



GOVERNO DO  
ESTADO DO CEARÁ  
Secretaria do Planejamento  
e Gestão

# ECONOMIA do CEARÁ em DEBATE 2015



GOVERNO DO  
ESTADO DO CEARÁ  
*Secretaria do Planejamento  
e Gestão*

# **ECONOMIA do CEARÁ em DEBATE 2015**

## **GOVERNO DO ESTADO DO CEARÁ**

Camilo Sobreira de Santana – Governador

Maria Izolda Cela – Vice Governadora

## **SECRETARIO DO PLANEJAMENTO E GESTÃO - SEPLAG**

Hugo Santana de Figueirêdo Junior – Secretário

## **INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ (IPECE)**

Flávio Ataliba F. D. Barreto – Diretor Geral

Adriano Sarquis B. de Menezes – Diretor de Estudos Econômicos

Décio N. Chaves de Assis (Gerente de Estatística do IPECE)

### **Organizadores**

Flávio Ataliba Flexa Daltro Barreto

Adriano Sarquis B. de Menezes

Décio N. Chaves de Assis

Fátima Juvenal de Sousa

### **Conselho Editorial**

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

Adriano Sarquis Bezerra de Menezes (IPECE / ETENE-BNB)

Almir Bittencourt da Silva (CAEN-UFC)

Marcelo Lettieri Siqueira (UFC / Receita Federal)

Marcos Sena (UNIFOR)

#### ENCONTRO DE ECONOMIA CEARÁ EM DEBATE 2015

v-1 - 2015

Flávio Ataliba Flexa Daltro Barreto, Adriano Sarquis B. de Menezes,  
Décio N. Chaves de Assis, Fátima Juvenal de Sousa (organizadores).

Fortaleza: IPECE, 2015. 390 p.

ISBN: 978-85-98664-34-7

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

II - Título

CDU 330 (813.1) Copyright © 2016 - IPECE - *Impresso no Brasil*

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.

INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ (IPECE)  
Av. Gal. Afonso Albuquerque Lima, s/nº - Edifício SEPLAG - 2º Andar  
Centro Administrativo Governador Virgílio Távora – Cambéba CEP: 60830-120 – Fortaleza-CE.  
Tel. (85) 3101-3496 - ouvidoria@ipece.ce.gov.br - www.ipece.ce.gov.br

## APRESENTAÇÃO

O livro Economia do Ceará em Debate 2015 é uma coletânea dos quinze artigos apresentados por ocasião do XI Encontro Economia do Ceará em Debate, realizado pelo Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE), no dia 08 de dezembro de 2015, no auditório da Secretaria do Planejamento e Gestão (SEPLAG). A seleção dos artigos coube à banca formada pelos professores doutores e pesquisadores Almir Bittencourt (CAEN-UFC); Marcos Sena (UNIFOR) e Marcelo Lettieri (UFC/Receita Federal), aos quais, em nome do IPECE, registramos os mais sinceros agradecimentos. Participaram do Encontro e contribuíram assim para o saudável debate os presidentes das mesas, analistas Paulo Pontes (IPECE) e Nicolino Trompieri (IPECE). Coube aos professores Flávio Ataliba (IPECE) e Ricardo Paes de Barros (INSPER) presidir as mesas temáticas respectivamente: MESA 1 - A IMPORTÂNCIA DA GESTÃO E DO DESENHO DE POLÍTICAS PÚBLICAS EFICIENTES PARA O DESENVOLVIMENTO ECONÔMICO – Moderador: Emerson Marinho (CAEN-UFC) e Mesa 2 – O DESAFIO DO DESENVOLVIMENTO INCLUSIVO NO BRASIL E OS AVANÇOS RECENTES NO CEARÁ – Moderador: Hugo Figueiredo (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. 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 dos artigos que lograram os três primeiros lugares, aqui publicados, e a Associação dos Auditores de Controle Interno do Estado do Ceará (AACI). Todos colaboraram e contribuíram com sua participação para o sucesso do evento, cujos textos apresentados permitiram a publicação deste livro, que vem com reflexões valiosas para o entendimento de diversas dimensões da economia do Ceará.

Flávio Ataliba Flexa Daltro Barreto

**Diretor Geral do**

**Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE)**

## SUMÁRIO

### **VIOLÊNCIA, DESENVOLVIMENTO E DEMOGRAFIA: UMA ANÁLISE ESPACIAL PARA A CIDADE DE FORTALEZA EM ANOS RECENTES**

Cleyber Nascimento de Medeiros, José Raimundo Carvalho, Victor Hugo de Oliveira

8

### **EXTREMA POBREZA INFANTIL, DESIGUALDADE E CRESCIMENTO NOS MUNICÍPIOS CEARENSES**

Dércio Nonato Chaves de Assis, Cleyber Nascimento de Medeiros, Cláudio André Gondim Nogueira

42

### **DESEMPENHO ESCOLAR E INSERÇÃO NO MERCADO DE TRABALHO: UMA AVALIAÇÃO DAS ESCOLAS ESTADUAL DE ENSINO PROFISSIONALIZANTES (EEEP) DO CEARÁ**

Francisca Zilania Mariano, Ronaldo A. Arraes, Natália de Olivindo Souza

75

### **A PERDA DE BEM ESTAR SOCIAL DA POPULAÇÃO DE JUAZEIRO DO NORTE COM O DESPERDÍCIO DE ÁGUA POTÁVEL**

Joelice Alexandre de Oliveira, Wellington Ribeiro Justo

108

### **ANÁLISE COMPARATIVA DA ESPECIALIZAÇÃO DA AGROPECUÁRIA NORDESTINA E CEARENSE A PARTIR DO ÍNDICE DE CONCENTRAÇÃO NORMALIZADO (ICN)**

Maria Josiell Nascimento da Silva, Maria Daniele Cruz dos Santos, Otácio Pereira Gomes, Camila Pereira Brígido Rodrigues

140

### **ANÁLISE DO IMPACTO ECONÔMICO E SOCIAL DA ÁREA LIVRE DE MOSCA DAS FRUTAS NO ESTADO DO CEARÁ**

Eliane Pinheiro de Sousa, Sílvia Helena Galvão de Miranda

165

**ANÁLISE DO PERFIL SOCIOECONÔMICO E TECNOLÓGICO DOS PRODUTORES DE FRUTICULTURA IRRIGADA NA REGIÃO SUL CEARENSE**

**Otácio Pereira Gomes, Kilmer Coelho Campos, Wellington Ribeiro Justo, Guillermo Gamarra Rojas, Maria Daniele Cruz dos Santos**

.....192

**AVALIAÇÃO DE FATORES SOCIOECONÔMICOS E COMPORTAMENTAIS NOS RESULTADOS DOS ALUNOS NA PROVA BRASIL DE 2011 NOS MUNICÍPIOS CEARENSES**

**Pedro Fernando Damasceno Rocha, Francisco José Silva Tabosa, Ricardo Brito Soares, Andrei Gomes Simonassi**

.....216

**COMPORTAMENTO RELIGIOSO DA POPULAÇÃO CEARENSE A PARTIR DE UMA PERSPECTIVA ECONÔMICA E SOCIAL**

**Eveline Carvalho, Luan Falcão, Guilherme Irffi**

.....243

**DECOMPOSIÇÃO DOS DIFERENCIAIS DE RENDIMENTOS POR COR E GÊNERO NO MERCADO DE TRABALHO BRASILEIRO: UMA ANÁLISE POR NÍVEIS OCUPACIONAIS NAS REGIÕES METROPOLITANAS**

**Edward Martins Costa, Daniel Tomaz de Sousa, Daniel Barbosa Guimarães**

.....261

**ESTIMANDO RETORNOS DA EDUCAÇÃO NO CEARÁ: UMA APLICAÇÃO DO MÉTODO DO LASSO PARA SELEÇÃO DE VARIÁVEIS INSTRUMENTAIS FRACAS**

**Rafael B. Barbosa, Daniel Barboza Guimarães**

.....287

**GERAÇÃO CANGURU? FATORES ASSOCIADOS À PERMANÊNCIA DOS JOVENS CEARENSES NO AMBIENTE FAMILIAR DE ORIGEM**

**Juliane da Silva Ciríaco, Otoniel Rodrigues dos Anjos Júnior, Priscila Silva Rodrigues, Nárdia Costa Alves**

.....308

**IMPACTOS DAS RENDAS NÃO-AGRÍCOLAS SOBRE OS INDICADORES DE POBREZA FOSTER-GREER-THORBECKE (FGT) PARA AS FAMÍLIAS RURAIS DO ESTADO DE CEARÁ**

**Alan Francisco Carvalho Pereira, Wellington Ribeiro Justo, João Ricardo Ferreira de Lima**

.....325

**MODELOS ECONOMÉTRICOS PARA PREVISÃO DA INFLAÇÃO NA REGIÃO METROPOLITANA DE FORTALEZA**

**Elano Ferreira Arruda, Antônio Clécio de Brito**

.....349

**PRINCIPAIS FATORES QUE IMPACTAM NA REPROVAÇÃO E EVASÃO DOS ALUNOS DOS CURSOS TECNOLÓGICOS NO INSTITUTO FEDERAL DE CIÊNCIA E TECNOLOGIA DO CEARÁ**

**José Nailton Martins de Sousa, Francisco José Tabosa, Andrei Simonassi, Pablo Urano de Carvalho Castelar**

.....369

# **VIOLÊNCIA, DESENVOLVIMENTO E DEMOGRAFIA: UMA ANÁLISE ESPACIAL PARA A CIDADE DE FORTALEZA EM ANOS RECENTES**

Cleyber Nascimento de Medeiros\*

José Raimundo Carvalho\*\*

Victor Hugo de Oliveira\*\*\*

## **Resumo**

O presente estudo teve por objetivo investigar a relação entre violência, demografia e desenvolvimento humano (IDH) em Fortaleza. A partir de dados georreferenciados de crimes violentos letais e intencionais (CVLI), foi possível calcular a incidência de tal fenômeno social para as 247 Unidades de Desenvolvimento Humano (UDH) da capital. A análise exploratória dos dados revela que os crimes violentos são espacialmente concentrados em Fortaleza, onde UDHs com elevada taxa de CVLI possuem vizinhos com a mesma característica. Entre 2012 e 2014, observou-se também um processo de difusão dos crimes violentos no território a partir das áreas de maior incidência. Diante de tais evidências, buscou-se averiguar se a incidência de crimes violentos estaria condicionada ao nível de desenvolvimento humano da população, uma vez que áreas com baixo (alto) IDH apresentam taxas elevadas (baixas) de CVLI. As estimações econométricas incorporando a dependência espacial, indicam que o incremento de 0,1 scores no IDH reduz, em média, 0,54% a taxa de CVLI na capital cearense. A demografia também desempenha papel relevante na determinação das taxas de crimes violentos em Fortaleza, uma vez que a densidade demográfica afeta de maneira não linear a taxa de CVLI e a população que vive

---

\*Doutor em Geografia (Universidade Estadual do Ceará - UECE). Filiação Institucional: Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE). Endereço: Av. General Afonso Albuquerque Lima, S/N, 2º andar, Edifício SEPLAG. Cambéa/Fortaleza/Ceará. CEP: 60.839-900. E-mail: cleyber.medeiros@ipece.ce.gov.br.

\*\* PhD em Economia (Penn State University). Filiação Institucional: Curso de Pós-Graduação em Economia – CAEN/UFC, Laboratório de Econometria e Otimização (LECO/CAEN/UFC). Endereço: Av. da Universidade, 2.700, 2º Andar, Benfica. E-mail: josecarv@ufc.br.

\*\*\* PhD em Economia (Universidad de Alicante - Espanha). Filiação Institucional: Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE), Universidade de Fortaleza (UNIFOR), Laboratório de Econometria e Otimização (LECO/CAEN/UFC). Endereço: Av. General Afonso Albuquerque Lima, S/N, 2º andar, Edifício SEPLAG. Cambéa/Fortaleza/Ceará. CEP: 60.839-900. E-mail: victor.hugo@ipece.ce.gov.br.



em condomínios ou apartamentos têm uma menor incidência de crimes violentos.

**Palavras-chave:** Violência, Desenvolvimento Humano, Demografia, Dependência Espacial, Fortaleza.

## **Abstract**

The current study aims to investigate the relationship between violence, demography and human development (HDI) in Fortaleza. Using georeferenced data of lethal violent crimes, it was possible to obtain the incidence of such social phenomenon for all 247 unities of human development (HDU) of the capital. The exploratory analysis of the data suggests that the violent crimes are spatially concentrated in Fortaleza, where the HDUs with high violent crime rate have neighbors with the same characteristic. From 2012 to 2014, there was a spatial diffusion of violent crimes from places with high incidence. Facing such evidence, we sought to determine whether the incidence of violent crime is conditioned to the level of human development of the population, since areas with low (high) HDI have high rates (low) of CVLI. The estimations from spatial models indicate that an increase of 0.1 scores in the HDI decreases, in average, the violent crime rate in 0.54% in the city. Moreover, the demographic aspects of the areas also play a role in the determination of violent crime rates, once the population density has non-linear effects on the CVLI rate and population living in condominiums or apartments has less violent crime rates.

**Key words:** Violence, Human Development, Demography, Spatial Dependence, Fortaleza.

## **1 INTRODUÇÃO**

Para se alcançar o desenvolvimento é necessário que se removam as principais fontes de privação da liberdade em uma sociedade (SEN, 1999). A violência é, sem dúvidas, um importante entrave para tal fim. Além de atentar contra a vida, a violência inibe o exercício pleno da cidadania (WHELLER, 2014), reduz a competitividade e o ambiente para negócios (GOLDBERG et al., 2014), dissuade o investimento e impede o crescimento econômico (LOUREIRO e SILVA, 2010). Em resumo, a violência perpetua um ambiente de incerteza e medo na sociedade (UNDOC, 2014).

Fortaleza tornou-se, recentemente, a capital brasileira mais violenta, onde as estatísticas de crimes violentos letais e intencionais (CVLI) superam, em termos absolutos e relativos, as de cidades como São Paulo e Rio de Janeiro<sup>1</sup>. Compreender seus condicionantes é de extrema relevância para a formação de políticas públicas mais eficazes no combate a tal fenômeno social e econômico. Particularmente, o presente estudo objetiva investigar como o desenvolvimento humano e a demografia se relacionam com a violência em Fortaleza, sendo esta a principal contribuição desta pesquisa.

Este trabalho possui ainda duas contribuições adicionais para a literatura. Uma das referidas contribuições diz respeito ao uso de duas bases de dados inéditas em estudos empíricos sobre violência no Brasil. A primeira delas é à base de dados georreferenciada de crimes violentos letais e intencionais<sup>2</sup> na capital cearense, permitindo explorar aspectos espaciais relacionados à distribuição das ocorrências. A segunda base de dados consiste no Atlas de Desenvolvimento Humano das Regiões Metropolitanas, cujos indicadores foram calculados para *cluters* de setores censitários homogêneos do ponto de vista socioeconômico.

Na literatura econômica, o estudo clássico de Becker (1968) sugere que os incentivos econômicos são parte fundamental para o engajamento do indivíduo na atividade criminal em detrimento ao mercado de trabalho, uma vez que a decisão de cometer um crime é função das expectativas de benefícios e custos envolvidos. Nessa perspectiva, poder-se-ia esperar uma correlação positiva entre desenvolvimento e criminalidade, uma vez que em sociedades desenvolvidas o ganho líquido da criminalidade seria muito pequeno em comparação a sociedades pobres ou em vias de desenvolvimento.

Todavia, Ehrlich (1973) sugere que o desenvolvimento, neste caso, representado pelo nível de renda, possui efeito indeterminado sobre a criminalidade, pois a relação entre a aversão ao risco e a renda pode influenciar a direção do efeito<sup>3</sup>. Além da aversão ao risco, a criminalidade se relaciona com variáveis que passam por profundas mudanças durante o processo de desenvolvimento econômico tal como a distribuição de renda, urbanização, renda *per capita*, desenvolvimento institucional (SOARES, 2004).

---

1. Segundo dados do Anuário Brasileiro de Segurança Pública 2015.

2. A taxa de CVLI por 100.000 habitantes é um dos indicadores mais adequados para mensurar a violência em virtude de sua comparabilidade e precisão, fornecendo informações relevantes como o local de ocorrência e a intensidade do fenômeno (UNDOC, 2014).

3. Estudos subsequentes confirmam a ambiguidade na relação entre desenvolvimento e crime (BENOIT e OSBORNE, 1995; ZHANG, 1997; IMROHOROGLU et al., 2000).

No presente estudo, o Índice de Desenvolvimento Humano (IDH) é a principal medida de desenvolvimento, pois captura não somente a dimensão renda como também as duas principais dimensões do capital humano: saúde e educação.

Não menos importante é a desigualdade de renda, uma vez que tal fenômeno econômico pode ser interpretado como a distância entre os ganhos do crime e custos de oportunidades entre ricos e pobres (EHRlich, 1973; FAJNZYLBER et al., 2002).

A demografia também influencia as taxas de crimes violentos contra a vida (O'Flaherty e Sethi, 2015). Áreas com alta densidade populacional estão positivamente associadas a um maior número potencial de vítimas, proporcionando um alto retorno para crimes patrimoniais (Glaeser e SACERDOTE, 1999).

No entanto, o comportamento agressivo pode ser menor em ambientes densamente povoados em virtude da crescente interação social entre os indivíduos<sup>4</sup>. Cita-se ainda que em níveis elevados de densidade populacional pode haver, por exemplo, disputa por recursos escassos, fomentando uma relação positiva entre densidade demográfica e violência. Neste caso, os benefícios sociais de uma maior densidade populacional em certa área pode estar limitada a determinado nível, sugerindo uma potencial relação não-linear com a violência (REGOECZI, 2002).

Outro aspecto relevante da demografia é a capacidade de monitoramento da população quanto ao crime. A Hipótese de Jacob (1961) postula que o adensamento populacional em determinado espaço pode inibir o comportamento criminoso em áreas censitárias, em virtude de um potencial efeito vigilância (CHRISTENS e SPEER, 2005).

Tal hipótese tem sido testada através do tipo de moradia predominante na localidade (por exemplo, moradias horizontais vis-à-vis moradias verticais). Faria et al. (2013) encontram que a concentração das taxas de crimes, especialmente os patrimoniais, em Brasília (no plano piloto), está associada à áreas com elevada proporção de moradias verticais, dentre outras características urbanas da capital brasileira.

Ademais, a maioria dos estudos nacionais que testam os determinantes das taxas de crimes com base no modelo de racionalidade microeconômica de Becker utilizam dados agregados regionalmente em virtude da indisponibilidade

4. Glaeser et al. (1996) mostram que o comportamento agressivo pode ser influenciado pelas interações sociais experimentadas pelo indivíduo.

de de dados em nível de indivíduos. Esse fato cria um problema de agregação, pois quanto maior a unidade geográfica maior será a perda de informação relativa ao problema econômico em si. Por outro lado, quanto menor a unidade geográfica maior será a possibilidade de influência da vizinhança sobre a violência em determinado local. Nesse contexto, por exemplo, a hipótese de que o criminoso reside na mesma área em que vive perde força (CERQUEIRA e LOBÃO, 2004). Logo, a dependência espacial não deve ser negligenciada em análise dos determinantes da violência, sob pena de gerar resultados viesados.

Existe uma extensa literatura em Sociologia e em Criminologia com particular interesse em explicar as razões da concentração espacial das ocorrências criminais. Anselin et al. (2000) apresenta diversos estudos empíricos que sugerem o espaço como elemento crucial na análise dos determinantes da criminalidade<sup>5</sup>.

Um exemplo de fenômeno social capaz de moldar a criminalidade em determinado espaço é a formação de gangs de rua. Este fenômeno tem sido associado a áreas de alta incidência criminal (TITA et al., 2005).

Uma das principais razões é o fato das gangs de rua se mostrarem bastante territoriais (TITA e RIDGEWAY, 2007). Além disso, a formação de gangs promove a difusão de crimes violentos (COHEN e TITA, 1999), em especial, quando elas estão associadas ao mercado de drogas (TITA e COHEN, 2004). A difusão também ocorre por meio da formação de redes sociais entre gangs rivais, onde a violência é perpetrada por, e contra, seus membros (TITA e RADIL, 2011).

As características urbanas e da atividade local são outros mecanismos que podem influenciar a distribuição espacial da violência. Os tipos de construções urbanas podem limitar a capacidade de controle social, por exemplo, da força policial. As atividades de rotina realizadas também podem favorecer a criminalidade. Este é o caso de áreas comerciais e áreas onde o consumo de álcool é elevado (ANSELIN et, al. 2000).

Além disso, há evidências que relacionam a precarização da infraestrutura local a uma maior incidência de criminalidade. Por exemplo, vizinhanças com edifícios abandonados chegam a apresentar índices de criminalidade (consumo de drogas ilícitas, crimes patrimoniais, etc.) duas vezes maior do que vizinhanças sem infraestrutura precária (SPELMAN, 1993). A precária infraestrutura

---

5. Não se pretende discutir neste estudo as diferentes linhas de pensamento referente à relação entre espaço e crime. O intuito é o de apenas justificar o uso de modelos econométricos que incorporam os potenciais efeitos espaciais sobre as ocorrências criminais.

local também atraem os comerciantes de drogas que buscam locais estáveis para sua atividade ilícita (ECK, 1994).

Outro aspecto relevante da espacialização da violência refere-se ao fato de que a polícia reduz crime quando a alocação de seus recursos é realizada com base nos “*hotspots*” de ocorrências criminais (WEISBURD e ECK, 2004). Entretanto, evidências têm mostrado que tal abordagem traz consigo um risco de migração da violência e/ou da desordem para áreas vizinhas que não eram originalmente objetivo da intervenção policial (REPPETTO, 1976). Nestes casos, a migração da violência é inevitável, porém possui um escopo limitado (HESSSELING, 1994). Por outro lado, estudos mostram que políticas de prevenção ao crime orientadas para determinada área (ou “*hotspot*”) resulta em benefícios para as áreas vizinhas em virtude do controle da difusão da violência e do aumento do custo de oportunidade da prática da violência (WEISBURD et al., 2006).

Na literatura nacional, há certa escassez de estudos sobre determinantes da violência que incorporem aspectos relacionados à distribuição espacial da violência<sup>6</sup>. Além disso, a maioria dos trabalhos que o fazem são realizados tendo como unidade de observação os municípios, quando na realidade a criminalidade tende a se concentrar nas grandes cidades (GLAESER e SACERDOTE, 1999).

Tais estudos encontram uma associação espacial positiva das taxas de homicídios por 100 mil habitantes, além de evidenciarem a importância do mercado de trabalho, da pobreza e, em especial, da desigualdade de renda como preditores dos crimes violentos (PUECH, 2004; CARVALHO et al., 2005; DE LIMA et al., 2005; DE OLIVEIRA, 2008).

Dos Santos e Dos Santos Filho (2011), controlando a dependência espacial positiva das taxas de homicídios das microrregiões brasileiras, sugerem a existência de convergência das mesmas para um nível de violência cada vez mais elevado no país. Recentemente, Menezes et al. (2014) também analisou os determinantes socioeconômicos da taxa de homicídios em Recife-PE, incorporando a dependência espacial entre os bairros e encontrando que a mesma é positivamente influenciada pela desigualdade de renda.

Os resultados do presente estudo sugerem que a violência, representada pela taxa de CVLI (que inclui homicídios, latrocínios e lesão corporal seguida

---

6. Sachsida e Mendonça (2013) apresentam uma profunda revisão da literatura nacional e internacional sobre determinantes da criminalidade. Cerqueira e Lobão (2003) fazem um apanhado dos principais modelos teóricos e resultados empíricos dos determinantes da criminalidade.

de morte por 100 mil habitantes), está substancialmente concentrada no lado Oeste da capital cearense, a qual sofreu um processo de difusão espacial entre 2012 e 2014. Nesse contexto, encontra-se uma dependência espacial positiva, onde UDHs com elevada taxa de CVLI são rodeadas por vizinhos na mesma condição.

Ao incorporar tais características espaciais por meio de modelos de regressão espacial, evidencia-se que a violência está inversamente associada ao nível de desenvolvimento humano local, mas não é sensível a desigualdade de renda local.

Além disso, os crimes violentos exibem uma relação não linear com a densidade demográfica, além de incidir menos em áreas onde há uma maior predominância da população residente em apartamentos/condomínios vis-à-vis a população residente em casas. Por fim, identifica-se efeitos heterogêneos do desenvolvimento humano sobre a taxa de CVLI no território, sugerindo que políticas públicas de melhorias das condições de vida poderiam contribuir para uma redução maior da violência em determinadas localidades.

O presente trabalho está organizado nas seguintes seções. A primeira corresponde a esta Introdução. A Seção 2 exhibe a base de dados utilizada no estudo. A Seção 3 descreve a abordagem empírica realizada. A Seção 4 traz os resultados. Finalmente, a Seção 5 conclui a investigação.

## **2 BASE DE DADOS**

Duas bases de dados são utilizadas neste estudo. A primeira refere-se à malha georreferenciada das Unidades de Desenvolvimento Humano (UDH) para a Região Metropolitana de Fortaleza contendo o cálculo do IDH e de diversos outros indicadores socioeconômicos e demográficos produzidos pelo PNUD (2014). Essa base de dados compõe o Atlas de Desenvolvimento Humano nas Regiões Metropolitanas Brasileiras.

As UDHs foram geradas com base na homogeneidade socioeconômica, se constituindo em clusters de setores censitários do censo demográfico 2010, sendo, portanto, áreas geográficas com um nível de detalhe maior que o de bairros. Neste caso, o estudo baseia-se em 247 áreas territoriais de Fortaleza em detrimento de 119 bairros.

Desta forma, as UDHs consentem captar a diversidade de situações relacionadas com o desenvolvimento humano que ocorrem no interior dos espaços

metropolitanos, desvendando o que é oculto pelas médias municipais. Destaca-se, em especial, a exigência de que as áreas criadas devem ter, pelo menos, 400 domicílios particulares permanentes amostrados, o que garante precisão estatística dos indicadores (PNUD, 2014).

A segunda base de dados consiste nos casos georreferenciados de crimes violentos letais e intencionais (i.e. homicídio doloso, latrocínio e lesão corporal seguida de morte) ocorridos durante os anos de 2012 a 2014, tendo como fonte a Secretaria de Segurança Pública e Defesa Social (SSPDS) do Governo do Estado do Ceará.

Os dados das ocorrências possuem data, hora, e coordenadas do local do delito. Estas informações são utilizadas no monitoramento diário realizado pela SSPDS, sendo fundamentais para a tomada de decisão quanto à política de segurança pública do Estado.

## 2.1 A Distribuição Espacial de Crimes Violentos Letais e Intencionais em Fortaleza

O Estado do Ceará tem se destacado negativamente pelo rápido crescimento dos crimes violentos letais e intencionais nos anos recentes. A Tabela 1 mostra um aumento de 58,4% no número de tais delitos entre 2010 e 2014.

**TABELA 1: Número e taxa de CVLI por cem mil habitantes para o Estado do Ceará e a cidade de Fortaleza – 2010/2014**

Ano	Número de CVLI			Taxa de CVLI por 100.000 habitantes	
	Ceará (A)	Fortaleza (B)	(B)/(A)	Ceará	Fortaleza
2010	2.803	1.233	44,0%	33,4	50,3
2011	2.807	1.238	44,1%	32,9	50,0
2012	3.735	1.689	45,2%	43,4	67,6
2013	4.395	1.993	45,3%	50,1	78,1
2014	4.439	1.989	44,8%	50,2	77,3

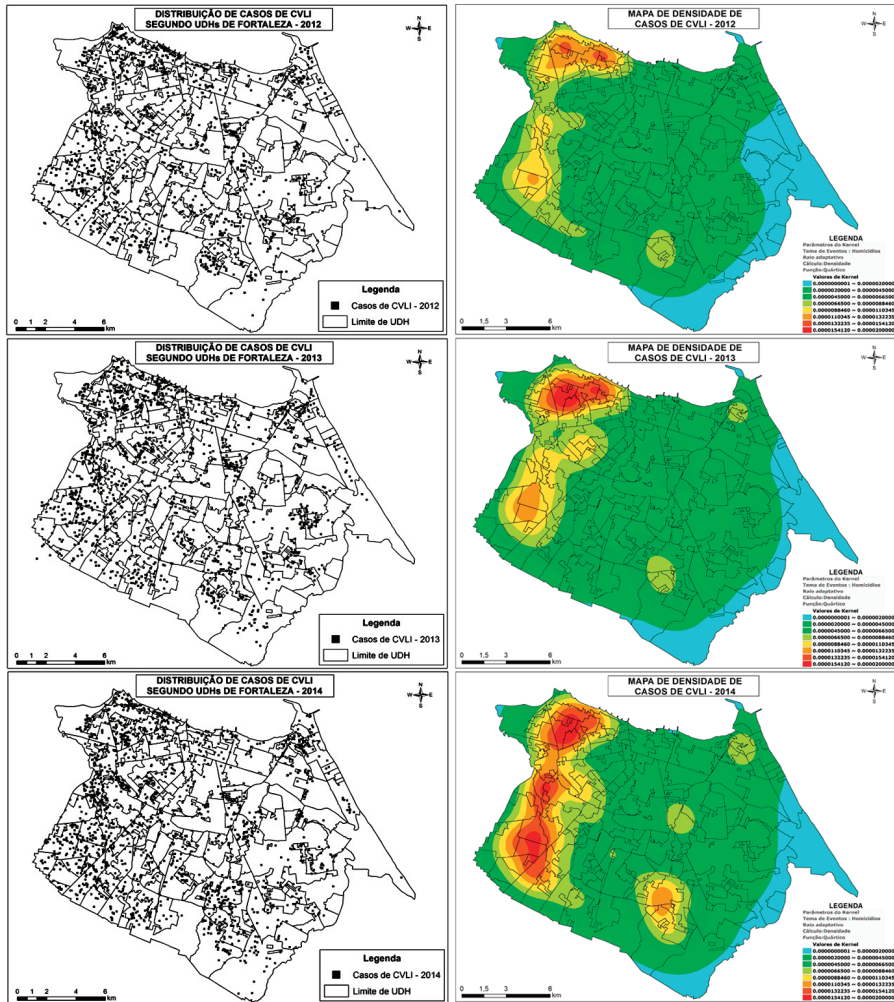
Fonte: Secretaria de Segurança Pública e Defesa Social (SSPDS/CE). Elaboração dos autores.

Nota. O cálculo da taxa de CVLI por 100.000 habitantes utiliza em 2010 dados do Censo Demográfico. Para os anos subsequentes, utiliza-se a estimativa populacional do IBGE.

Em termos de participação, Fortaleza possui um grande peso na contabilidade dos crimes violentos letais e intencionais do Ceará. Sua participação era de 44% em 2010, chegou a 45,3% em 2013, ficando com 44,8% em 2014.

Ao levar em conta a população residente, percebe-se que o cenário de violência em Fortaleza se intensificou no período. Em 2010, a taxa de CVLI por 100.000 habitantes era de 50,3, saltando para 77,3 em 2014. Essa variação corresponde a um crescimento de 54% no período, influenciando o aumento taxa de CVLI do Estado. Entretanto, as ocorrências de crimes contra a vida em Fortaleza não se distribuíram homogeneamente no espaço geográfico como mostram os mapas da Figura 1, a seguir.

**FIGURA 1: Mapas de distribuição de casos de CVLI e Densidade de kernel segundo UDhS de Fortaleza - 2012 a 2014**



Fonte: Secretaria de Segurança Pública e Defesa Social (SSPDS/CE). Elaboração dos autores.



Bailey e Gatrell (1995) sugerem o uso do método de estimação não-paramétrica de *Kernel* para observar se a distribuição dos eventos é aleatória ou se possui uma distribuição regular no espaço, formando agrupamento (ou clusters).

