polo Otero (2012); Sol (2013); Dados da Nova Zelândia foram levados em conta por Laswad, Fisher e Oyelere (2005); Os estados brasileiros foram analisados por Zuccolotto e Teixeira (2014). Por outro lado, as análises ao nível dos países (Ríos et al. (2014); Andreula e Chong (2015); Wehner e De Renzio (2013); De Renzio e Angemi (2012)) são definitivamente menos difundidas.

Paralelamente ao lado de fatores econômicos, fatores institucionais, riqueza econômica, situação da autoridade pública, tipo do sistema jurídico, considera-se também fatores políticos como principais determinantes da transparência fiscal. A literatura existente sugere que três características principais do sistema político tem impacto mais significativo na transferência fiscal: orientação ideológica do governo, controle governamental sobre a legislatura e a competição política.

A evidência sobre o efeito da orientação política sobre a transparência fiscal é mista. Enquanto algumas contribuições não encontram qualquer efeito da ideologia governamental sobre a divulgação fiscal (Alt et al., 2006), outros apoiam o oposto, Guillamón et. al. (2011) relatam que se o partido apresentar ideologia de direita a transparência será afetada negativamente. Por outro lado, segundo Ferejohn (1999) governos de esquerda reforçam a divulgação de informações fiscais para defender um setor público mais transparente.

Quando os líderes tem maior influência sobre os parlamentares acabam por ocasionar o enfraquecimento do nível da transparência, enquanto os líderes não possuem o apoio de parlamentares, apresenta-se maior desencadeamento de informações (Michener, 2015). Na perspectiva dos governos, os benefícios decorrentes do sigilo podem superar o custo da informação (Roberts (2006); Hazell & Worthy (2010); Michener, (2015)).

A competição política também pode ser um forte impulsionador da transparência fiscal. Na verdade, as partes que compõem o governo podem pedir mais transparência para ter controle sobre as atividades governamentais. Não surpreendentemente, as análises empíricas em relação aos efeitos da competição política sobre a divulgação fiscal mostrarem-se significantes, embora com algumas exceções (Zuccolotto & Teixeira, 2014). Desse modo, Alt et al. (2006), bem como Ríos et al. (2014), observaram que a competição política tende a aumentar a transparência fiscal. Wehner e de Renzio (2013) acham que a fragmentação partidária influencia positivamente a divulgação de informações orçamentais quando ocorrem eleições livres e justas. Andreula e Chong (2015) mostram que a competição política

exerce um papel positivo na transparência fiscal.

Em busca dos determinantes da transparência fiscal em 36 países de 2003 a 2013 (Cicatiello & Gaeta, 2016), corroboraram com Alt et al. (2006), relatando que maior controle do governo sobre o parlamento levará a uma maior transparência fiscal. Além disso, contrariando achados anteriores (Andreula & Chong, 2015). Segundo Leachman et al. (2007) legislaturas fracionários têm menos interesse na transparência fiscal, o que sugere que a disponibilização de informações financeiras é afetada pelo “problema de fragmentação”.

Outro ponto a ser frisado é a dimensão populacional, pois é frequentemente considerada como um determinante das práticas de transparência (Christianes, 1999). Geralmente, os municípios com maior dimensão apresentam maior grau de divulgação de informações fiscais, resultando em uma procura mais acentuada por partes dos cidadãos por informações (Girou e McLelland 2003). Desse modo, as administrações públicas com maior dimensão são obrigadas a divulgar informações nos websites, sendo este meio o que preenche melhor às necessidades dos cidadãos (Laswad et al., 2005), o que fomenta a transparência segundo Caba e López (2009) é que, quanto maior for o município, maior a transparência financeira.

A hipótese considerada tem por fundamento a argumentação de Guillamón et al. (2011) que defendem que práticas de disponibilização de informações fiscais requerem materiais e recursos humanos apropriados os quais tenderão a estarem presentes, sobretudo, em municípios com maior população. Referem-se ainda que municípios mais populosos gerem orçamentos mais elevados, existindo maior pressão sobre a sua gestão por parte da sociedade civil. Caamaño et al. (2013) apresentam o mesmo entendimento, defendendo que municípios com maior número de residentes tendem a apresentar maior número de trabalhadores municipais o que é favorável ao cumprimento dos requisitos de transparência.

Estudos apontam que a idade dos residentes tem impacto sobre a transparência, no entanto, não há conformidade do impacto, podendo ser positivo ou negativo. Enquanto estudos como o de Reddick (2005) apontam para um efeito negativo da idade relatando que um maior percentual de idosos implica em menor predisposição na utilização de tecnologias de informação, no qual, constituem um veículo para a obtenção de informação da administração pública, outros apresentam um efeito inverso defendendo que residentes mais idosos tenderão a ser mais participativos e exercem uma maior pressão por transparência (Esteller-Moré & Otero, 2012).

Além disso, grandes quantidades de trabalhos têm buscado identificar o impacto do endividamento na transparência, argumentando o efeito favorável na transparência de níveis de endividamento mais reduzidos. Os resultados empíricos obtidos por Alt et al.

(2006) evidenciam o impacto desfavorável do endividamento, entretanto, em diversos outros estudos o endividamento parece não apresentar qualquer relevância estatística (Albalade del Sol, 2013; Guillamón et al., 2011).

Segundo Zuccolotto e Teixeira (2014) países com instituições supremas de auditoria com maior independência e poder, constitucionalmente definido, apresentam maior transparência, uma vez que essas instituições, por deterem maior poder e independência do executivo, irão agir a partir de interesses distintos, reduzindo, dessa forma, os conflitos de interesses entre os governantes e os governados.

2.3.2 Lei de Acesso à Informação e o Índice de Transparência

A preocupação com a transparência do setor público tem início em meados de 1950 na sequência da implantação da Nova Gestão Pública (NGP) no Reino Unido. A partir de 1990 esta preocupação vem intensificando-se, relacionada à intervenção de programas internacionais na elaboração de difusões e orientações de boas práticas e ao surgimento de estudos que destacam os seus efeitos positivos.

O tema tem sido amplamente aprofundado por entidades internacionais, principalmente, com enfoque na disponibilização de informações orçamentais, no intuito de promover o controle, a sustentabilidade e a responsabilização da despesa pública. O Fundo Monetário Internacional (FMI) elaborou o Código de Boas Práticas para a Transparência Orçamental (2007)⁵, segundo o qual, a transparência se enquadra em quatro pilares: definição clara das funções (política vs. gestão), responsabilidades (dos diferentes níveis de governo), acesso à informação e existência de garantias de integridade. A Organização de Cooperação e de Desenvolvimento Econômico (OCDE, 2002)⁶ elaborou o Manual das Melhores Práticas para a Transparência Orçamental do qual fornecem coordenadas para os países que visam implantar ações destinadas à melhoria dos seus níveis de transparência.

O Brasil com a aprovação da Lei 12.527 em 18 de novembro de 2011⁷ dá um importante passo para a sua transparência pública, ampliando a democracia e fortalecendo os instrumentos de controle da gestão pública. Desse modo, foi garantido o direito de acesso à informação. O país já é destaque no âmbito de divulgação espontânea de informações

5. A primeira versão surge em 1998. O Código foi alvo de revisões em 1999, 2001 e 2007. Encontra-se em fase de aprovação uma nova revisão do código iniciada em 2013.

6. O manual encontra-se estruturado em três partes principais, a saber, documentos e processos a disponibilizar pelos governos, tipo de informação a figurar nos documentos e processos que asseguram a qualidade e integridade da informação constante dos documentos.

7. Entrou em vigor 180 (cento e oitenta) dias após essa data, ou seja, em 16 de maio de 2012.

governamentais: o Portal da Transparência do Governo Federal, criado e gerenciado pela CGU já foi diversas vezes premiado, nacional e internacionalmente, sendo considerado um dos mais completos sites de transparência no mundo. Entretanto, faltava uma lei que regulasse o acesso amplo a qualquer documento ou informação específica buscada pelo cidadão.

Desse modo, em comemoração aos três anos de vigência da Lei de Acesso à Informação, a CGU lança o índice de transparência pública, o qual busca avaliar o grau de cumprimento às normas da nova lei por parte dos estados e municípios brasileiros.

Por outro lado, o Ministério Público Federal dá início a elaboração de um ranking sobre a transparência nacional, abordando tanto o âmbito estadual como municipal, com o intuito de combater a corrupção, inicialmente foram realizadas duas avaliações, no qual, a média nacional aumentou cerca de 30 pontos percentuais da primeira (outubro de 2015) para a segunda (entre abril e maio de 2016) avaliação.

O questionário aplicado pelas unidades do Ministério Público Federal no Brasil inteiro foi elaborado tendo como base a ação nº 4 da Estratégia Nacional de Combate à Corrupção e Lavagem de Dinheiro (ENCCLA) do ano de 2015, por representantes do Ministério Público Federal (MPF), Conselho Nacional do Ministério Público (CNMP), Controladoria Geral da União (CGU), Tribunal de Contas da União (TCU), Secretaria do Tesouro Nacional (STN), Associação dos Membros dos Tribunais de Contas (ATRICON), Banco Central, entre outras instituições de controle e fiscalização.

O questionário formulado (ANEXO A) a várias mãos pelas instituições de controle é essencialmente baseado nas exigências legais, à exceção dos dois itens finais que são considerados “boas práticas de transparência”.

Todas as questões respondidas pelos avaliadores entram na nota, mas com peso diferenciado de acordo com a sua importância. O peso de cada nota foi dado por meio de votação entre as diversas instituições de controle que participaram da sua elaboração no bojo da ENCCLA. Desse modo, o município pode obter nota de 0 a 10, de acordo com o sua avaliação, quanto mais próximo de dez maior a transparência do município. As questões e seus pesos podem ser encontrados no ANEXO A.

3. ESTRATÉGIA EMPÍRICA

Nesta seção, são apresentadas as bases de dados empregadas neste trabalho, bem como as estatísticas descritivas das variáveis utilizadas e por fim, o modelo econométrico.

3.1. Base de Dados e Estatísticas Descritivas

No presente trabalho, faz-se uso de múltiplas bases de dados, no qual, referem-se aos municípios cearenses. Ao final, tendo em vista as limitações impostas pela disponibilidade de informações, quais sejam, valor da transferência municipal, índice de governança e índice de gestão fiscal, a amostra contou com um total de 166 municípios do Ceará.

As variáveis dependentes do modelo simultâneo aqui estimado são desenvolvimento, desigualdade e índice de transparência. No qual, foi realizado uma transformação dos dados, passando de dados contínuos para binários, a saber, o valor que estiver acima da média da variável passou a ser um e zero caso contrário. A Tabela 1 traz a média e a descrição das variáveis dependentes antes da transformação. A Tabela 2 apresenta as estatísticas descritivas das variáveis explicadas após a transformação e das variáveis explicativas.

Tabela 1 - Descrição das variáveis dependentes

Variável	Média	Descrição
Desenvolvimento	0,616	Índice de Desenvolvimento Humano Municipal (IDHM) referente ao ano de 2010, no qual, é a medida de três dimensões básicas do desenvolvimento humano: educação, longevidade e saúde. Obtido junto ao Atlas de Desenvolvimento Humano do Brasil.
Desigualdade	0,536	Índice de GINI de 2010, que é uma medida da desigualdade de renda. Foi coletado através do departamento de informática do Sistema Único de Saúde do Brasil (DATASUS).
Índice de Transparência Municipal	4,413	Índice de Transparência Municipal de 2016, no qual, mede o grau de transparência do site municipal. Obtido junto ao Observatório de Informações Municipais. O questionário utilizado pode ser visualizado no anexo A.