Desta forma, a partir da localização georreferenciada das ocorrências de CVLI foram gerados mapas de superfície contínua onde os tons em azul claro demonstram menor densidade e os tons em vermelho correspondem aos locais com maior densidade de CVLI, verificando-se que a ocorrência de CVLI seguiu um padrão espacialmente não aleatório<sup>8</sup>, segundo os mapas da Figura 1 (ver mapas ampliados da Figura 1 no Apêndice).

Averigua-se também um aumento da área de maior densidade deste tipo de delito entre os anos de 2012 a 2014 (ampliação da cor vermelha nos mapas), sugerindo um processo de difusão dos crimes violentos em UDHs situadas nos bairros da Barra do Ceará, Vila Velha, Genibaú, Bom Jardim, Jangurussu, Conjunto Palmeiras, Aerolândia e Vicente Pinzon. Há diversas causas potenciais para a difusão espacial dos crimes violentos letais e intencionais em Fortaleza (exemplo, tráfico de drogas, conflito entre gangs, etc.), porém sua investigação não fez parte do escopo do presente estudo.

Além disso, calculou-se o Índice de Moran Global para a média aritmética das taxas de CVLI (anos de 2012 a 2014), observando-se uma associação espacial positiva da violência. Em outras palavras, UDHs com elevada taxa de CVLI são, em média, rodeadas por UDHs com altas taxas de CVLI (Figura 2). Esse resultado foi obtido a partir da matriz de contiguidade do tipo *Rook*<sup>9</sup> (Moran-I 0,105, *p-valor*=0,008), a matriz do tipo *Queen* (Moran-I 0,095, *p-valor*=0,013).

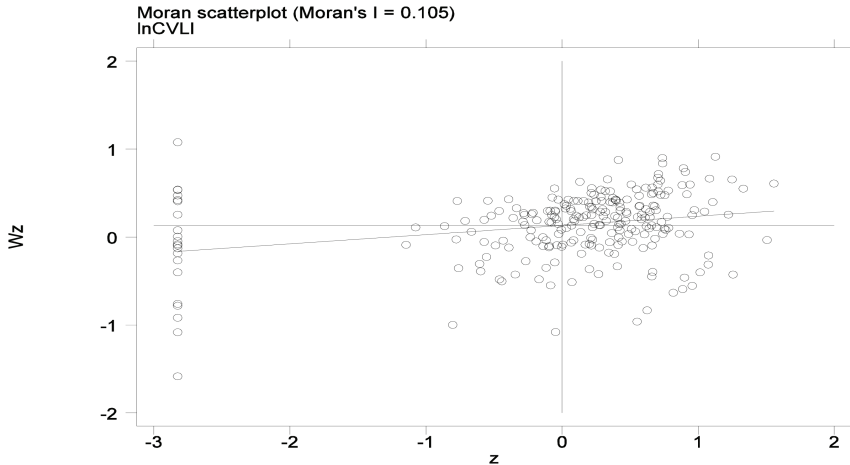
---

7. A técnica de Kernel gera uma superfície contínua onde o valor da densidade será proporcional ao das amostras (casos de CVLI) por unidade de área dentro de uma região de influência, ponderando-os pela distância de cada um à localização de interesse (CÂMARA e CARVALHO, 2002). O cálculo da densidade de kernel nesta pesquisa utilizou a função quártica, usando o mesmo raio adaptativo e legenda para os três anos.

8. Tal fenômeno já foi evidenciado para a própria capital cearense por MEDEIROS et al. (2013) utilizando dados de homicídios coletados de maneira informal pelo Jornal Diário do Nordeste no ano de 2011, e para outras cidades brasileiras como Campinas-SP (De MELO et al., 2015).

9. Usaram-se as matrizes de contiguidade do tipo Queen e Rook. A primeira considera vizinhas duas regiões (UDHs) que possuam fronteiras comuns, analisando os nós (vértices do polígono georreferenciado). Já a segunda considera vizinha apenas às regiões que tenham fronteira física (lado do polígono) em comum.

**FIGURA 2: Gráfico de Dispersão de Moran para o Logaritmo Natural do Valor Médio da Taxa de CVLI em Fortaleza – 2012/2014**



Fonte: Elaborado pelos autores.

Portanto, as Figuras 1 e 2 sugerem que as ocorrências de crimes violentos letais e intencionais em Fortaleza se concentram em áreas específicas da cidade, e ignorar tal característica em modelos de regressão pode levar a resultados enviesados.

## 2.2 A Violência e o Desenvolvimento Humano em Fortaleza

Conforme PNUD (2014), o desenvolvimento humano corresponde ao processo de ampliação das liberdades dos indivíduos com relação às suas capacidades e às oportunidades a seu dispor, de forma que cada pessoa possa escolher a vida que deseja ter visando à melhora ou a manutenção de seu bem-estar.

Neste contexto, o IDH foi concebido de forma a contemplar três importantes dimensões para a expansão da liberdade das pessoas: a oportunidade de se levar uma vida longa e saudável (Longevidade), o acesso ao conhecimento (Educação), e a possibilidade de ter um padrão de vida digno (Renda)<sup>10</sup>.

A Tabela 2 apresenta o IDH, e seus subíndices, para o município de Fortaleza relativo aos anos de 2000 e 2010. Verifica-se uma elevação do indicador no derradeiro decênio, saindo de um valor de 0,652 para 0,754, equivalendo a um crescimento relativo de 15,64%.

10. O IDH varia no intervalo de 0 a 1, sendo classificado em cinco classes: Muito baixo (0,000 - 0,499), Baixo (0,500 - 0,599), Médio (0,600 - 0,699), Alto (0,700 - 0,799) e Muito alto (0,800 - 1,000).

**TABELA 2: IDH do município de Fortaleza - 2000 e 2010**

Indicador	2000	2010	Taxa de Variação (%)
IDH	0,652	0,754	15,64
IDH – Educação	0,534	0,695	30,15
IDH – Longevidade	0,744	0,824	10,75
IDH – Renda	0,697	0,749	7,46

Fonte: PNUD (2014). Elaboração dos autores.

Este dado evidencia uma melhora na qualidade de vida dos fortalezenses na última década. A dimensão da Longevidade possuía no ano 2000 o maior valor, permanecendo na mesma posição em 2010, sendo acompanhada, respectivamente, das dimensões da Renda e da Educação no último ano. Em síntese, a capital cearense estava qualificada na faixa de Médio desenvolvimento humano no ano 2000 passando para a faixa de Alto desenvolvimento humano em 2010.

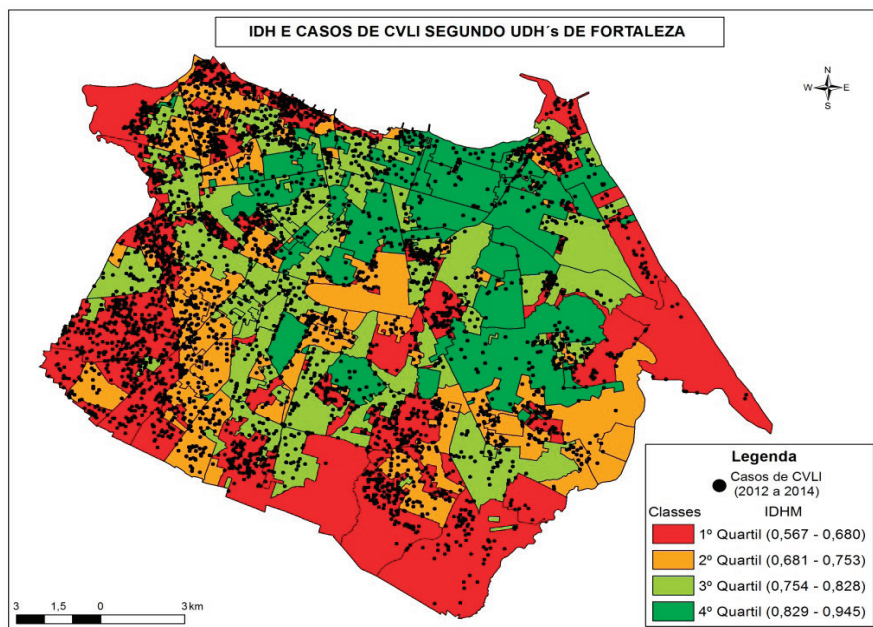
A distribuição espacial do desenvolvimento humano em Fortaleza pode ser visualizada na Figura 3, abaixo, a qual exhibe as UDHs segundo a classificação por quartis. Nessa perspectiva, podem-se identificar as seguintes classes: 0,567-0,680 (cor vermelha), 0,681-0,753 (laranja), 0,754-0,828 (verde claro), e 0,829-0,945 (verde escuro).

Evidencia-se pela análise do mapa do IDH, segundo a classificação dos quartis, que as áreas com os piores índices de desenvolvimento humano situam-se em zonas periféricas da cidade, estando às mesmas destacadas nas cores vermelha e laranja na legenda do mapa.

Por sua vez, os locais com maiores IDH encontram-se em áreas reconhecidas como mais nobres da cidade, como por exemplo, nos bairros da Aldeota, Meireles, Dionísio Torres e Cocó.

Na mesma figura, observa-se a localização georreferenciada dos casos de CVLI para o período em análise (2012 a 2014), revelando que a maioria deste tipo de crime concentra-se em setores da cidade que possuem os menores índices de desenvolvimento humano (1º e 2º quartil), ou seja, locais com condições relativas desfavoráveis de saúde (longevidade), educação e renda.

**FIGURA 3: IDH em 2010 e registro de CVLI em Fortaleza no período 2012-2014**



Fonte: Elaboração dos autores.

Para reforçar os resultados evidenciados na Figura 3, apresenta-se na Tabela 3 o quantitativo de ocorrências e a estimativa da taxa de CVLI por 100.000 habitantes para cada um dos quatro grupos de quartis, verificando-se que pelos menos 70% dos casos ocorreram no 1º e 2º quartil (nos três anos averiguados), que são às áreas da cidade com menores índices de desenvolvimento humano.

**Tabela 3:** Distribuição do número e da taxa de CVLI segundo quartil do IDH das Unidades de Desenvolvimento Humano - 2012 a 2014

Ano	1º Quartil			2º Quartil			3º Quartil			4º Quartil		
	Nº de casos	%	Taxa de CVLI	Nº de casos	%	Taxa de CVLI	Nº de casos	%	Taxa de CVLI	Nº de casos	%	Taxa de CVLI
2012	723	52,0	82,0	320	23,0	54,7	238	17,1	41,6	109	7,8	26,5
2013	755	49,7	85,6	347	22,8	59,3	285	18,8	49,8	133	8,8	32,3
2014	973	53,0	110,3	409	22,3	69,9	343	18,7	59,9	110	6,0	26,7
Média	817,0	51,6	92,6	358,7	22,7	61,3	288,7	18,2	50,4	117,3	7,5	28,5

Fonte: PNUD, IBGE e SSPDS. Elaboração dos autores.

Em relação às estimativas das taxas por 100.000 habitantes, constata-se também que estes locais possuem índices bem superiores às UDHs que estão inseridas nos 3º e 4º quartil (lugares com maiores IDH). Dentro deste contexto, sinaliza-se que as áreas mais violentas da capital, no tocante a crimes violentos contra a vida, são também as áreas mais vulneráveis em termos de desenvolvimento humano.

É importante salientar que essa relação não é causal. Na realidade, a distribuição espacial do desenvolvimento humano em Fortaleza também não é aleatória, e pode compartilhar dos mesmos fatores espaciais que influenciam as ocorrências de crimes violentos letais e intencionais. Segundo Peres et al. (2008), as áreas nas grandes cidades onde há maiores carências de serviços públicos relacionados à educação, saúde, lazer, cultura e segurança constituem-se em locais favoráveis à criminalidade, devido, sobretudo, ao abandono pelo poder público destes territórios.

Cardia et al. (2003) corrobora a afirmação anterior ao mencionar que com frequência os crimes contra a vida se concentraram em determinados espaços. Nesses espaços ocorrem também concentrações de desigualdades econômicas e sociais, aliadas à sobreposição de carências. Ou seja, o espaço revela a desigualdade de condições de vida e possibilita que determinados grupos sejam alvos preferenciais da mortalidade por causas externas.

Para Kleinschmitt et al. (2011), esses lugares, com graves violações de direitos humanos e sobreposições de múltiplas carências econômicas e sociais, se tornariam terreno fértil para a violência se desenvolver.

Destaca-se, entretanto, que, conforme Azevedo (2003), não existe uma relação direta entre pobreza e criminalidade. Há, muito mais, uma vulnerabilidade dos pobres frente à falta da oferta com qualidade de serviços públicos e, muito menos, uma suposta tendência ou motivação criminosa imposta por sua condição social.

No entanto, os resultados da Figura 3 e da Tabela 3 sugerem que há uma relação espacial inversa entre violência e desenvolvimento humano em Fortaleza, motivando o uso de métodos de regressão espacial para quantificar tal relação.

### **2.3 Demais Fatores Determinantes da Violência**

Para analisar a relação entre o desenvolvimento humano e a violência em Fortaleza, torna-se fundamental levar em conta outros potenciais fatores

capazes de confundir tal relação. A Tabela 4 apresenta as estatísticas descritivas das variáveis explicativas adicionais consideradas no estudo. Tais variáveis têm como fonte de dados o próprio Atlas do Desenvolvimento Humano das Regiões Metropolitanas e o Governo do Estado do Ceará.

**TABELA 4: Estatísticas Descritivas das Variáveis Explicativas**

Variável	Descrição	Média	Desvio Padrão	Mín.	Máx.
Atlas do Desenvolvimento Humano nas Regiões Metropolitanas					
IDH	Índice de Desenvolvimento Humano	0,75	0,09	0,57	0,95
GINI	Índice de Gini	0,46	0,05	0,37	0,62
DEMOG	Densidade demográfica (população por hectare)	165,8	214,8	1,48	222,0
CONDOM	% de pessoas vivendo em apartamento ou condomínios em relação à população total da UDH	25,45	27,66	0,00	100,00
JOVEM	% de pessoas com idade entre 10 e 24 anos em relação à população total	27,59	3,52	16,19	36,39
MCHEFE	% de Mulheres chefe de família com filhos menores de 15 anos e fundamental incompleto	32,87	15,95	0,00	63,72
DESOC	% população economicamente ativa de 10 anos ou mais de idade desocupada na semana de referência	8,12	2,03	4,28	15,99
AGUAESG	% de pessoas em domicílios com abastecimento de água e esgotamento sanitário inadequados	1,06	1,17	0,00	5,06
PAREDE	% de pessoas em domicílios com paredes inadequadas	0,31	0,42	0,00	2,11
DENS	% de pessoas vivendo em domicílios com mais de 2 moradores por dormitório	28,89	13,27	2,24	56,70
Governo do Estado do Ceará					
INFPOL	= 1, se a UDH possui pelo menos uma delegacia da Polícia Civil ou batalhão da Polícia Militar no raio de 100 metros	0,17	0,380	0	1

Fonte: Elaboração dos autores.

Na Tabela 4, observa-se um valor médio do IDH de 0,75 considerando todas as 247 UDHs. Embora, em média, as unidades observacionais possam ser consideradas com alto desenvolvimento humano, ainda se percebe uma desigualdade na capital cearense, uma vez que existem UDHs variando seu índice entre 0,57 e 0,95. Quanto às características populacionais, nota-se que as UDHs possuem aproximadamente 166 habitantes por hectare, e um quarto

de seus residentes vivem em apartamentos ou condomínios. Além disso, a proporção de jovens com idade entre 10 e 24 anos é de aproximadamente 27,6%.

Outra variável testada frequentemente nos estudos empíricos sobre criminalidade é a taxa de desemprego da população. Em 2010, a taxa média de desocupação nas UDHs era de aproximadamente 8%. Do ponto de vista do modelo de Becker (1968), uma alta taxa de desemprego local reduz o custo de oportunidade do indivíduo de se tornar um criminoso.

Uma grande parte da literatura econômica tem se dedicado a estimar os efeitos da força policial sobre as taxas de crimes. Intuitivamente, quanto maior for o efetivo policial maior será probabilidade de encarceramento, elevando o custo de oportunidade de entrar na atividade criminal no sentido de Becker (1968). Todavia, a maioria dos estudos empíricos que tentam analisar o efeito do efetivo policial sobre a criminalidade não levam em consideração a potencial causalidade reversa da relação. Uma exceção é Chalfin e McCrary (2013), os quais mostram que o efeito polícia é maior sobre crimes contra a vida do que para crimes contra o patrimônio em cidades de médio e grande porte dos Estados Unidos. Sachsida e Mendonça (2013) usam dados em painel e mostram que um aumento do efetivo policial contribui para a redução das taxas de homicídios no Brasil.

No presente estudo, a força policial é representada por uma variável binária a qual indica 1 caso a UDH possua pelo menos uma delegacia da Polícia Civil ou batalhão da Polícia Militar em um raio de até 100 metros. Aproximadamente 17% das UDHs possuem infraestrutura policial nas proximidades.

No intuito de controlar potenciais fatores de desorganização social<sup>11</sup>, as seguintes variáveis também foram incluídas na análise: proporção de mulheres chefe de família com crianças menores de 15 anos e com fundamental incompleto<sup>12</sup>, representando quase 1/3 da população feminina nas UDHs, proporção da população com abastecimento de água ou esgoto inadequado (em média, 1% da população das UDHs), proporção da população vivendo em domicílios com paredes inadequadas (em média, 0,3% da população), e proporção

---

11. Essa teoria parte da ideia de que a criminalidade emergiria como consequência de efeitos indesejáveis na organização de relações sociais em nível comunitário e de vizinhanças (ENTORF e SPENGLER, 2002). Tais relações são condicionadas não somente pela desagregação familiar, como também pelo status socioeconômico, heterogeneidade étnica, urbanização, condições de moradia e mobilidade residencial (CERQUEIRA e LOBÃO, 2003).

12. Glaeser e Sacerdote (1999) mostram que entre um terço e metade do efeito urbano sobre o crime pode ser explicado pela presença de mais famílias chefiadas por mulheres nas cidades. Essa variável representa a desagregação familiar, e tem sido abordada em estudos que assume os pressupostos da teoria da desorganização social.

da população vivendo em domicílios com mais de 2 pessoas por dormitório (em média, 29% da população das UDHs).

### 3 ABORDAGEM ECONOMETRICA

#### 3.1 Regressão Espacial

O objetivo desta subseção é o de estimar a associação entre a violência e o desenvolvimento humano em Fortaleza. Para testar tal relação, estimam-se os seguintes modelos espaciais: Defasagem Espacial (Spatial Autoregressive Model – SAR), Erro Espacial (Spatial Error Model – SEM) e, finalmente, o modelo de Autocorrelação Espacial (Spatial Autocorrelation Model – SAC), por meio do programa Stata 12<sup>o</sup>.

Tais modelos são discutidos em Anselin (1988), LeSage e Peace (2009) e Almeida (2012). Particularmente, o modelo SAC é uma forma aninhada dos modelos SAR e SEM. Assim, representa-se a estratégia de estimação através de uma única especificação no formato *cross-section* que é dado por:

$$Y = \rho WY + X\beta + u \tag{1}$$

$$u = \lambda Wu + \varepsilon$$

onde  $Y$  é logaritmo natural da média da taxa de homicídios das UDHs entre 2012 e 2014.<sup>13</sup> O vetor  $X$  contém as variáveis explicativas, incluindo o IDH de cada área e as demais variáveis controles (ver Tabela 4) cujos parâmetros associados são representados pelo vetor  $\beta$ . O erro aleatório é dado por  $\varepsilon$  o qual é assumido ser normalmente distribuído com média zero e variância finita. O termo  $W$  é a matriz de contiguidade que indica os vizinhos de cada UDH.

Finalmente, os parâmetros espaciais são  $\rho$  e  $\lambda$ , que medem respectivamente a autoregressividade espacial da taxa de CVLI e a autocorrelação espacial dos erros. Note que o modelo se reduz ao de defasagem espacial quando  $\lambda = 0$ , ou ao de erro espacial quando  $\rho = 0$ . Logo, a significância das estimativas dos parâmetros espaciais indicará que tipo de dependência espacial existe. Compara-se cada versão do modelo espacial com o de Mínimos Quadrados Ordinários com base no teste razão de verossimilhança ( $LR$ ).

13. O uso da média aritmética se faz necessário como forma de suavizar a variável em virtude da sazonalidade das ocorrências em determinadas áreas. Além disso, o logaritmo natural da taxa de homicídios proporciona uma interpretação dos coeficientes em termos de elasticidades e/ou semi-elasticidades.



Ademais, outros testes são implementados para ajudar a decidir a especificação mais adequada para a análise ora proposta. Além de estatísticas como o log likelihood, critério de Akaike e de Schwarz, apresenta-se os testes (versão simples e robusta) para o Multiplicador de Lagrange com o intuito de identificar a presença de defasagem espacial ( $LM_\rho$  ou LM SAR), ou de dependência espacial do erro ( $LM_\lambda$  ou LM SEM), ou de ambas dependências espaciais ( $LM_\rho + LM_\lambda$  ou LM SAC).

André e Carvalho (2014) utilizam essa mesma abordagem espacial para analisar os determinantes espaciais da demanda residencial por água em bairros de Fortaleza. Os autores sugerem que negligenciar os efeitos espaciais pode afetar sobremaneira a estimativa da elasticidade preço-demanda por água.

### 3.2 Regressão Ponderada Geograficamente

De acordo com a Figura 3, há uma substancial heterogeneidade espacial do desenvolvimento humano na cidade de Fortaleza. Enquanto UDHs localizadas mais ao centro da cidade possuem níveis de desenvolvimento humano mais elevado, à medida que se desloca para as áreas periféricas os níveis de IDH decrescem. Desta forma, questiona-se se a sensibilidade das taxas de crimes violentos letais e intencionais varia quanto aos diferentes níveis de desenvolvimento humano na capital cearense<sup>14</sup>.

Para observar a heterogeneidade na relação entre violência e desenvolvimento humano em Fortaleza emprega-se o método de Regressão Ponderada Geograficamente (RPG), segundo Fotheringham et al. (2000).

Diferentemente da subseção 3.1, a RPG a ser estimada não leva em consideração um conjunto de variáveis controles além do IDH. Obviamente, isso leva a um questionamento relativo à correta especificação do modelo.

Contudo, ao considerarmos a hipótese de efeitos heterogêneos do IDH sobre a taxa de CVLI do ponto de vista geográfico, é admissível que o modelo RPG possua uma especificação distinta do que foi realizado na subseção anterior, uma vez que o objetivo é estimar a variação espacial do IDH sobre a taxa de CVLI, gerando estimativas locais para cada uma das 247 UDHs. Portanto, o modelo empírico de RPG a ser estimado é dado por:

$$Y_i = \beta_0(\mu_i, v_i) + \beta_1(\mu_i, v_i)X_i + \varepsilon_i \quad (2)$$

onde  $(\mu_i, v_i)$  representa as coordenadas do ponto  $i$  no espaço, e  $\beta_1(\mu_i, v_i)$  é o

---

14. Conforme Almeida (2012), analisar apenas a resposta média ou global para um fenômeno pode não ser suficiente para compreensão do mesmo, uma vez que os fenômenos socioeconômicos são propensos serem não constantes através do espaço.

coeficiente local do ponto  $i$ . Seguindo a Lei de Tobler<sup>15</sup>, a RPG admite que os dados mais próximos do ponto da regressão têm uma probabilidade maior de influenciá-lo, pressupondo assim uma ponderação geográfica, usando para tanto, a função de densidade de *kernel*.

Nessa perspectiva, para a estimação do modelo RPG utilizou-se o programa ArcGis 9.3<sup>6</sup> empregando a densidade de *kernel* adaptativo. Para a delimitação do comprimento de banda usou-se a menor distância capaz de minimizar a variância das estimativas locais, recorrendo-se a pelo menos 30 observações em cada ponto da regressão.

Conforme Almeida (2012), a escolha do *kernel* adaptativo justifica-se pela heterogeneidade das áreas a serem analisadas, uma vez que o filtro adaptativo possibilita que se expandam as áreas em que as observações sejam escassas e diminua em áreas em que as observações tenham alta densidade de dados.

## 4 RESULTADOS

### 4.1 Resultado dos Modelos de Regressão Espacial

Nesta seção, apresenta-se os resultados estimados dos modelos de Mínimos Quadrados Ordinários (MQO), Defasagem Espacial (Spatial Autoregressive Model – SAR), Erro Espacial (Spatial Error Model – SEM) e, finalmente, o modelo de Autocorrelação Espacial (Spatial Autocorrelation Model – SAC), conforme disposto na Tabela 5.

Os modelos espaciais estimados utilizam uma matriz de contiguidade do tipo *Rook* (0,138,  $p$ -valor= 0,001), uma vez que o Índice de Moran Global estimado a partir dos resíduos do modelo de MQO se mostrou superior ao da matriz de contiguidade do tipo *Queen* (0,124,  $p$ -valor= 0,002). Portanto, as estimativas advindas do modelo de MQO indicam a presença de dependência espacial, revelando que esse modelo é inadequado para analisar os efeitos do desenvolvimento humano sobre a violência em Fortaleza. Essa evidência é corroborada pela significância da estatística LR em todos os modelos espaciais estimados.

Já o Índice de Moran Global calculado para os resíduos dos modelos espaciais mostra ausência de dependência espacial, indicando que tal problema é devidamente controlado nas três especificações. No entanto, há a necessidade de se escolher a especificação mais adequada para a interpretação dos resultados. Neste caso, em qualquer das especificações espaciais estimadas, a autoregressividade espacial e a autocorrelação espacial dos erros se mostram presentes de

15. Corresponde a “primeira lei da geografia”, segundo a qual todas as coisas são parecidas, mas coisas mais próximas se parecem mais que coisas mais distantes.

acordo com as estatísticas LM. Logo, a especificação mais adequada é o modelo SAC, o qual incorpora a autoregressividade espacial da variável dependente e a autocorrelação espacial do erro.

**TABELA 5: Resultados dos modelos de MQO e de Dependência Espacial**

	MQO		SAR		SEM		SAC	
	<i>Estimativa</i>	<i>E.P.</i>	<i>Estimativa</i>	<i>E.P.</i>	<i>Estimativa</i>	<i>E.P.</i>	<i>Estimativa</i>	<i>E.P.</i>
IDH	-1,068	4,348	-4,751 **	2,159	-5,634 **	2,223	-5,394 **	2,223
GINI (0-1)	1,400	1,743	1,429	0,971	1,486	0,964	1,353	0,964
ln(DEMOG)	-0,078	0,319	-0,934 ***	0,294	-0,913 ***	0,299	-0,871 ***	0,299
ln(DEMOG)2	-0,050	0,040	0,096 ***	0,034	0,089 ***	0,034	0,080 **	0,034
CONDOM	-0,008	0,005	-0,004 **	0,002	-0,004 **	0,002	-0,004 **	0,002
JOVEM	-0,006	0,046	0,008	0,020	0,008	0,020	0,010	0,020
MCHEFE	0,016	0,011	-0,001	0,006	-0,002	0,006	-0,002	0,006
DESOC	0,061	0,047	0,017	0,020	0,020	0,019	0,022	0,019
AGUAESG	-0,188 **	0,086	-0,046	0,039	-0,054	0,037	-0,056	0,037
PARADE	-0,219	0,162	-0,043	0,092	-0,060	0,087	-0,069	0,087
DENS	0,012	0,019	-0,010	0,009	-0,014	0,009	-0,013	0,009
INFPOL	0,142	0,123	-0,130 *	0,077	-0,156 **	0,074	-0,158 **	0,074
<b>Moran-I (Resíduos)<sup>a</sup></b>	0,138 ***		0,531		0,593		0,669	
<b>Efeitos Espaciais</b>								
$\rho$			0,149 *	0,076			-0,264 **	0,127
$\lambda$					0,299 ***	0,092	0,517 ***	0,121
<b>Teste LR</b>								
$H_0: \rho=0$			3,829 **				4,330 **	
$H_0: \lambda=0$					10,650 ***		18,414 ***	
$H_0: \rho+\lambda=0$							22,755 ***	
<b>Crítérios de Escolha</b>								
Log. Likelihood			-177,1		-173,3		-171,4	
Crítério de Akaike			1,900		1,880		1,891	
Crítério de Schwarz			2,285		2,262		2,275	
<b>Teste LM</b>								
LM SAR (Anselin)			14,312 ***		14,951 ***		12,960 **	
LM SAR Robusto			26,180 ***		23,022 ***		19,221 ***	
LM SEM (Burridge)			2,987 *		4,022 **		5,537 **	
LM SEM Robusto			11,854 ***		12,093 ***		11,798 ***	
LM SAC			26,167 ***		27,045 ***		27,757 ***	

Nota: Nível de significância: \*\*\* 1%, \*\* 5% e \* 10%. Elaboração dos autores.

<sup>a</sup> Índice de Moran Global obtido a partir da matriz de contiguidade do tipo Rook.

Na versão SAR estimada, o parâmetro  $\rho$  se mostrou positivo embora

marginalmente significativa<sup>16</sup>. Porém, ao incorporar a autoregressividade espacial da taxa de CVLI e a correlação espacial dos erros,  $\rho$  passa a ser negativo, -0,264. Isso significa dizer que, em média, a queda de 1% da taxa de CVLI da vizinhança eleva a taxa de CVLI da UDH sob análise em aproximadamente 0,26%. Já o efeito espacial nos erros, capturado pelo parâmetro  $\lambda$ , é positivo e significativo, com um valor de 0,517.

Esse resultado sugere que a direção da correlação espacial da taxa de CVLI segundo o Índice de Moran (ver Figura 2) advém da autocorrelação espacial do erro ao invés da autoregressividade taxa de CVLI.

Muito provavelmente, a autocorrelação espacial do erro possa estar capturando efeitos espaciais relacionados, por exemplo, à difusão de crimes violentos no território (seja por formação de gangs e redes sociais com gangs rivais e/ou expansão de mercados de drogas, etc.). Todavia, o conjunto de informações disponíveis é insuficiente para testar tal mecanismo de efeito espacial, sendo uma das limitações do presente estudo.

Analisando os parâmetros estimados das variáveis explicativas do modelo SAC, tem-se que o aumento de 0,1 score do IDH, em média, reduz em 0,54% (-0,54 x 0,1) a taxa de CVLI em Fortaleza. Este resultado, portanto, confirma a relação inversa entre a incidência da violência e o nível de desenvolvimento humano na capital cearense. Puech (2004) também encontra uma relação inversa entre violência e desenvolvimento humano, onde um aumento de 0,1 scores no IDH dos municípios do Estado de Minas Gerais em 1991 reduziria a taxa de crimes violentos em 0,21% em 2000, porém, marginalmente significativa.

O índice de Gini, que mensura o grau de desigualdade de renda, se mostrou positivamente associado à taxa de crimes violentos letais e intencionais por 100 mil habitantes. No modelo SAC, a variação de 0,1 scores do Gini eleva a taxa de CVLI em aproximadamente 0,14%, mas tal estimativa é estatisticamente insignificante.

Percebe-se também que a densidade demográfica e a violência possuem uma relação não linear no formato de U, corroborando Regoeczi (2002). Segundo as estimativas a violência é inversamente relacionada à densidade demográfica até o nível de densidade de 5,44 em termos de log natural (ou 230 habitantes por hectare), quando a partir de então a relação se torna positiva.

Outro importante resultado é a relação inversa entre a proporção da população vivendo em apartamento ou condomínios e a taxa de CVLI. Um

16. Usando um modelo SAR, Menezes et al. (2013) também obtiveram uma estimativa negativa para o parâmetro espacial autoregressivo ao analisar os determinantes da taxa de homicídios em Recife-PE.

aumento de 1% em tal indicador reduz a taxa de crimes violentos em 0,004%. Esse resultado corrobora a hipótese de Jacobs (1961), indicando uma maior capacidade de monitoramento dos crimes violentos por parte da população que vive em apartamentos e condomínios vis-à-vis a população residente em casas.