Fonte: Elaboração do autor.

Tabela 2 – Descrição Estatística

	Sigla	Obs.	Média	DP	Mínimo	Máximo
Variáveis de interesse						
<i>Desenvolvimento</i>	<i>desen</i>	166	0,439	0,497	0	1
<i>Índice de transparência municipal</i>	<i>itm</i>	166	0,536	0,500	0	1
<i>Desigualdade</i>	<i>desig</i>	166	0,445	0,498	0	1
Variáveis explicativas						
<i>Média de anos municipal</i>	<i>idade</i>	166	30,526	1,598	27,540	36,199
<i>Log do número de residentes</i>	<i>lpop</i>	166	10,010	0,834	8,490	14,576
<i>Capital social</i>	<i>caps</i>	166	0,178	0,460	0,087	0,376
<i>Log do valor da transferência</i>	<i>lrdpc</i>	166	5,750	0,236	5,156	6,884
<i>Taxa do número de empregados</i>	<i>txe</i>	166	0,322	0,070	0,153	0,501
<i>Índice de qualidade institucional</i>	<i>iqim</i>	166	2,774	0,462	1,3	4
<i>Índice de gestão fiscal</i>	<i>igf</i>	166	0,383	0,146	0	0,773
<i>Índice de governança</i>	<i>ig</i>	166	3,987	1,832	1	8
<i>Persistência política</i>	<i>pp</i>	166	0,530	0,500	0	1
<i>Partido de centro</i>	<i>pc</i>	166	0,469	0,500	0	1
<i>Partido de esquerda</i>	<i>pe</i>	166	0,234	0,425	0	1

Fonte: Elaboração do autor.

A transformação das variáveis dependentes se dar decorrente do seguinte fato, quando se examina quantitativamente uma variável dependente contínua, não se extrai o real significado do acréscimo/decrécimo na variação de uma variável explicativa qualitativamente relevante. Desse modo, afirmar que houve determinado impacto quantitativo no Gini, no IDHM ou no ITM não exprime se houve mudança em seu status. Ao medi-las de forma qualitativa tem-se um efeito sobre as chances de ocorrência de mudança de tal status. Portanto, nesse estudo utiliza-se as variáveis explicadas de forma qualitativa para capturar a chance de mudança em seu status.

Com relação a descrição estatística, 70 municípios apresentaram-se acima da média da variável desenvolvimento, isto é, aproximadamente, 43,9% da amostra. Todavia, 74 apresentaram-se abaixo da média da variável desigualdade e 89 acima da média da variável de transparência, significando, respectivamente, 44,5% e 53,6% da amostra.

Em busca de medir o impacto da variável idade sobre as variáveis de interesse, foi utilizada a idade média municipal, no qual, foi obtida através do Censo Demográfico de

2010 (Instituto Brasileiro de Geografia e Estatística). Na média, a idade municipal é de 30,5, valor que vai de 27,53 no município que apresenta menor média (Marco) a 36,2 naquele que apresenta maior (São João do Jaguaribe).

Em relação à população foi empregado o log do total do número de residentes, procurando mensurar se o tamanho da população influencia na administração dos municípios, no qual, foi extraída do tesouro nacional para o ano de 2010. Por sua vez, a medida utilizada para o capital social foi à proporção dos votos brancos ou nulos em relação aos votos úteis. Os dados foram obtidos junto ao IPEA (Instituto de Pesquisa Econômica) para o ano de 2008.

Com o intuito de auferir o nível de ocupação do mercado de trabalho, empregou-se a taxa de pessoas empregadas, a variável foi obtida através do Censo Demográfico de 2010. Na média, a taxa de empregados é de 32,2%, valor que vai de, aproximadamente, 15,3% no município que apresenta menor nível de emprego (Choró) a 50,1% naquele que apresenta maior (Fortaleza).

Para mensurar as transferências municipais foi utilizado a variável log da transferência do bolsa família em relação à transferência total. No qual, foi obtida junto ao Tesouro Nacional sendo utilizado o ano de 2010.

As medidas utilizadas para a qualidade institucional dos municípios foram a Qualidade Institucional Municipal, Índice de Gestão Fiscal e Índice de Governança. A primeira é calculada pelo Ministério do Planejamento Orçamento e Gestão (MPOG). O mesmo varia de um a seis (sendo que valores mais próximos de seis representam maior qualidade das instituições) e é composto a partir de três conjuntos de indicadores, no qual, recebem pesos iguais, a saber: (i) grau de participação, (ii) capacidade financeira, (iii) capacidade gerencial.

O Índice de Gestão Fiscal é composto por cinco indicadores: receita própria, gastos com pessoas, investimentos, liquidez e custo de vida. Nesse sentido, varia de 0 a 1, sendo que, quanto mais próximo de 1, melhor a gestão fiscal do município no ano de observação. Foi extraído do Observatório de informações municipais para o ano de 2010.

O Índice de Governança mensura a qualidade das práticas governamentais dos municipais. Sendo a soma de dois indicadores, e o valor desta variável pode ir de zero a oito, quanto mais próximo de oito, melhores são as práticas governamentais que apresenta o município. Os dois indicadores são: número de instrumentos administrativos (a soma de variáveis binárias que indicam a existência de distritos administrativos das cidades, do

Plano Diretor, da Lei de Parcelamento do Solo, da Lei de Zoneamento, do Código de Obras e do Código de Posturas) e o número de instrumentos de planejamento (que é calculada também através da soma das seguintes variáveis binárias: existência de Plano de Governo e Planejamento Municipal).

A forma de mensurar a influência política sobre o índice de transparência foi através de três variáveis, quais sejam, persistência política, partido de centro e partido de esquerda. A primeira é uma variável binária, a saber, apresenta valor 1 se o prefeito for reeleito e 0 caso contrário, no qual, foi obtido junto ao IPEA e utilizado o ano 2006. Os prefeitos foram reeleitos em 53% dos municípios, como pode ser visto na Tabela 2.

Em relação à ideologia do partido, no qual, pode ser de centro ou de esquerda, foram mensuradas duas variáveis binárias. Desse modo, apresentam valor 1 se pertencer ao tipo do partido e 0 caso contrário. Na média, aproximadamente, 47% dos partidos são de centro e 23,4% são de esquerda. Coletadas através do IPEA para o ano de 2006.

3.2. MODELO ECONOMÉTRICO

O modelo teórico proposto diz respeito ao elenco de possíveis explicações para o desenvolvimento, transparência e desigualdade. Tal modelo, de acordo com as variáveis selecionadas, relaciona-se com quatro conjuntos de variáveis: i) as que estariam ligadas às características socioeconômicas da população; ii) as que estariam ligadas às qualidades institucionais municipais; iii) as com intuito econômico; iv) as que estariam ligadas a classificação e permanência política.

O modelo teórico e estrutural é expresso na forma de um sistema de equações, onde, as siglas e sua descrição encontram-se na Tabela 2, no qual, é sintetizado nas funções a seguir:

$$desen = itm + desig + idade + lpop + lt + iqim \quad (1)$$

$$itim = desen + desig + caps + iqim + igf + ig + pp + pc + pe \quad (2)$$

$$desig = desen + itm + idade + lpop + txe \quad (3)$$

Para que o modelo possa obter estimativas numéricas únicas dos coeficientes estruturais, a partir das equações reduzidas, é necessário que seja identificado, para tanto, deve satisfazer os seguintes pressupostos:

- i) Condição de ordem para identificação: Em um modelo com M equações simultâ-

neas, para que uma equação seja identificada, o número de variáveis predeterminadas excluídas da equação não deve ser menor do que o número de variáveis endógenas incluídas da equação menos 1, isto é,

$$K - k \geq m - 1 \quad (4)$$

Onde:

K = número de variáveis predeterminadas no modelo, incluindo o intercepto

k = número de variáveis predeterminadas em uma dada equação

M = número de variáveis endógenas no modelo

m = número de variáveis endógenas em uma dada equação

Se $K - k = m - 1$, segundo Gujarati (2006), a equação é, assim, identificada, mas, se $K - k > m - 1$, ela é sobreidentificada.

ii) Condição de posto para identificação: Em um modelo com M equações em M variáveis endógenas, uma equação é identificada se, apenas se, ao menos um determinante diferente de zero de ordem $(M-1)(M-1)$ puder ser construído por meio dos coeficientes das variáveis (tanto endógenas quanto predeterminadas) excluídas da equação especificada, mas incluídas em outras equações do modelo.

De acordo com Gujarati (2006), embora a condição de ordem seja fácil de aplicar, ela fornece apenas uma condição necessária para identificação. Por outro lado, a condição de posto é uma condição suficiente e se for satisfeita, a condição de ordem necessariamente também o é, entretanto, o inverso não é verdadeiro.

O método utilizado nesse estudo está baseado nos modelos de escolha qualitativas, no qual, o objetivo é explicar a escolha da alternativa j, em que, P_j é função das características do município. Especificamente, a probabilidade é determinada por $P(Y=j/x)$, com $j = 0, 1$. Um aspecto de suma importância é o fato de que a variável dependente apresenta-se na forma discreta. O modelo probit pode ser apresentado como:

$$P_{ij} = \text{Pr ob}(Y_i = j / x_i) = \phi(\beta' X) = \int_{-\infty}^{\beta' X} \frac{1}{\sqrt{2\pi}} e^{-(\beta' X)^2/2} dt \quad (5)$$

Em que:

Y_i é a variável aleatória que indica a escolha;

P_{ij} = $\text{Pr ob}(Y_i = j / x_i)$ é a probabilidade do município i optar pela escolha j;

x_i é a matriz de atributos observáveis para os municípios;

β é o vetor de parâmetros a serem estimados.

A equação (4) representa a função de distribuição do probit. De acordo com Gujarati (2006), o modelo assegura que $0 < P_{ij} < 1$ e $\sum_{j=1}^m P_{ij} = 1$, para identificação do modelo, o β_j é fixado para uma das categorias e os parâmetros são interpretados com respeito àquela categoria, chamada de categoria base.