Vale salientar que Faria et al. (2013) encontra efeitos negativos e insignificantes para a proporção de pessoas vivendo em casas em relação aos crimes violentos, e negativamente significativa para crimes patrimoniais em Brasília-DF. Percebe-se ainda que as estimativas dos autores provêm de estimações via MQO, ignorando a potencial dependência espacial existente.

A proporção de jovens com idade entre 10 e 24 anos e a taxa de desocupação se mostraram insignificantes, embora os efeitos sejam positivos corroborando a literatura. A proporção de mulheres chefes de família também se mostra insignificante, assim como a proporção de pessoas vivendo em domicílios com abastecimento de água ou coleta de esgoto irregulares, proporção de pessoas vivendo em domicílios com paredes inadequadas, e proporção de pessoas vivendo em domicílios com mais de dois moradores por dormitório.

Por fim, nota-se que UDHs com infraestrutura policial nas imediações apresentam menor taxa de crimes violentos. Em Média, unidades observacionais com delegacias da Polícia Civil ou batalhões da Polícia Militar possuem uma taxa de CVLI por 100 mil habitantes 0,16% menor do que UDHs sem a presença de tal infraestrutura policial. Embora essa variável não seja uma *proxy* adequada para o efetivo policial alocado na respectiva UDH, ela controla o efeito espacial da infraestrutura policial uma vez que a ocorrência de crimes violentos é menos provável de ocorrer nas suas imediações.

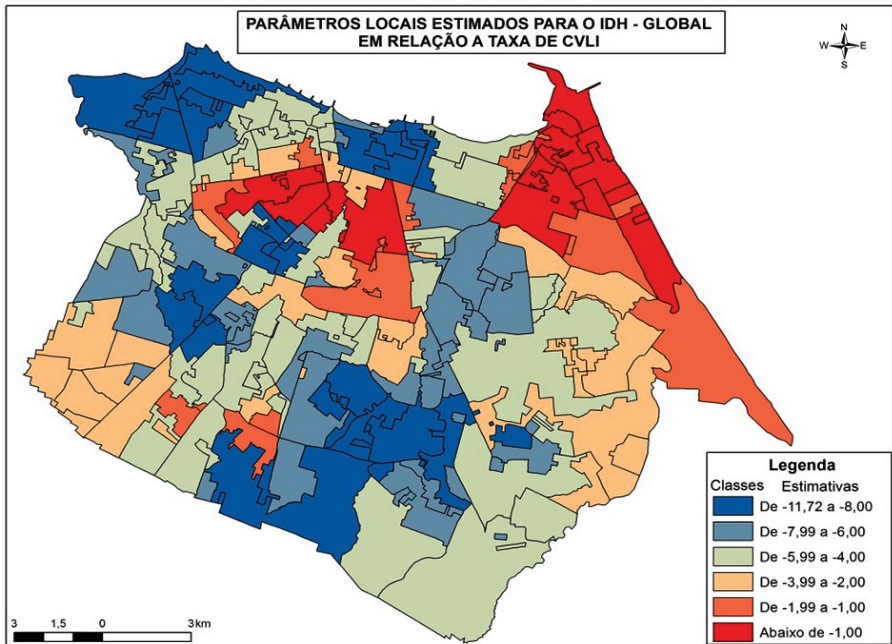
## 4.2 Resultados da Regressão Ponderada Geograficamente

Após a análise econométrica global é importante realizar uma análise local para se melhor compreender o fenômeno estudado. Assim, para cada uma das 247 UDHs presentes em Fortaleza foram obtidas, por meio do modelo RPG, estimativas dos parâmetros que quando apresentadas visualmente na forma de mapas facilitam a avaliação da variação espacial da associação entre desenvolvimento humano e violência na capital cearense.

Deste modo, conforme apresentado na Figura 4, percebe-se que as estimativas de diminuição da taxa de CVLI a partir de um aumento no IDH são diferentes nas UDHs de Fortaleza indicando a presença de instabilidade estrutural, ou seja, às relações entre às variáveis não são constantes para todas as

unidades geográficas.

**FIGURA 4: Heterogeneidade na Sensibilidade da taxa de CVLI em relação ao IDH**



Fonte: Elaboração dos autores.

Observando a legenda do mapa, verifica-se que a variação de 0,1 score do índice de desenvolvimento humano implica em uma redução de até -1,17% nas áreas na cor azul. Esse impacto é menor à medida que a cor das áreas se aproximam da cor vermelha. No modelo econométrico o efeito médio estimado foi de -0,54%.

No tocante às áreas de Fortaleza, evidenciam-se maiores impactos do IDH na diminuição da taxa de CVLI em UDHs presentes nos bairros da Barra do Ceará, Vila Velha, Cristo Redentor, Pirambú, Jardim Iracema, Passaré, Barroso, Jangurussu, Messejana, Centro, Praia de Iracema, Granja Portugal, Bon-sucesso e Vila Pery. Estas UDHs são caracterizadas por possuírem população com níveis menores de longevidade, educação e renda *per capita*.

Já os locais com menores impactos referem-se a áreas com melhor estrutura de serviços públicos e onde residem pessoas com melhores condições de desenvolvimento humano na cidade, citando, por exemplo, as UDHs situadas

nos bairros de Papicu, Varjota, Cocó, Meireles, Fátima, Benfica, Gentilândia, entre outros.

Não obstante, vale mencionar que mesmo nestes bairros existem UDHs com baixos índices de desenvolvimento humano, caracterizando uma desigualdade territorial presente em Fortaleza.

Portanto, a análise da regressão ponderada geograficamente vem complementar os resultados da análise global realizada, sendo bastante útil para o planejamento e a concepção de políticas públicas por indicar que áreas possuem maior retorno para o investimento em desenvolvimento humano no combate à criminalidade.

## 5 CONCLUSÃO

O presente estudo teve por objetivo investigar a relação entre violência, demografia e desenvolvimento humano em Fortaleza. A partir de dados georreferenciados de crimes violentos letais e intencionais, que inclui homicídios, latrocínios e lesão corporal seguida de morte, foi possível calcular a incidência de tal fenômeno social para as 247 Unidades de Desenvolvimento Humano da capital.

Dois resultados emergem de imediato. O primeiro sugere que entre 2012 e 2014, houve um processo de difusão espacial dos casos de CVLI por 100 mil habitantes em Fortaleza, especialmente nas UDHs que possuem interseção com os bairros: Barra do Ceará, Genibaú, Bom Jardim, Jangurussu, Aerolândia e Vicente Pinzon. Além disso, observa-se a existência de dependência espacial positiva, ou seja, áreas com elevada taxa de CVLI possuem vizinhos, também, com elevadas taxas de CVLI. Estas duas evidências sugerem que a distribuição dos crimes violentos em Fortaleza não ocorre de forma aleatória em seu espaço geográfico.

Diante de tal evidência, buscou-se averiguar se a incidência de crimes violentos contra a vida estaria condicionada ao nível de desenvolvimento humano da população, uma vez que áreas com baixo (alto) IDH apresentam taxas elevadas (baixas) de CVLI. As estimações econométricas incorporando a dependência espacial, indicam que o incremento de 0,1 scores no IDH reduz, em média, 0,54% a taxa de CVLI.

Embora o efeito médio estimado seja uma informação importante para balizar as políticas públicas de enfrentamento a violência em Fortaleza, é útil

complementar esta informação com a identificação da localização das áreas mais ou menos sensíveis a investimentos no desenvolvimento humano.

Neste contexto, através do método de Regressão Ponderada Geograficamente foi possível mapear as áreas com elevado retorno ao IDH como é o caso de UDHs presentes nos bairros da Barra do Ceará, Vila Velha, Cristo Redentor, Pirambú, Jardim Iracema, Passaré, Barroso, Jangurussu, Messejana, Centro, Praia de Iracema, Granja Portugal, Bonsucesso e Vila Pery. Desta forma, recomenda-se como política pública, priorizar tais áreas com o intuito de reduzir a violência local.

Como exemplos de ações específicas que podem ser desenvolvidas nestes locais, citam-se à melhoria dos equipamentos públicos como praças e parques, a pavimentação de ruas e a iluminação pública, o saneamento básico e a coleta de lixo, o maior acesso aos serviços de saúde, a ampliação de políticas de geração de empregos, a oferta de educação em tempo integral, entre outras ações.

Não menos importantes são os efeitos das variáveis demográficas sobre a violência na capital cearense. A densidade demográfica apresenta uma relação não linear com a taxa de CVLI, indicando que a relação entre a violência e a densidade demográfica muda de decrescente para crescente quando o número de habitantes por hectare atinge o nível de 230 pessoas/ha. Também se observa que quanto maior a proporção da população da UDH vivendo em condomínios ou apartamentos menor é a taxa de CVLI, o que sugere um maior poder de vigilância desse grupo demográfico em relação à população que vive em casas.

Vale ainda mencionar que a política pública de desenvolvimento local deve ser acompanhada também de uma eficiente e eficaz política de segurança pública, uma vez que a simples presença da infraestrutura policial também está associada a níveis menores de violência local. Extensões futuras deste estudo priorizarão a incorporação de variáveis que possam capturar melhor o efeito policial, incluindo as áreas de atuação dos programas do Governo Estadual Ronda do Quarteirão e Em Defesa da Vida.



## REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, E. S. **Econometria Espacial Aplicada**. Editora Alínea. Campinas. 498 p. 2012.

ANDRÉ, D. M.; CARVALHO, J. R. Spatial Determinants of Urban Residential Water Demand in Fortaleza, Brazil. **Water Resource Management**, v. 28, p. 2401-2414, 2014.

ANSELIN, L. **Spatial Econometrics: Methods and Models**. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.

ANSELIN, L.; COHEN, J.; COOK, D.; GORR, W.; TITA, G. **Spatial Analysis of Crime**, In: Measurement and Analysis of Crime and Justice, Criminal Justice, v. 4, p. 213-262, 2000.

AZEVEDO, M. A. Concepções sobre criminalidade e modelos de policiamento. **Revista Psicologia, Ciência e Profissão**, Belo Horizonte, v. 23, n. 3. p. 18-25, 2003.

BAILEY, T.C.; GATRELL, A.C. **Interactive Spatial Data Analysis**. Essex: Longman Scientific and Technical, 1995.

BECKER, G. Crime and Punishment: An Economic Approach. **Journal of Political Economy**, v. 76, n. 2, 169– 217, 1968.

BENOIT, J.-P.; OSBORNE, J. M. Crime, Punishment, and Social Expenditure. **Journal of Institutional and Theoretical Economics**, v. 151, p. 326–347, 1995.

CÂMARA, G.; CARVALHO, M.S **Análise Espacial de Eventos**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais - INPE. 2002. Disponível em: <http://www.dpi.inpe.br/gilberto/livro/analise/cap2-eventos.pdf>. Acesso em 13/04/2015.

CÁRDIA, N.; ADORNO, S.; POLETO, F. Homicídio e violação de direitos humanos em São Paulo. **Revista Estudos Avançados**, São Paulo, v. 17, n. 47, p. 43-73, 2003.

CARVALHO, A; CERQUEIRA, D; LOBAO, W. **Socioeconomic Structure , Self-fulfilment, Homicides and Spatial DEpendence in Brazil**, IPEA, Rio de Janeiro, Texto para Discussão, n. 1105, 2005.

CERQUEIRA, D.; LOBÃO, W. Determinantes da Criminalidade: Arcabouços

Teóricos e Resultados Empíricos, **DADOS – Revista de Ciências Sociais**, v. 47, n. 2, p. 233-269.

CHALFIN, A; McCRARY, J. **The Effect of Police on Crime: New Evidence from US Cities, 1960-2010**. NBER Working Paper Series, n. 18815, 2013.

CHRISTEN, B; SPEER, P. Predicting Violent Crime Using Urban and Suburban Densities. **Behavior and Social Issues**, v.14, p.113-127, 2005.

COHEN, J.; TITA, G. Diffusion in Homicide: Exploring a General Method for Detecting Spatial Diffusion Processes. **Journal of Quantitative Criminology**, v. 15, n. 4, 1999.

DE LIMA, M. L.; XIMENES, R.; DE SOUZA, E.; LUNA, C.; ALBUQUERQUE, M. F. Spatial Analysis of Socioeconomic Determinants of Homicides in Brazil. **Revista de Saúde Pública**, v. 39, n. 2, p.1-6.

DE MELO, S. N.; MATIAS, L. F.; ANDRESEN, M. A. Crime Concentrations and Similarities in Spatial Crime Patterns in a Brazilian Context. **Applied Geography**, v. 62, p. 314-324, 2015.

DE OLIVEIRA, C. Análise Espacial da Criminalidade no Rio Grande do Sul. **Revista de Economia**, v. 34, n. 3, p.35-60, 2008.

DOS SANTOS, M. J.; DOS SANTOS FILHO, J. I. Convergência das Taxas de Crimes no Território Brasileiro. **Revista EconomiA**, v.12, n.1, p.131-147, 2011.

ECK, J. E. 1994. **Drug Markets and Drug Places: A Case Control Study of the Spatial Structure of Illicit Drug Dealing**. Ph.D. diss., University of Maryland.

EHRLICH, I. Participation in Illegitimate Activities: A Theoretical and Empirical Investigation. **Journal of Political Economy**, v. 81, n. 3, p. 521– 565, 1973

ENTORF, H.; SPENGLER, H. Socioeconomic and Demographic Factors of Crime in Germany Evidence from Panel Data of the German States. **International Review of Law and Economics**, v. 20, p. 75-106, 2000.

FANJZYLBER, P.; LEDERMAN, D.; LOYAZA, N. What Causes Violent Crime. **European Economic Review**, v. 46, p. 1323-1357, 2002.

FARIA, J. R.; OGURA, L; M.; SACHSIDA, A. Crime in a Planned City: The Case of Brasília. **Cities**, v.32, p. 80-87, 2013.

FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA. **Anuário Brasileiro de Segurança Pública 2015**. São Paulo, 2015. Disponível na internet: <http://www.forumseguranca.org.br/>. Acesso em 06/10/2015.

FOTHERINGHAM, A. S.; BRUNSDON, C.; CHARLTON, M. **Quantitative Geography: Perspectives on Spatial Data Analysis**. London: Sage Publications, 2000. 282p.

GLAESER, E. L.; SACERDOTE, B. Why Is There More Crime in Cities? **Journal of Political Economy**, v. 107, n. 6, p. S225-S258, 1999.

GLAESER, E. L.; SACERDOTE, B.; SCHEINKMAN, J. A. Crime and Social Interactions. **Quarterly Journal of Economics**, p. 507-548, 1996.

GOLDBERG, M.; KIM, K.W.; ARIANO, M. **How Firms Cope with Crime and Violence**, The World Bank, Washington DC, 2014.

HESSELING, R. **Displacement: A review of the empirical literature**. In *Crime Prevention Studies*, vol. 3, ed. Ronald V. Clarke. Monsey, NY: Criminal Justice Press, 1994.

IBGE – Instituto Brasileiro de Geografia e Estatística. **Produto Interno Bruto dos municípios brasileiros em 2012**. 2014. Disponível na internet: <http://www.ibge.gov.br/>. Acesso em 16/09/2015.

IMROHOROGLU, A.; MERLO, A.; RUPERT, P. On the Political Economy of Income Redistribution and Crime, **International Economic Review**, v. 41, n. 1, p. 1–25, 2000.

IPECE - Instituto de Pesquisa e Estratégia Econômica do Ceará. IPECE INFORME. **Caracterização espacial dos homicídios dolosos em Fortaleza**. 2012. Disponível na internet: <http://www.ipece.ce.gov.br>. Acesso em 02/04/2015.

JACOBS, J. **The life and death of great American cities**. New York: Random House, 1961.

KLEINSCHMITT, S. C.; LIMA, J. F.; WADI, Y. M. Relação entre o crescimento da desigualdade social e dos homicídios no Brasil: o que demonstram os indicadores. **Revista Interseções**, v.13, n. 1, p. 65-90, 2011.

LESAGE, J. P.; PACE, R. K. **Introduction to Spatial Econometrics**, CRC Press. 2009.

LOUREIRO, P. R. A.; SILVA, E. C. D. Does Violence Deter Investment and

Hinder Economic Growth? **Brazilian Review of Econometrics**, v. 30, n. 1, p. 53-67, 2010.

MEDEIROS, C. N.; PINHEIRO, F. S. A.; SOUZA, G. M.; PINHEIRO, D. R. C. Correlações Espaciais entre Homicídios e Concentração de Aglomerados Subnormais em Fortaleza/CE/Brasil. **Revista GeoUECE**, Fortaleza, v. 2, n. 1, p. 92-110, 2013.

MENEZES, T.; SILVEIRA-NETO, R.; MONTEIRO, C.; RATTON, J. L. Spatial Correlation between Homicide Rates and Inequality: Evidence from Urban Neighborhoods. **Economic Letters**, v. 120, p.97-99, 2013.

O'FLAHERTY, B.; SETHI, R. Urban Crime. **Handbook of Regional and Urban Economics**, v.5B, ch.23, p.1519-1621, 2015.

PERES, M. F. T.; CARDIA, N.; MESQUITA NETO, P.; SANTOS, P. C.; ADORNO, S. Homicídios, desenvolvimento socioeconômico e violência policial no município de São Paulo, Brasil. **Revista Panam Salud Pública**, São Paulo, v. 23, n. 4, p. 268-273, 2008.

PERRY, J. D.; SIMPSON, M. E. Violent Crime in a City: Environmental Determinants. **Environment and Behavior**, v. 19, n. 1, p. 77-90, 1987.

PNUD - Programa das Nações Unidas. **Atlas do Desenvolvimento Humano nas Regiões Metropolitanas Brasileiras**. 2014. Disponível na internet: <http://www.pnud.org.br/Default.aspx>. Acesso em 02/04/2015.

PUECH, F. **How do Criminals Locate? Crime and Spatial Dependence in Minas Gerais**, Mimeo, 2004. Disponível na internet: <http://core.ac.uk/download/pdf/9309879.pdf>. Acesso em 20/09/2015.

REPPETTO, T. A. Crime Prevention and the Displacement Phenomenon. **Crime & Delinquency**, v. 22, p.166-67, 1976.

SACHSIDA, A.; MENDONÇA, M. J. **Evolução dos Determinantes da Taxa de Homicídios no Brasil**. IPEA, Rio de Janeiro, Texto para Discussão, n. 1808, 2013.

SANTOS, M. J. Dinâmica temporal da criminalidade: mais evidências sobre o efeito inércia nas taxas de crimes letais nos estados brasileiros. **Revista Economia**, v. 10, n. 1, p. 169-194, 2009.

SEN, A. **Development as Freedom**, *Oxford New York: Oxford University Press, 1999*.

SPELMAN, W. Abandoned Buildings: Magnets for crime? **Journal of Criminal Justice**, v. 21, n. 5, p. 481-495, 1993.

SOARES, R. Development, Crime and Punishment: Accounting for International Differences in Crimes Rates. **Journal Development Economics**, v.73, p.155-184, 2004.

TITA, G.; COHEN, J.; ENGBERG, J. An Ecological Study of the Location of Gang “Set Space”, **Social Problems**, v. 52, n. 2, p. 272-299, 2005.

TITA, G.; RADIL, S. Spatializing the Social Networks of Gangs to Explore Patterns of Violence. **Journal of Quantitative Criminology**, v. 27, p. 251-545, 2011.

TITA, G.; RIDGEWAY, G. The Impact of Gang Formation on Local Patterns of Crime. **Journal of Research in Crime and Delinquency**, v. 44, n. 2, p. 208-237, 2007.

TITA, G.; COHEN, J. **Measuring Spatial Diffusion of Shots Fired Activity Across City Neighborhoods**. In: Goodchild MF, Janelle DG eds Spatially Integrated Social Science. Oxford Press, New York, p. 171–204, 2004.

UNDOC. **Global Study on Homicide 2013: Trends, Context, Data. United Nations Office on Drugs and Crime**. Vienna, 2004. Disponível na internet: <http://www.unodc.org/gsh/>. Acesso em 11/04/2015.

WEISBURD, D.; ECK, J. What Can Police Do to Reduce Crime, Disorder, and Fear? **Annals of the American Academy of Political and Social Science**, v. 593, p. 42-65, 2004.

WEISBURD, D.; LAURA A. WYCKOFF, L. A.; READY, J.; JOHN E. ECK, J. E.; HINKLE, J. C.; GAJEWSKI, F. Does Crime Just Move Around the Corner? A Controlled Study of Spatial Displacement and Diffusion of Crime Control Benefits, **Criminology**, v. 44, n. 3, p. 549-592, 2006.

WHEELER, J. Citizenship, Violence and Authority in Rio’s Favelas. **Accord Series**, n. 25, p. 95-99, 2014.

ZHANG, J. The Effect of Welfare Programs on Criminal Behavior: A Theoretical and Empirical Analysis, **Economic Inquiry**, v. 35, n. 1, p. 120–37, 1997.

APÊNDICE

FIGURA A1: IDH segundo UDHs de Fortaleza e Bairros – 2010

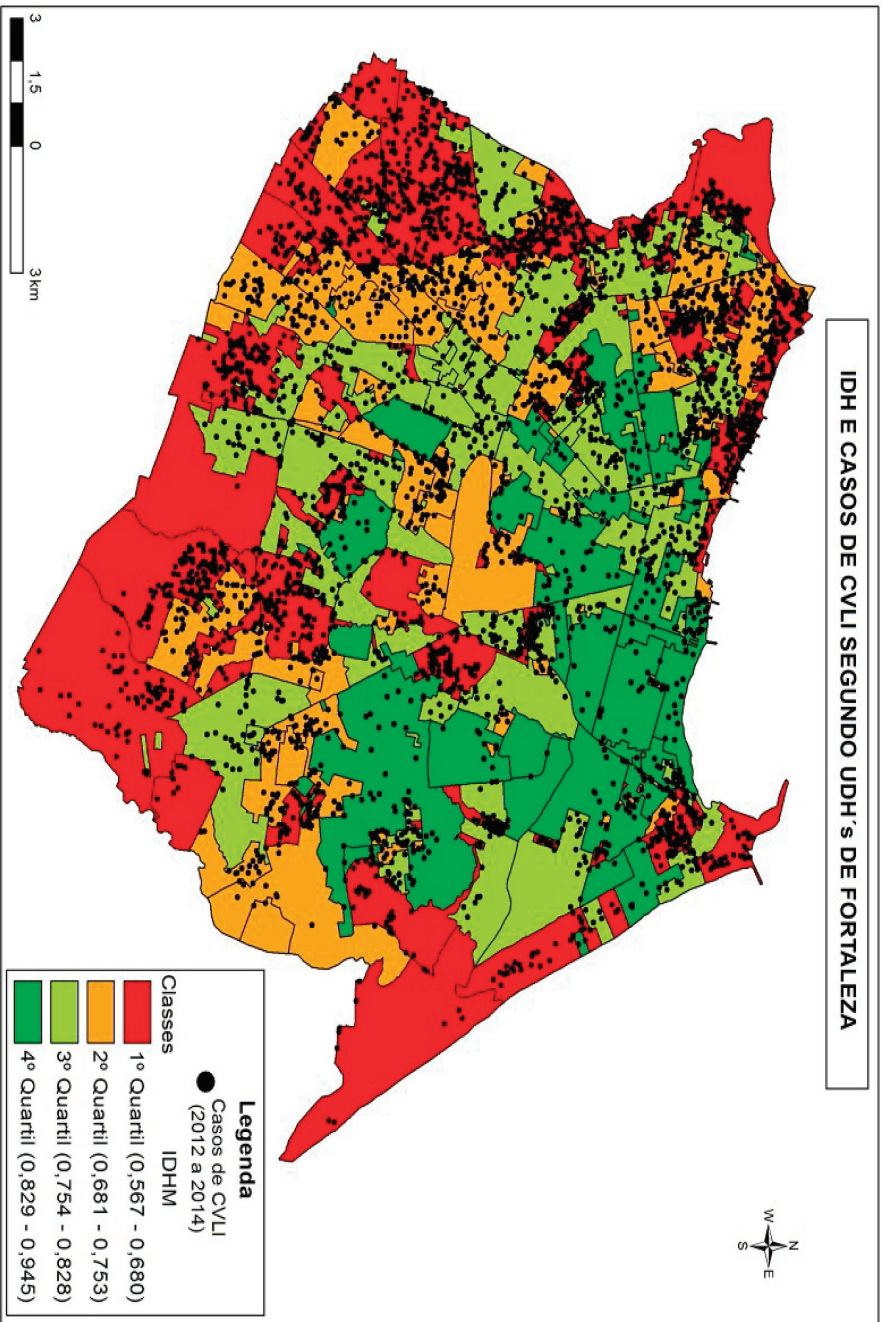


FIGURA A2: Mapa de densidade de casos de CVLI em Fortaleza – 2012

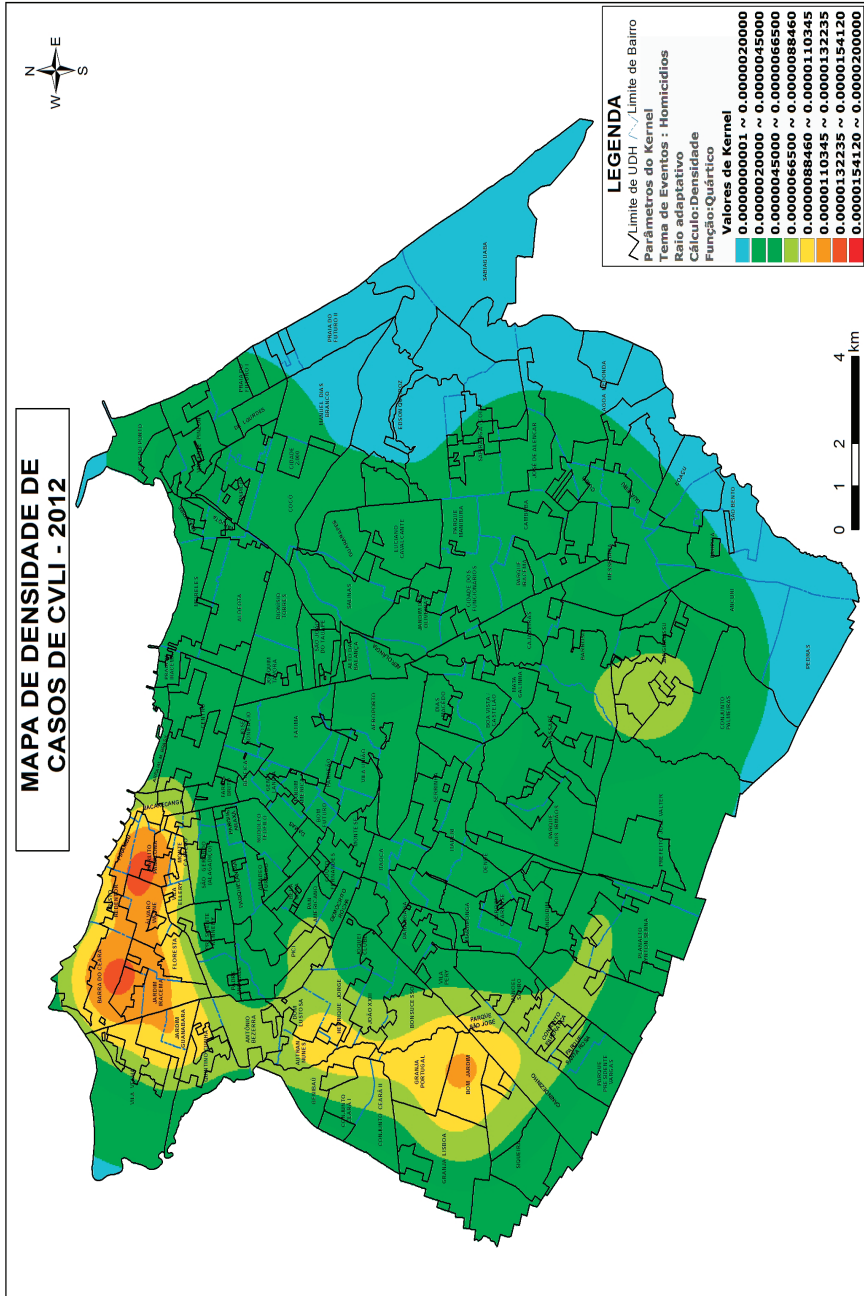


FIGURA A3: Mapa de densidade de casos de CVLI em Fortaleza – 2013

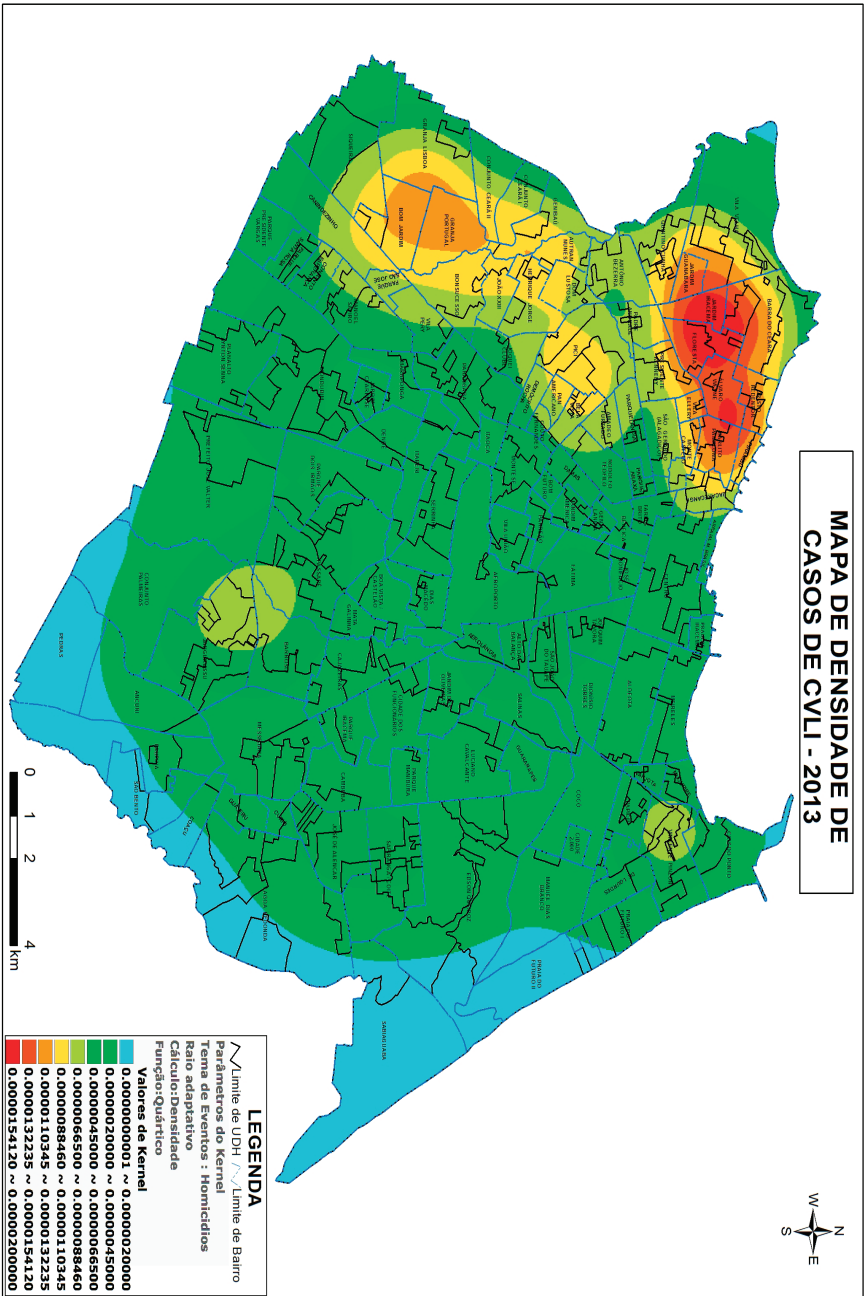
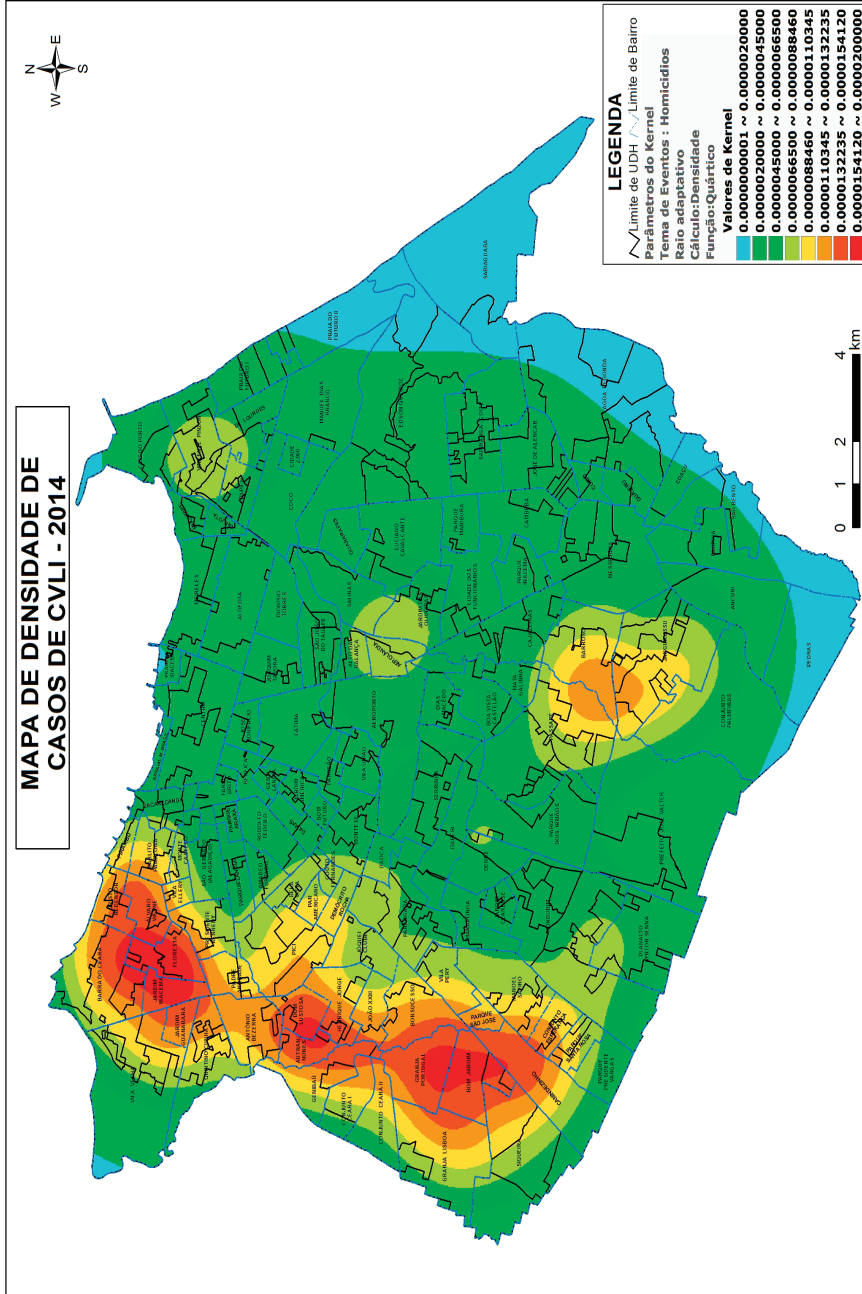




FIGURA A4: Mapa de densidade de casos de CVLI em Fortaleza – 2014



# EXTREMA POBREZA INFANTIL, DESIGUALDADE E CRESCIMENTO NOS MUNICÍPIOS CEARENSES

Décio Nonato Chaves de Assis\*

Cleyber Nascimento de Medeiros\*\*

Cláudio André Gondim Nogueira\*\*\*

## Resumo

O presente artigo analisa a evolução da pobreza extrema para a população total e infantil (grupo etário de 0 a 14 anos) nos municípios cearenses ao longo das últimas duas décadas. Seguindo a abordagem proposta por Bourguignon (2003), buscou-se investigar a relação triangular *pobreza-crescimento-desigualdade* e verificar como o crescimento econômico e a desigualdade de renda influenciaram a dinâmica da extrema pobreza em grupos etários distintos. Os resultados ratificaram que ambas as medidas de pobreza extrema são fortemente sensíveis aos níveis de desigualdade inicial. Além disso, verificou-se que a extrema pobreza infantil é menos sensível a sofrer reduções, a partir de diminuições na desigualdade de renda e de aumentos na renda média dos municípios cearenses, do que a extrema pobreza total, evidenciando uma maior persistência da extrema pobreza infantil. Adicionalmente, constatou-se que políticas que visem à redução das desigualdades terão, em média, maior impacto em diminuir tanto a extrema pobreza total como a infantil. Por fim, às informações produzidas fornecem subsídios para a alocação de diferentes políticas públicas de combate à extrema pobreza nos diversos municípios cearenses. Além de destacar a necessidade de um olhar específico para a infância como prerrogativa para o fim do ciclo vicioso de extrema pobreza no Ceará.

**Palavras-Chave:** Extrema Pobreza Infantil, Desigualdade, Crescimento Econômico.

\* Mestre em Economia pelo CAEN/UFC. Filiação Institucional: Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE). Endereço: Av. General Afonso Albuquerque Lima, S/N, 2º andar, Edifício SEPLAG. Cambéba/Fortaleza/Ceará. CEP: 60.839-900. E-mail: decio.chaves@ipece.ce.gov.br

\*\* Doutor em Geografia (PropGeo/UECE). Filiação Institucional: Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE). Endereço: Av. General Afonso Albuquerque Lima, S/N, 2º andar, Edifício SEPLAG. Cambéba/Fortaleza/Ceará. CEP: 60.839-900. E-mail: cleyber.medeiros@ipece.ce.gov.br.

\*\*\* Mestre em Economia (CAEN/UFC e Pennsylvania State University, EUA). Filiação Institucional: Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE). Endereço: Av. General Afonso Albuquerque Lima, S/N, 2º andar, Edifício SEPLAG. Cambéba/Fortaleza/Ceará. CEP: 60.839-900. E-mail: claudio.nogueira@ipece.ce.gov.br.

## Abstract

This article analyzes the evolution of extreme poverty in the total and child population (0-14 years old) in Ceara's municipalities over the past two decades. Following the approach proposed by Bourguignon (2003), the aim was to investigate the triangular relationship between *poverty, growth, and inequality* and then evaluate how economic growth and income inequality influenced the dynamics of extreme poverty in different age groups. The results obtained confirm that both extreme poverty measures are strongly sensitive to initial levels of inequality. In addition, it was found that extreme child poverty is less prone to decrease from reductions in income inequality or from increases in the average income of Ceara's municipalities, than extreme absolute poverty, showing a greater persistence of extreme poverty in the child population. Additionally, it was found that policies aimed at reducing inequalities have on average greater impact in decreasing extreme poverty both in the total and in the child population. Finally, the information produced provides subsidies for the allocation of different public policies, highlighting the need for a specific look at childhood as prerogative to end the vicious cycle of extreme poverty in the State.

**Key words:** Extreme Child Poverty, Inequality, Economic Growth.

## 1 INTRODUÇÃO

O Brasil caracteriza-se historicamente por possuir um número elevado de indivíduos em estado de pobreza extrema. Segundo relatório<sup>1</sup> da Secretaria Geral da Organização das Nações Unidas (ONU, 2005) sobre o Projeto do Milênio, a região nordeste do Brasil enquadra-se, conjuntamente com o oeste da China, norte da Índia e sul do México, como algumas das regiões notáveis em nível mundial por apresentarem bolsões de pobreza.

Nesse contexto, o estado do Ceará também é destaque negativo apresentando 17,8% da sua população na situação de extrema pobreza, com base no parâmetro estabelecido pelo Ministério do Desenvolvimento Social e Combate à Fome<sup>2</sup> (MDS). Não obstante esse fato, conforme estudo da ONU (2010), cinco cidades brasileiras estão entre as 20 mais desiguais do mundo, sendo a

1. Projeto do Milênio das Nações Unidas 2005. Investindo no Desenvolvimento: Um plano prático para atingir os Objetivos de Desenvolvimento do Milênio. Visão Geral.

2. No ano de 2011, o MDS estipulou o valor de renda familiar mensal de R\$ 70,00 por pessoa, como linha de miséria, quando do lançamento do Programa de Erradicação da Extrema Pobreza do Governo Federal. De acordo com esse critério, o estado do Ceará possuía 1.502.924 moradores residentes em domicílios com rendimentos mensais por pessoa que não ultrapassavam esse valor (IBGE, 2010).

capital cearense a 13<sup>a</sup> nesse grupo e, a 1<sup>a</sup> na região Nordeste<sup>3</sup>.

De uma maneira geral, ainda não há um completo consenso em nossa sociedade e nas diversas esferas de Governo em tratar esses temas de forma absolutamente prioritárias. As desigualdades sejam elas étnicas, etárias, de gênero ou espacial não só impedem que determinada população tenha acesso de forma democrática aos bens e serviços, como também determinam o sucesso das gerações futuras especialmente as crianças e jovens menos favorecidos. Na verdade, o custo social que se impõe com a presença de elevadas taxas de pobreza extrema e desigualdade em uma sociedade, acaba traduzindo-se também em um expressivo custo econômico.

Desse modo, o grande desafio em termos de proposição de políticas de combate a pobreza tem sido identificar ações que possam reverter seus níveis de forma mais acelerada. As políticas adotadas com essa finalidade concentram-se em dois grandes grupos: naquelas que possam estimular o crescimento econômico (no sentido do aumento da renda média da economia), e naquelas que visam reduzir as disparidades de renda entre os indivíduos.

Nesse sentido, vários estudos analisam a repercussão de alterações na renda média e na sua distribuição sobre medidas monetárias de pobreza, usando como medida, por exemplo, a proporção de indivíduos abaixo de um nível predeterminado de consumo e/ou renda. A título de exemplo, Meng et al. (2005), Adams (2004), Bruno, Ravallion e Squire (1998), e Chen e Ravallion (1997) estimam que em média um acréscimo de 10% na renda das economias levaria a um decréscimo na pobreza absoluta no intervalo de 20% a 30%. Ademais, diversos outros trabalhos indicam que o impacto do crescimento econômico sobre os níveis de pobreza seria tanto maior quanto menor fosse a desigualdade de renda existente nesses países (Lopez e Serven (2004), Son e Kakwani (2003), Bourguignon (2003) e Ravallion (1997)).

Para além desse escopo, a literatura em nível internacional vem recentemente apresentando consenso no desejo de combater a extrema pobreza infantil, o que tornou essa matéria de grande interesse entre os formuladores de políticas e pesquisadores. No entanto, não há relato de qualquer esforço anterior para comparar o grau e evolução da extrema pobreza infantil no Brasil ou em seus estados de baixa renda média, sendo essa, portanto, a primeira motivação para este estudo.

A segunda motivação é que a questão de como o crescimento econômico e a desigualdade de renda afetam a pobreza infantil encontra-se aberta ao debate. Assim, este trabalho busca contribuir para a literatura que investiga a relação

3. O documento "O Estado das Cidades do Mundo 2010/2011: Unindo o Urbano Dividido" também informou que o Brasil é o país com a maior distância social na América Latina.

triangular pobreza-crescimento-desigualdade fornecendo uma ampla investigação do estado do Ceará, que é um excelente caso a ser analisado. Pois, além de ser um Estado com pluralidades étnicas, raciais e culturais, apresentando regiões com distintas condições geográficas e econômicas, detêm taxas de extrema pobreza infantil bem superiores aos demais grupos etários.

Diante do exposto, a proposta dessa pesquisa é avaliar, a partir da modelagem proposta por Bourguignon (2003), a influência do crescimento econômico e da desigualdade de renda sobre a dinâmica da extrema pobreza para a população total, bem como para o grupo etário de 0 a 14 anos, no estado do Ceará ao longo das últimas duas décadas. Para tanto, utilizar-se-á os dados do Atlas do Desenvolvimento Humano no Brasil para cada um dos 184 municípios cearenses relativo aos anos de 1991, 2000 e 2010.

Além desta introdução, este artigo divide-se em mais seis seções. A seção seguinte faz uma apresentação da relação teórica entre pobreza, crescimento e desigualdade. A terceira seção exhibe os fatos estilizados da extrema pobreza total e infantil. Na quarta expõe-se a metodologia. Na quinta temos a análise dos resultados da pesquisa. Por fim, na sexta, são feitas as considerações finais do trabalho.

## 2 REFERENCIAL TEÓRICO

O aumento da renda média e a redução da desigualdade de distribuição em uma sociedade são reconhecidamente importantes mecanismos para combater a pobreza e/ou extrema pobreza. Essas relações podem ser mais perceptíveis quando consideramos a possibilidade de mensurar as privações dos indivíduos. Dentre os indicadores monetários, o consumo e a renda são geralmente os mais utilizados. O consumo tem suas vantagens na medida em que é mais fidedigno as condições reais de bem-estar de um indivíduo, uma vez que atende mais diretamente as suas necessidades básicas.

Por outro lado, a renda muitas vezes prevalece nos estudos de pobreza na medida em que as coletas de informações através de censos e pesquisas são feitas de forma periódica utilizando esse critério, já que além de apresentar menor custo, ela permite fazer comparações entre regiões. Além do mais, a renda permite fazer decomposições entre seus diversos componentes (renda do trabalho, juros, transferências, etc.) que são peças essenciais para análise econômica.

Os trabalhos seminais de Datt e Ravallion (1992) e Kakwani (1993) apontam como as alterações no nível de renda e na desigualdade podem afetar os níveis de privação de renda dos indivíduos. Isso permite, por exemplo,

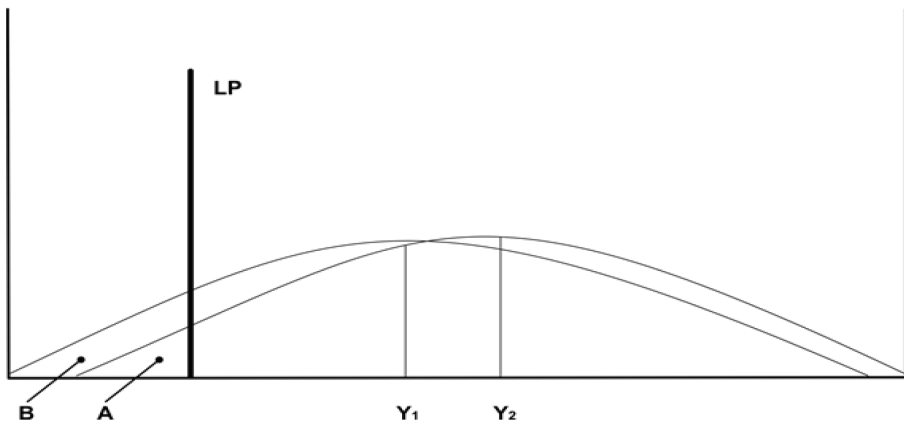
considerar como pobres indivíduos cuja renda é inferior a determinado valor monetário conhecido como *linha de pobreza*. Essa linha pode ser especificada a partir de uma determinada distribuição, cujos parâmetros estariam relacionados à renda média, variância e medida de assimetria. Assim, pode-se denotar  $f(y)$  como a função de distribuição da renda, onde  $y$  é o nível de renda do indivíduo e LP a linha de pobreza especificada.

Na Figura 1, mostra-se o deslocamento de  $f(y)$  provocado pela variação na renda média de  $Y_1$  para  $Y_2$ . Pode-se verificar que no primeiro período a medida de pobreza corresponde à soma das áreas (A + B), enquanto que (A) representa sua magnitude correspondente no segundo período. Considerando que o termo:

$$\int_0^{LP} f(y) dy \quad (1)$$

é a diferença entre as áreas dada por  $(A - B) < 0$ , percebe-se que há uma correlação negativa entre crescimento da renda e nível de pobreza o que implicaria evidentemente num impacto negativo do aumento da renda média na redução dos níveis de pobreza.

FIGURA 1 – Efeito de um aumento da renda média sobre a medida de pobreza

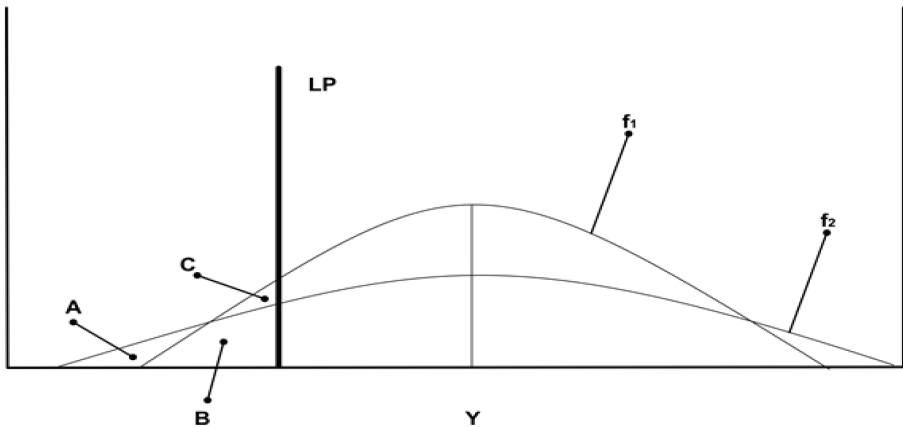


Fonte: Elaboração própria.

O efeito da variação da desigualdade graficamente ilustrado (Figura 2). As curvas  $f_1$  e sobre a pobreza pode também ser  $f_2$  representam as distribuições de renda antes e depois do aumento da dispersão, mantendo-se a renda média  $Y$  no mesmo nível. De acordo com a distribuição  $f_1$ , a medida de pobreza corresponderia à soma das áreas (C + B). No entanto, considerando,  $f_2$ , a incidência da pobreza seria a soma das áreas (A + B), representando então

num acréscimo no montante ( $A - C$ ). Nesse caso, para um dado nível de renda, maior desigualdade implicaria em um maior nível de pobreza absoluta.

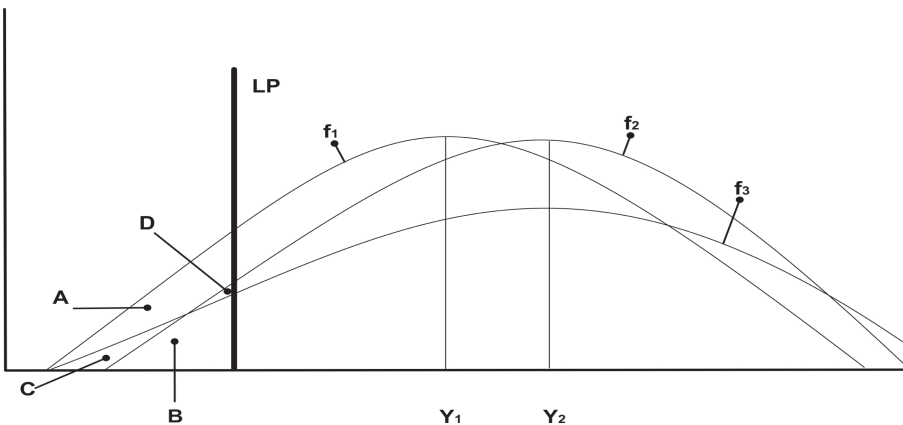
**FIGURA 2 – Efeito do aumento da concentração da renda sobre a pobreza**



Fonte: Elaboração própria.

Uma terceira abordagem seria verificar o efeito do nível da concentração de renda sobre o potencial que a expansão da renda teria na redução da pobreza absoluta. Esse efeito pode ser investigado no diagrama apresentado na Figura 3, onde  $f_1$ ,  $f_2$  e  $f_3$  são respectivamente, a distribuição original, a distribuição após a variação da renda média de  $Y_1$  para  $Y_2$  e a distribuição após a variação da renda média acompanhada de um aumento da concentração.

**FIGURA 3 – Efeito de um aumento da renda seguido de um aumento da concentração da renda sobre a medida de pobreza**



Fonte: Elaboração própria.

Constata-se que em  $f_1$ , a medida de pobreza corresponde a soma das áreas (A + B + C + D). No caso de  $f_2$ , essa medida é dada apenas por (B + D), de modo que a variação na pobreza corresponderia à área dada por (A + C). Examinando a distribuição  $f_3$ , verifica-se que o nível de pobreza no segundo período seria dado por (C + B), de modo que sua variação seria representada pelas somas das áreas (A + D). Na medida em que a área C é maior que a área D, pode-se verificar que o efeito da renda sobre a pobreza é atenuado quando a desigualdade aumenta.

Assim, utilizando uma função de distribuição da renda, podem-se construir algumas medidas de pobreza muito comuns nessa literatura. Estas medidas pertencem à classe proposta por Foster, Greer e Thorbecke (1984), podendo ser definida a partir da seguinte expressão:

$$P(\alpha) = \int_0^{LP} \left( \frac{LP-y}{LP} \right)^\alpha f(y) dy \quad (2)$$

onde, LP é a linha de pobreza.

O certo é que a partir de (2), pode-se derivar as medidas de pobreza conhecidas como P(0), P(1) e P(2), quando assume os valores 0, 1 e 2, respectivamente. Para o primeiro caso, tem-se a medida de incidência da pobreza que é simplesmente a razão entre o número de indivíduos vivendo em situação de pobreza e a população total, ou simplesmente a *proporção de pobres* P(0). No caso de  $\alpha=1$  tem-se o hiato da pobreza ou índice de Insuficiência de Renda, P(1), que mede a intensidade da pobreza e para  $\alpha = 2$ , o hiato da pobreza ao quadrado, P(2), que representa o grau de severidade. Entretanto, examinando com maior ênfase a proporção de pobres P(0), vê-se que esta depende exclusivamente dos parâmetros da função  $f(y)$ . Uma variação em P(0) pode ser resultado de uma variação na média e/ou nos demais parâmetros da distribuição. Considerando que a concentração de renda é função da variância e da assimetria de  $f(y)$ , pode-se assegurar que  $f(y)$  dependa da renda média e da medida de concentração. Desse modo, a magnitude do impacto da expansão da renda média de uma economia nos seus níveis de pobreza absoluta, definida neste contexto como elasticidade renda-pobreza, pode ser dado por:

$$\varepsilon = \frac{\partial P}{\partial \bar{y}} \frac{\bar{y}}{P} = \frac{\bar{y}}{P} \int_0^{LP} \left( f_{\bar{y}}(y) + f_G(y) \frac{dG}{d\bar{y}} \right) dy \quad (3)$$

onde,  $\bar{y}$  é a renda média, G é a medida de concentração da renda,  $f_y(y)$  é a



derivada da distribuição da renda em relação à  $\bar{y}$ ,  $f_G(y)$  é a derivada de  $f(y)$  em relação à desigualdade,  $G$ . Os valores de  $f_y(y)$  dependem dos valores de  $y$ . Caso assuma-se a validade da hipótese de Kuznets<sup>4</sup> (1955),  $(dG / dy)$  seria diferente de zero e deste modo não se conseguiria isolar o efeito exclusivo da renda sobre os níveis de pobreza, como imaginado na Figura 1. Todavia, tal hipótese não tem sido confirmada pela literatura empírica recente, como visto nos trabalhos de Deininger e Squire (1998), Chen e Ravallion (1997), Easterly (2000) e Dollar e Kraay (2002).

Assumindo assim a inexistência de uma relação a priori entre  $P(0)$  e (1), o valor absoluto do impacto da renda sobre a pobreza tende a ser negativo. Como visto na Figura 3, a magnitude desse impacto depende inversamente do nível da desigualdade de renda existente, uma vez que o valor absoluto de (1) é uma função decrescente da medida de concentração. Assim, o impacto do crescimento da renda sobre a pobreza é potencializado quanto menor os níveis de concentração de renda de uma sociedade.

No caso da influência da desigualdade sobre a pobreza absoluta esta pode ser medida na equação (4) a partir de (2).

$$\eta = \frac{\partial P}{\partial G} \frac{G}{P} = \frac{G}{P} \int_0^P \left( f_y(y) \frac{dy}{dG} + f_G(y) \right) dy \quad (4)$$

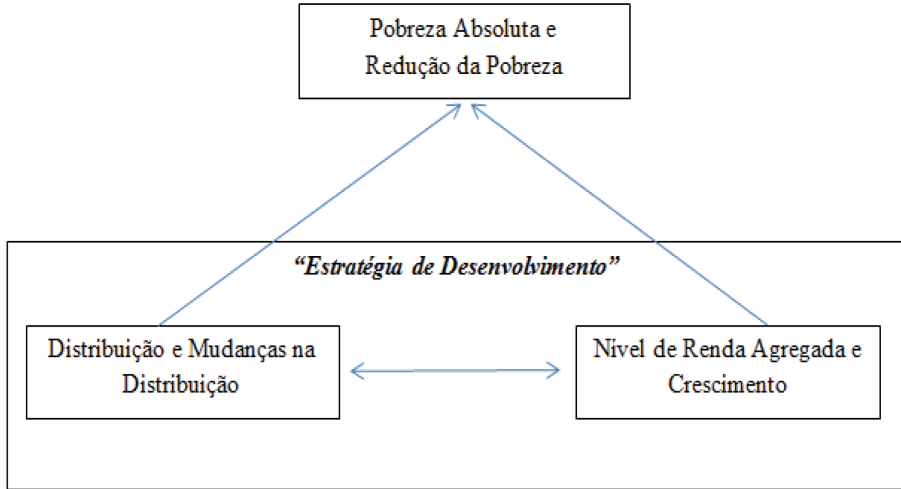
Supondo não haver influência da desigualdade na renda média, pode-se verificar a partir da expressão acima que a concentração de renda tem uma relação direta com os níveis de pobreza. Isto sugere, por exemplo, que um aumento da concentração de renda em uma economia levaria a maiores níveis de pobreza<sup>5</sup>.

Recentemente, o trabalho de Bourguignon (2003, 2004) ganhou notoriedade por formalizar matematicamente tais relações, e por propor o chamado “*Triângulo Pobreza-Crescimento-Desigualdade*”. A intuição de sua proposição é exposta abaixo na Figura 4.

4. A Hipótese de Kuznets considera que a distribuição de renda piora no estágio inicial de desenvolvimento quando a economia se movimenta da área rural em direção à industrialização, mas depois melhora devido à redução da participação do setor rural na economia.

5. Neste aspecto, a literatura também não tem apresentado uma relação conclusiva entre o efeito da desigualdade sobre o crescimento da renda. Por exemplo, enquanto Alesina e Rodrik (1994) e Perotti (1996) encontram que a desigualdade de renda é inversamente relacionada com o crescimento subsequente, Alesina e Perotti (1996), Li e Zou (1998) e Forbes (2000) mostram que a desigualdade de renda, medida pelo coeficiente de Gini, é positivamente relacionada com o crescimento econômico. Por outro lado, Barro (2000) encontra que essa relação pode ser positiva para economias ricas e, prejudicial para economias pobres.

FIGURA 4: O Triângulo Pobreza-Crescimento-Desigualdade



Fonte: Elaborado pelos autores. Adaptado de Bourguignon (2004).

Segundo o autor, o verdadeiro desafio, para o estabelecimento de uma estratégia de desenvolvimento para a redução da pobreza, está nas interações entre distribuição e crescimento, e não na relação entre pobreza e crescimento. Para ele, há pouca controvérsia entre os economistas de que o crescimento é essencial para a redução da pobreza (de renda) sob o pressuposto de que a distribuição de renda permaneça mais ou menos constante. Da mesma forma, muitas evidências sugerem que o agravamento da distribuição tende a aumentar a pobreza (Por exemplo, Ravallion, 1997; Easterly, 2000; Adams, 2004). No entanto, o autor afirma que a verdadeira questão no estabelecimento de uma estratégia de desenvolvimento é se o crescimento e a distribuição são independentes ou, fortemente inter-relacionadas.

Como extensão dessa análise, pode-se identificar também como a expansão da renda média da economia e a redução da desigualdade afetam as taxas de extrema pobreza para a população como um todo, bem como por faixa etária<sup>6</sup>. Mais especificamente, a faixa etária infantil. É conceitualmente apelativo que as melhorias na distribuição da renda de uma sociedade, e/ou o aumento da sua renda média, possibilitem, por exemplo, as crianças adquirirem importantes atributos de longo prazo em termos de produtividade quando na sua fase adulta.

6. Batana, Bussolo e Cockburn (2013) lançam a temática de estimar as taxas de extrema pobreza global por grupos etários: crianças, adultos e idosos. Os autores destacam que as taxas de extrema pobreza para crianças são muito superiores às estimadas para adultos e idosos.

A garantia de uma renda mínima para as crianças constitui-se numa meta universal para se atingir o desenvolvimento econômico de uma região. A falta de renda está diretamente ligada à insegurança alimentar, as más condições de higiene, a inexistência de água potável, a baixa ingestão de calorias e a más condições sanitárias em uma família.

Nesse sentido, Begum, Deng e Gustafsson (2012) descrevem a pobreza infantil em Bangladesh e na China, e investigam as razões para alterações ao longo do tempo, bem como as diferenças entre os dois países. Os autores definem uma criança como pobre se ele/ela vive em um domicílio com um rendimento de menos de um dólar por dia. O trabalho decompõe a pobreza infantil em três efeitos: renda média, distribuição de renda e circunstâncias demográficas. Os resultados mostram que o crescimento econômico pode reduzir a pobreza infantil, mas nem sempre o faz. Segundo os autores para entender as mudanças no tempo e entre países da medida de pobreza infantil, pode ser necessário considerar também mudanças/diferenças na distribuição da renda infantil, bem como na composição demográfica.

Assim, para compreender a dinâmica da extrema pobreza infantil nos municípios cearenses, deve-se levar em consideração a perspectiva do crescimento econômico bem como da desigualdade de renda. Desse modo, torna-se importante avaliar em que medida a queda recente da desigualdade e do aumento da renda média no Ceará, traduziu-se na redução da extrema pobreza (total e infantil) da população nos diversos municípios cearenses<sup>7</sup>.

### 3 A EXTREMA POBREZA NOS MUNICÍPIOS CEARENSES

Esta seção tem por objetivo analisar a distribuição municipal da extrema pobreza (total e infantil) no estado do Ceará referente ao período de 1991 a 2010, correlacionando-a com indicadores relativos à renda domiciliar *per capita* e a desigualdade de renda dos municípios cearenses, recorrendo-se para tanto aos dados disponíveis no Atlas do Desenvolvimento Humano no Brasil.

Nesse sentido, a Tabela 1 mostra o percentual da população total em situação de extrema pobreza, assim como para o grupo demográfico com idade

7. Apesar de alguns trabalhos estimarem elasticidades renda e desigualdade da pobreza infantil considerando apenas a renda média e a desigualdade de famílias com crianças, este trabalho optou pelo uso da renda média total e desigualdade total da população por considerar a existência de externalidades. Por exemplo, o incremento da renda de famílias sem crianças pode gerar efeitos transbordamento para famílias com crianças através do mercado de trabalho pelo aumento da demanda por bens e serviços. Outro ponto é que o aumento da renda média pode contribuir para uma maior arrecadação de tributos, ampliando a capacidade do Estado oferecer serviços de educação e saúde para as crianças mais pobres.

entre 0 e 14 anos, nas duas últimas décadas. Ressalta-se que a taxa de extrema pobreza total corresponde à proporção de pessoas com renda domiciliar<sup>8</sup> *per capita* de até R\$ 70,00 mensais, em reais de agosto de 2010. Por sua vez, a taxa de extrema pobreza infantil refere-se à proporção dos indivíduos com até 14 anos de idade que possuem essa renda domiciliar per capita mensal.

**TABELA 1: Percentual de população em situação de extrema pobreza total e extrema pobreza infantil – Ceará – 1991/2010**

ANO	Taxa de extrema pobreza Total	Taxa de extrema pobreza Infantil
1991	39,76	50,76
2000	28,11	39,85
2010	14,69	22,38

Fonte: PNUD (2013). Elaboração dos autores.