O modelo probit é estimado pela técnica *Limited-Information Maximum Likelihood* (LIML). Desse modo, de L regressores, z , são variáveis endógenas ou predeterminadas. Seja y o vetor de M variáveis endógenas e x o vetor de K variáveis predeterminadas, com $M + K = L$. A partição δ_0 segue o seguinte conjunto $\delta_0 = (\gamma'_0, \beta'_0)'$. Onde m equações podem ser escritas como:

$$y = z' \delta_0 + e = \underset{(1 \times L) \quad (L \times 1)}{y'} \underset{(1 \times M) \quad (M \times 1)}{\gamma'_0} + \underset{(1 \times K)}{x} \underset{(1 \times K)}{\beta_0} + e \quad (6)$$

Obviamente, está simples equação é um sistema incompleto devido a existência de M variáveis endógenas. Portanto, para a solução do sistema de equações é necessário encontrar a forma reduzida. Com isso, seja Π_0 associada com a forma reduzida da matriz de coeficientes, temos:

$$\underset{(M \times 1)}{y} = \underset{(M \times K)}{\Pi_0} \underset{(K \times 1)}{x} + \underset{(M \times 1)}{v} \quad (7)$$

Combinando as equações (5) e (6), obtemos o sistema de 1+M equações:

$$\underset{((1+M) \times (1+M))}{\bar{\Gamma}_0} \underset{((1+M) \times 1)}{\bar{y}} + \underset{((1+M) \times K)}{\bar{B}_0} \underset{(K \times 1)}{x} = \underset{((1+M) \times 1)}{\bar{e}} \quad (8)$$

Onde,

$$\bar{y} = \begin{bmatrix} y \\ \bar{y} \\ \underset{(M \times 1)}{y} \end{bmatrix}, \bar{e} = \begin{bmatrix} e \\ \bar{v} \\ \underset{(M \times 1)}{v} \end{bmatrix}, \bar{\Gamma}_0 \equiv \begin{bmatrix} 1 & -\gamma'_0 \\ 0 & I_M \end{bmatrix} \text{ e } \bar{B}_0 \equiv \begin{bmatrix} -\beta'_0 & 0' \\ \underset{(1 \times K)}{(1 \times K)} & \underset{(1 \times (K-K_m))}{(1 \times (K-K_m))} \\ -\underset{(M \times K)}{\Pi_0'} & \end{bmatrix}$$

Desse modo, a equação (7) é um sistema completo de 1 + M equações simultâneas devido: $\left| \bar{\Gamma}_0 \right| = 1 \neq 0$

Os coeficientes estimados não representam diretamente as respostas marginais das variáveis explicativas e são de difícil interpretação. Assim, é necessário calcular os efeitos marginais para analisar corretamente os resultados. Estes são obtidos pela diferenciação de (4), tal que:

$$\delta_{ij} = \frac{\partial p_{ij}}{\partial x_i} = p_{ij} \left[\beta_j - \sum_{j=1}^m p_{ij} \beta_k \right] = p_{ij} \left[\beta_j - \bar{\beta}_i \right] \quad (9)$$

Em que, $\bar{\beta}_i = \sum_j^m p_{ij} \beta_j$ é a probabilidade média de β_j .

O efeito marginal varia com o ponto de estimação, x_i , uma vez que p_{ij} varia com x_i . Para cada x_i , particular, o efeito marginal, $\partial p_{ij} / \partial x_k$, não necessariamente terá o mesmo sinal de β_{jk} , sendo que, o efeito marginal será positivo se $\beta_j > \bar{\beta}_i$. O erro-padrão pode ser estimado usando o método delta.

4 ANÁLISE DOS RESULTADOS

Após apresentação do arcabouço teórico e metodológico a etapa seguinte consistiu em verificar se o modelo simultâneo obtém solução única para a matriz de coeficientes, desse modo, como satisfaz a condição de ordem e de posto o sistema é identificado. Além disso, em verificar as estimativas obtidas pela aplicação do modelo probit, no qual, busca analisar como a relação entre desenvolvimento, índice de transparência e desigualdade, e seus fatores determinantes, estão conectados. As estimativas do referido modelo simultâneo seguem a interpretação dos parâmetros no modelo probit, haja vista que não são de fácil interpretação, pois não representam diretamente as repostas marginais como no método de MQO (Mínimos Quadrados Ordinários). Nesse sentido, as análises foram realizadas a partir dos coeficientes marginais de cada variável explicativa.

Ratificando o ajustamento e a robustez do modelo simultâneo, temos o teste de significância global, no qual, rejeita-se a hipótese nula de que os coeficientes conjuntamente são iguais à zero, confirmando a relevância das variáveis explicativas. Com relação a endogeneidade realizou-se o teste de Wald, no qual, nos três testes realizados, não rejeitou-se a hipótese nula de exogeneidade. Para tanto, o teste é realizado entre duas equações e, por fim, é gerado um p-valor, que no caso de estudo não aceitou-se a hipótese alternativa a 5% de significância nos três testes realizados, como pode ser visto na Tabela 3.

No intuito de examinar a relação proposta são apresentadas as estimativas dos efeitos marginais das variáveis (Tabela 3) considerando as três equações analisadas. Percebe-se pelos resultados dos coeficientes que o modelo simultâneo possui um bom poder explicativo para as variáveis de interesse. Na equação de desenvolvimento as variáveis log da do número de residentes, média de anos dos indivíduos e índice de qualidade institucional, não são significantes. Enquanto, para a equação do índice de transparência municipal a variável insignificante foi o partido de esquerda. Além disso, temos a estimativa da desigualdade, no qual, a variável que não apresentou relevância ao modelo foi índice de transparência municipal.

A primeira implicação desse resultado é a comprovação da simultaneidade entre o

desenvolvimento e a desigualdade, onde, a variação de uma dessas variáveis influenciará na mudança da outra e, além disso, na variação do índice de transparência municipal. Por outro lado, foi constatado que mudança no grau de transparência implica na variação do nível de desenvolvimento, entretanto, não foi visualizado impacto da mesma sobre a desigualdade. Nesse sentido, se o índice de transparência estiver acima da média implicará em redução de 14,1% da chance do município apresentar melhor desenvolvimento. Todavia, se o município estiver acima da média da variável desenvolvido, resultará no aumento de 27,7% e 36,2%, de apresentar maior, respectivamente, transparência e desigualdade. Por sua vez, se o município estiver abaixo da média da variável desigualdade, as chances de transparência aumentam em 27,6% e reduzem a probabilidade do desenvolvimento em 10,4%.

Observando cada estimativa individualmente, a desigualdade e o desenvolvimento são afetados negativamente pelo aumento da média de idade municipal, isto é, se a média de anos populacional aumenta em uma unidade, as chances do município em questão ser menos desigual e mais desenvolvido reduzem-se, respectivamente, 21,7%, e 4,1%.

O log populacional é estatisticamente significativo a 1% sobre a estimativa da desigualdade, apresentando relação positiva com com a má distribuição de renda. A saber, com o aumento de uma unidade do log populacional implicará no aumento de 34,5% sobre as chances de o município ser mais desigual. Por outro lado, o log da transferência municipal correlaciona-se positivamente com o desenvolvimento, desse modo, com o aumento da transferência municipal, sua probabilidade de ser classificado como desenvolvido aumenta.

Com relação ao aumento do capital social, em uma unidade, as chances de o município ser classificado como transparente aumentam em 2,6% sua probabilidade. Por outro lado, a taxa de empregado foi a variável que apresentou maior poder explicativo sobre a equação de desenvolvimento. Desse modo, com o aumento da taxa de emprego, haverá queda na probabilidade do município ser classificado como desigual.

Tabela 3 – Estimação dos coeficientes e efeitos marginais do modelo simultâneo

Variáveis Explicativas	Variáveis Resposta					
	<i>desenv</i>		<i>itm</i>		<i>desig</i>	
	Coeficientes	Marginal effects	Coeficientes	Marginal effects	Coeficientes	Marginal effects
<i>desenv</i>	-	-	0,613*** (0,159)	0,277	10,046*** (0,324)	0,362
<i>itm</i>	-0,681** (0,339)	-0,141	-	-	0,027 (0,316)	0,001
<i>desig</i>	-0,355* (0,186)	-0,104	0,743*** (0,203)	0,276	-	-
<i>idade</i>	-0,092 (0,078)	-0,041	-	-	-0,445*** (0,094)	-0,217
<i>lpop</i>	-0,253 (0,175)	-0,112	-	-	0,705*** (0,223)	0,345
<i>lt</i>	4,085*** (0,827)	1,823	-	-	-	-
<i>iqim</i>	0,361 (0,222)	0,161	-	-	-	-
<i>caps</i>	-	-	0,074*** (0,024)	0,026	-	-
<i>igf</i>	-	-	2,365*** (0,618)	0,841	-3,041*** (0,922)	-1,488
<i>ig</i>	-	-	-0,196*** (0,056)	-0,069	-	-
<i>pp</i>	-	-	-0,766*** (0,231)	-0,088	-	-
<i>pe</i>	-	-	-0,430 (0,268)	-0,101	-	-
<i>pc</i>	-	-	-0,512*** (0,196)	-0,106	-	-
<i>txe</i>	-	-	-	-	-20,21*** (3,474)	-9,892
<i>constante</i>	-19,28*** (4,508)	-	-	-	14,17*** (3,78)	-
P – valor						
Test de Wald entre as equações (1) e (2)				-0,370		
Test de Wald entre as equações (1) e (3)				-0,588		
Test de Wald entre as equações (2) e (3)				-0,440		
Nº de observações			166			
Wald chi2			80,07			
Prob > chi2			0,000			

Fonte: Elaboração do próprio autor.

Nota 1: O desvio padrão estar logo abaixo da variável.

Nota 2: *** significância a 1%; ** significância a 5%; * significância a 10%.

Já as variáveis de cunho classificatório dos partidos, isto é, se o partido é de centro ou de esquerda, apresentaram significância estatística sobre a disponibilização de informações financeiras. Se o partido apresenta ideologia de centro ou de esquerda, as chances reduzem

em, respectivamente, 10,1% ou 10,6% a probabilidade do município ser mais transparente.

Observa-se que a variável *igf* relaciona-se positivamente com a transparência, apresentando significância a 1%, indicando que, quanto maior for o índice de gestão fiscal mais elevadas serão as chances de o município apresentar maior transparência, a saber, se o município obtiver nota máxima com relação ao índice de gestão fiscal, que neste caso é 1, a sua probabilidade eleva-se em 84,5% do município ser mais transparente.

Por sua vez, o índice de governança impacta negativamente sobre o índice de transparência, como o índice varia de 1 a 8, quanto mais próximo do topo melhor será a responsabilidade administrativa, gestão fiscal e decisão dos gestores sobre os tributos arrecadados, ocasionando maior transparência. Com a variação do índice para mais, a probabilidade do município ser mais transparente reduz em 6,9%.

Em relação a variável *pp*, no qual, significa persistência política, tem impacto como esperado, a saber, se o prefeito foi reeleito reduz em 8,8% em sua probabilidade de o município ser mais transparente. Isso ocorre devido ao fato de que, com a reeleição, o político acaba por se acomodar, diminuindo a disponibilização de informações necessárias à sociedade.

5 CONCLUSÃO

O presente estudo teve como objetivo averiguar como a relação entre desenvolvimento, índice de transparência e desigualdade, e seus fatores determinantes, estão conectados, com o intuito de entender as disparidades que assolam o território cearense.

Um dos resultados encontrados que merece destaque é a relação de causalidade entre as variáveis de interesse. O desenvolvimento e a desigualdade dos municípios cearenses apresentam causalidade entre si. Já o impacto das mesmas sobre o nível de transparência, apresentou-se positivo. Entretanto, a variação da transparência implicou mudança apenas no desenvolvimento, não sendo significativa na determinação da desigualdade.

No âmbito das variáveis que denotam características socioeconômicas e financeiras, as variáveis de maior relevância foram a idade e a taxa de empregados, no qual, apresentaram impacto negativo na explicação da desigualdade. Todavia, o impacto passa a ser positivo quando consideramos a variável log da população. Por sua vez, o aumento do valor de transferência afeta positivamente o desenvolvimento.

No grupo das variáveis que mensuram a qualidade institucional, há de se ressaltar

quanto maior o índice de gestão fiscal maior será a transparência, todavia, menor a desigualdade. Por outro lado, quanto melhor o nível de governança menor será a disponibilização de informações financeiras.

No âmbito das variáveis políticas, chama-se atenção para a persistência política, isto é, se o prefeito foi reeleito reduz as chances de transparência. Além disso, se o partido apresentar ideologia de centro ou de esquerda, o município será menos transparente.

A partir dos resultados aqui apresentados, fica evidente a necessidade de políticas direcionadas a redução da desigualdade como forma de melhorar o nível de desenvolvimento econômico e transparência municipal, buscando o aumento do bem estar da sociedade.