Observando a mencionada tabela, verifica-se que o percentual da população em situação de miséria diminuiu no período estudado no Ceará, especificamente tinha-se um valor de 39,76% em 1991 passando para 28,11% no ano 2000 e 14,69% em 2010, representando uma redução relativa de 63,05% entre 1991 e 2010.

No tocante à extrema pobreza infantil também se constatou um arrefecimento na referida taxa, saindo de 50,76% em 1991 para 39,85% em 2000, atingindo a marca de 22,38% no ano de 2010, equivalendo a uma diminuição relativa de 55,91% entre o ano inicial e final da série pesquisada.

Deste modo, estes resultados evidenciam uma melhora nas condições de vida da população do Estado nos últimos decênios. Não obstante, percebe-se que as taxas de extrema pobreza continuavam elevadas em 2010, sobretudo a extrema pobreza infantil que deteve um valor bem superior ao índice registrado para o total da população<sup>9</sup>.

Conforme IPECE (2011), parcela significativa dos indivíduos em situação de miséria no Estado é composta por pessoas com idade entre 0 e 14 anos, mais precisamente tinha-se em 2010 um montante de 583.865 crianças nesta condição, significando uma participação de 38,84% do total de extremamente pobres neste ano. França et al. (2015) corrobora ao citar que domicílios que possuem uma maior quantidade de crianças são mais vulneráveis a vivenciar uma situação de extrema pobreza.

Segundo Souza (2011), a significativa diminuição da extrema pobreza

8. O universo de indivíduos é limitado àqueles que vivem em domicílios particulares permanentes, que são destinados exclusivamente para habitação de uma ou mais pessoas.

9. Ressalta-se que o grupo etário de 0 a 14 anos perdeu participação ao total populacional do Estado na última década, passando de 33,5% para 25,9% (IBGE, 2010).

no país, e consequentemente no Ceará, pode está associada a políticas de assistência à família, como por exemplo, o programa Bolsa Família. Para o autor, o Bolsa Família tem sido efetivo em focalizar as transferências de renda para as famílias mais pobres, mas, por outro lado, não tão efetivo em estimular de maneira significativa à acumulação de capital humano das novas gerações, no caso crianças de 0 a 14 anos.

Neste contexto, torna-se importante empreender um olhar específico para este grupo demográfico, ou seja, avaliar a taxa de extrema pobreza infantil em complemento a análise da extrema pobreza total.

Destaca-se que quando são estudados os índices de extrema pobreza em nível municipal constata-se que a distribuição não é homogênea, tendo cidades com altos valores e outras com baixas taxas. Nesse sentido, a Tabela 2 apresenta os municípios com as dez maiores e dez menores taxas de extrema pobreza (total e infantil) para o ano de 2010.

**Tabela 2:** Municípios com as dez maiores e menores proporções de pessoas extremamente pobres (total e infantil) - 2010

Município	% de extremamente pobres - Total	Município	% de extremamente pobres - Infantil
<b>10 Maiores</b>			
Granja	43,63	Ipaporanga	59,56
Ipaporanga	42,11	Granja	57,16
Miraíma	41,87	Croatá	55,34
Choró	41,75	Choró	54,00
Croatá	41,08	Miraíma	53,91
Santana do Acaraú	39,10	Tarrafas	53,19
Graça	37,99	Graça	51,72
Amontada	37,41	Santana do Acaraú	50,50
Viçosa do Ceará	36,84	Araripe	50,23
Araripe	36,68	Ararendá	50,15
<b>10 menores</b>			
Limoeiro do Norte	7,98	Pacajus	11,83
Pacajus	7,96	Maranguape	11,45
Russas	7,55	Limoeiro do Norte	11,29
Horizonte	7,51	Horizonte	10,64
Juazeiro do Norte	6,73	Juazeiro do Norte	10,55
Pacatuba	6,66	Pacatuba	10,44
Caucaia	6,53	Caucaia	10,17
Eusébio	5,60	Eusébio	8,78
Maracanaú	4,59	Maracanaú	7,29
Fortaleza	3,36	Fortaleza	5,73

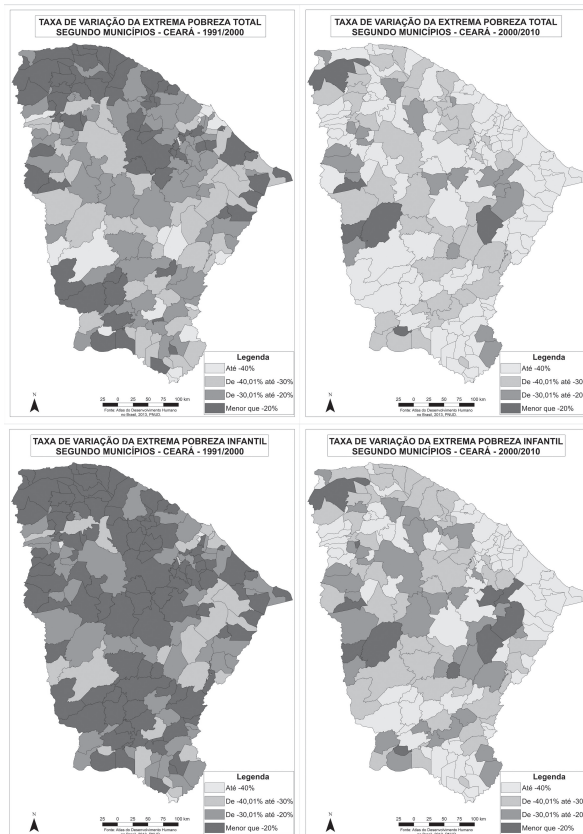
Fonte: PNUD (2013). Elaboração dos autores.

A Figura 5 exibe a taxa de variação da extrema pobreza (total e infantil) concernente aos períodos 1991/2000 e 2000/2010 para os municípios do Estado, observando-se que na década 2000/2010 houve uma maior redução das taxas, uma vez que na última década anotou-se um maior quantitativo de municípios que diminuíram a taxa de extrema pobreza em mais de 40% (cor cinza clara na legenda dos mapas).

Este resultado provavelmente é oriundo da diminuição da desigualdade de renda registrada no período 2000/2010 aferida por meio do índice de Gini, enquanto que na década 1991/2000 houve elevação da mesma.

De acordo com Oliveira et al. (2014), os programas sociais de transferência de renda sofreram expressiva expansão no decênio de 2000 no Ceará, contribuindo para a redução da desigualdade de renda no Estado.

**FIGURA 5: Mapas com a variação da taxa de extrema pobreza total e infantil dos municípios cearenses para os períodos de 1991/2000 e 2000/2010**

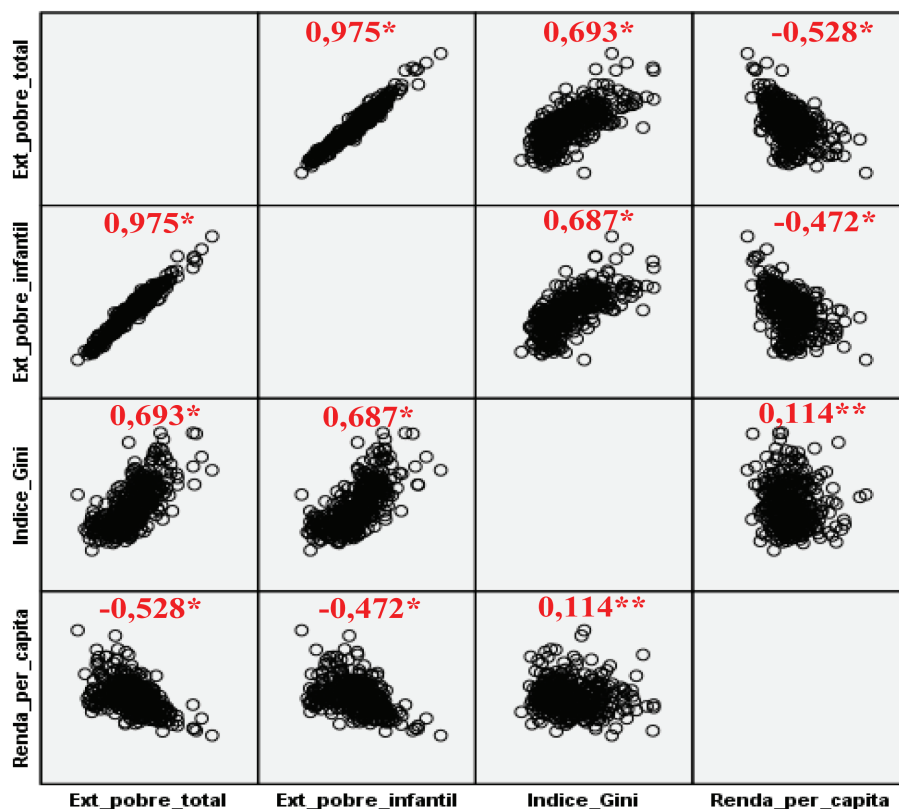


Fonte: PNUD (2013). Elaboração dos autores.

Nessa perspectiva, tendo por escopo avaliar a associação linear entre a taxa de variação, para os intervalos temporais de 1991/2000 e 2000/2010, da extrema pobreza dos municípios cearenses (total e infantil), com a renda domiciliar *per capita* e o índice de desigualdade de renda de Gini gerou-se o gráfico de dispersão e as correlações de Pearson, apresentadas na Figura 6.

Ressalta-se que Bourguignon (2004) traz uma ampla discussão da relação entre essas três variáveis ao apresentar o “*Triângulo Pobreza-Crescimento-Desigualdade*”, a qual foi detalhada anteriormente na seção atinente ao referencial teórico.

FIGURA 6: Matriz de correlação de Pearson para a taxa de variação nos períodos 1991/2000 e 2000/2010 dos indicadores relacionados à taxa de extrema pobreza (total e infantil), índice de Gini e renda domiciliar *per capita*



Fonte: Elaboração dos autores. \* Significativo a 1%. \*\* Significativo a 5%.

Analisando a mencionada figura, verifica-se que há uma forte correlação positiva e significativa para as variáveis da taxa de extrema pobreza total e infantil, ou seja, municípios com elevada taxa de extrema pobreza total também possuem alto índice para a taxa de extrema pobreza infantil. Averiguou-se, ainda, uma correlação negativa entre a taxa de extrema pobreza (total e infantil) e a renda domiciliar *per capita*, isto é, municípios com altos valores de renda *per capita* tendem a possuir baixa proporção de indigentes.

Outro resultado evidenciado é a associação positiva entre o índice de Gini e as taxas de extrema pobreza, sinalizando que políticas que visem à redução da desigualdade de renda podem contribuir para a diminuição da extrema pobreza dos municípios.

Nesta conjuntura, cabe o questionamento de quais são as razões que poderiam contribuir para esse resultado, bem como definir o grau de importância destas variáveis, em cada um dos municípios do Estado, em explicar a extrema pobreza das cidades cearenses, sendo fundamental o emprego de técnicas econométricas às quais serão abordadas na próxima seção.

#### 4.1 Dados

A base de dados usada nesta pesquisa corresponde ao Atlas do Desenvolvimento Humano no Brasil<sup>10</sup>, elaborado pelo Programa das Nações Unidas para o Desenvolvimento (PNUD) em parceria com a Fundação João Pinheiro (FJP) e o Instituto de Pesquisa Econômica Aplicada (IPEA). Foram utilizados os seguintes indicadores para cada um dos 184 municípios cearenses atinente aos anos de 1991, 2000 e 2010:

- **Taxa de extrema pobreza total:** Consiste no percentual dos indivíduos com renda domiciliar *per capita* igual ou inferior a R\$ 70,00 mensais, em reais de agosto de 2010, caracterizando a população extremamente pobre ou em situação de miséria. O universo de indivíduos é limitado àqueles que vivem em domicílios particulares permanentes.

- **Taxa de extrema pobreza infantil:** Corresponde ao percentual dos indivíduos com até 14 anos de idade que têm renda domiciliar *per capita* igual ou inferior a R\$ 70,00 mensais, em reais de agosto de 2010, distinguindo a população infantil extremamente pobre ou em situação de miséria. O universo de indivíduos é limitado àqueles que vivem em domicílios particulares permanentes.

---

10. Acesse em: (<http://www.atlasbrasil.org.br/2013/>).



**Renda *per capita* (R\$):** Razão entre o somatório da renda de todos os indivíduos residentes em domicílios particulares permanentes e o número total desses indivíduos. Valores em reais de agosto de 2010.

**Índice de Gini:** Mede o grau de desigualdade existente na distribuição de indivíduos segundo a renda domiciliar *per capita*. Seu valor varia de 0 (quando não há desigualdade) a 1 (quando a desigualdade é máxima). O universo de indivíduos é limitado àqueles que vivem em domicílios particulares permanentes.

A variável dependente corresponde à taxa de extrema pobreza (total e infantil) enquanto que as demais são variáveis explicativas. Ressalta-se que no modelo econométrico foi utilizada a taxa de variação destas variáveis, alusiva aos períodos de 1991/2000 e 2000/2010.

O Quadro 1 apresenta as estatísticas descritivas para os indicadores concernente aos anos de 1991, 2000 e 2010, tendo como unidade de observação os 184 municípios cearenses. Verifica-se uma evolução da taxa de extrema pobreza no período estudado, uma vez que a média reduziu-se significativamente tanto para a extrema pobreza total quanto para à infantil.

**QUADRO 1: Estatísticas descritivas para a taxa de extrema pobreza (total e infantil) dos municípios cearenses - 1991, 2000 e 2010**

Ano	Média	Mínimo	Máximo	Desvio-Padrão
<b>Taxa de extrema pobreza total (%)</b>				
1991	53,98	15,25	82,72	12,81
2000	41,05	9,02	60,95	10,63
2010	23,48	3,36	43,63	8,83
<b>Taxa de extrema pobreza Infantil (%)</b>				
1991	64,88	22,24	88,78	12,77
2000	54,27	14,69	75,85	12,36
2010	33,38	5,73	59,56	11,64
<b>Renda domiciliar <i>per capita</i> (R\$)</b>				
1991	113,29	44,65	457,04	42,78
2000	163,46	86,97	610,48	55,64
2010	267,64	171,62	846,36	76,34
<b>Índice de Gini (0 – 1)</b>				
1991	0,5252	0,3800	0,7100	0,0638
2000	0,5896	0,4700	0,7100	0,0499
2010	0,5278	0,4200	0,6600	0,0459

Fonte: PNUD (2013). Elaboração dos autores.

Não obstante, ainda há uma grande variabilidade nos municípios, conforme as estatísticas do desvio-padrão, valor máximo e mínimo. Tinha-se, por exemplo, cidades no ano 2010 com taxa de extrema pobreza total superior a 40% enquanto que outras possuíam menos de 5%.

Quanto à renda *per capita*, a mesma registrou aumento no valor médio nas duas últimas décadas, embora também tenha ocorrido um incremento da variabilidade, evidenciando que o crescimento da renda não foi homogêneo nos municípios.

Por sua vez, o Índice de Gini teve um comportamento peculiar. Constatou-se um aumento da desigualdade dos municípios no período 1991/2000, passando a diminuir no intervalo 2000/2010, atingindo em 2010 um valor médio do índice de Gini ainda superior ao ano de 1991.

#### 4.2 Modelo Econométrico

A partir do nexos *Pobreza-Crescimento-Desigualdade* derivado por Bourguignon (2003), vários trabalhos usaram sua intuição para estimar a relação entre essas variáveis (Kalwij e Verschoor, 2007; Fosu, 2009, 2015). Este modelo de “identidade”, com base em uma aproximação para uma distribuição de renda assumida lognormal, nos permite explicar a heterogeneidade desse nexos entre municípios e períodos de tempo. O modelo de Bourguignon completo pode ser apresentado como:

$$\Delta \ln P_{it} = \beta_1 + \beta_2 \Delta \ln Y_{it} + \beta_3 \Delta \ln Y_{it} G^I + \beta_4 \Delta \ln Y_{it} \left(\frac{Z}{Y}\right) + \beta_5 \Delta \ln G_{it} + \beta_6 \Delta \ln G_{it} G^I + \beta_7 \Delta \ln G_{it} \left(\frac{Z}{Y}\right) + \alpha_i + \varepsilon_{it} \quad (5)$$

onde:

$$i = 1, 2, \dots, 184.$$

$$t = 1991, 2000, 2010.$$

A variável dependente,  $\Delta \ln P_{it}$ , é representada pela variação logarítmica da taxa de pobreza (total e infantil). As variáveis explicativas são:  $\Delta \ln Y_{it}$ , a variação logarítmica da renda média domiciliar per capita;  $\Delta \ln G_{it}$ , a variação logarítmica do coeficiente de Gini;  $G^I$ , o índice de Gini do período inicial; e  $\left(\frac{Z}{Y}\right)$ , a proporção da linha de pobreza  $Z$  em relação à renda  $Y_t$ .

O termo  $\alpha_i$  são os efeitos idiossincráticos (heterogeneidade não observada de cada município que não variam no tempo) associados a cada município e

que podem afetar de alguma maneira os níveis de pobreza, como por exemplo, as características naturais, as condições climáticas, estrutura econômica, etc. Assim,  $\alpha_i$  pode ser entendido como um representante de variáveis omitidas, que apesar de não observáveis, afetam a pobreza local. Estes efeitos podem ser tratados como fixos ou como resultados de um processo aleatório. Dito isto, o viés de omissão de variáveis não observáveis pode ser tratado via estimação de efeitos aleatórios (EA) ou efeitos fixos (EF), conseqüentemente, o termo  $\alpha_i$  pode ser visto como uma variável aleatória, ou como um parâmetro a ser estimado para cada observação  $i$ .

A técnica de efeitos fixos remove  $\alpha_i$  antes da estimação, sendo uma aplicação de Mínimos Quadrados Ordinários sobre dados transformados (desvios da média no tempo).

A transformação se dá subtraindo de cada variável a sua própria média temporal, dessa forma acaba-se eliminando a heterogeneidade não observada do município, pois esta não varia no tempo, e as estimativas assim obtidas serão eficientes e consistentes.

Outra possibilidade seria supor a não correlação entre a heterogeneidade não observada e as variáveis explicativas do modelo, nesse caso deve-se utilizar a técnica de efeitos aleatórios, este caso ocorreria quando a diferença entre os  $\alpha_i$  fosse estatisticamente insignificante. Por fim, o componente  $\varepsilon_{it}$  é um termo aleatório com média zero e variância constante. Os subscritos  $i$  e  $t$  correspondem ao município e ao período de tempo da observação, respectivamente.

Há uma série de vantagens em se utilizar dados em painel, entre elas: o número maior de observações (graus de liberdade), elevando o nível de informação da amostra, reduz-se o risco de multicolinearidade, aumenta a eficiência, permite ajustamentos dinâmicos, etc. Com respeito aos métodos de estimação, uma das desvantagens do uso do método de efeitos fixos (EF) está associada à impossibilidade de se considerar variáveis que não variam no tempo, pois estas são removidas antes do processo de estimação. Por outro lado, se existirem razões para acreditar que os efeitos individuais dos municípios estão correlacionados com as variáveis explicativas o método de efeitos fixos é mais apropriado, uma vez que, os estimadores dos efeitos aleatórios são inconsistentes nessa situação. Devido o fato de ambas as técnicas apresentarem vantagens e desvantagens uma sobre a outra, os resultados serão apresentados a partir das duas técnicas, acompanhados do teste de Hausman, que é um teste formal que permite escolher o método de estimação mais adequado para cada situação.

#### 4.2.1 Elasticidades Renda-Pobreza Extrema e Desigualdade-Pobreza Extrema

Com base na equação (5), pode-se obter as elasticidades parciais renda-pobreza extrema e desigualdade-pobreza extrema dos municípios, respectivamente como:

$$E_{yp}^{it} = \frac{\partial \Delta \ln P_{it}}{\partial \Delta \ln Y_{it}} = \beta_2 + \beta_3 G^I + \beta_4 \left( \frac{Z}{Y} \right) \quad (6)$$

$$E_{gp}^{it} = \frac{\partial \Delta \ln P_{it}}{\partial \Delta \ln G_{it}} = \beta_5 + \beta_6 G^I + \beta_7 \left( \frac{Z}{Y} \right) \quad (7)$$

A equação (6) mostra a sensibilidade da extrema pobreza com respeito à renda. O sinal de  $\beta_2$  previsto ser negativo: um aumento no crescimento da renda deve reduzir o crescimento da extrema pobreza, *ceteris paribus*. Em contraste,  $\beta_3$  é esperado ser positivo, para um nível mais elevado de desigualdade inicial deve diminuir a taxa em que a aceleração do crescimento é transformada em redução da pobreza. O sinal de  $\beta_4$  deve ser positivo, bem como, consistente com a hipótese que uma renda maior (em relação à linha de pobreza) teria associado com ela uma elasticidade de renda mais elevada.

Já a equação (7) nos permite verificar a sensibilidade da extrema pobreza com respeito à desigualdade, O sinal de  $\beta_5$  é teoricamente positivo, pois se espera que para uma piora na distribuição de renda tenha-se um aumento da pobreza, *ceteris paribus*. Em contrapartida,  $\beta_6$  deverá ser negativo, dado um efeito cada vez menor da desigualdade inicial proveniente de aumentos da pobreza em decorrência de aumentos da desigualdade. O sinal de  $\beta_7$  também deverá ser negativo, em uma economia de relativa baixa renda (alto  $15 Z/Y$ ), ao melhorar a distribuição de renda (reduzindo G) pode-se aumentar a probabilidade de que mais pessoas caiam na pobreza.

Assim, espera-se que a redução da pobreza seja maior quando: (i) o crescimento da renda seja maior, (ii) o declínio na desigualdade seja maior, (iii) a desigualdade inicial seja inferior ou (iv) a proporção da linha de pobreza em relação à renda seja menor (maior renda). Além disso, os efeitos sobre a pobreza, tanto do crescimento da renda, como da diminuição da desigualdade são esperados serem mais elevados quando a desigualdade inicial for menor e a proporção da linha de pobreza em relação à renda seja menor. Estes dois últimos efeitos, portanto, trabalham através das elasticidades renda e desigualdade da extrema pobreza. Isto é, ambas as elasticidades diminuiria com o aumento da desigualdade inicial e com o aumento da proporção da linha de pobreza em relação à renda; daí, a desigualdade inicial desempenha um papel a parte neste

processo de transformação de crescimento em redução da pobreza.

## 5 ESTIMAÇÕES E RESULTADOS

Nesta seção expõem-se os resultados das estimativas realizadas a partir da equação (5) para as medidas de extrema pobreza total e infantil. As estimativas estão dispostas nas Tabelas 3 e 4, respectivamente. Pelos motivos já expostos as regressões foram realizadas por efeitos fixos como por efeitos aleatórios. Os resultados do teste de Hausman encontram-se na última linha das tabelas, caso o valor  $p$  seja baixo (inferior a 5%), como ocorreu, indica que o método mais adequado é o de efeitos fixos, uma vez que se rejeita a hipótese nula do teste, de que o modelo deve ser estimado considerando efeitos aleatórios. A Tabela 4 mostra que os resultados para a extrema pobreza infantil não diferem, no que se refere aos sinais dos coeficientes, aos da extrema pobreza total.

Verifica-se que todos os coeficientes de inclinação parciais são significativos a um nível de significância consideravelmente pequeno, além de apresentarem sinais de acordo com o esperado. Com respeito às elasticidades não se pode inferir diretamente dos resultados da regressão, pois como explicado, esta depende tanto do nível de desigualdade no período inicial do município que está sendo considerado, como da relação entre a linha de pobreza e a renda média. Pode-se perceber isso pelos termos interativos que entraram na regressão para captar essa não linearidade do efeito da renda e da desigualdade sobre a extrema pobreza.

**TABELA 3: Resultados das estimações por Efeitos Fixos e Aleatórios para a Extrema Pobreza Total**

Extrema Pobreza Total	Efeitos Fixos			Efeitos Aleatórios		
	Coef.	t	P>t	Coef.	z	P>z
Renda	<b>-2,247</b>	-6,05	0,000	-1,138	-7,30	0,000
Renda*Gini <sub>t0</sub>	<b>2,232</b>	3,44	0,001	-0,076	-0,28	0,780
Renda*(z/y)	<b>0,524</b>	2,50	0,013	1,524	9,35	0,000
Gini	<b>4,173</b>	9,24	0,000	2,702	5,56	0,000
Gini*Gini <sub>t0</sub>	<b>-2,298</b>	-3,09	0,002	-1,552	-1,92	0,055
Gini*(z/y)	<b>-5,597</b>	-13,57	0,000	-2,865	-7,36	0,000
Constante	<b>0,013</b>	0,59	0,553	-0,132	-6,05	0,000
N	368			368		
R <sup>2</sup> : within	0,88			0,82		
Prob>F	0,000			Prob>chi <sup>2</sup> = 0,000		
Hausman (Prob>chi <sup>2</sup> )	0,000					

Fonte: Elaboração dos autores. Obs.: O negrito na segunda coluna indica o melhor método com base no teste de Hausman.

Em termos dos efeitos interativos entre Renda e Gini, os coeficientes estimados nos permitem observar que a extrema pobreza infantil (2,36) mostrou-se mais sensível que a extrema pobreza total (2,23) aos níveis de desigualdade inicial nos municípios cearenses, demonstrando que para um nível mais elevado de desigualdade inicial a taxa em que a aceleração do crescimento é transformada em redução da pobreza deve diminuir.

**TABELA 3: Resultados das estimações por Efeitos Fixos e Aleatórios para a Extrema Pobreza Infantil**

Extrema Pobreza Infantil	Efeitos Fixos			Efeitos Aleatórios		
	Coef.	t	P>t	Coef.	z	P>z
<b>Renda</b>	<b>-2,371</b>	-6,140	0,000	-1,331	-8,370	0,000
<b>Renda*Gini<sub>10</sub></b>	<b>2,362</b>	3,500	0,001	0,319	1,140	0,253
<b>Renda*(z/y)</b>	<b>1,052</b>	4,820	0,000	1,787	10,92	0,000
<b>Gini</b>	<b>3,928</b>	8,360	0,000	2,590	5,400	0,000
<b>Gini*Gini<sub>10</sub></b>	<b>-1,879</b>	-2,430	0,016	-1,257	-1,570	0,116
<b>Gini*(z/y)</b>	<b>-5,694</b>	-13,27	0,000	-3,099	-8,070	0,000
<b>Constante</b>	<b>0,044</b>	1,890	0,060	-0,090	-4,130	0,000
<b>N</b>	368			368		
<b>R<sup>2</sup>: within</b>	0,87			0,82		
<b>Prob&gt;F</b>	0,000			<b>Prob&gt;chi<sup>2</sup> = 0,000</b>		
<b>Hausman (Prob&gt;chi<sup>2</sup>)</b>	0,000					

Fonte: Elaboração dos autores. Obs.: O negrito na segunda coluna indica o melhor método com base no teste de Hausman.

Com base nas estimativas dos modelos propostos, foi possível calcular para todos os municípios cearenses, as seguintes elasticidades: [1] Elasticidade Renda da Extrema Pobreza Total (ER\_TOT); [2] Elasticidade Desigualdade da Extrema Pobreza Total (ED\_TOT); [3] Elasticidade Renda da Extrema Pobreza Infantil (ER\_INF); e [4] Elasticidade Desigualdade da Extrema Pobreza Infantil (ED\_INF).

Em conformidade com Bourguignon (2003), todos os valores da Elasticidade Renda da Extrema Pobreza foram negativos, indicando que elevações na renda domiciliar *per capita* provocam reduções no percentual de extremamente pobres (seja considerando a população total ou a infantil). Por outro lado, todos os valores da Elasticidade Desigualdade da Extrema Pobreza foram positivos, denotando que reduções na desigualdade (medida pelo Índice de Gini) geram reduções na proporção de indigentes.

O Quadro 2 exhibe estatísticas descritivas referentes às elasticidades esti-

madras para os municípios cearenses, verificando-se que os sinais das médias e das medianas são compatíveis com o que seria esperado, além de evidenciar uma simetria e normalidade da distribuição.

Ademais, as dispersões dos valores dos municípios não foram muito elevadas, conforme indicam os coeficientes de variação calculados. Assim, há diferenças entre os municípios em termos dos valores das elasticidades, mas essas diferenças não tenderam a ser extremamente altas (i.e., não há *outliers*).

#### QUADRO 2: Estatísticas descritivas referentes às elasticidades estimadas para os municípios cearenses

Estatística	ER_TOT	ER_INF	ED_TOT	ED_INF
Média	-0,9304	-0,8395	1,4199	1,3680
Mediana	-0,9362	-0,8388	1,4285	1,3683
Desvio Padrão	0,1367	0,1452	0,3151	0,3170
Coefficiente de Variação	0,1470*	0,1730*	0,2219	0,2317

Fonte: PNUD (2013). Elaborado pelos autores. Notas: \* Calculado considerando o valor absoluto da média.

As elasticidades estimadas também se mostraram significativamente correlacionadas entre si (Quadro 3). Mais especificamente, de forma mais notável, tanto as elasticidades de desigualdade como as de renda entre a população total e a infantil apresentaram coeficientes de correlação positivos e muito próximos a 1, indicando que os municípios que apresentaram valores mais elevados (em módulo) das elasticidades no caso da população total, também, tenderam a apresentar (com elevado grau de previsibilidade) os mais relevantes no caso da população infantil. Já no caso das comparações entre diferentes tipos de elasticidades para os mesmos grupos populacionais, os coeficientes foram negativos (o que já seria esperado) e a intensidade das relações foi bem menos relevante.

#### QUADRO 3: Matriz de correlação de Pearson entre as elasticidades estimadas para os municípios cearenses

	ED_INF	ED_TOT	ER_INF	ER_TOT
ED_INF	1,000	0,996*	-0,469*	-
ED_TOT	0,996*	1,000	-	-0,371*
ER_INF	-0,469*	-	1,000	0,980*
ER_TOT	-	-0,371*	0,980*	1,000

Fonte: PNUD (2013). Elaborado pelos autores. Notas: \* Correlação significativa a 1% (teste bicaudal). \*\* As correlações irrelevantes para a pesquisa não foram reportadas.

Então, considerando novamente as médias das elasticidades é possível obter duas conclusões importantes. Primeiramente, constatou-se que tanto para a Elasticidade Renda como para a Elasticidade Desigualdade, o módulo de suas médias foi maior para a população total do que para a infantil. Isso foi confirmado por meio de testes de diferença de médias (com amostras pareadas), o que produziu estatísticas *t* respectivamente iguais a 42,034 (Renda) e 24,326 (Desigualdade), significantes a 1%. Mais especificamente, na comparação das Elasticidades Renda, todos os municípios apresentaram maiores valores (em módulo) quando se considera a população total em relação à infantil. Já na comparação das Elasticidades de Desigualdade, apenas 10 municípios apresentaram maiores valores considerando a população infantil.

Outra conclusão importante diz respeito à comparação entre os valores dos diferentes tipos de elasticidade considerando os mesmos grupos populacionais. No caso, verifica-se que, em módulo, as médias das Elasticidades de Desigualdade são maiores que as das Elasticidades de Renda, tanto para a população total como para a infantil. Mais uma vez, testes de diferença de médias (com amostras pareadas) confirmaram essa percepção, uma vez que produziram estatísticas *t* respectivamente iguais a 22,643 (população total) e 25,606 (população infantil), significantes a 1%. Considerando-se a população total, tem-se de forma mais específica que apenas 8 municípios apresentaram Elasticidade Renda maior (em módulo). Concomitantemente, isto ocorreu em apenas 7 municípios no caso da população infantil.