É importante destacar que o presente trabalho traz contribuições importantes para a literatura nacional. Uma delas é a mensuração do desenvolvimento de forma multidimensional, no qual, é mais alinhado com estudos internacionais. Em geral, estudos realizados para o Brasil até então se limitaram ao uso de variáveis unidimensionais para a mensuração de crescimento.

Desse modo, tendo como base o coeficiente de Gini, o IDHM e o índice de transparência municipal aqui utilizados, entende-se a necessidade de políticas públicas voltadas à sociedade, procurando a conscientização da população sobre o seu papel nas escolhas tomadas pela administração municipal. Nesse sentido, com a participação efetiva dos cidadãos aumentam as chances de as políticas desenvolvidas pelo poder local voltar-se a aspectos que necessariamente melhore a qualidade de vida.

Entre as fragilidades deste estudo, destacam-se as limitações da base de dados existentes para o âmbito municipal, com destaque para as duas variáveis adotadas para a mensuração do desenvolvimento e da desigualdade, a saber, respectivamente, foram empregados o IDHM e o índice de Gini, no qual, encontrou-se disponível para o ano de 2010 ou anteriores. O ideal é que os dados sejam mais atuais, de modo que possam capturar todo o apanhado de informações referentes ao desenvolvimento e desigualdade sobre a sociedade.

Por fim, destaca-se que desenvolvimento e desigualdade são temas bastante amplos e com muitas faces, existindo oportunidade e abrangência para estudos em inúmeras escalas com o envolvimento de diversas variáveis, no qual, podem auxiliar na formulação de cenários e no entendimento da realidade que nos cerca. Cabe, portanto, como sugestão para estudos futuros, a investigação de como o desenvolvimento e a desigualdade se relacionam com outras medidas sociais, financeiras, etc. Além disso, a literatura em relação à transparência no âmbito nacional é muita escassa, sendo interessante, estudos que abordem os determinantes da má disponibilização de informações financeiras.

REFERÊNCIAS

ACEMOGLU, D.; JOHNSON, S.; ROBINSON, J. A. The colonial origins of comparative development: an empirical investigation. **American Economic**

Review, v. 91, n. 5, p. 1369-1402, 2001. _____ Reversal of Fortune : Geography and Institutions in the making of the Modern world Income distribution. **Quarterly Journal of Economics**, n.

117, p. 1231-1294, 2002. _____ Institutions as the fundamental cause of longrun growth. **National Bureau of Economic Research**, Cambridge, 2004. Disponível em: < [http:// www.nber.org/papers/w10481](http://www.nber.org/papers/w10481)>. Acesso em: 24/06/2017.

ALBALATE DE SOL, D. (2013). The institutional, economic and social determinants of local government transparency. **Journal of Economic Policy Reform**, 16(1), 90-107.

ANDREULA, N.; CHONG, A. (2015). Do good institutions improve fiscal transparency? **Economics of Governance**, 1-23.

ALT, J.; LASSEN, D.; ROSE, S. (2006). The causes of fiscal transparency: evidence from the American States. **IMF Staff Papers** 53, 30–57.

BANERJEE, A.; DUFLO E. (2011). Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. **Penguin Books**.

BARRO, R. J. (2000). Inequality and growth in a panel of countries. **Journal of economic growth**, v. 5, p. 5-32.

BLUHM, R.; SZIRMAI, A. Institutions, inequality and growth: a review of theory and evidence on the institutional determinants of growth and inequality. **Innocenti Working Paper** (2011).

CAAMAÑO-ALEGRE, J.; LAGO-PEÑAS, S.; REYES-SANTIAS, F.; SANTIAGO-BOUBETA, A. (2013). Budget Transparency in Local Governments: An Empirical Analysis. **Local Government Studies** 39(2), 182-207.

CABA, C.; LÓPEZ, A. M. (2009). Governmental financial transparency in MERCOSUR member countries. **International Review of Administrative Sciences**, Vol.75(1), pp. 169-181.

CICATIELLO, L.; GAETA, E. G. L. (2016). Political determinants of fiscal transparency: a panel data empirical investigation. **Munich Personal RePEc Archive**. No. 72609.

CHANG, H.-J. (2009). **Hamlet without the prince of Denmark**: How development has disappeared from today's 'development' discourse. Routledge. Abingdon, p.1-11. 15.

CHONG. A.; GRADSTEIN, M. Inequality and Institutions. **Inter-American Development Bank**. (2004).

CHONG, A.; GRADSTEIN, M.; Inequality and informality. **J. Public Econ.** 91 (1–2),

159–179. 2007a.

CHONG, A.; GRADSTEIN, M. Inequality and institutions. **Rev. Econ. Stat.** 89 (3), 454–465. 2007b.

CHRISTIAENS, J. (1999). Financial accounting reform in flemish municipalities: an empirical investigation. **Financial Accountability & Management**, Vol.15(1), pp. 21-40.

DE RENZIO, P.; ANGEMI, D. (2012). COMRADES OR CULPRITS? DONOR ENGAGEMENT AND BUDGET TRANSPARENCY IN AID-DEPENDENT COUNTRIES. **Public Administration and Development**, 32(2), 167-180.

EASTERLY, W.; LEVINE, R. Tropics, germs and crops: How endowments influence economic development. **NBER Working Paper** No.9106, 2002.

ENGERMAN, S.; SOKOLOFF, K. Factor endowments, Inequality, and Paths of development among New World Economies. **Economia**, v. 3, p. 41-109, 2002.

ESTELLER-MORÉ, A.; OTERO, J. P. (2012). **Fiscal Transparency, Public Management Review**. 14(8), 1153-1173.

FEREJOHN, J. (1999). Accountability and authority: toward a theory of political accountability. **Democracy, accountability, and representation**, 131, 133.

GANDÍA, J. L.; ARCHIDONA, M. C. (2008). Determinants of web site information by Spanish city councils. **Online Information Review**, 32(1), 35-57.

GARCÍA, A. C.; GARCÍA-GARCÍA, J. (2010). Determinants of Online Reporting of Accounting Information by Spanish Local Government Authorities. **Local Government Studies**, 36(5), 679-695.

GIROUX, G.; MCLELLAND, A. (2003). Governance structure and accounting at large municipalities. **Journal of Accounting and Public Policy**, Vol.22, pp. 203-230.

GUILLAMÓN, M.; BASTIDA, F.; BENITO, B. (2011). The Determinants of Local Government's Financial Transparency. **Local Government Studies**, 37(4), 391-406.

GUJARATI, Damodar. N. **Econometria básica**. 4ª edição. Rio de Janeiro: Elsevier, 2006.

HALL, R.E.; JONES, C. I. Why do some countries produce so much more than others?. **Quarterly Journal of Economics**, v. 114, n.1, p. 83-116, 1999.

HAZELL, R.; WORTHY, B. (2010). Assessing the performance of freedom of information. **Government Information Quarterly**, 27(4), 352-359.

HODGSON, G. M. (1998). The approach of institutional economics. **Journal of Economic Literature**. Vol. 36, No. 1, p. 166-192.

KIM, J.H. (2016). A Study on the Effect of Financial Inclusion on the Relationship Between Income Inequality and Economic Growth. **Emerging Markets Finance and Trade**, v. 52, n. 2, p. 498-512.

KOTSCHY, R.; SUNDE, U. Democracy, inequality, and institutional quality. **European Economic Review**. No. 91, 209–228 (2017).

KUZNETS, S., (1955), 'Economic Growth and Income Inequality', *The American Economic Review*, vol. 65, pp. 1-28.

LASWAD, F.; FISHER, R.; OYELERE, P. (2005). Determinants of voluntary Internet financial reporting by local government authorities. **Journal of Accounting and Public Policy**, 24(2), 101-121.

LEACHMAN, L. L.; ROSAS, G.; LANGE, P.; BESTER, A. (2007). The political economy of budget deficits. **Economics & Politics**, 19(3), 369-420.

Lei n. 12.527 de 18 de novembro de 2011. (2011). Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei no 8.112, de 11 de dezembro de 1990; revoga a Lei no 11.111, de 5 de maio de 2005, e dispositivos da Lei no 8.159, de 8 de janeiro de 1991; e dá outras providências. Recuperado em 1 maio, 2012, de

<http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm>.

MICHENER, G. (2015). How Cabinet Size and Legislative Control Shape the Strength of Transparency Laws. **Governance**, 28(1), 77-94.

MIZRAHI, R. S. (2010). O impacto da desigualdade sobre o desenvolvimento econômico. **Opinion sur**. Argentina, n. 84.

NARITOMI, J. Herança Colonial, Instituições & Desenvolvimento: Um estudo sobre a desigualdade entre os municípios Brasileiros. **Dissertação de mestrado**. (2007).

NORTH, D. (1991). Institutions. **The Journal of Economic Perspectives**, 5 (1), p. 97–112.

OKUN, A. M. (2015). Equality and efficiency: The big tradeoff. **Brookings Institution Press**.

PNUD - Programa das Nações Unidas para o Desenvolvimento. **Atlas do desenvolvimento humano no Brasil**: atlas Brasil, 2013. Disponível em: <www.atlasbrasil.org.br/2013/>.

Reddick, C. (2005). Citizen interaction with e-government: From the streets to servers? **Government Information Quarterly**, 22(1), 38–57.

RÍOS, A.-M.; BASTIDA, F.; BENITO, B. (2014). Budget Transparency and Legislative Budgetary Oversight An International Approach. **The American Review of Public Administration**.

ROBERTS, A. (2006). *Blacked out: Government secrecy in the information age*. Cambridge University Press.

SAVOIA, A.; EASAW, J.; MCKAY, A. Inequality, Democracy, and Institutions: A Critical Review of Recent Research. *World Development*, Vol. 38, No. 2, pp. 142–154, 2010.

SCARPIM, J. E.; SLOMSKI, V. (2007). Estudos dos fatores condicionantes do índice de desenvolvimento humano nos municípios do estado do Paraná: instrumento de controladoria para tomada de decisões na gestão governamental. *Revista de administração pública*, vol 41, num 5, p 909-933. Rio de Janeiro.

SETH, S.; VILLAR, A. (2015). Human Development, Inequality and Poverty: empirical findings. *Working papers series*.

Sol, D. A. D. (2013). The institutional, economic and social determinants of local government transparency. *Journal of Economic Policy Reform*.

WEHNER, J.; DE RENZIO, P. (2013). Citizens, Legislators, and Executive Disclosure: The Political Determinants of Fiscal Transparency. *World Development*, 41, 96-108.

ZUCCOLOTTO, R.; TEIXEIRA, M. A. (2014). Budgetary Transparency and Democracy: The Effectiveness of Control Institutions. *International Business Research*, 7(6), 83.

ANEXO A - Aspectos avaliados pelo Ministério Público Federal para a elaboração do Índice de Transparência Municipal.