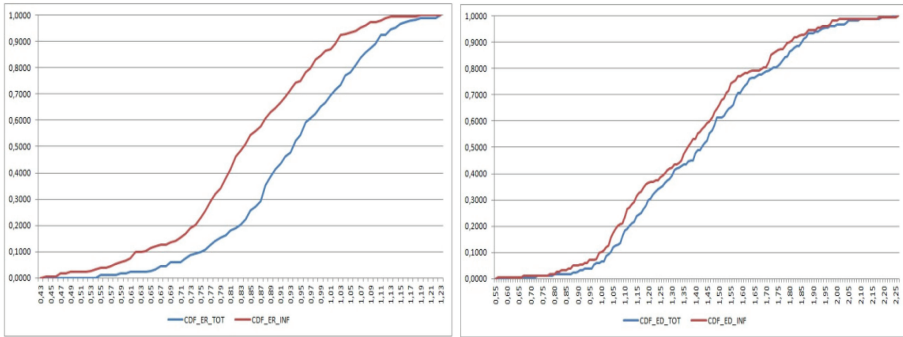
Uma maneira alternativa de se considerar essas mesmas questões seria por meio do conceito de dominância estocástica (LEVY e KROLL, 1979). No caso, a Figura 7, ao se efetuar a comparação das funções de densidade cumulativas (CDFs) das elasticidades, mostra-se que houve dominância estocástica de primeira ordem no caso da renda e de segunda ordem<sup>11</sup> no caso da desigualdade por parte das distribuições da população total em relação à infantil. Já na comparação entre as CDFs para os mesmos grupos populacionais (ver a Figura 8), verificou-se que existe dominância estocástica de primeira ordem em ambos os casos em favor da elasticidade de desigualdade.

---

11. Considerando a metodologia delineada por Levy e Kroll (1979), a dominância só não foi de primeira ordem por causa do quantil mais elevado, em que o valor da população infantil superou o da total.

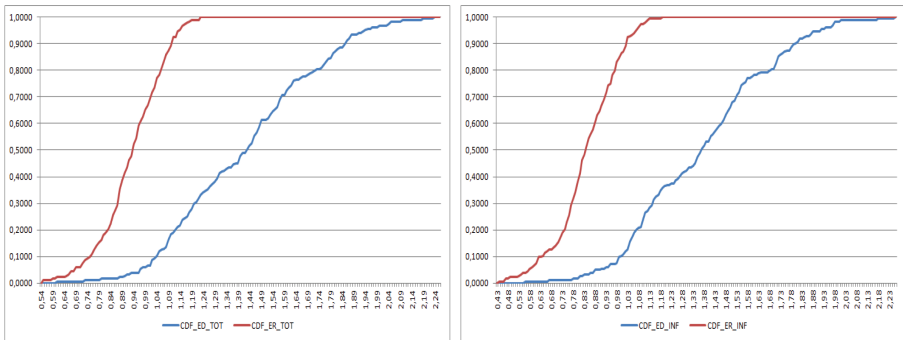


**FIGURA 7: Funções de densidade cumulativas (CDFs) das elasticidades para os municípios cearenses (2010) – Comparação entre a população total e a infantil**



Fonte: PNUD (2013). Elaborado pelos autores.

**FIGURA 8: Funções de densidade cumulativas (CDFs) das elasticidades para os municípios cearenses (2010) – Comparação entre desigualdade e renda**



Fonte: PNUD (2013). Elaborado pelos autores.

Então, como seria possível qualificar melhor essas conclusões? No caso, tem-se em primeiro lugar que os efeitos das políticas públicas, tanto de crescimento econômico como as que visam à redução da desigualdade, tendem a ter um efeito mais significativo sobre a redução da extrema pobreza total. Ou seja, para as mesmas variações da renda domiciliar *per capita* e do Índice de Gini, os dados sugerem que a redução da extrema pobreza infantil seria menor que a total, i.e., a extrema pobreza infantil seria mais persistente e difícil de mitigar.

O outro *insight* fundamental advindo das análises efetuadas seria que as

políticas que focam na diminuição da desigualdade tenderiam a ser mais efetivas na redução da extrema pobreza, tanto quando se considera a população total como a infantil. Em outras palavras, mantendo-se tudo mais constante, uma redução de 1% no Índice de Gini de certo município traria, em geral, uma maior redução da extrema pobreza (total ou infantil) que o aumento de 1% na renda domiciliar *per capita*.

Por exemplo, se o município de Fortaleza, que possui a menor taxa de extrema pobreza total (3,36%), tiver uma redução de 10% na desigualdade de renda isso acarretaria em uma diminuição de 22,4% na extrema pobreza. Em relação à renda *per capita* média, um crescimento de 10% implicaria em um arrefecimento de 7,8% na extrema pobreza total. Em outro extremo, o município de Granja, que têm a maior proporção de extremamente pobres (43,63%), teria reduções de 8,8% e 7,3%, respectivamente, nos mesmos contextos. Ressalta-se que estas estimativas podem ser realizadas para todos os municípios.

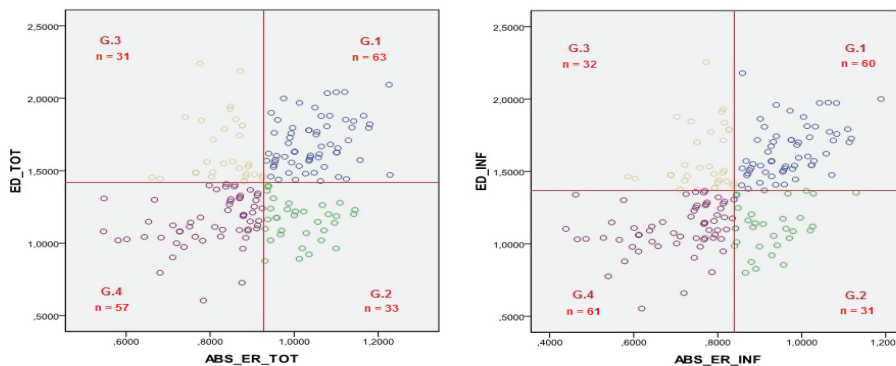
Continuando a análise, com base nos valores das elasticidades e das suas respectivas médias, propõe-se, tanto considerando a população total como a infantil, a divisão dos municípios cearenses em quatro *clusters* distintos, quais sejam:

- **Grupo 1 [G.1]:** ER Alta (maior ou igual à média) e ED Alta (maior ou igual à média);
- **Grupo 2 [G.2]:** ER Alta (maior ou igual à média) e ED Baixa (menor que a média);
- **Grupo 3 [G.3]:** ER Baixa (menor que a média) e ED Alta (maior ou igual à média);
- **Grupo 4 [G.4]:** ER Baixa (menor que a média) e ED Baixa (menor que a média)

No caso, tem-se que o G.1 concentra os municípios que apresentam as melhores respostas tanto para as políticas que estimulam o crescimento econômico (e conseqüentemente a renda domiciliar *per capita*) e que reduzem a concentração da renda. Já os municípios que compõem o G.4 seriam aqueles que apresentam respostas relativamente mais baixas para ambas as políticas. Os outros dois grupos, por sua vez, apresentam municípios que apresentam respostas mais relevantes para um tipo de política (crescimento econômico no caso do G.2 e de distribuição de renda no caso do G.3). A dispersão e classificação

em grupos dos municípios cearenses podem ser visualizadas com o auxílio da Figura 9.

**FIGURA 9: Diagramas de dispersão e divisão por grupos de municípios, conforme os valores das elasticidades estimadas**



Fonte: PNUD (2013). Elaborado pelos autores.

Nota: A Elasticidade Renda foi ilustrada considerando valores absolutos (ABS).

De forma a confirmar a legitimidade, esses agrupamentos foram validados por meio de análises discriminantes (PEREIRA, 2004), utilizando as elasticidades de renda e de desigualdade como variáveis independentes.

No caso, considerando-se o Lambda de Wilks, verificou-se que, tanto para a população total como para a infantil, ambas as elasticidades consideradas apresentaram significativo poder discriminante. Ademais, o modelo conseguiu prever corretamente o posicionamento de 95,1% dos municípios no caso da população total e 92,9% no caso da população infantil. Como os grupos foram considerados suficientemente diferentes, o Quadro 4 apresenta as médias das elasticidades em cada um deles.

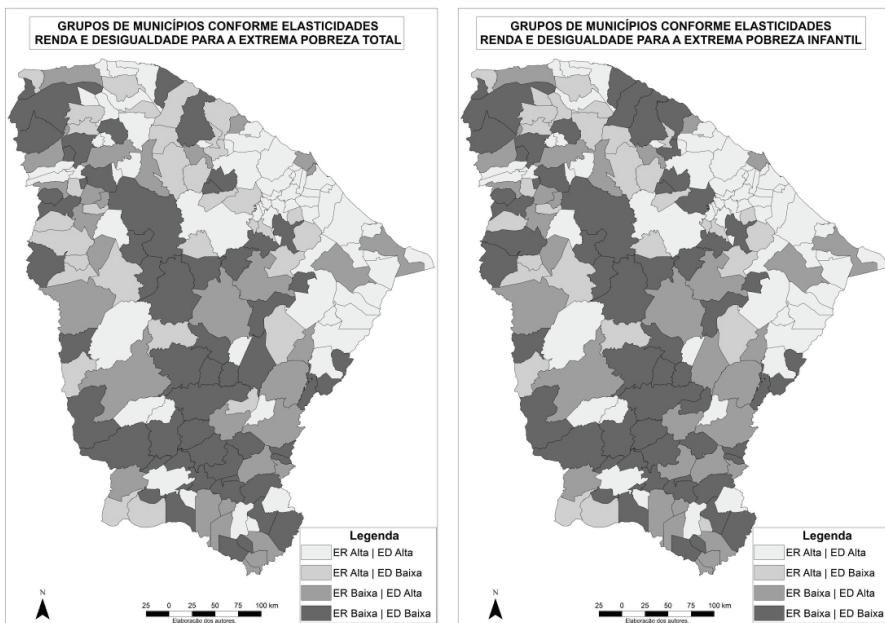
**QUADRO 4: Média das elasticidades de acordo com os grupos definidos – População Total e Infantil – 2010**

DISCRIMINAÇÃO	MÉDIA DAS ELASTICIDADES			
	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4
ABS_ER_TOT	1,0464	1,0131	0,8352	0,8060
ED_TOT	1,6886	1,1530	1,6458	1,1545
ABS_ER_INF	0,9713	0,9253	0,7654	0,7053
ED_INF	1,6479	1,1104	1,5915	1,1066

Fonte: PNUD (2013). Elaborado pelos autores.

Do ponto de vista geográfico, as distribuições dos municípios cearenses nos grupos previamente definidos podem ser verificadas com a ajuda da Figura 10, apresentada a seguir. De maneira geral, é possível perceber a partir da comparação entre os dois mapas que os municípios, em sua maioria, tenderam a ser classificados nos mesmos grupos seja para quando as elasticidades foram calculadas para a população total ou para a infantil. De fato, apenas 7 municípios no total foram classificados em grupos diferentes de acordo com a população considerada. Mais especificamente, a tabulação cruzada entre esses agrupamentos gerou um coeficiente V de Cramer<sup>12</sup> igual a 0,902, que pode ser considerado bastante elevado e estatisticamente significativo.

**FIGURA 10: Mapas dos grupos de municípios conforme elasticidades à renda e desigualdade – 2010**



Fonte: PNUD (2013). Elaborado pelos autores.

Nota: A Elasticidade Renda foi ilustrada considerando valores absolutos (ABS).

Portanto, por meio da avaliação dos grupos consente-se visualizar os municípios que são mais sensíveis a redução da extrema pobreza (total e infantil) por meio de políticas públicas relacionadas à redução da desigualdade e/ou de crescimento econômico.

12. Uma medida de associação entre duas variáveis qualitativas nominais (PEREIRA, 2004).

## 6 CONSIDERAÇÕES FINAIS

O presente artigo teve como objetivo analisar a evolução da extrema pobreza total e infantil no Ceará; tanto do ponto de vista temporal, ao longo das últimas duas décadas; como do ponto de vista espacial, considerando aspectos dos 184 municípios cearenses. Para tanto, fez-se uso das taxas de extrema pobreza para a população total, bem como para o grupo etário de 0 a 14 anos, calculadas pelo Programa das Nações Unidas para o Desenvolvimento (PNUD) em parceria com a Fundação João Pinheiro (FJP) e o Instituto de Pesquisa Econômica Aplicada (IPEA).

Conforme discutido na seção de fatos estilizados, há uma nítida diferença de magnitude das taxas de extrema pobreza total e infantil nos municípios cearenses. Assim, seguindo a modelagem proposta por Bourguignon (2003), buscou-se investigar a relação triangular pobreza-crescimento-desigualdade e verificar como o crescimento econômico e a desigualdade de renda afetaram as citadas medidas de pobreza extrema.

Os resultados ratificaram que tanto a extrema pobreza total como a extrema pobreza infantil são fortemente sensíveis aos níveis de desigualdade inicial nos municípios cearenses, demonstrando que para um nível mais elevado de desigualdade inicial a taxa em que a aceleração do crescimento é transformada em redução da pobreza deve diminuir. Além disso, verificou-se que a extrema pobreza infantil é menos sensível a sofrer reduções, a partir de diminuições na desigualdade de renda e de aumentos na renda média dos municípios, do que a extrema pobreza total, evidenciando, portanto, uma maior persistência da extrema pobreza infantil.

Outra evidência constatada foi à diferença entre as elasticidades estimadas, já que as médias das elasticidades pobreza-desigualdade foram superiores as de renda, tanto para a extrema pobreza total como para a infantil. Quando se considera a extrema pobreza total, tem-se que apenas os municípios de Amonatada, Choró, Croatá, Granja, Ipaporanga, Moraújo, Salitre e Uruoca apresentaram maiores elasticidade pobreza-renda (em módulo), totalizando oito cidades. Concomitantemente, isto ocorreu em somente sete municípios no caso da extrema pobreza infantil, sendo eles: Choró, Croatá, Granja, Ipaporanga, Moraújo, Salitre e Uruoca.

Mapearam-se quatro grupos de municípios com base nos valores das elasticidades renda e desigualdade, identificando no Grupo 1 os que detêm melhores respostas para políticas públicas atinentes ao crescimento econômico

e diminuição da desigualdade, enquanto o Grupo 4 foi composto por cidades que têm, relativamente, menos impacto na redução da extrema pobreza (total e infantil) a crescimento e redistribuição. Por sua vez, os Grupos 2 e 3 correspondem aos municípios que têm respostas mais significativas para o crescimento econômico e a desigualdade de renda, respectivamente.

Geograficamente, o Grupo 1 é abrangido por municípios situados na Região Metropolitana de Fortaleza (com exceção da capital, que ficou no Grupo 3), Litoral Leste e Vale do Jaguaribe. Já o Grupo 4 têm a maior parte de seus municípios localizados nas regiões do Sertão Central, Inhamuns, Crateús, Centro Sul e Litoral Oeste.

Neste contexto, os resultados encontrados nesse documento apontam que, caso seja utilizado apenas o instrumento “crescimento econômico”, o Ceará pode levar, ainda, muito tempo para extinguir por completo sua pobreza. Desse modo, essa tendência poderia ser acelerada se fossem utilizadas de forma mais efetiva políticas que possam alterar o perfil distributivo de nossa sociedade. A utilização de estratégias voltadas para a geração de renda com foco nos mais pobres é de extrema importância.

Ademais, programas de transferências de renda condicionados e aporte de ativos financeiros, como os programas de microcréditos, são também de grande valia. Vale citar também que a melhoria da educação e saúde pública para os mais necessitados são instrumentos fundamentais para alterar definitivamente a desigualdade de renda no Estado, apesar dos seus impactos serem obtidos apenas no longo prazo.

Vale lembrar que, o estado do Ceará detém um Fundo Estadual de Combate à Pobreza<sup>13</sup> (FECOP) que poderia ter os seus recursos empreendidos em um programa específico e focalizado de transferência condicionada de renda as famílias extremamente pobres. Já que, a tecnologia de focalização do Programa Bolsa Família não tem sido tão eficiente em chegar a esse grupo, e esforços maiores devem ser empregados a isso. O Cadastro Único Ceará é uma excelente base para a identificação desses indivíduos.

Outro desafio, como dito anteriormente, é que a extrema pobreza infantil apresentou reduções nos últimos anos em menor magnitude que a extrema pobreza total, indicando uma maior dificuldade de melhoria das condições de vida nesse grupo demográfico.

13. O Fundo Estadual de Combate à Pobreza - FECOP foi instituído pela Lei complementar Nº 37, de 26 de novembro de 2003 e regulamentado pelo Decreto Nº 27.379 de 1º de março de 2004.

Desse modo, torna-se importante empreender políticas públicas voltadas à acumulação de capital humano desse público, citando, por exemplo, a educação em tempo integral com a universalização de creches e escolas do ensino fundamental, a focalização e melhoria dos serviços de saúde especificamente para crianças, a garantia à nutrição e à alimentação adequada, criação de espaços públicos apropriados para atividades de lazer, assistência social, etc.

No contexto do Ceará, cita-se que o Governo do Estado lançou em 2015 o programa Mais Infância, que tem por objetivo o fortalecimento dos vínculos familiares, assim como proporcionar o desenvolvimento infantil através da oferta de educação de qualidade. O referido programa objetiva ainda a revitalização de espaços públicos, criação de ambientes de lazer, construção de centros de educação infantil, entre outras ações. Tais medidas, se implementadas com focalização, podem vir a gerar um impacto positivo na redução da pobreza extrema infantil no Estado.

Finalmente, menciona-se que às informações produzidas nesta pesquisa, tanto motivam a busca da redução das desigualdades, como fornecem subsídios para a alocação de políticas públicas de combate à extrema pobreza nos diversos municípios cearenses. Além de destacar a necessidade de um olhar específico para a infância como prerrogativa para o fim do ciclo vicioso de pobreza no Ceará.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMS, R. H. Economic growth, inequality and poverty: estimating the growth elasticity of poverty. **World Development**, v. 32, n. 12, p. 1989-2014, 2004.
- ALESINA, A.; RODRICK, D. Distributive politics and economic growth. **Quarterly Journal of Economics**, v. 109, p. 465-490, 1994.
- BARRO, R. J. Inequality and growth in a panel of countries. **Journal of Economic Growth**, v. 5, n. 1, p. 87-120, 2000.
- BATANA, Y.; BUSSOLO, M.; COCKBURN, J. Global extreme poverty rates for children, adults and the elderly. **Economics Letters**, v. 120, p. 405-407, 2013.
- BEGUM, S. S.; DENG, Q.; GUSTAFSSON, B. Economic growth and child poverty reduction in Bangladesh and China. **Journal of Asian Economics**, v. 23, p. 73-85, 2012.
- BOURGUIGNON, F. The growth elasticity of poverty reduction: explaining heterogeneity across countries and time periods. In: EICHER, T.; TURNOVSKY, S. (eds.) **Inequality and growth, theory and policy implications**. Cambridge: The MIT Press, 2003.
- BOURGUIGNON, F. **The poverty-growth-inequality triangle**. Washington, DC: The World Bank, 2004.
- BRUNO, M.; RAVALLION, M.; SQUIRE, L. Equity and growth in developing countries: old and new perspectives on the policy issues. In: TANI, V.; CHU, K. Y. (eds.) **Income distribution and high growth**. Cambridge, MA: MIT Press, 1998.
- CHEN, S.; RAVALLION, M. What can new survey data tell us about recent changes in distribution and poverty? **The World Bank Economic Review**, v. 11, n. 2, 1997.
- DATT, G.; RAVALLION, M. Growth and redistribution components of changes in poverty measure: a decomposition with application to Brazil and India in the 1980s. **Journal of Development Economics**, v. 38, p. 275-295, 1992.
- DEININGER, K.; SQUIRE, L. New ways of looking at old issues: inequality and growth.



**Journal of Development Economics**, v. 57, p. 259-287, 1998.

DOLLAR, D.; KRAAY, A. Growth is good for the poor. **Journal of Economic Growth**, v. 7, n. 3, p. 195–225, 2002.

EASTERLY, W. **The effect of IMF and World Bank programs on poverty**. Washington, DC: The World Bank, 2000 (mimeo.).

FORBES, K. J. A Reassessment of the relationship between inequality and growth.

**American Economic Review**, v. 90, p. 869-887, 2000.

FOSTER, J. E.; GREER, J.; THORBECKE, E. A Class of decomposable poverty measures. **Econometrica**, v. 52, p. 761-776, 1984.

FOSU, A. K. Inequality and the impact of growth on poverty: comparative evidence for Sub-Saharan Africa. **Journal of Development Studies**, v. 45, p. 726-745, 2009.

FOSU, A. K. Growth, inequality and poverty in Sub-Saharan Africa: recent progress in a global context. **Oxford Development Studies**, v. 43, p. 44-59, 2015.

FRANÇA, J. M. S.; FRANÇA, N. C.; CARVALHO, J. R. **Pobreza multidimensional e pobreza monetária no Nordeste**: análise comparativa da focalização do programa Bolsa Família em 2012. Fortaleza: LEP/CAEN/UFC, 2015 (Ensaio sobre pobreza e desigualdade n. 24). Disponível em: <http://www.caen.ufc.br/attachments/article/113/esp24-.pdf>.

IPECE. **A caracterização da extrema pobreza no estado do Ceará**. Fortaleza: IPECE, 2011 (Informe Econômico n. 09). Disponível em: [http://www.ipece.ce.gov.br/publicacoes/ipeceinforme/lpece\\_Informe\\_09\\_junho\\_2011.pdf](http://www.ipece.ce.gov.br/publicacoes/ipeceinforme/lpece_Informe_09_junho_2011.pdf)

LEVY, H.; KROLL, Y. Efficiency analysis with borrowing and lending: criteria and their effectiveness. **Review of Economics and Statistics**, Fev. 1979.

KAKWANI, N. Poverty and economic growth with application to Cote d'Ivoire. **Review of Income and Wealth**, v. 39, p. 121–139; 1993.

KALWIJ, A.; VERSCHOOR, A. Not by growth alone: the role of the distribution of income in regional diversity in poverty reduction. **European Economic Review**, v. 51, p. 805–829, 2007.

KUZNETS, S. Economic growth and income inequality. **American Economic Review**, v. 45, p.1-28, 1955.

LI, H.; ZOU, H. Income inequality is not harmful for growth: theory and evidence. **Review of Development Economics**, v. 2, p. 318-334, 1998.

LOPEZ, H.; SERVEN, L. **The mechanics of growth-poverty-inequality relationship**. The World Bank, 2004 (mimeo.).

MENG, X.; GREGORY, R.; WANG, Y. Poverty, inequality, and growth in urban China, 1986-2000. **Journal of Comparative Economics**, v. 33, p. 710-729, 2005.

OLIVEIRA, J. L.; MANSO, C. A.; ASSIS, D. N. C. Renda, pobreza e desigualdade. In: BARRETO, F. A.; MENEZES, A. S. B. **Desenvolvimento econômico do Ceará: evidências recentes e reflexões**. Fortaleza: IPECE, 2014, p. 209-225.

PEREIRA, A. **Guia prático de utilização do SPSS – análise de dados para Ciências Sociais e Psicologia**. 5.ed. Lisboa: Edições Sílabo, 2004. 243 p.

PEROTTI, R. Growth, income distribution, and democracy: what can the data say.

**Journal of Economic Growth**, v. 1, p. 149-187, 1996.

RAVALLION, M. Can high-inequality developing countries escape absolute poverty?

**Economic Letters**, v. 56, 1997.

SON, H.; KAKWANI, N. **Poverty reduction: do initial conditions matter?** Washington, DC: The World Bank, 2003 (mimeo.).

SOUZA, A. P. Políticas de Distribuição de Renda no Brasil e o Bolsa Família. In: BACHA, E. L.; SCHWARTZMAN, S. (Orgs.). **Brasil: a nova agenda social**. Rio de Janeiro: LTC, cap. 5, p.166-186. 2011.

# DESEMPENHO ESCOLAR E INSERÇÃO NO MERCADO DE TRABALHO: UMA AVALIAÇÃO DAS ESCOLAS ESTADUAL DE ENSINO PROFISSIONALIZANTES (EEEP) DO CEARÁ

Francisca Zilania Mariano\*

Ronaldo A. Arraes\*\*

Natália de Olivindo Souza\*\*\*

## Resumo

O Governo do Ceará criou uma política de implementar as Escolas Estadual de Ensino Profissionalizantes (EEEP) no estado, cujo propósito são propiciar aos alunos uma educação básica voltada para o ensino médio, juntamente com uma profissionalização técnica direcionada ao mercado de trabalho. Dessa forma, este trabalho tem como objetivo elaborar uma avaliação sobre esse tipo de política e encontrar o efeito das EEEP sobre o desempenho no Exame Nacional do Ensino Médio, bem como o impacto dessas sobre as chances de os alunos estarem trabalhando. Informações disponibilizadas pela SEDUC referentes a lista dessas escolas e pelo INEP relacionados ao ENEM 2013 e ao censo escolar 2013, possibilitaram a construção de um modelo baseado no escore de propensão. Após o *matching* realizado através do método *nearest-neighbor*, observou-se que os alunos das escolas profissionalizantes apresentam diferenciais significativos em todas as áreas de conhecimento do ENEM e possuem maiores chances de ultrapassarem as notas de cortes referentes aos cursos da UFC, porém estes possuem menor probabilidade de estar trabalhando no período que prestaram o exame, em relação aos alunos das escolas regulares. Esses resultados mostram que as EEEP estão rompendo com a dualidade estrutural entre formação para o mercado de trabalho e preparação para a universidade, ao integrar as duas redes de ensino, profissional e geral.

**Palavras-Chave:** Ensino Profissionalizante; Inserção Produtiva, Desempenho Escolar.

---

\* Professora UFC/Campus de Sobral, Doutoranda em Economia, CAEN/UFC, Contato:(85)988177364, zilania@ufc.br.

\*\* PhD, Professor CAEN/UFC, ronald@ufc.br.

\*\*\* Estudante do curso de Finanças, UFC/Campus de Sobral.

Endereço: Av. da Universidade, 2700 – 2º andar CAEN/UFC, 60020-181 Fortaleza – CE.

**Abstract**

School Achievement and Insertion in the Labour Market: An evaluation of Ceará EEEP

**Abstract:** The Government of Ceará created a policy to implement Schools of Vocational Education (EEEP) whose purpose is providing students a basic education at high school level, along with a technical professionalization directed to the labor market. Thus, this paper aims to draw up an assessment of such policy and find the effect of Ceara EEEP on performance in the National Secondary Education Examination as well as the impact of these on the chances of the students are working. Information provided by SEDUC regarding the list of these schools and by INEP related to ENEM 2013 and the school census in 2013 made it possible to build up a model based on the propensity score. After matching performed by the nearest-neighbor method, it was observed that students of vocational schools show significant differences in all areas of knowledge ENEM and have a better chance of overcoming the notes cuts related to UFC courses, but these have less likely to be working in the period that they took examination in relation to students from regular schools. These results show that EEEP are breaking with the structural duality of training for the labor market and preparation for university, to integrate the two school systems, professional and general.

**Key words:** Vocational Education; Insertion in the Labour Market, ENEM.

**JEL:** C21, I00, I2, J1

**INTRODUÇÃO**

A análise de políticas públicas é relevante para informar a sociedade e aos gestores se os objetivos propostos por elas estão sendo alcançados e serve como base para auxiliar e melhor direcionar a gestão na tomada de decisões. Dessa forma, avaliações de programas governamentais voltados à Educação são importantes por ser uma das áreas mais suscetíveis de críticas, haja vista o baixo desempenho educacional no Brasil relativo a outros países muito mais pobres economicamente. Sob a hipótese de que magnitude econômica de um país segue sua grandeza educacional, regiões como a América Latina se saem muito pior do que outras pelo que seus níveis de renda poderiam prever, conforme atestam Ganimian e Rocha (2011), Hanushek e Woessmann (2012) e Levy e Schady (2013). Assim, existe um debate sobre como melhorar a qualidade do

ensino e, especificamente, verificar o papel de programas e ações governamentais nesse processo.

Dentre esses tipos de políticas encontra-se a concepção do ensino profissionalizante. Esse ensino era caracterizado por uma “dualidade”, onde, segundo Kuenzer (2005), existia tipos diferentes de escola para classes sociais distintas, profissional e técnico voltado para níveis socioeconômicos mais baixo e uma educação básica voltada para cursos universitários que priorizava a classe social média e alta.

Historicamente, segundo Tavares (2012), não há como avaliar a educação com profissionalização no Brasil sem fazer referência a “dualidade” existente, onde o ensino preparatório para o ingresso na universidade era voltado para a elite, e o profissionalizante direcionado aos filhos dos trabalhadores (CASTRO; TIEZZI, 2005; MARTINS, 2012; MOEHLECKE, 2012). Embora a criação do ensino profissionalizante tenha sido guiada para atender jovens que viviam à margem da sociedade (TAVARES, 2012), após a criação da primeira Lei de Diretrizes e Bases da Educação Nacional, ele passou a ser constituído juntamente com o nível médio, e acrescentou o foco no ensino superior como um de seus propósitos. Dessa forma, estudantes de escolas profissionais não seriam prejudicados, pois, além de estarem preparados para o mercado de trabalho, também poderiam ingressar numa universidade.

Assim, através da integração das duas redes de ensino, profissional e geral, pode-se romper a dualidade estrutural entre formação para o trabalho e preparação para a universidade (CASTRO; TIEZZI, 2005; MARTINS, 2012; MOEHLECKE, 2012). Diversos autores buscam verificar os efeitos desse tipo de ensino sobre a inserção no mercado de trabalho (NEUMAN; ZIDERMAN 1989; CHEN; WEKO, 2009; SEVERNINI; ORELLANO, 2010; ASSUNÇÃO; GONZAGA, 2010; ARAÚJO et.al, 2014), todavia, ainda, são poucos os estudos que avaliam os impactos do ensino profissionalizante sobre a educação pública brasileira, destaque para Araújo et al (2014), o qual avaliou, com base em informações do ENEM 2009, se os alunos que realizam o currículo específico da Educação Profissional e Tecnológica (EPT) apresentam aumento na proficiência em disciplinas básicas. Por outro lado, a revisão da literatura aqui feita não constatou algum trabalho em nível estadual com foco de avaliação de política sobre escolas profissionalizantes, lacuna esta a ser preenchida com o estudo de caso para o Ceará.

O Governo Estadual do Ceará iniciou em 2008 um projeto de criação de Escolas Estaduais de Educação Profissional (EEEP), que incluem, além do ensi-

no com as disciplinas básicas do currículo do ensino médio, os cursos técnicos e profissionalizantes voltados para diversas áreas (saúde, agricultura, computação, finanças, etc.), cuja escolha, pela Secretaria de Educação, atende a realidade local da cidade onde a instituição é implantada (SEDUC, 2015).

Essa ação possibilita ao aluno concluir o ensino médio e se profissionalizar para o mercado de trabalho. Com isso, além de formar o jovem para o mercado de trabalho, o ensino das escolas profissionalizantes também o capacita a ingressar em universidades. Silva (2013) observa o desempenho das escolas estaduais no Exame Nacional do Ensino Médio (ENEM), através da comparação entre as médias escolares nos anos 2009 a 2011 e constatou que o estado do Ceará melhorou a performance com a participação dos alunos das EEEP. Além disso, segundo informações disponibilizadas pelo INEP (2015) referentes a nota média por escola no ENEM/2014, 80% das melhores escolas públicas do Ceará são de ensino profissional.

Com isso, esse trabalho procura contribuir com o debate existente na literatura acerca do ensino profissionalizante em sua dupla função de preparar os estudantes do nível médio tanto com educação formal visando exames para ingressar em universidades, como para o mercado de trabalho. A condução da análise é feita através de modelos econométricos apropriados para avaliação de políticas públicas, de onde se extrairá os efeitos das EEEP sobre o desempenho dos estudantes nas cinco áreas e na média geral do ENEM, bem como o impacto sobre as chances de os alunos que cursaram EEEP ultrapassarem as notas de cortes referentes aos cursos da UFC e sobre as chances de estarem trabalhando. Para a verificação empírica serão utilizadas informações da SEDUC referentes à lista de escolas profissionalizantes, e pelo INEP relacionados ao ENEM 2013 e ao censo escolar 2013, para a construção de um modelo baseado no score de propensão. Após o *matching* realizado através do método *nearest-neighbor*, constatou-se diferenciais significativos e positivos sobre o desempenho escolar, entre os que fizeram EEEP e aqueles que cursaram escola regular apenas.

Em sequência, o artigo está organizado com as seguintes seções: aspectos teóricos da literatura sobre as escolas profissionalizante e as EEEP no estado do Ceará; a abordagem metodológica; resultados e conclusões.

## 2 REVISÃO DA LITERATURA

### 2.1 Ensino Profissionalizante

Para Crouch et al. (1999), existem três justificativas para o desenvolvimento de políticas públicas voltadas para a qualificação mais especializada da mão de obra, oposição a educação formal. Primeira, melhores condições de um país participar de um mercado internacional competitivo; segunda, estas políticas tendem a contribuir para a redução do desemprego, principalmente entre os jovens; terceira, a qualificação da mão de obra permitiria que as economias de países em desenvolvimento evoluíssem de uma situação equilíbrio gerado pela produção de commodities de valor agregado inferior, resultado da utilização de mão de obra barata pouco qualificada e de baixa produtividade, para situações de equilíbrio com a produção de mercadorias de maior valor agregado, proporcionando assim, melhor distribuição da renda.

Dentre as políticas públicas voltadas para a educação, a implantação do ensino profissionalizante é analisada por diversos autores que buscam verificar os fatores históricos (KUENZER, 2005; FOLEY, 2007; TAVARES, 2012; MARTINS, 2012), os efeitos sobre a inserção no mercado de trabalho (NEUMAN; ZIDERMAN 1989; CHEN; WEKO, 2009; SEVERNINI; ORELLANO, 2010; ASSUNÇÃO; GONZAGA, 2010; ARAÚJO et.al, 2014) e sobre o desempenho em exames utilizados para ingressar em universidades (ARAÚJO et.al, 2014).

Segundo Kuenzer (2005), existia uma dualidade estrutural na educação que se caracterizava pela existência de tipos diferentes de escola para classes sociais distintas, onde a trajetória escolar de estudantes situados nas classes de renda média e alta se constituía de uma educação básica voltada para cursos universitários, enquanto que para os demais restava como alternativa uma base educacional associada à formação em cursos técnicos.

A maior participação de jovens com nível socioeconômico mais baixo no ensino profissionalizante foi comprovada por Foley (2007), na Austrália, o qual objetivava estudar o perfil dos participantes do sistema de ensino e formação profissional (VET). Para tanto, o autor utilizou dados sobre os estudantes matriculados na rede pública em 2001 e constatou que os alunos de origens socioeconômicas mais elevadas são maioria na participação dos setores de ensino superior e secundário, o que não ocorre no caso do setor de ensino profissionalizante.

Buscando avaliar o efeito da política educacional sobre o mercado de trabalho em Israel, Neuman e Ziderman (1989) comprovam que os salários de egressos do ensino profissional são, em média, 10% superiores aos daqueles formados pela educação formal. Ao contrário, Chen e Weko (2009) demonstram não haver efeitos da formação profissional sobre o mercado de trabalho na Indonésia em 1997 e 2000.

No Brasil, Tavares (2012) afirma que o ensino profissionalizante foi criado para atender crianças, jovens e adultos que viviam à margem da sociedade, onde as primeiras escolas que constituíram a Rede Federal de Educação Profissional tinham a função de instruir tais indivíduos através do ensino de um ofício ou profissão e prepara-los para o mercado de trabalho.

Nesse esforço de preparação para o mercado de trabalho, Severnini e Orellano (2010) questionou se os programas de qualificação profissional realmente contribuem para os treinados obterem uma melhoria de bem-estar e se esse ensino contribui para aumentar a probabilidade de inserção do indivíduo no mercado de trabalho brasileiro. Para tanto, os autores investigaram se este tipo de ensino aumentou a probabilidade de inserção no mercado de trabalho e a elevação de renda dos egressos até meados da década de 1990, comparativamente aos que não cursaram esse tipo de ensino. Através de microdados da Pesquisa sobre Padrões de Vida (PPV) de 1996 do IBGE encontraram que os egressos de cursos profissionalizantes de nível básico tinham renda esperada 37% maior que a de indivíduos que não fizeram esse tipo de curso no ensino fundamental. Por outro lado, para os egressos do ensino profissional de nível tecnológico, observa-se uma redução de 27% da renda esperada, comparativamente aos que não participaram desse tipo de curso no ensino superior.

Assunção e Gonzaga (2010) apresentam os resultados de uma análise realizada com base nos microdados do suplemento especial sobre educação profissional da PNAD/2007, cujo objetivo reside em analisar a inserção da população brasileira nos cursos de educação profissional e verificar o impacto da educação profissional sobre o rendimento dos trabalhadores brasileiros. Dos resultados, observaram que a inserção da educação profissional em famílias com renda per capita inferior a dois salários mínimos é bem menor que nas demais faixas de renda e que a educação profissional aumenta a produtividade dos trabalhadores. Dessa forma, os autores sugerem que essa educação mereça atenção no desenho de políticas públicas.

Tavares (2012) além de fazer referência a “dualidade” existente no ensino, também diagnosticou que, depois da criação da primeira Lei de Diretrizes e



Bases da Educação Nacional, isso viria a mudar, uma vez que esta passou a constituir o ensino profissionalizante juntamente com o ensino médio e estabeleceu que este tivesse como um de seus propósitos o foco no ensino superior. Logo, já se poderia esperar que os jovens participantes do ensino profissionalizante não seriam prejudicados, pois além de estarem preparados para o mercado de trabalho também podem ingressar numa universidade. Além disso, a escola de ensino profissionalizante também é colocada como um meio de promover as habilidades pessoais do jovem, a vivência com as outras pessoas, a inserção no mercado de trabalho e o melhor desempenho nas disciplinas regulares (ARAÚJO et. al, 2014).

Buscando analisar o desempenho escolar e inserção no mercado de trabalho dos alunos das escolas técnicas profissionalizantes do Brasil, Araújo et. al (2014) usou os resultados do ENEM 2009 divulgados pelo INEP e através de um modelo de pareamento com escore de propensão para escolas federais e privadas encontrou uma relação positiva entre participar de ensino técnico e profissionalizante e melhor desempenho escolar, além de maiores chances de entrar no mercado de trabalho.

Segundo Martins (2012), através da integração das duas redes de ensino, profissional e geral, pode-se romper a dualidade estrutural entre formação para o trabalho e preparação para a universidade. É, pois, nesse contexto que se enquadram as escolas profissionalizantes de ensino implementadas pelo Governo do Ceará.

## **2.2 Ensino Profissionalizante no Ceará**

Segundo Nibon (2014), no Ceará, a oferta de cursos técnicos era realizada pelo Instituto Centro de Ensino Tecnológico – CENTEC, pelo Instituto Federal do Ceará (IFCE) e pelo Sistema S, onde havia uma oferta de cento e quinze cursos, nos quais estavam matriculados 14.606 jovens e adultos, segundo dados da Secretaria da Ciência e Tecnologia do Ceará - SECITECE, em 2005, representando apenas três e meio por cento dos alunos matriculados no ensino médio, portanto, uma oferta insignificante, considerando o percentual de atendimento da rede de educação no ensino médio. Dessa forma, tinha-se, assim, um cenário estadual que evidenciava a necessidade de desenvolvimento de uma política de educação profissional que potencializasse a articulação entre o ensino e o mercado de trabalho e, ao mesmo tempo, ampliasse as possibilidades de qualificação do processo de aprendizagem (RELATÓRIO DE GESTÃO,

2014).

Em 2007, a SEDUC aderiu ao Programa Brasil Profissionalizado e no segundo semestre de 2008, foram instituídos os Centros Educacionais para a Juventude – CEJOVEM, posteriormente denominados de Escolas Estaduais de Educação Profissional (EEEP), o qual, passou a disponibilizar, além do ensino básico do currículo do ensino médio, cursos técnicos e profissionalizantes voltados para diversas áreas e disciplinas que buscam à formação pessoal e convívio com a sociedade. Essa ação proporciona ao aluno concluir o ensino médio e se profissionalizar em áreas direcionadas para o mercado de trabalho. Para tanto, o estudante adquire, como primeira experiência profissional, o estágio obrigatório na área cursada, o qual se realiza no último ano de participação na escola. Este deve ser monitorado por um acompanhamento técnico-pedagógico por profissionais diplomados de cada área específica, visando avaliar e orientar os estudantes. Esse monitoramento é realizado por meio de um **Sistema Informatizado de Captação e Estágios – SICE** para cada curso de cada escola, o qual permite gerenciar a captação de vagas, encaminhar os alunos às concedentes, acompanhar os processos de formalização jurídica do estágio e da frequência mensal do estagiário.

Segundo informações divulgadas pela SEDUC, em 2008, eram 25 escolas em 20 municípios, em 2014 esse número aumentou para 106 e 112, respectivamente. Em 2015, existem 112 Escolas Estaduais de Educação Profissional (EEEP) no estado, perfazendo um atendimento a cerca de 40 mil alunos, e oferecendo 53 cursos técnicos em diversas áreas, tais como: saúde, agricultura, computação, finanças, dentre outras, as quais são selecionadas e implantadas na EEEP conforme a realidade local da cidade onde esta é implantada. Em relação aos recursos gastos na construção, manutenção e funcionamento das escolas já foram gastos mais de um bilhão de reais em parcerias do governo estadual e federal (SEDUC, 2015). No período de 2008 a 2014, foram investidos R\$1.036.097.010,22 (Um bilhão, trinta e seis milhões, noventa e sete mil, dez reais e vinte e dois centavos), 71% provenientes do Governo do Estado e 29% do Fundo Nacional de Desenvolvimento da Educação do Ministério da Educação (FNDE)/MEC, como pode ser observado na tabela A1 em anexo. Segundo o Relatório de Gestão (2014) o modelo cearense foi inspirado no Programa de Desenvolvimento dos Centros de Ensino Experimental – PRO-CENTRO - implementado no Estado de Pernambuco. Os Centros de Ensino Experimental (CEE), também conhecidos como ginásios experimentais, eram escolas modelo cuja diretriz pedagógica é prover o ensino integral a estudantes

de ensino médio, e atualmente implantada em São Paulo, Goiás e Rio de Janeiro. Diferentemente, as EEEP optaram por incluir no ensino médio a formação profissionalizante.

Esse tipo de ensino está voltado não só para a inserção do jovem no mercado de trabalho, mas também para a aprovação em universidades, pois a partir da implantação da primeira Lei de Diretrizes e Bases da Educação Nacional, um dos focos da escola de educação profissional também passou a ser à entrada do aluno na universidade (CASTRO; TIEZZI, 2005; MARTINS, 2012; MOEHLECKE, 2012). Segundo um Informe do IPECE, Silva (2013) mostra que a melhora no desempenho do estado do Ceará na participação no ENEM dos anos 2009-2011 pode estar relacionada a maior participação dos alunos das EEEP do estado. Para isso, a autora utilizou as médias no exame nesse período e observou que 29 escolas profissionais tiveram médias melhores ou iguais à nacional. Além disso, no ranking das melhores escolas do Ceará, as profissionalizantes ocuparam 33 posições das 50.

Segundo o Relatório de Gestão (2014), houve uma ampliação da aprovação de egressos das EEEP na universidade e uma redução da inserção no mercado de trabalho, como pode ser observado no gráfico A1 em anexo. Este aumento ocorreu, principalmente, nos grupos de municípios de médio e pequeno porte e em relação à Fortaleza e Região Metropolitana, o relatório verificou a ampliação da aprovação em universidades privadas.

Além disso, segundo informações do INEP (2015) referentes ao exame ENEM/2014, 8 (oito) das 10 (dez) melhores escolas são de ensino profissional. Além disso, entre as escolas públicas, a escola que apresentou melhor desempenho no CE foi escola Adriano Nobre EEEP, na cidade de Itapajé, no interior do Ceará, o qual obteve média geral de 588.16.

A partir desses indicadores e da hipótese que a utilização de políticas públicas é uma das formas para tentar se chegar ao caminho mais coerente para a educação no Brasil (CASTRO; TIEZZI, 2005), faz-se necessário que sejam realizados trabalhos sobre o impacto dessas políticas para um melhor direcionamento da gestão governamental. Com isso, esse trabalho procura promover uma avaliação sobre esse tipo de política e encontrar o efeito das EEEP do Ceará sobre o desempenho no ENEM, sobre as chances de os alunos ultrapassarem as notas de cortes referentes aos cursos da UFC e sobre as chances de estarem trabalhando.

### 3 METODOLOGIA

#### 3.1 Modelo

Segundo a teoria do Contrafactual, o ideal para se avaliar uma política seria observar uma mesma unidade em duas circunstâncias distintas, com e sem a política simultaneamente, porém, como isso não é possível, procura-se comparar dois grupos estatisticamente idênticos, onde, um estaria sobre a intervenção da política e o outro não. Assim, Rubin (1974) desenvolveu uma forma de construção através da randomização entre os grupos que receberá a política (grupo de tratamento) e o grupo que não a receberá (grupo de controle). Esse procedimento ficou conhecido como *Rubin Causal Model* (RCM) ou *Randomized Controlled Trial* (RCT).

Quando o processo de randomização não é possível, busca-se construir os grupos de controle e tratamento como uma imitação do RCT. Assim, se a definição dos grupos não ocorrer de forma aleatória, então devem existir outras características das unidades que determinem a designação para cada grupo, ou seja, o pressuposto da hipótese de identificação requer que existam unidades de ambos os grupos, tratamento e controle, para cada característica  $X_i$  para o qual se deseja comparar. Para esse tipo de modelo, chamados de não-experimentais, haverá o problema de autosseleção (BECKER; ICHINO, 2002; ANGRIST; PISCHKE, 2009). Assim, algumas hipóteses são necessárias para garantir a estimação da relação causal, se as chances para o tratamento não forem aleatórias, como a Hipótese da Independência Condicional (HIC), o qual assume que, os resultados potenciais são independentes da variável binária de tratamento ao se condicionar às variáveis observáveis,  $X_i$ . Além disso, assume-se que fatores não observáveis não são fontes de viés.

Através da diferença entre as médias dos resultados dos escores das unidades do grupo de tratamento e do grupo controle, pode-se estimar o efeito médio dos indivíduos do grupo tratado (EMT). Dessa forma, o valor da variável dependente ( $Y$ ) passa a ser independente da condição (tratado ou controle), uma vez que os indivíduos com características observáveis idênticas possuem a mesma chance de receber o tratamento. Logo,

$$\{Y_{0i}, Y_{1i}\} \perp d_i | X_i \quad (1)$$

A variável binária  $d_i = 1$  indica que a unidade  $i$  recebeu o tratamento;  $Y_{11}$  e  $Y_{i0}$  são os resultados de uma unidade após a adesão ao programa e caso esta não fosse beneficiada pelo programa, respectivamente.

Condicionando às variáveis observáveis,  $X_i$ , e assumindo HIC, tem-se

que:

$$\underbrace{E[Y_i|X_i, d_i = 1] - E[Y_i|X_i, d_i = 0]}_{\text{Diferença Observada}} = \underbrace{E[Y_{1i}|X_i] - E[Y_{0i}|X_i]}_{\substack{\text{Efeito médio do Tratamento} \\ \text{nos tratados (ATT)}}} + \underbrace{E[Y_{0i}|X_i] - E[Y_{0i}|X_i]}_{\text{Viés de Seleção}}$$

$$= E[Y_{1i} - Y_{0i} | X_i] \quad (2)$$

Em termos gerais, o método consiste em especificar, a priori, um grupo de unidades que foram sujeitas a um determinado tratamento, e compará-las com outras isentas do tratamento (grupo de controle), embora com características semelhantes. Em seguida, estimam-se os efeitos do tratamento por meio da diferença entre os resultados médios dos grupos de tratamento e controle.

### 3.1.1 Pareamento por Escore de Propensão (PEP)

O Pareamento por Escore de Propensão (PEP), desenvolvido por Rosenbaum e Rubin (1983), busca lidar com o problema da dimensionalidade, aproximando-se a características da estimação da relação causal de um experimento aleatório. Para tanto, a hipótese da independência condicional, descrita anteriormente, precisa ser satisfeita para que se possa construir um grupo de controle similar ao grupo de tratamento tomando com base na distribuição de variáveis observadas.

De forma geral, o PEP corresponde à probabilidade condicional de uma unidade receber o tratamento em virtude de um conjunto das características observáveis  $X$ , o qual pode ser calculado conforme a equação (3):

$$P(X) = P(D = 1 | X) \quad (3)$$

Esse método apresenta uma solução prática para o problema da multidimensionalidade no pareamento e o efeito de tratamento pode ser determinado de acordo com a equação (4):

$$E(Y_1 - Y_0 | D = 1, P(X)) = E(Y_1 | D = 1, P(X)) - E(Y_0 | D = 0, P(X)) \quad (4)$$

Assim, o viés associado às diferenças do nível das variáveis observadas é eliminado ajustando as diferenças entre as unidades de tratamento e controle, gerando um estimador que permite que a distribuição das características observadas ( $X$ ) do grupo de controle e do grupo de tratamento seja idêntica. Portanto, o contrafactual pode ser construído através dos resultados do grupo de tratamento como explicita a equação (5):  $E(Y_0 | P(X), D = 1) = E(Y_0 | P(X), D = 0) = E(Y_0 | P(X))$  (5)

O processo de seleção ocorre através de características observáveis, onde as unidades que as possuam de forma idêntica tenham a mesma pro-

babilidade de serem alocadas como tratamento ou controle. Segundo Heckman, Lalonde e Smith (1999), a probabilidade das unidades serem participantes ou não participantes do tratamento pode ser descrita na equação

$$0 < P(D_i = 1|X_i) < 1 \quad (6)$$

Essa condição irá assegurar que as observações do grupo de tratamento tenham observações comparáveis do grupo de controle quanto às características  $X_i$  (KHANDKER; KOOLWAL; SAMAD, 2010). Rosenbaum e Rubin (1983) propõe o Teorema do Escore de Propensão, o qual torna prático o pareamento, reduzindo o número de variáveis do vetor  $X_i$  a um único escalar, a probabilidade de recebimento do tratamento, dado as características observadas. Angrist e Pischke (2009) mostra que esse teorema pode ser enunciado da seguinte forma: Supondo que a Hipótese de Independência Condicional (HIC) seja satisfeita tal que  $\{Y_{0i}, Y_{1i}\} \perp d_i | X_i$ , então  $\{Y_{0i}, Y_{1i}\} \perp d_i | P(X_i)$ , ou seja, se os resultados potenciais são independentes da variável de tratamento condicionada a um vetor multivariado  $X_i$ , então os resultados potenciais são independentes da variável de tratamento condicionada a uma função escalar desse mesmo vetor, que é o escore de propensão, definido como  $P(X_i) \equiv E[d_i | X_i] = P[d_i = 1 | X_i]$ .

Segundo Lee (2006), outra hipótese necessária para a estimação dos efeitos de tratamento utilizando métodos de pareamento é o *critério de balanceamento*, onde este é satisfeito quando para cada valor do escore de propensão,  $X$  tem a distribuição similar para os grupos de tratamento e controle, ou seja,  $D \perp X | P(X)$ .

A estimação por escore de propensão ocorre da seguinte forma: estima-se  $P(X_i)$  com algum modelo paramétrico, como probit ou logit e em seguida, a estimação do efeito do tratamento pode ser encontrada ou pelo pareamento do valor encontrado no primeiro passo ou utilizando algum esquema de pesos (CALIENDO; KOPEINIG, 2005). Essa estimativa tem por objetivo encontrar um grupo de controle que seja o mais semelhante possível ao grupo de tratamento em termos do escore de propensão, dadas às características observadas. Vale ressaltar que, esse procedimento se dá dentro da região de suporte comum. Assim, pelo teorema do escore de propensão e supondo a HIC, o EMT, resultante do pareamento direto dos valores de propensão entre tratados e não tratados, aplicando a lei de expectativas iteradas sobre  $X_i$ , é dado por:

$$EMT = E\{E[Y_i | P(X_i), D_i = 1] - E[Y_i | P(X_i), D_i = 0] | D_i = 1\} \quad (7)$$

Smith e Todd (2005) mostram que o estimador típico PEP possui a seguinte forma:

$$EMT_{PSM} = \frac{1}{N_T} \left[ \sum_{i \in T} Y_{1,i} - \sum_{j \in C} \omega(i,j) Y_{0,j} \right] \quad (8)$$

Onde  $\omega(i, j)$  é o tipo de pesos utilizado para agregar o resultado potencial dos indivíduos do grupo de controle e depende dos escores de propensão do participante  $[P(X_i)]$ , e não-participante  $[P(X_j)]$ , e  $n$  é o número de indivíduos tratados pertencentes à região de suporte comum.

Dentre os métodos de pesos utilizados para a estimação do EMT utilizando PEP, Caliendo e Kopeinig (2005) citam o *nearest-neighbor*, *radius*, *stratification*, *kernel*, os quais são brevemente definidos a seguir.

- *Stratification* consiste na divisão da distribuição de escore de propensão estimada em intervalos, onde, em cada bloco unidades de tratamento e de controle possuam em média o mesmo escore de propensão. Além disso, em cada intervalo, computa-se a diferença das médias das variáveis de efeito de tratados e não-tratados e assim, o EMT do tratamento é obtido como uma média dos EMT de cada bloco ponderada pelos pesos extraídos da distribuição dos tratados. De acordo com esse método, caso haja blocos que não contenha observações de tratados e de não-tratados, esse bloco é descartado.

O método de *nearest-neighbor*, vizinho mais próximo, resolve, em parte, o problema de exclusão de observações, pois este compara cada unidade de tratamento com a unidade de controle com o escore de propensão mais próximo. Assim, o EMT é obtido pela média das diferenças nas variáveis de interesse entre tratados e não-tratados. Dessa forma, toda unidade de tratamento é pareada com uma unidade de controle, porém, sua ocorrência pode se dá entre valores muito distantes de escore de propensão, o que pode ocasionar em um contrafactual não bem construído.

- Já o método *kernel*, utiliza pesos inversamente proporcionais à distância entre os valores do escore de propensão dos tratados e dos não-tratados e assim todas as unidades de tratamento são pareadas com uma média ponderada de todas as unidades de controle, onde este pode ser visualizado no equação (9):

$$\omega(i, j)_{kernel} = \frac{K\left(\frac{P_j - P_i}{\alpha_n}\right)}{\sum_{k \in C} K\left(\frac{P_k - P_i}{\alpha_n}\right)} \quad (9)$$

Onde  $K(\cdot)$  representa uma função kernel e  $\alpha_n$  é o um parâmetro denominado *janela*.

- O método *radius* determina uma vizinhança do escore de propensão de cada unidade de tratamento e parear com unidades de controle que pertençam a essa vizinhança, onde quanto menor o raio da vizinhança, melhor o pareamento, porém, isso pode resultar que algumas unidades de tratamento não sejam pareadas.

Após o pareamento por estes métodos será feito o teste de balanceamento para cada um deles para verificar qual se mostra o mais adequado. Para que esses estimadores baseados no escore de propensão possam ser considerados sem viés é necessário que as condições de receber o tratamento sejam exógenas. Além disso, outro ponto importante a ser analisado é quanto às quais variáveis incluir no modelo como vetor de controle,  $X$ . Ravaillon (2008) sugere elas devem escolhidas tanto pela relevância na designação e participação no tratamento, levando em consideração os fatores econômicos, sociais e políticos do programa, quanto nas variáveis de resultado, cujos impactos se desejam mensurar. Assim, o método de PEP depende do grau em que as características observadas explicam a participação no programa.

### 3.2 Variáveis e Bases de Dados

Para avaliar o impacto da política de profissionalização no ensino médio das EEEP do Ceará sobre o desempenho dos alunos no ENEM e sobre a inserção no mercado de trabalho foram utilizados como indicadores de impacto as notas dos alunos nas cinco áreas e na média geral do ENEM, variáveis binárias construídas a partir dos quantis das notas de corte<sup>1</sup> relativas ao processo de seleção da Universidade Federal do Ceará – UFC em 2013 e uma variável que identifica se o aluno está trabalhando. Além de observar os efeitos das EEEP sobre o desempenho dos alunos nas provas do ENEM, esse trabalho se preocupou em buscar verificar se esses efeitos estão ajudando aos alunos a ultrapassarem as notas de corte dos cursos da UFC e assim terem a possibilidade de ingressarem em uma faculdade. Vale ressaltar que, o ideal seria observar a tomada de decisão de todos os alunos após o ENEM, ou seja, identificar se eles ingressaram no ensino superior, público ou privado, e qual curso escolheram, porém por indisponibilidade de informações, optou-se por construir diferentes variáveis binárias elaboradas com base na distribuição das notas de corte, somente para os cursos da UFC, os quais buscam verificar em diversos quantis, aqueles alunos que atingiram esses valores e assim, poderiam escolher cursos que possuem nota de corte inferior a estes. Essas variáveis juntamente com as demais estão descritas no quadro a seguir.

---

1. Ver notas disponíveis em: [http://www.sisu.ufc.br/images/arquivos/notas\\_corte/notas\\_de\\_corte\\_2013.pdf](http://www.sisu.ufc.br/images/arquivos/notas_corte/notas_de_corte_2013.pdf)



## QUADRO 01 – Indicadores de Impacto da Política das EEEPs

Nota_CN	Nota do aluno i na área Ciências da Natureza e suas tecnologias
Nota_CH	Nota do aluno i na área Ciências Humanas e suas tecnologias
Nota_LC	Nota do aluno i na área Linguagem, Códigos e suas tecnologias
Nota_MT	Nota do aluno i na área Matemática e suas tecnologias
Nota_R	Nota do aluno i em Redação
Nota_MG	Nota do aluno i na Média Geral
Corte_0,01	Assume valor igual a 1 se o aluno obteve média geral superior a 385.56 e 0 c.c
Corte_0,05	Assume valor igual a 1 se o aluno obteve média geral superior a 470.98 e 0 c.c
Corte_0,1	Assume valor igual a 1 se o aluno obteve média geral superior a 503.62 e 0 c.c
Corte_0,25	Assume valor igual a 1 se o aluno obteve média geral superior a 579.58 e 0 c.c
Corte_0,5	Assume valor igual a 1 se o aluno obteve média geral superior a 625.04 e 0 c.c
Corte_0,75	Assume valor igual a 1 se o aluno obteve média geral superior a 664.04 e 0 c.c
Corte_0,9	Assume valor igual a 1 se o aluno obteve média geral superior a 697.14 e 0 c.c
Corte_0,95	Assume valor igual a 1 se o aluno obteve média geral superior a 714.68 e 0 c.c
Corte_0,99	Assume valor igual a 1 se o aluno obteve média geral superior a 765.37 e 0 c.c
Trabalho	Assume valor igual a 1 se o aluno estiver trabalhando e 0 caso contrário

Fonte: Elaboração Própria

As informações referentes ao desempenho nas cinco áreas de conhecimento do ENEM 2013<sup>2</sup> e às condições socioeconômicas dos alunos das escolas da rede estadual de educação do Ceará foram retiradas do INEP. A identificação das escolas profissionalizantes e suas dotações de infraestrutura foram obtidas da SEDUC e Censo Escolar/2013, respectivamente.

Para atender o objetivo proposto por este trabalho, precisa-se construir um grupo que pertence ao programa (grupo de tratados) e outro grupo que não foi beneficiado (grupo de controle). A partir da lista de escolas disponibilizadas pela SEDUC, criou-se uma variável qualitativa binária que especifica a escola EEEP e identifica o grupo de tratamento.

2. Último ano com informações disponíveis sobre o desempenho individual no ENEM.

### 3.3 Descrição das Variáveis de Controle

Buscando garantir que o modelo capte o efeito do programa, fez-se necessário controlar características dos alunos e das escolas de forma a isolar esse efeito. Estas variáveis, descritas no quadro (2), foram escolhidas com o intuito de captar as características pessoais dos alunos, status educacional dos pais e as condições de infraestrutura das escolas.

#### QUADRO 02 – Variáveis de Controle

Variáveis	Descrição
<b>Alunos</b>	
Idade	Anos de idade
Sexo	1 se masculino; 0 caso contrário
Raça	1 se branca; 0 caso contrário
Estado civil	1 se solteiro; 0 caso contrário
<b>Escolaridade Pai</b>	
E_fund	1 se o pai possui escolaridade até o ensino fundamental; 0 caso contrário
E_sup	1 se o pai possui de escolaridade no mínimo o ensino superior incompleto; 0 caso contrário
<b>Escolaridade Mãe</b>	
E_fund	1 se o mãe possui escolaridade até o ensino fundamental; 0 caso contrário
E_sup	1 se o mãe possui de escolaridade no mínimo o ensino superior incompleto; 0 caso contrário
<b>Condições Socioeconômicas</b>	
Ate2SM	1 se a renda familiar for até 2 salários mínimos; 0 caso contrário
CSE*	Índice de condição socioeconômico
<b>Escolas</b>	
Infra**	Índice de infraestrutura da escola

Fonte: Elaboração própria.

Nota (\*): Índice criado a partir da análise de componente principal que engloba a existência das seguintes características dos alunos obtidas pelo ENEM 2013: Se possui TV, DVD, radio, microcomputador, automóvel, máquina de lavar, geladeira, freezer, telefone fixo, celular, acesso à internet, TV por assinatura, aspirador de pó, empregada, banheiro. (\*\*): Índice gerado pela análise de componentes principais que engloba a existência das seguintes características das escolas obtidas no Censo Escolar 2013: abastecimento de água por rede pública, de esgoto sanitário por rede pública, de coleta periódica de lixo, de sala de diretoria, de sala de professores, de laboratório de informática e de ciências, de quadra de esportes, de cozinha, de biblioteca, de parque infantil, de sanitário adequado a aluno portador de necessidades especiais (PNEs), de dependências e vias adequadas a PNEs, de televisão, de videocassete, de DVD, de parabólica, de copiadora, de retroprojeter, de impressora, de computadores, de acesso à internet, e de alimentação escolar para os alunos

Vale ressaltar que, como a variável que reflete a renda familiar está disponível em intervalos de salário mínimo, optou-se por fazer um corte de dois salários, tendo em vista que são pais de alunos de escolas públicas, pressupondo-se assim, não possuírem renda familiar muito elevada. Para facilitar a inter-

pretação da medida dos índices CSE e Infra, as variáveis foram normalizadas, de forma a assumir valores entre 0 a 1, sendo que quanto mais próximo de 1, maior os índices CSE do aluno e Infra da escola.

## 4 RESULTADOS

### 4.1 Análise Descritiva

Estatísticas descritivas das variáveis dependentes e controles relacionadas aos alunos de EEEP e escolas regulares do Ceará em 2013 estão explicitadas na Tabela 1. Observa-se que embora as médias das notas dos alunos das EEEP serem superiores às de alunos de escolas regulares em todas as provas, há que se considerar também a variabilidade dessas notas. Em Ciências da Natureza (CN), o diferencial entre os alunos dessas escolas é 39 pontos e apresenta uma variação relativa de 13,5% no grupo do ensino profissional e 12,4% no ensino tradicional. As áreas Ciências Humanas (CH) e Linguagem e Códigos (LC) apresentam diferença semelhante, 52,9 e 52,6 pontos, respectivamente, com dispersão relativa de 14% e 13% para as EEEP e 14,6% e 15% para as regulares, nas duas áreas.