Nº	Pergunta da avaliação	Peso	%
1	O ente possui informações sobre transparência na internet?	2	2%
2	O site contém ferramenta de pesquisa de conteúdo que permita o acesso à informação?	2	2%
3	Há informações sobre a receita nos últimos 6 meses, incluindo natureza, valor de previsão e valor arrecadado?	10	10%
4a	As despesas apresentam dados dos últimos 6 meses contendo: valor do empenho	4	4%
4b	As despesas apresentam dados dos últimos 6 meses contendo: valor da liquidação	3	3%
4c	As despesas apresentam dados dos últimos 6 meses contendo: favorecido	4	4%
4d	As despesas apresentam dados dos últimos 6 meses contendo: valor do pagamento	4	4%
5a	O site apresenta dados dos últimos 6 meses contendo: íntegra dos editais de licitação	4	4%
5b	O site apresenta dados dos últimos 6 meses contendo: resultado dos editais de licitação (vencedor é suficiente)	3	3%
5c	O site apresenta dados dos últimos 6 meses contendo: contratos na íntegra	4	4%
6a	O ente divulga as seguintes informações concernentes a processos licitatórios com dados dos últimos 6 meses: modalidade	1	1%
6b	O ente divulga as seguintes informações concernentes a processos licitatórios com dados dos últimos 6 meses: data	1	1%
6c	O ente divulga as seguintes informações concernentes a processos licitatórios com dados dos últimos 6 meses: valor	1	1%
6d	O ente divulga as seguintes informações concernentes a processos licitatórios com dados dos últimos 6 meses: número / ano de edital	1	1%
6e	O ente divulga as seguintes informações concernentes a processos licitatórios com dados dos últimos 6 meses: objeto	1	1%
7a	O site apresenta: a prestação de contas (Relatório de Gestão) do ano anterior	2	2%
7b	O site apresenta: Relatório resumido da Execução Orçamentária (RRO) dos últimos 6 meses	3	3%
7c	O site apresenta: Relatório de Gestão Fiscal (RGF) dos últimos 6 meses	3	3%
7d	O site apresenta: Relatório estatístico contendo a quantidade de pedidos de informação recebidos, atendidos e indeferidos, bem como informações genéricas sobre os solicitantes.	2	2%
8	O site possibilita a gravação de relatórios em diversos formatos eletrônicos, abertos não proprietários, tais como planilhas e texto (CSV), de modo a facilitar a análise das informações.	2	2%
9a	Possibilidade de entrega de um pedido de acesso de forma presencial: existe indicação precisa no site de funcionamento de um SIC físico	1	1%
9b	Possibilidade de entrega de um pedido de acesso de forma presencial: há indicação do órgão	1	1%
9c	Possibilidade de entrega de um pedido de acesso de forma presencial: há indicação de endereço	1	1%
9d	Possibilidade de entrega de um pedido de acesso de forma presencial: há indicação de telefone	1	1%
9e	Possibilidade de entrega de um pedido de acesso de forma presencial: há indicação de horários	1	1%
10	Há possibilidade de envio de pedidos de informação de forma eletrônica (E SIC)?	8	8%
11	Apresenta possibilidade de acompanhamento posterior da solicitação?	7	7%
12	A solicitação por meio do E SIC é feita de forma fácil e simples sem a exigência de pelo menos um item de identificação do requerente que dificultem ou impossibilitem o acesso	5	5%
13	No site está disponibilizado o registro das competências e estrutura organizacional do ente?	2	2%
14	O portal disponibiliza endereços e telefones das respectivas unidades e horários de atendimento ao público?	2	2%
15	Há divulgação de remuneração individualizada por nome do agente público?	10	10%
16	Há divulgação de diárias e passagens por nome de favorecido e constando: data, cargo e motivo de viagem?	5	5%

Fonte: Elaboração do próprio autor.

DETERMINANTES DA PARTICIPAÇÃO DOS CASAIS NO MERCADO DE TRABALHO: ANÁLISE PARA A REGIÃO NORDESTE E O ESTADO DO CEARÁ

Daniel Tomaz de Sousa*

Maria Adreciana Silva de Aguiar**

Francisco Germano Carvalho Lúcio***

Priscila Silva Rodrigues****

RESUMO

Este estudo visa identificar os determinantes das decisões de entrada dos casais no mercado de trabalho, especificamente nas regiões urbanas do estado do Ceará e da Região Nordeste. Este objetivo baseia-se nas mudanças da configuração familiar na qual o homem deixa de ser o único provedor, e ambos os cônjuges passam a contribuir para a renda familiar e tomam decisões conjuntamente de entrada e saída no mercado de trabalho. A partir dos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2015, utilizou-se um modelo probit bivariado aparentemente não relacionado para a análise da ocupação dos casais no mercado de trabalho de acordo com as características individuais e familiares. Alguns resultados apontam similaridades na comparação entre a Região Nordeste e o estado do Ceará. Um desses sugere que quanto maior o nível de escolaridade maior é a chance de participação no mercado de trabalho tanto do marido quanto da esposa; o contrário é observado para a educação do cônjuge que impacta negativamente sobre essa chance. A presença de filhos aumenta a probabilidade de ocupação dos maridos. No entanto, a fecundidade reduz as chances da esposa estar inserida no mercado de trabalho, com maior efeito para aquelas que residem no Nordeste, em comparação com as residentes do estado do Ceará.

Palavras-Chave: Mercado de trabalho, casais, região urbana.

ABSTRACT

This study aims to identify the determinants of couples' entry decisions in the labor market, specifically in the urban regions of the state of Ceará and the Northeast Region.

* Doutorando em Economia – PPGE/UFPB, daniel25tomaz@gmail.com, (83) 99835-5120.

** Doutoranda em Economia - CAEN/UFC, adreciane@gmail.com.

*** Doutorando em Economia - CAEN/UFC, germanocarvalho15@hotmail.com.

**** Doutoranda em Economia - CAEN/UFC, psilinhag12@gmail.com.

This goal is based on changes in the family configuration in which men cease to be the sole provider, and both spouses contribute to family income and make joint decisions of entry and exit in the labor market. Based on data from *Pesquisa Nacional por Amostra de Domicílios (PNAD)* of 2015, we used a Seemingly unrelated bivariate probit model to the analysis of the occupation of couples in the labor market according to individual and family characteristics. Some results point similarities in the comparison between the Northeast Region and the state of Ceará. One of these suggests that the higher the level of schooling the greater the chance of participation in the labor market of both husband and wife; the opposite is observed for the education of the spouse who negatively impacts on this chance. The presence of children increases the likelihood of the husbands' occupation. However, fertility reduces the chances of the wife being inserted in the labor market, with greater effect for those residing in the Northeast, compared to the residents of the state of Ceará.

Keywords: Labor market, couples, urban region.

1 INTRODUÇÃO

Quase totalidade dos lares brasileiros sempre tiveram os homens como principais provedores de renda, fato decorrente da formação sociocultural das famílias no país. Porém, dado a relativamente recente inserção da mulher no mercado de trabalho, a participação feminina na composição da renda familiar vem aumentando ao longo dos anos. Atualmente, em muitos casos as mulheres são as principais provedoras de suas famílias, muitas das quais fornecem a única renda da casa.

Salvo exceções, há uma tendência mundial de aumento da escolaridade feminina. Assim, nas últimas décadas, mulheres mais escolarizadas tiveram a oportunidade de ingressar no mercado de trabalho, ascendendo a cargos antes nunca imaginados e podendo ter maior participação na renda familiar.

Blau e Kahn (2007) observaram que desde a segunda guerra mundial há uma crescente participação das mulheres no mercado de trabalho dos Estados Unidos. Considerando que em 1947 cerca de 32% das mulheres estavam inseridas no mercado de trabalho no país, os autores revelaram que a partir dos anos 2000 a taxa de participação da força de trabalho feminina aumentou em cerca de 60%.

Essa mudança na estrutura dos rendimentos familiares influenciou diversos estudos a analisarem a inserção da mulher no mercado de trabalho e sua participação nos rendimentos familiares e ainda trabalhos que investigam diferenças salariais e de rendimentos

femininos¹.

A partir do exposto acima, tem-se que famílias nas quais ambos os cônjuges contribuem para a renda familiar e tomam decisões quanto a entrada e saída no mercado de trabalho conjuntamente já são parte de um número representativo na maioria dos países desenvolvidos. Estes casais, ao tomarem suas decisões, escolhem entre trabalho remunerado, trabalho doméstico e lazer. Essa decisão é influenciada tanto pela condição do cônjuge quanto pelas características ou atributos individuais².

Dado esse contexto e as mudanças observadas com relação ao mercado de trabalho dos casais brasileiros, o presente trabalho tem como objetivo traçar os determinantes das tomadas de decisão de ingresso no mercado de trabalho por parte dos casais em 2015, especificamente nas áreas urbanas do estado do Ceará e da Região Nordeste. Deseja-se, portanto, verificar a probabilidade da ocupação de ambos os cônjuges, observando os fatores que podem influenciar nessa decisão. Para tanto, utilizou-se o modelo Probit bivariado aparentemente não relacionado considerando o plano amostral complexo da PNAD de 2015.

O trabalho foi desenvolvido em mais 4 seções, além desta introdução. A próxima seção traz uma breve revisão de literatura contendo uma explanação dos principais trabalhos que tratam sobre a temática da oferta de trabalho de casais. A terceira seção trata da metodologia adotada e dos dados utilizados. Na quarta seção este trabalho revela os principais resultados encontrados. Na quinta e última seção constam as principais conclusões.

2 REFERENCIAL TEÓRICO

Dado as mudanças ocorridas nas últimas décadas com relação a formação de renda das famílias e a participação da mulher no mercado de trabalho a partir da segunda metade do século XX na grande maioria dos países no mundo, estudos diversos³ buscaram analisar o mercado de trabalho feminino ou o mercado de trabalho da mulher cônjuge, observando as condições de trabalho, bem como diferenças salariais e segmentação no mercado de trabalho e a geração de renda familiar.

Além de pesquisas no âmbito específico da mulher surgiu ainda o interesse em analisar a família e as decisões tomadas pelos casais. Jatobá (1989), por exemplo, afirmou que a decisão do indivíduo, seja ele o chefe da família ou a mulher cônjuge, é tomada dentro

1. São exemplos: Hoffmann e Leone (2009), Gomes (2010); Madalozzo, Martins e Shiratori (2010); Blau & Kahn (2007); Marri e Wajzman (2007).

2. Para maiores detalhes ver Batista (2016); Guimarães (2009); Jatobá (1994).

3. São exemplos: Becker (1991); Batista (2016); Madalozzo, Martins e Shiratori (2010); Sedlacek e Santos (1991).

do núcleo familiar e que essa decisão sofre influência de diversas características de cada família. Assim, dado esses atributos, as decisões quanto à entrada ou a saída do mercado de trabalho são tomadas em conjunto.

Jatobá (1989) segue a hipótese de que a família está mais sujeita a distorções da economia e do mercado de trabalho de forma dependente ao sexo do indivíduo considerado chefe da família. Assim, procurou encontrar a taxa de participação da família na força de trabalho (TPFFT) e afirmou que essa taxa tem relação com o sexo do chefe de família.

A literatura que trata as mudanças ocorridas com relação à divisão do trabalho entre homens e mulheres e a análise dos novos padrões de escolha de trabalho de maridos e esposas está inserida em um novo campo de estudos econômicos conhecido como *New Home Economics*. Isso conduziu a pesquisas voltadas para as decisões no âmbito da família e o trabalho doméstico (MADALOZZO, MARTINS E SHIRATORI, 2010).

Parte da literatura que analisa o mercado de trabalho e rendimentos dos casais foca na questão do duplo rendimento, em outras palavras, famílias nas quais ambos os cônjuges trabalham. Em geral, os resultados apontam que a inserção da mulher no mercado de trabalho e a questão de os dois cônjuges estarem trabalhando e contribuirão nos gastos familiares leva a uma maior estabilidade em momentos de turbulência econômica. Porém, haverá um aumento do poder de barganha entre os cônjuges, principalmente para a mulher⁴. Na mesma linha, Madalozzo, Martins e Shiratori (2010) afirmam que o poder de barganha adquirido pela esposa é uma função dos ganhos que ela adquire no mercado.

Outros trabalhos como o de Suliano, Irffi e Veras (2014) e de Silva e Santos (2015) analisaram o mercado de trabalho de casais, porém de forma diferente dos demais trabalhos citados e dos objetivos deste artigo. Tais estudos focaram na comparação entre casais homoafetivos e heterossexuais. Silva e Santos (2015) concluíram que casais homoafetivos, se comparados a casais heterossexuais, possuem um maior nível de bem-estar apresentando uma discriminação positiva no mercado de trabalho. Suliano, Irffi e Veras (2014), por sua vez, geram conclusões análogas, que os casais homoafetivos possuem maiores ganhos vis-à-vis os casais heterossexuais. A semelhança das conclusões deve-se ao fato de maiores ganhos geralmente estão associados a maiores níveis de bem-estar.

A teoria da oferta de trabalho introduziu um conceito advindo da observação da mudança ocorrida com a inserção da mulher no mercado de trabalho, dado que muitas mulheres acabam entrando no mercado de trabalho após seus maridos ingressarem em uma situação de desemprego. Assim a literatura estimou o chamado efeito trabalhador adicio-

4. Ver Batista (2016) e Deutsch, Roksa, Meeske (2003), por exemplo.

nal, ou seja, aquele indivíduo na família que após o desemprego do trabalhador principal passa a inserir-se no mercado de trabalho com o propósito de manter o rendimento familiar e o consumo (SPETZER, 1997).

Dado o conceito de trabalho adicional, sob a ótica da teoria tradicional da oferta de trabalho cada indivíduo, seja o marido ou a esposa, toma a decisão entre trabalhar de forma remunerada, trabalhar no âmbito doméstico ou ter lazer. Tal decisão considera a maximização da utilidade da família. Assim, um “trabalhador adicional”, ao decidir entrar no mercado de trabalho, compara os ganhos de utilidade da entrada dado que o trabalho é remunerado, e as perdas, dado o menor tempo para lazer e a questão do trabalho doméstico.

Em relação a literatura nacional que trata sobre o mercado de trabalho dos casais, Guimarães e Santos (2009) analisaram a probabilidade de ocupação no mercado de trabalho de maridos e esposas. Buscaram compreender os aspectos que afetam as decisões da ocupação desses indivíduos a partir de uma perspectiva que considera gênero e raça.

Na tentativa de explicar os determinantes da ocupação familiar na área rural do Brasil, Figueiredo, Marta e Guimarães (2010) analisaram apenas o meio rural e examinaram a probabilidade da ocupação no mercado de trabalho de maridos e esposas.

O trabalho de Marri e Wajnman (2007) caminha nesse sentido e também traz a discussão acerca do trabalho de maridos e esposas. Porém, o intuito foi responder como as características observadas estão correlacionadas com a probabilidade de as famílias terem as mulheres como principais provedoras e ainda observar diferenciais de salários entre os cônjuges.

Outros trabalhos na literatura, assim como Da Silva e Da Cunha (2017), observaram a oferta de trabalho de casais sondando como se deu a evolução da participação dos casais no mercado de trabalho e ainda como a oferta de trabalho de maridos e esposas sofreram mudanças dados os ciclos econômicos. Oshita *et al.* (2016) verificaram o impacto da aquisição de capital humano sobre oferta de trabalho dos casais entre os anos de 2003 e 2015.

Dada a amplitude de abordagens expostas acima, frisa-se que este trabalho busca analisar e verificar os determinantes da ocupação dos casais, cearenses e nordestinos. Na próxima seção encontra-se a metodologia adotada para o exercício econométrico proposto.

3 METODOLOGIA

3.1 Dados e Tratamento

O presente artigo utiliza informações da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2015, considerando apenas as regiões urbanas do Nordeste e do estado do Ceará. Considerou-se o fato de que a PNAD consiste de uma pesquisa por amostragem complexa, por envolver estratificação, conglomeração e probabilidades desiguais de seleção. Pois, segundo Carvalho, Néri e Silva (2006) as estimativas das variâncias são influenciadas pelo plano amostral.

Neste trabalho considerou-se apenas as áreas urbanas do Nordeste e do estado do Ceará. Este recorte fora tomado sob o argumento de formar uma amostra mais homogênea e evitar discrepâncias entre as áreas. Além disso, foram consideradas apenas as pessoas com idade entre 15 e 65 anos e que possuíam cônjuge de sexo diferente⁵. Para captar a probabilidade de participação no mercado de trabalho, construiu-se uma variável binária que assume valor igual a 1 se o marido está no mercado de trabalho e 0, caso contrário. Construção análoga foi desenvolvida para a esposa.

No modelo considerou-se ainda como variáveis binárias os seguintes aspectos: (i) cor/raça, (ii) se o casal tem filhos menores de 14 anos, (iii) se a mulher é a chefe da família (pessoa de referência) e (iv) se reside em região metropolitana. Já as variáveis contínuas utilizadas foram referentes a educação do indivíduo e do cônjuge, ambas computadas como anos de escolaridade, a idade, ao tamanho da família e a renda familiar *per capita*. Essas variáveis são usualmente empregadas nos estudos empíricos sobre a participação no mercado de trabalho de homens e mulheres.

Após a eliminação das observações com pelo menos uma das informações supracitadas não declarada, chegou-se a uma população de 9.935.817 de famílias no Nordeste e 1.520.414 no Ceará.

3.2 Probit Bivariado

O modelo Probit bivariado, uma extensão do modelo Probit, utiliza duas equações com erros correlacionados (GREENE, 2012).

Caso as decisões de participação no mercado de trabalho tanto do marido quanto da esposa sejam dependentes, então, tem-se $cov(\varepsilon_1, \varepsilon_2 | x_1, x_2) = \rho$. Neste caso, o modelo mais adequado seria o Probit bivariado aparentemente não relacionado, pois permite

5. Na tentativa de deixar a amostra mais homogênea e devido a pequena proporção de casais homossexuais optou-se por retirá-los.

que diferentes variáveis independentes sejam colocadas nas equações de participação dos maridos e esposas. Para o caso de não haver correlação, pode-se estimar o modelo Probit univariado.

Este trabalho adota o modelo Probit bivariado para estimar a relação entre a probabilidade de participação do marido no mercado de trabalho (y_1^*) e a participação da esposa no mercado de trabalho (y_2^*). A especificação geral para as duas equações é dada por:

$$\begin{aligned} y_1^* &= \mathbf{x}'_1 \boldsymbol{\beta}_1 + \varepsilon_{1i} & y_1 &= 1 \text{ se } y_1^* > 0; y_1 = 0 \text{ caso contrário} \\ y_2^* &= \mathbf{x}'_2 \boldsymbol{\beta}_2 + \varepsilon_{2i} & y_2 &= 1 \text{ se } y_2^* > 0; y_2 = 0 \text{ caso contrário} \\ \left(\begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \end{array} \middle| \mathbf{x}_1, \mathbf{x}_2 \right) &\sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right] \end{aligned}$$

Vale ressaltar que a escolha simultânea dos casais gera quatro possibilidades diferentes e excludentes entre si, com relação a participação no mercado de trabalho. Assim, tem-se as seguintes situações:

- i) Marido e esposa participam [$y_1 = 1, y_2 = 1$];
- ii) Marido participa e a esposa não participa [$y_1 = 1, y_2 = 0$];
- iii) Marido não participa e a esposa participa [$y_1 = 0, y_2 = 1$] e
- iv) Ambos não participam [$y_1 = 0, y_2 = 0$].

A distribuição de probabilidade conjunta do resultado [$y_1 = 1, y_2 = 1$] pode ser escrita em termos das probabilidades condicional e marginal:

$$\text{Prob} [y_1 = 1, y_2 = 1 | \mathbf{x}] = \text{Prob} [y_2 = 1 | y_1 = 1 | \mathbf{x}] \cdot \text{Prob} [y_1 = 1 | \mathbf{x}]$$

$$\text{Prob} [y_1 = 1, y_2 = 1 | \mathbf{x}] = \Phi(x'_1 \boldsymbol{\beta}_1, x'_2 \boldsymbol{\beta}_2, \rho) = \Phi_{(y_1 = 1, y_2 = 1)}$$

Onde \mathbf{x} é o vetor composto por elementos de \mathbf{x}_1 e \mathbf{x}_2 .

Similarmente, as demais probabilidades podem ser descritas como:

$$\text{Prob} [y_1 = 1, y_2 = 0 | \mathbf{x}] = \Phi(x'_1 \boldsymbol{\beta}_1, -x'_2 \boldsymbol{\beta}_2, \rho) = \Phi_{(y_1 = 1, y_2 = 0)}$$

$$\text{Prob} [y_1 = 0, y_2 = 1 | \mathbf{x}] = \Phi(-x'_1 \boldsymbol{\beta}_1, x'_2 \boldsymbol{\beta}_2, \rho) = \Phi_{(y_1 = 0, y_2 = 1)}$$

$$\text{Prob} [y_1 = 0, y_2 = 0 | \mathbf{x}] = \Phi(-x'_1 \boldsymbol{\beta}_1, -x'_2 \boldsymbol{\beta}_2, \rho) = \Phi_{(y_1 = 0, y_2 = 0)}$$

O efeito de uma mudança em uma variável contínua, α_k , pode ser analisado considerando diferentes níveis. Considerando inicialmente que α_k esteja tanto em \mathbf{x}_1 como em \mathbf{x}_2 , o efeito marginal sobre a probabilidade conjunta de ambos estarem empregados,

$\Phi_{(y_1=1, y_2=1)}$, é dado por:

$$\begin{aligned} \frac{\partial \Phi_{(y_1=1, y_2=1)}}{\partial \alpha_k} &= \frac{\partial \Phi_{(y_2=1)} \Phi_{(y_1=1|y_2=1)}}{\partial \alpha_k} \\ &= \Phi_{(y_1=1|y_2=1)} \Phi_{(y_2=1)} \beta_{2k} + \Phi_{(y_2=1|y_1=1)} \Phi_{(y_1=1)} \beta_{1k} \end{aligned}$$

De maneira análoga, podem ser considerados os efeitos de uma variação de α_k nas demais probabilidades conjuntas. Já o efeito marginal na probabilidade de a esposa estar no mercado de trabalho $\Phi_{(y_2=1)}$ pode ser expresso como:

$$\frac{\partial \Phi_{(y_2=1)}}{\partial \alpha_k} = \Phi_{(y_2=1)} \beta_{2k} + \Phi_{(y_2=1|y_1=1)} \Phi_{(y_1=1)} \beta_{1k} - \Phi_{(y_2=1|y_1=0)} \Phi_{(y_1=0)} \beta_{1k}$$

Se $\rho = 0$ a expressão acima reduz-se a $\Phi_{(y_1=1)} \beta_{1k}$, sendo o mesmo resultado do modelo Probit univariado. O mesmo ocorre para $\Phi_{(y_1=1)}$.

4 RESULTADOS

4.1 Características individuais e familiares dos casais

Os resultados obtidos apontam que o percentual de indivíduos de cor branca é maior no grupo dos ocupados, tanto para Nordeste quanto para Ceará, fato encontrado tanto para maridos quanto esposas. Por exemplo, no Nordeste 26% dos maridos ocupados são de cor branca, já os que se apresentam desocupados são 22,3%. Para o Ceará, ainda no que compete aos maridos, 26,4% dos ocupados e 20,7% dos desocupados são de cor branca.

Os maridos apresentam, em média, idade superior à das mulheres, independente da condição de ocupação e nível regional. Os homens ocupados apresentam cerca de 41 anos de idade no Nordeste. As mulheres, nesta mesma condição e região, apresentam-se com algo em torno 38 anos de idade.

Em média, o nível educacional das esposas é mais elevado do que o dos homens, em todos os níveis de ocupação e região. As mulheres possuem mais de 9 anos de estudo em contraponto aos homens que apresentam cerca de 7,7 anos de estudo. Verifica-se que os homens casados no Ceará têm uma média de anos de estudo maior se comparados aos homens casados de toda a região Nordeste, o mesmo acontece para as mulheres casadas desocupadas.

Ainda acerca das características individuais obteve-se que os maridos são os que apresentam maior percentual de chefes de família. Contudo, é importante destacar que

o número de mulheres que estão passando a ser chefes de família segue uma tendência crescente, de acordo com dados do Censo Demográfico⁶ do IBGE (2010). O estado do Ceará tem maior proporção de esposas chefes da família (35,3%) do que a média da região Nordeste (28,8%).

No que compete às características da família, cerca de 36,5% dos casais do Nordeste apresentam filhos com idade inferior a 14 anos. No Ceará, tem-se um comportamento semelhante, 36% dos casais possuem essa característica.

O tamanho da família aproxima-se de 3,5 membros em ambos os casos. A renda familiar per capita foi de R\$ 813,30 no Nordeste de forma agregada e de R\$ 753,96 quando se considera apenas o estado do Ceará. Em relação a residência dos casais, 27,4% estão em regiões metropolitanas no Nordeste. No Ceará, 57,9% dos casais estão localizados em regiões metropolitanas.

Dado que a diferença do nível educacional se dá apenas de forma marginal e, ainda, que o tamanho médio da família é idêntico em ambos escopos de análise, há algo intrigante no diferencial salarial médio supracitado. A análise não nos permite afirmar se o maior percentual observado de casais residentes em regiões metropolitanas impacta nessa diferença. Pretende-se, portanto, investigar esse achado numa abordagem direcionada/específica futura.

Os resultados relatados acima, além de outros resultados, são concentrados na tabela 1, abaixo.

Tabela 1: Características dos casais segundo condição no mercado de trabalho na Região Nordeste (NE) e Ceará

Características	Marido				Esposa			
	Ocupados (%)		Desocupados (%)		Ocupados (%)		Desocupados (%)	
Individuais	NE	Ceará	NE	Ceará	NE	Ceará	NE	Ceará
Branco	25,99	26,45	22,33	20,72	29,32	30,04	21,03	16,03
Idade (média)	41,04	40,65	37,96	36,03	38,08	38,19	32,66	31,53
Educação (média)	7,85	7,97	7,32	7,75	9,72	9,69	9,19	9,65
Chefe	72,84	69,45	68,06	69,13	28,64	31,91	28,94	38,61
Características da Família	Nordeste				Ceará			
Filhos 14	36,53				36,04			
Tamanho da Família (média)	3,48				3,48			
Renda Familiar (média)	813,30				753,96			
Região metropolitana	27,42				57,95			

Fonte: Elaboração própria a partir dos dados da PNAD de 2015.

Nota: Resultados expandidos para a população.

6. Em 2000, apenas 19,5% das famílias com a presença do cônjuge, eram chefiadas por mulheres, já em 2010 este percentual passou para 46,4%.

4.2 Determinantes da ocupação dos casais no mercado de trabalho

Na tabela 2 são dispostos os resultados do modelo Probit bivariado e os efeitos marginais das variáveis explicativas sobre a participação dos casais no mercado de trabalho urbano no Nordeste e no Ceará, em 2015. O coeficiente de correlação (ρ) entre as estimativas mostrou-se negativo (-0,8206 e -0,7020 para o Nordeste e Ceará, respectivamente) e significativo. Isso indica que há características não observáveis inversamente relacionadas nas equações de participação das esposas e maridos. Além disso, prova que o modelo utilizado é o mais apropriado dado o escopo deste trabalho.

Tabela 2: Probit Bivariado Aparentemente Não Relacionado da Ocupação dos Casais no Nordeste e Ceará - 2015

Marido	Coefficiente	DP	dy/dx	Coefficiente	DP	dy/dx
Branco	-0,0754	0,0800	-0,0090	-0,0174	0,0326	-0,0022
Idade	0,0927***	0,0219	0,0111***	0,0928***	0,0075	0,0119***
Idade2	-0,0013***	0,0002	-0,0002***	-0,0013***	0,00009	-0,0002***
Educação	0,0737***	0,0091	0,0088***	0,0672***	0,0040	0,0086***
Educação do Cônjuge	-7,1550***	0,1250	-0,8558***	-7,4060***	0,0452	-0,9490***
Filhos 14	0,0092	0,0839	0,0011	0,1410***	0,0332	0,0181***
Tamanho da Família	0,0614**	0,0297	0,0073**	0,0285**	0,0122	0,0037**
Mulher Chefe	-6,9800***	0,0837	-0,8350***	-7,4370***	0,0400	-0,9530***
Renda Familiar	0,0006***	0,0002	7,11e-05***	0,0002***	0,00005	3,43e-05***
Região Metropolitana	0,1520**	0,0737	0,0182**	-0,0548**	0,0276	-0,0070**
Constante	-1,5470***	0,4450	-	-1,3120***	0,1610	-
Esposa						
Branco	0,0002	0,0675	3,59e-05	-0,0173	0,0279	-0,0032
Idade	0,0895***	0,0180	0,0162***	0,1250***	0,0068	0,0229***
Idade2	-0,0012***	0,0002	-0,0002***	-0,0016***	0,00008	-0,0003***
Educação	0,1039***	0,0084	0,0188***	0,0993***	0,0032	0,0182***
Educação do Cônjuge	-8,1510***	0,7723	-1,4726***	-7,1280***	0,4000	-1,3060***
Filhos 14	-0,2042***	0,0537	-0,0369***	-0,1020***	0,0289	-0,0187***
Tamanho da Família	-0,0373**	0,0190	-0,0067**	-0,0531***	0,0109	-0,0097***
Mulher Chefe	0,2240***	0,0566	0,0405***	0,2140***	0,0265	0,0392***
Renda Familiar	0,0003***	0,00008	4,85e-05***	0,0002***	0,00003	3,18e-05***
Região Metropolitana	0,0471	0,0757	0,0085	-0,0605**	0,0247	-0,0111**
Constante	-2,5460***	0,3490	-	-3,0770***	0,1320	-
athrho	-1,1590***	-3,380	-	-0,8710***	0,1070	-
ρ	-0,8206***	0,1118	-	-0,7020***	0,0540	-

Fonte: Elaboração própria a partir dos dados da PNAD de 2015.

Notas: Níveis de significância: ***1%, **5% e *10%. DP=Desvio-Padrão.

Os resultados apontam que a cor/raça dos casais (marido e esposa) não influencia na probabilidade de oferta de trabalho dos mesmos. Fato constatado pela análise do coeficiente da variável 'Branco'. Resultado intuitivo, uma vez que a oferta de trabalho é necessidade intrínseca da grande maioria das pessoas.

Constata-se ainda que a medida que os casais envelhecem as chances de estarem ocupados aumentam, porém a taxas decrescentes. Isso ocorre tanto para a região Nordeste quanto para o estado do Ceará, expresso pelo efeito positivo variável 'Idade'.

Em todas as estimativas, a educação apresentou coeficiente positivo e significativo. Isso indica que quanto maior o nível de escolaridade maior é a probabilidade de participação no mercado de trabalho tanto do marido quanto da esposa. No entanto, o efeito da educação foi maior para as esposas, 1,9% na região Nordeste e 1,8% no Ceará contra, aproximadamente, 0,8% para os maridos. Este resultado corrobora com o achado de Guimarães e Santos (2010).

Em relação a educação do cônjuge, os resultados apontam que quanto maior a escolaridade do cônjuge menor a chance de participar do mercado de trabalho. Na região Nordeste uma maior educação da esposa reduz a probabilidade de participação do marido no mercado em 85,6% enquanto que no Ceará esse resultado alcança 95%. Já para as esposas, uma maior educação do marido tem impacto negativo para inserção feminina ainda maior do que a do caso anterior (147% e 131% para o Nordeste e Ceará, respectivamente).

O resultado referente às mulheres apontado no parágrafo anterior é destacado também por Soares e Izaki (2002). Segundo os autores, devido a escolaridade do marido ter impacto negativo sobre a inserção feminina, isso mostra indícios de trabalho secundário, onde a oferta de trabalho das esposas depende das características familiares.

No caso das variáveis relacionadas a família, a presença de filhos menores de 14 anos aumenta a probabilidade de ocupação dos maridos. Já para as esposas o contrário é observado, com maior efeito para aquelas que residem no Nordeste, 3,7% em comparação com 1,9% das residentes do estado do Ceará. Assim como destacado por Gomes (2010), Smock (2000), Sedlacek e Santos (1991), mulheres com filhos têm menos chance de estar na força de trabalho, pois, a maioria é responsável pelos cuidados com os filhos.

Assim como a variável anterior, o número de componentes da família impacta positivamente sobre a probabilidade de ocupação dos maridos. Enquanto que esta mesma variável diminui as chances de inserção da mulher no mercado. Este achado também está relacionado a maior presença de filhos em famílias maiores.

O fato de a mulher ser chefe da família reduz as chances de inserção do marido em 83,5% na região Nordeste e 95,3% no Ceará. Quando a mulher é a pessoa de referência na família, tanto no Nordeste quanto no Ceará, tem-se um aumento de cerca de 4% na sua chance de ocupação no mercado.

A renda familiar *per capita* aumenta as chances de ocupação tanto dos maridos como das esposas, sendo que a renda do trabalhador está incluída nessa variável. Portanto, quanto maior a renda maior a chance de participação no mercado de trabalho dos casais para manter o padrão de vida alcançado⁷.

Para finalizar a análise da tabela 2, obteve-se que residir em região metropolitana do Nordeste mostrou-se positiva e significativa para participação dos maridos, assim como para o estado do Ceará. Por outro lado, não se pode apontar efeito para o caso das esposas na mesma região. Já para o estado do Ceará a região metropolitana reduz a probabilidade de ocupação das mulheres em 1,1%.

4.3 Análise dos Efeitos Marginais das Probabilidades Bivariadas dos Casais

Agora, consideremos as tabelas 3 e 4, dispostas ao final desta subseção, as quais fornecem os efeitos marginais de cada variável sobre a decisão conjunta dos casais de participar do mercado de trabalho no Nordeste e Ceará, respectivamente. As categorias referentes às famílias são classificadas como quando apenas um cônjuge trabalha, sendo o único provedor (Mulher/Homem); quando ambos participam do mercado de trabalho; e, por último, o caso em que nenhum se encontra ocupado.

A idade impacta positivamente na probabilidade tanto de que apenas um membro seja provedor, mulher ou homem, quanto de os dois cônjuges estarem inseridos no mercado de trabalho. Este resultado pode estar relacionado a maior experiência no mercado. Pois, para Menezes, Fernandez e Dedecca (2005) a experiência tem papel fundamental na inserção no mercado de trabalho, principalmente no caso das mulheres, com mais importância até do que a escolaridade.

Quanto ao nível educacional da esposa, obteve-se que um ano a mais de estudo amplia as chances de a mulher ser a única provedora. Por outro lado, um ano a mais de estudo do marido reduz a probabilidade de a família dispor apenas da contribuição monetária feminina em 126% e 145% para o Nordeste e Ceará, respectivamente.

Nos casos em que a mulher é a única provedora, a fecundidade reduz a chances dessa configuração familiar. Essa evidência atua no sentido de reafirmar que o cuidado com os

7. Ver, por exemplo, Menezes, Fernandez e Dedecca (2005) e Figueiredo, Marta e Guimarães (2010).

filhos ainda está muito vinculado a figura feminina. Resultado intuitivo, uma vez que os primeiros cuidados em quase totalidade das vezes são insubstituíveis e que as mães não abrem mão de tal responsabilidade e vivência. Além do fato de que, uma vez inserida no mercado de trabalho, nem sempre é possível um afastamento do trabalho com manutenção do vínculo empregatício. Devido ao mesmo motivo supracitado e/ou intensificado por ele, quanto maior o tamanho da família menor é a probabilidade de a mulher estar ativa no mercado de trabalho.

Com relação ao fato de a mulher ser chefe da família (pessoa de referência), essa configuração aumenta as chances de ser a única provedora na família. O contrário é observado para o caso de ter apenas o homem como provedor.

O aumento da renda familiar *per capita* aumenta as chances de famílias com duplo rendimento (ambos trabalham), além das participações da mulher ou do homem como únicos provedores. Portanto, as mulheres muitas vezes não só contribuem com parte significativa da renda familiar como também podem ser a única provedora, tal como destaca Marri e Wajnman (2007).

Por fim, residir em Região Metropolitana na região Nordeste, assim como no modelo Probit bivariado, reduz a probabilidade de oferta. No entanto, residir na região metropolitana do estado do Ceará aumenta as chances de o homem ser o único provedor.

Tabela 3: Efeitos marginais obtidos com base nas probabilidades bivariadas para o Nordeste em 2015

Variável	Mulher Provedora	Homem Provedor	Ambos trabalham	Nenhum trabalha
Branco	-0,00301	-0,00205	-0,00016	0,00523
Idade	0,02180***	0,01081***	0,00108***	-0,03370***
Idade2	-0,00027***	-0,00015***	-0,00001***	0,00044***
Educação da Esposa	0,02919***	-0,9199***	-0,0292***	0,9200***
Educação do Marido	-1,2650***	0,04075**	-0,04073***	1,2649***
Filhos 14	-0,01867***	0,01811***	-0,0003***	0,00058
Mulher Chefe	0,06730***	-0,92511***	-0,02811***	0,8860***
Tamanho da Família	-0,00952***	0,00385**	-0,00019**	0,00587**
Renda Familiar	0,00002**	0,00003***	0,00002***	-0,00006***
Região Metropolitana	-0,01052***	-0,0064*	-0,00056**	0,01750***

Fonte: Elaboração própria a partir dos dados da PNAD de 2015.

Notas: Níveis de significância: ***1%, **5% e *10%.

Tabela 4: Efeitos marginais obtidos com base nas probabilidades bivariadas para o Ceará em 2015

Variável	Mulher Provedora	Homem Provedor	Ambos trabalham	Nenhum trabalha
Branco	0,00019	-0,00886	-0,00015	0,00883
Idade	0,01570***	0,01064***	0,00044**	-0,02680***
Idade2	-0,00021***	-0,00015***	-0,00001**	0,00037***
Educação da Esposa	0,01440***	-0,84142***	-0,01440**	0,8410***
Educação do Marido	-1,4500**	0,02297**	-0,02290*	1,4498***
Filhos 14	-0,03640***	0,00165	-0,00055*	0,03530**
Mulher Chefe	0,05390***	-0,82147***	-0,01350***	0,7810***
Tamanho da Família	-0,00676**	0,00732**	0,00002	-0,00059
Renda Familiar	0,00005***	0,00006***	0,000002**	-0,00012***
Região Metropolitana	0,00807	0,01771**	0,00044	-0,02620

Fonte: Elaboração própria a partir dos dados da PNAD de 2015.

Notas: Níveis de significância: ***1%, **5% e *10%.

5 CONSIDERAÇÕES FINAIS

Na maioria dos países desenvolvidos um número considerável de famílias possui uma configuração na qual ambos os cônjuges contribuem para a renda familiar e tomam decisões conjuntamente de entrada e saída no mercado de trabalho.

A partir disso e da detecção recente das mudanças observadas com relação ao mercado de trabalho dos casais brasileiros, de seguirem a tendência desses países, este estudo objetivou identificar os determinantes das decisões de entrada dos casais no mercado de trabalho. Especificamente nas regiões urbanas do estado do Ceará e da Região Nordeste.

Para tanto, utilizou-se um modelo Probit bivariado aparentemente não relacionado, com dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2015, considerando o plano amostral. Além disso, analisou-se tanto as características dos casais segundo condição no mercado de trabalho para o escopo definido acima quanto os efeitos marginais obtidos com base nas probabilidades bivariadas para o Nordeste e para o Ceará.

Outros resultados mostram que a medida que a idade aumenta as chances de estarem os casais ocupados também aumenta. Isso segue a literatura da economia do trabalho, pois a idade representa uma *proxy* para experiência. Outro exemplo é a educação, quanto maior o nível de escolaridade maior é a probabilidade de participação no mercado de trabalho de ambos os cônjuges. Frisa-se que esse efeito foi maior para as esposas.

Em relação a educação do cônjuge, os resultados apontam que quanto maior a escolaridade do cônjuge menor a chance de participar do mercado de trabalho. Na região Nordeste uma maior educação da esposa reduz a probabilidade de participação do marido no mercado em 85,6% enquanto que no Ceará esse resultado alcança 95%. Já para as esposas, uma maior educação do marido tem impacto negativo para inserção feminina ainda maior do que a do caso anterior (147% e 131% para o Nordeste e Ceará, respectivamente).

Obtiveram-se, ainda, resultados interligados. Ao que diz respeito a família, por exemplo, se a mulher é a única provedora de renda na família, a fecundidade reduz a chances dessa configuração familiar. Diretamente ligado a isso, tem-se que a presença de filhos que aumenta a probabilidade de ocupação dos maridos.

Espera-se que os resultados encontrados auxiliem, de alguma forma, no entendimento desse mercado de trabalho específico/conjugado e, conseqüentemente, em análises governamentais para desenho e/ou adequação de políticas públicas específicas/direcionadas. Pode-se, ainda, gerar uma visão alternativa/relativa sobre o desemprego, uma análise do desemprego por unidade de habitação, por exemplo.

Para estudos futuros, pretende-se investigar a evolução dessa participação dos casais no mercado de trabalho nos estados da região Nordeste. Além disso, averiguar as configurações familiares com relação as estratégias de mulher/homem como único provedor e inserção de ambos no mercado de trabalho.

REFERÊNCIAS

- BATISTA, N. F. Evolução da distribuição dos casais brasileiros de acordo com a geração da renda familiar por gênero. *In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, Anais... ABEP*, p. 1-25, 2016.
- BLAU, Francine D.; KAHN, Lawrence M. Changes in the labor supply behavior of married women: 1980 – 2000. **Journal of Labor Economics**, v. 25, n. 3, p. 393-438, 2007.
- BECKER, G. S. **A Treatise on the Family**, 2nd edition, Harvard University Press, Cambridge, 1991.
- CARVALHO A. P.; NÉRI, M. C.; SILVA, D. B. do N. Diferenciais de salários por raça e gênero no Brasil: aplicação dos procedimentos de Oaxaca e Heckman em pesquisas amostrais complexas. *In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS*, 15, 2006, Caxambú-MG. **Anais... Minas Gerais: ABEP**, 2006.
- DA SILVA, C.; DA CUNHA, M. S. Oferta de trabalho de casais nas grandes regiões metropolitanas brasileira de 2002 a 2015: evolução e determinantes. XXII Encontro Regional de Economia. **Anais...**, 2017.
- DEUTSCH, F. M.; ROKSA, J.; MEESKE, C. How gender counts when couples count their money. **Sex Roles: a Journal of Research**, Vol. 48, Nos 7/8, 2003.
- FIGUEIREDO, A. M. R.; MARTA, J. M. C.; GUIMARÃES, P. W. Determinantes da ocupação do núcleo familiar composto na área rural do Brasil. **Economia e Desenvolvimento**, n. 22, 2010.
- GUIMARÃES, P. W.; SANTOS, C. M. Determinantes da ocupação no mercado de trabalho de maridos e esposas. **Revista Brasileira de Gestão e Desenvolvimento Regional**, v. 6, n. 2, 2010.
- GOMES, C. F. **Diferencial de salário por ocupação e atividade das mulheres brasileiras de acordo com seu estado conjugal**. 2010. Monografia (Ciências Econômicas) – Instituto de Ensino e Pesquisa – Insper, São Paulo, 2010.
- GREENE, W. H. **Econometric analysis**. Boston: Pearson, 7a. ed. 2012.
- HOFFMANN, R.; LEONE, E. T. Participação da mulher no mercado de trabalho e desigualdade da renda domiciliar per capita no Brasil: 1981 2002. **Nova economia**, v. 14, n. 2, 2009.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. Censo Demográfico 2010 – Famílias e Domicílios: Resultados da Amostra. Disponível em: <http://biblioteca.ibge.gov.br/visualizacao/periodicos/97/cd_2010_familias_domicilios_amostra.pdf>. Acesso em: 20 de agosto de 2017.
- JATOBÁ, J. A família na força de trabalho: Brasil metropolitano-1978 1986. Anais do VII Encontro Nacional de Estudos Populacionais, ABEP. **Anais...**, p. 147-173, 2016.

MADALOZZO, R.; MARTINS, S. R.; SHIRATORI, L. Participação no mercado de trabalho e no trabalho doméstico: homens e mulheres têm condições iguais?. **Estudos Feministas**, p. 547-566, 2010.

MARRI, I. G.; WAJNMAN, S. Esposas como principais provedoras de renda familiar. **Revista Brasileira de Estudos de População**, v. 24, n. 1, p. 19-35, 2007.

MENEZES, W. F.; FERNANDEZ, J. C.; DEDECCA, C. Diferenciações regionais de rendimentos do trabalho: uma análise das regiões metropolitanas de São Paulo e de Salvador. **Estudos Econômicos**, v.35, n.2, p.271-296, Abr./Jun.2005.

OSHITA, M. G. B.; PERETTI FILHO, V.; SIMÃO, C. H. M.; DA CUNHA, M. S. A oferta de trabalho dos casais: uma análise da transição da desocupação para o emprego no brasil metropolitano de 2003 a 2015. XX Encontro de economia da região Sul. **Anais...**, 2017.

SEDLACEK, G; SANTOS, E. A mulher cônjuge no mercado de trabalho como estratégia de geração de renda familiar. **Pesquisa e Planejamento Econômico**, v. 21, n. 3, p. 449-470, 1991.

SMOCK, P.J. Cohabitation in the United States: an appraisal of research themes, findings, and implications. **Annual Review of Sociology**, Palo Alto, CA, v.26, p.1-20, 2000.

SILVA, W. R., & SANTOS, D. D. (2015). Trabalho e bem-estar: uma comparação entre casais heterossexuais e homoafetivos brasileiros. In: 43º ENCONTRO NACIONAL DE ECONOMIA. **Anais...**, (p. 20). Costão do Santinho, SC.

SOARES, S.; IZAKI, R. S. **A participação feminina no mercado de trabalho**. Rio de Janeiro: IPEA, 2002, 22p. (Texto para discussão, n. 923).

SPLETZER, J. R. Reexamining the added worker effect. **Economic Inquiry**, v. 35, n. 2, p. 417-427, 1997.