TABELA 1 - Estatísticas Descritivas das variáveis da amostra – 2013

Variáveis	EEEP			Regular		
	Média	Desvio-padrão	Coefficiente de variação	Média	Desvio-padrão	Coefficiente de variação
Dependentes:						
Nota_CN	476,00	64,06	0,1345	436,99	54,23	0,1240
Nota_CH	524,93	73,04	0,1391	472,02	69,21	0,1466
Nota_MT	529,63	87,73	0,1656	463,30	80,41	0,1735
Nota_LC	502,35	66,10	0,1315	449,71	68,07	0,1513
Nota_Red	563,64	127,05	0,2254	470,33	130,65	0,2777
Nota_MG	519,31	64,10	0,1234	458,47	60,65	0,1322
Corte_0,01	0,9352	0,24	0,2566	0,7621	0,42	0,5511
Corte_0,05	0,7783	0,41	0,5267	0,4833	0,49	1,0138
Corte_0,1	0,6583	0,47	0,7139	0,3554	0,47	1,3224
Corte_0,25	0,3977	0,48	1,2069	0,1654	0,37	2,2370
Corte_0,5	0,2382	0,42	1,7632	0,0797	0,27	3,3877
Corte_0,75	0,1553	0,36	2,3180	0,0469	0,21	4,4776
Corte_0,9	0,1098	0,31	2,8233	0,0331	0,17	5,1359
Corte_0,95	0,0899	0,28	3,1145	0,0261	0,15	5,7471
Corte_0,99	0,0355	0,18	5,0704	0,0096	0,09	9,3750
Trabalho	0,0161	0,12	7,4534	0,1480	0,35	2,3648
Controles:						
Características_Alunos						
Idade	17	0,85	0,05	19	4,43	0,2331
Sexo	0,41	0,49	1,1951	0,43	0,49	1,1395
Estado Civil	0,99	0,08	0,0808	0,94	0,21	0,2234
Raça	0,22	0,41	1,8636	0,19	0,39	2,0526
Escolaridade_Pai						
E_fund	0,64	0,47	0,7343	0,80	0,39	0,4875
E_sup	0,05	0,23	4,6	0,02	0,16	8
Escolaridade_Mãe						
E_fund	0,53	0,49	0,9245	0,73	0,44	0,6024
E_sup	0,12	0,33	2,75	0,05	0,23	4,6
Condições						
Socioeconômicas						
Até2SM	0,90	0,28	0,3111	0,94	0,23	0,2446
NSE	0,45	0,13	0,2888	0,39	0,13	0,3333
Características_Escolas						
Infra	0,30	0,07	0,2333	0,27	0,11	0,4074

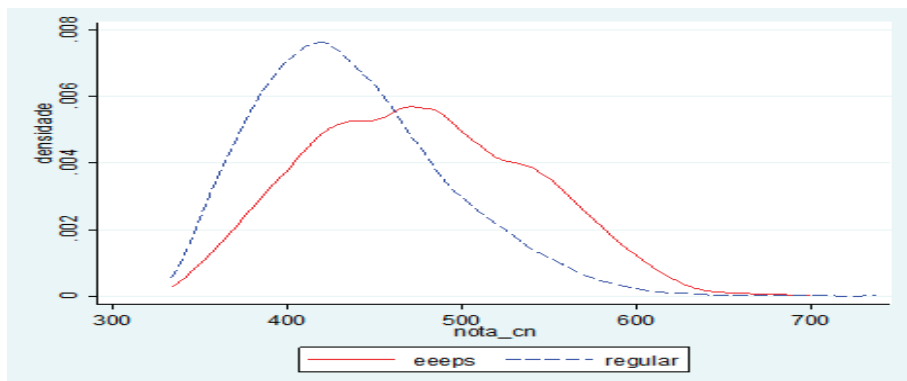
Fonte: Elaboração Própria com base nos Microdados do ENEM e do Censo Escolar de 2013

Apresentando diferenciais mais elevados e menos homogêneo, estão as provas de Matemática (MT) e Redação (Red), com diferença de 66,33 e 93,31 pontos, respectivamente, e variação relativa de, aproximadamente 17% em MT, nos dois tipos de escolas, e 22,5%, em Red, para as EEEP, e 27,8% para as regulares. Na Média Geral, os alunos profissionalizantes apresentaram desempenho superior em, aproximadamente, 61 pontos, 13% a mais que os alunos das escolares de ensino regular, porém, observa-se que, apesar da variação absoluta ser maior para o grupo das escolas profissionais, estes apresentam variação relativa menor aos do ensino regular, 12,3% e 13,2%, respectivamente. Além disso, optou-se por verificar se os alunos conseguiram ultrapassar as notas de cortes dos cursos da UFC. Definida a partir do primeiro e último quantil dessas notas, tem-se que 93,52% dos alunos da EEEP obtiverem média geral superior a 385,56 e 3,55% ultrapassaram 765.37 pontos, enquanto que nas regulares, apenas 76,21% e 0,96%, excederam esses valores, respectivamente.

Quanto a variável trabalho, o qual indica se o aluno está trabalhando no período que prestou ao exame, observa-se que, apenas 1,6% dos alunos das profissionalizantes estavam trabalhando, enquanto que nas escolas regulares esse percentual é de 14,8%, porém, vale ressaltar que a idade média dos alunos das EEEP é de 17 anos, menor do que a requerida para a entrada no mercado de trabalho.

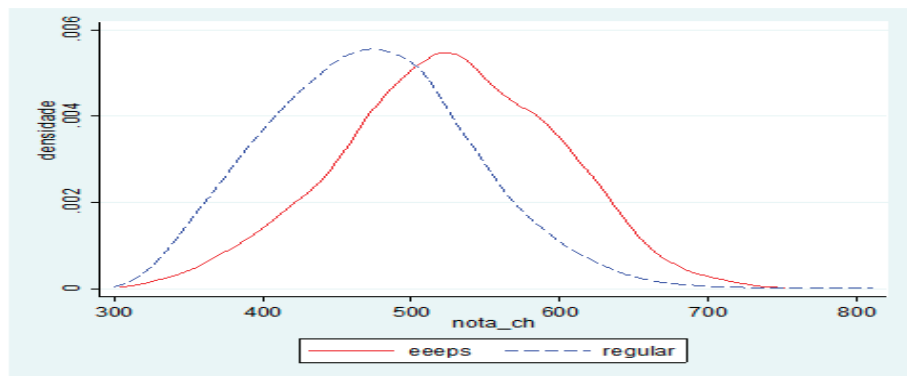
Os gráficos 3 a 8 mostram a distribuição das proficiências dos alunos nas cinco áreas e na média geral, para as escolas profissionalizantes, grupo de tratados, e para as regulares, grupo de controle, representados pelas linhas contínuas e pontilhadas, respectivamente. Através destes, pode-se observar que o gráfico para os alunos do ensino profissional encontra-se mais concentrado para a direita, indicando que a frequência de alunos com notas maiores é superior aos alunos pertencentes ao ensino tradicional.

**GRÁFICO 3 - Distribuição das notas em Ciências Naturais**



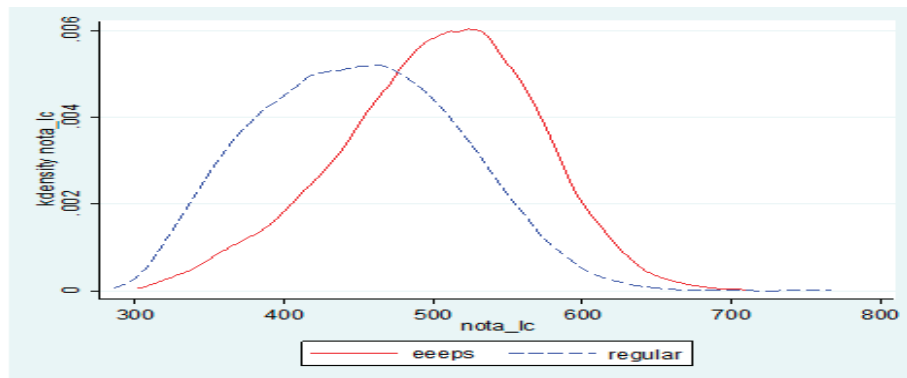
Fonte: Elaboração Própria com base no ENEM 2013.

**GRÁFICO 4 - Distribuição das notas em Ciências Humanas**



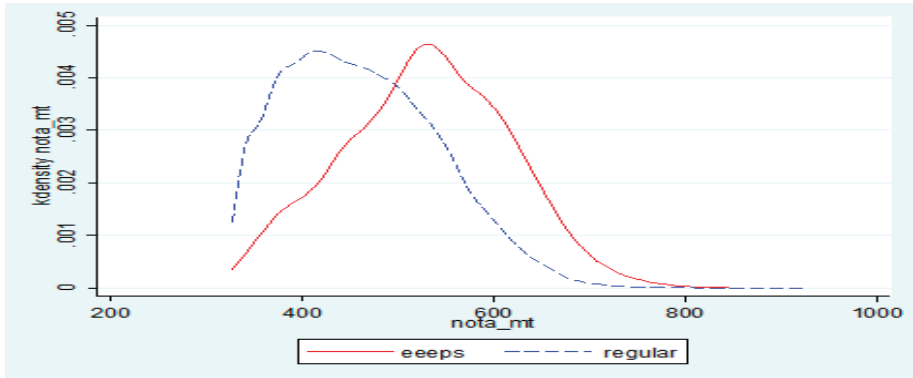
Fonte: Elaboração Própria com base no ENEM 2013.

**GRÁFICO 5 - Distribuição das notas em Linguagens e Códigos**



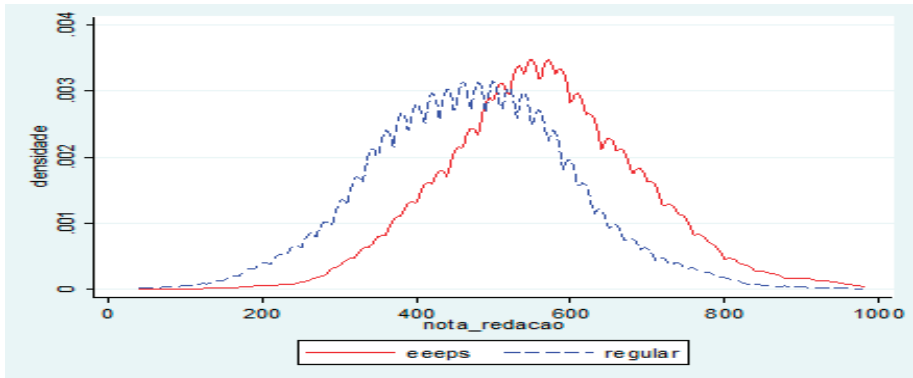
Fonte: Elaboração Própria com base no ENEM 2013.

**GRÁFICO 6 - Distribuição das notas em Matemática**



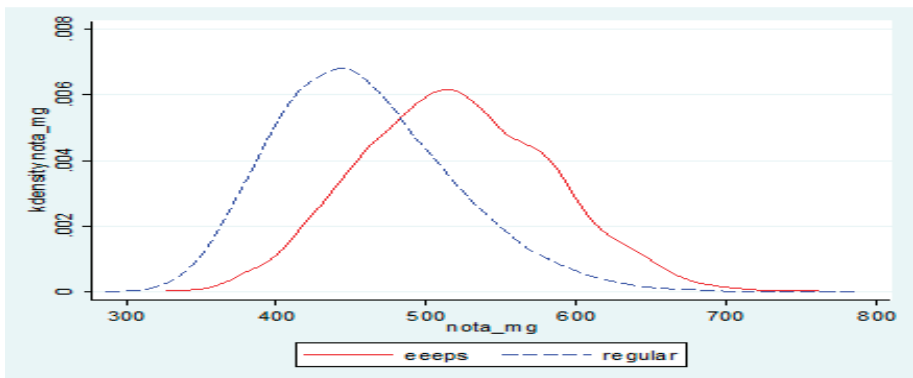
Fonte: Elaboração Própria com base no ENEM 2013.

**GRÁFICO 7 - Distribuição das notas em Redação**



Fonte: Elaboração Própria com base no ENEM 2013.

**GRÁFICO 8 - Distribuição das notas em Média Geral**



Fonte: Elaboração Própria com base no ENEM 2013.

A partir dessa análise descritiva, espera-se que o efeito da política do ensino profissionalizante seja significativo sobre a performance dos alunos no ENEM, porém somente através da comparação entre esses grupos não se pode afirmar isso. Essa é uma técnica “ingênua” comumente usada por não especialistas, pois apenas por não ter passado pela intervenção não significa que o grupo de não tratados representa bem o que ocorreria com o grupo de tratamento caso este não tivesse sido tratado, ou seja, a ausência do tratamento para alguns não gera automaticamente o contrafactual de não tratamento para outros. Dessa forma, são necessárias metodologias apropriadas que busquem isolar o efeito dos programas dos efeitos causados por outros fatores que afetam a(s) variável(eis) de resultado de interesse.

Referente as demais variáveis presentes na Tabela 1, pode-se afirmar que os alunos das escolas profissionalizantes possuem em média 17 anos e desvio padrão de meses, enquanto que os alunos das escolas regulares apresentam média de 19 anos e com variação absoluta de 4 anos. Isso mostra indícios que nas escolas tradicionais existe distorção idade-série e problemas de repetência, diferentemente das profissionalizantes.

Além disso, nas EEEP 41% são homens, 99% solteiros, 22% considerados brancos, 64% e 53% apresentam pais e mães com escolaridade até o fundamental, 5% e 12 % com pelo menos nível superior incompleto, respectivamente. Nessa amostra, possuem, em média, índice de condição socioeconômica (CSE) de 0,45 e 90% dos alunos possuem renda familiar até dois salários mínimos. Já nas regulares, os alunos são maioria mulheres, apresentam menores proporções de solteiros e brancos, de pais com no mínimo superior incompleto, índice CSE de 0,39, maiores proporções de pais e mães com escolaridade até o fundamental e 94% possuem renda familiar até dois salários mínimos e possuem índice de infraestrutura próximo às escolas de ensino profissional.

#### **4.2 Estimação do pareamento com escore de propensão**

Seguindo os procedimentos de Becker e Ichino (2002), inicialmente estima-se um modelo logit binário de o aluno pertencer ou não a uma escola profissionalizante utilizando, como características observáveis, as mesmas variáveis selecionadas para descrever os grupos de tratamento e controle, de forma a se encontrar um suporte comum de probabilidade entre os alunos dos dois tipos de escola, cujos resultados estão expostos na Tabela 2. Posteriormente, descarta-se da amostra aqueles alunos com escore de propensão fora do suporte. Em

seguida, o modelo escolhido deverá atender as seguintes propriedades: (1) teste da *balancing property*, ou seja, existir um número suficiente de alunos no ensino regular sob o mesmo intervalo de probabilidade de pertencer a profissional, de forma a se promover um pareamento (terem, em média, mesmas covariadas); e, (2) testes de robustez do Pseudo-R2 e da análise gráfica da densidade.

Os resultados mostram que, as variáveis raça e E\_sup do pai foram estatisticamente irrelevantes, considerando nível de significância a 5%. Observa-se que quanto maior a idade menor a probabilidade do aluno pertencer a EEEP, enquanto que, se for uma mulher, solteira, com pai e mãe de escolaridade acima do fundamental, possuir renda familiar até dois salários mínimos e quanto maior o índice de condição socioeconômica, maiores serão as chances desse aluno pertencer a escola profissional.

Tabela 2 – Estimação do Modelo Logit Binário

Variáveis	Coefficientes	Desvio-padrão	p-valor
<b>Características_Alunos</b>			
Idade	-0,3184	0,0115	0,000
Sexo	-0,1285	0,0263	0,000
Estado Civil	0,6569	0,1354	0,000
Raça	0,0250	0,0313	0,422
<b>Escolaridade_Pai</b>			
E_fund	-0,3351	0,0322	0,000
E_sup	-0,1055	0,0638	0,098
<b>Escolaridade_Mãe</b>			
E_fund	-0,4127	0,0313	0,000
E_sup	0,1518	0,0456	0,000
<b>Condições Socioeconômicas</b>			
Ate2SM	0,215	0,0498	0,000
CSE	2,0614	0,0965	0,000
<b>Características_escolas</b>			
Infra	1,8739	0,1105	0,000
Constante	1,9585	0,2654	0,000
Log likelihood	20657,973		
Pseudo R2	0.0869		
Obs.	62109		
Suporte Comum	[0.00152955, 0.60850558]		

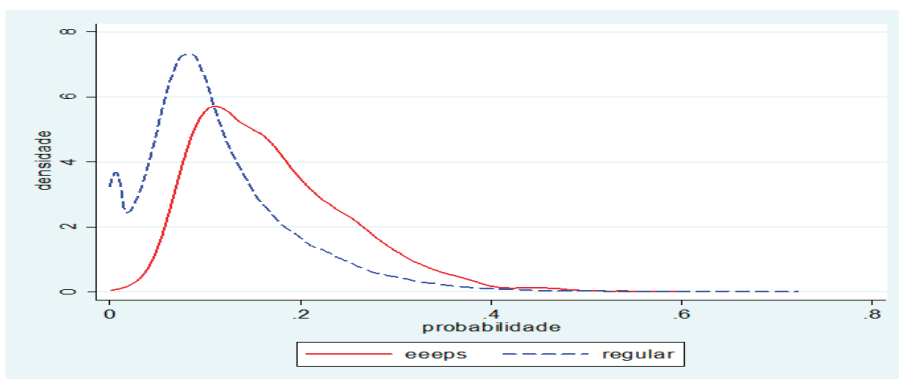
Fonte: Elaboração Própria.

Verifica-se que a região de suporte comum foi o intervalo, cujo escore de propensão varia de 0.00152955 a 0.60850558. Os gráficos 9 e 10 apresentam



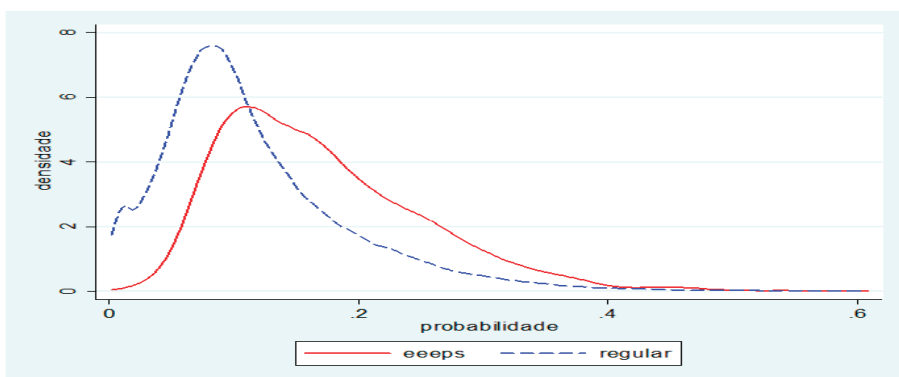
as densidades de probabilidade estimadas de pertencer ao ensino profissional entre os alunos para toda a amostra e somente os que estão no suporte comum, respectivamente. Dos 62.109 alunos para o ano de 2013, 60.207 (7370 da profissional) ficaram dentro intervalo. Essa região assegura que as observações do grupo de tratamento tenham observações comparáveis do grupo de controle quanto às características pessoais dos alunos, status educacional dos pais e as condições de infraestrutura das escolas.

GRÁFICO 9 - Distribuição do escore de propensão – toda amostra



Fonte: Elaboração Própria.

GRÁFICO 10 - Distribuição do escore de propensão – região suporte comum.



Fonte: Elaboração Própria.

Dessa forma, o *matching* para se obter o Efeito Médio de Tratamento (EMT) se dará dentro deste intervalo. Porém, outros métodos fazem-se neces-

sário para se estimar o EMT, além de se utilizar somente o escore de propensão, tais como: *nearest-neighbor*, *Radius* e *Kernel*. Um procedimento importante com relação à construção do escore de propensão e de implementação do pareamento é a checagem das condições de balanceamento. A Tabela A2 em anexo mostra o teste da diferença das médias das variáveis no tratamento e no controle antes e após o pareamento. A hipótese nula do teste é que a diferença das médias é zero, ou seja, as médias dos dois grupos são estatisticamente iguais. Antes do pareamento todas as médias foram estatisticamente diferentes, porém depois de pareados, percebe-se que em quase todos os casos tem-se um pareamento com um bom balanceamento das variáveis.

Através do pareamento pelo *nearest-neighbor* somente a variável infraestrutura da escola revelou coeficiente significativo ao nível de 5%. Os resultados para o *radius* (0,01) e *kernel* (0,01) foram semelhantes, tendo como significativas as variáveis idade, E\_sup para a mãe e Infra, porém para o pareamento por *radius*, a variável E\_fund para o pai também foi significativa. Além disso, o **Pseudo - R<sup>2</sup>** reduz de 0,087 para 0,002, demonstrando que as variáveis explicativas perdem poder para explicar as diferenças entre os alunos. Dessa forma, o pareamento pelo *nearest-neighbor* se mostrou o mais balanceado.

Os diferenciais de notas entre os dois tipos de escola estimados por pareamento com primeiro *nearest-neighbor*<sup>3</sup> estão dispostos na Tabela 3, a qual revela serem tais diferenciais significativos a favor das EEEP em todas as áreas de conhecimento. Em média, um aluno pertencente a escola profissional apresenta desempenho superior em Ciências humanas (CH) e em Linguagens e Códigos (LC), aproximadamente, 8% a mais do que um aluno da escola regular pertencente ao grupo de controle. Em Matemática (MT) o diferencial é de 10,43% em favor dos alunos de ensino profissional. Menor e Maior efeito foi observado na área de Ciências da Natureza (CN) e em Redação (Red), de 6,8% e 13,27%, respectivamente. Na Média Geral, o impacto das EEEP foi de 9,53%. A partir das variáveis binárias construídas com base nas notas de corte, pode-se afirmar que os alunos do ensino profissionalizante apresentam, em média, probabilidade maior de ultrapassar essas notas em todos os quantis estabelecidos e assim ingressarem em cursos mais concorridos, cuja nota de corte é maior. No primeiro corte, tem-se um diferencial de 10 pontos percentuais para os alunos da EEEP, chegando a uma diferença de 22 p.p ao considerar o terceiro corte, ou seja, as EEEP apresentam, em média, 22 p.p a mais de um aluno obter média geral superior a 503,62 em relação as escolas regulares e assim poderem escolher entre cursos como Engenharia da Computação em Sobral, Letras (Português –Espanhol) em Fortaleza, dentre outros. A partir deste corte,

3. Os diferenciais estimados com base nos métodos Radius, Kernel e Stratification também foram encontrados e podem ser verificados na tabela A3 em anexo.

o diferencial se torna decrescente, porém continua positivo, com valor mínimo no último corte de, aproximadamente 2 p.p, indicando assim, que esses alunos possuem probabilidade maior, embora pequena, de obterem notas equivalentes as estabelecidas para os cursos de medicina nas cidades de Fortaleza, Sobral e Cariri, por exemplo, se comparados aos alunos das escolas tradicionais. Em suma, o ensino profissional de nível médio no Ceará demonstra ser mais competente em potencializar habilidades do conhecimento cognitivo comparativamente ao ensino regular.

Quanto aos efeitos relacionados à inserção no mercado de trabalho, observa-se que os alunos de EEEP possuem, em média, 11,3 (*nearest-neighbor*) pontos percentuais a menos do aluno estar trabalhando em relação aos alunos das escolas regular. Resultado semelhante foi encontrado por Araújo et al (2014) ao mostrar que o efeito seria uma menor probabilidade de trabalhar do aluno de EPT (Educação Profissional e Tecnológica) em relação ao do ensino regular em aproximadamente 2 pontos percentuais.

Tal resultado pode estar associado ao perfil dos estudantes das EEEP, onde estes podem estar priorizando continuar os estudos e almejem alcançar um nível superior a começar a trabalhar.

**TABELA 3 - Efeito médio do tratamento nas áreas de conhecimento do ENEM, nas variáveis de corte e na inserção no mercado de trabalho por *nearest-neighbor***

Variáveis	EEEP	Regular	EMT	Estatística t
Notas_CN	476,00	445,67	30,32	21,37
Notas_CH	524,93	483,75	41,18	21,46
Notas_MT	529,63	479,58	50,05	25,29
Notas_LC	502,35	463,97	38,38	23,12
Notas_Red	563,64	497,57	66,06	17,74
Notas_MG	519,31	474,11	45,20	26,91
Corte_0,01	0,9352	0,8344	0,101	16,71
Corte_0,05	0,7783	0,5707	0,208	21,39
Corte_0,1	0,6583	0,4387	0,220	23,96
Corte_0,25	0,4894	0,2246	0,173	21,97
Corte_0,5	0,2382	0,1126	0,126	19,20
Corte_0,75	0,1553	0,0646	0,091	17,36
Corte_0,9	0,1098	0,0465	0,063	15,18
Corte_0,95	0,0899	0,0371	0,053	13,56
Corte_0,99	0,0355	0,0163	0,019	6,88
Trabalho	0,016	0,128	-0,113	-20,47

Fonte: Elaboração própria

Segundo o Relatório de Gestão (2014), o percentual de inserção de egressos das escolas profissionalizantes no mercado de trabalho reduziu de 27,7%

em 2011 para 20,4% em 2014, enquanto que o percentual de aprovação desses alunos em universidades aumentou de 16,9% para 37,4% nesse período. Nesse sentido, pode estar existindo um trade-off entre desempenho acadêmico e inserção no mercado de trabalho, porém, outros fatores devem ser levados em consideração, como a idade dos alunos, pois ao concluírem o ensino profissionalizante com idade inferior a 18 anos, impossibilita a entrada imediata no mercado de trabalho, além de outros fatores subjetivos aos alunos. Esses resultados apontam para uma necessidade em acompanhar esses jovens após o término do ensino médio para verificar de forma mais detalhada se estes estão trabalhando e/ou ingressaram em uma universidade.

## CONCLUSÃO

Este trabalho procura ampliar o debate existente na literatura da Educação sobre o ensino profissionalizante, com ênfase nos efeitos gerados no desempenho escolar e na inserção no mercado de trabalho. Para tanto, busca-se trazer novas evidências sobre o papel das escolas profissionais públicas do Ceará tanto em nível social de inserção produtiva quanto educacional. Esses dois aspectos são avaliados conjuntamente ao se comparar alunos dessas escolas com outros do ensino regular, testando-se a existência de diferenciais significativos entre eles.

A opção metodológica para a condução dessa avaliação baseou-se em um modelo de pareamento por score de propensão aplicado a dados oficiais de 2013, objetivando encontrar o efeito médio dessa política sobre o desempenho dos estudantes nas cinco áreas e na média geral do ENEM; sobre as chances desses alunos atingirem média geral mais elevada e ultrapassarem as notas de corte dos cursos da UFC; e sobre as chances de estarem trabalhando.

Após o *matching* realizado através do método *nearest-neighbor*, o qual se mostrou o mais balanceado, verificou-se que os alunos das escolas profissionalizantes apresentam diferenciais significativos em todas as áreas de conhecimento, entre os que fizeram EEEP e aqueles que cursaram escola regular apenas, cujo efeito mais expressivo foi em Redação, 13,27%, a mais para um aluno da escola profissional. Além disso, estes alunos possuem probabilidade maior de ultrapassar as notas de cortes em todos os quantis estabelecidos, tendo como destaque o terceiro, cujo diferencial foi, em média, 22 p.p a mais para os alunos obterem média geral superior a 503,62 pontos em relação às escolas regulares e

embora apresentando diferencial menor, esses alunos possuem maiores chances de obterem notas equivalentes as estabelecidas para os cursos de medicina nas cidades de Fortaleza, Sobral e Cariri, por exemplo. Isso mostra indícios que, essas escolas estão proporcionando um ensino de qualidade a estes jovens e ampliando a possibilidade de ingressarem em uma faculdade.

Dos resultados, observou-se ainda, que os alunos da rede profissional possuem menor probabilidade de estar trabalhando, no período que prestaram o exame, em relação aos alunos das escolas regular. Vale ressaltar, que esse resultado pode estar sendo afetado pela idade dos alunos, pois constatou-se que a idade média do aluno concluir o ensino profissionalizante é 17 anos, inferior a mínima para entrada imediata no mercado de trabalho. Contudo, o Relatório de gestão (2014) verificou que houve uma redução do ingresso de concludentes das EEEP no mercado de trabalho. Dessa forma, reitera o desafio de identificar, com maior precisão, os motivos que estão levando os alunos a esta escolha.

Diante do exposto, as EEEP do estado Ceará estão rompendo com a dualidade estrutural entre formação para o trabalho e preparação para a universidade, ao integrar as duas redes de ensino, profissional e geral, visando não somente o mercado de trabalho, mas também o ensino superior. Assim, essas escolas podem servir como exemplos e modelos para a continuidade dessa rede de ensino no estado e para implementação desse tipo de ensino em outros estados. Porém, ressalta-se a necessidade do acompanhamento desses jovens pós ensino médio, não só os do ensino profissionalizante, mas do regular também, para verificar se estes estão no mercado de trabalho e exercendo uma profissão dentro da área o qual se profissionalizou, caso das EEEP, e/ou ingressaram no nível superior, como forma melhorar e/ou redefinir as estratégias utilizadas no modelo integrado de eEnsino Médio à Educação pProfissional e da rede geral de ensino.

## REFERÊNCIAS BIBLIOGRÁFICAS

ANGRIST, J., PISCHKE, J.S. **Most Harmless Econometrics: an empiricist's companion.** Princeton, New Jersey: Princeton University Press. 2009

ARAÚJO, A. J. N; CHEIN, E.; PINTO, C. **Ensino Profissionalizante, Desempenho Escolar e Inserção Produtiva: Uma Análise com dados do ENEM.** 2014 Disponível: <<http://www.anpec.org.br/encontro/2014/submis->

sao/files\_I/i12-ac3a8f487db438fc6278e117ab468a01.pdf>

ASSUNÇÃO J; GONZAGA G. Educação Profissional no Brasil: Inserção e retorno. **Série Cenários**, n. 3, Brasília: SENAI.DN, 2010.

BECKER, S., ICHINO, A. Estimation of Average Treatment Effects based on Propensity Scores. **The Stata Journal**, v.2 n.4, p.358-377. 2002

CALIENDO, M.; KOPEINIG, S. Some Practical Guidance for the Implementation of Propensity Score Matching. **Journal of Economic Surveys**, v.22 n.1, p. 31-72. 2008

CASTRO, M.H.G.; TIEZZI, S. A reforma do ensino médio e a implantação do ENEM no Brasil. In: BROCK, C.; SCHWARTZMAN, S. (Org.). **Os desafios da educação no Brasil**. Rio de Janeiro: Nova Fronteira, 2005. p. 119-154. Disponível em: < <http://www.schwartzman.org.br/simon/desafios/4ensinomedio.pdf>> Acesso em 12 junho 2015

CHEN, X; WEKO, T. US Department of Education NCES 2009. **Aurora**, v. 202, p. 502-7334, 2009.

CROUCH, C., D.; SAKO, M. **Are skills the answer?** The political economy of skill creation in advanced industrial countries. New York: Oxford University Press. 1999

FOLEY, P. **The Socio-Economic Status of Vocational Education and Training Students in Australia**. National Centre for Vocational Education Research Ltd. PO Box 8288, Stational Arcade, Adelaide, SA 5000, Australia, 2007.

GANIMIAN, A. J.; ROCHA, A. S. **Measuring Up?. How Did Latin America and the Caribbean Perform on the 2009 Programme for International Student Assessment (PISA)?**. Partnership for Educational Revitalization in the Americas (PREAL), 60 p., 2011

HANUSHEK, E. A.; WOESSMANN, L. Schooling, Educational Achievement, and the Latin American Growth Puzzle. **Journal of Development Economics**, v.99, n.2, p.497-512, 2012.

HECKMAN, J.; LALONDE, R.; SMITH, J. The economics and econometrics of active labor market programs. *In*: ASHENFELTER, O., CARD, D. (Eds.) **The Handbook of Labor Economics**. Amsterdam: North Holland. v.3A, part.6, cap.31. 1999.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS – INEP. Disponível em: <http://www.inep.gov.br/> Acesso em: 10 julho 2015.

KUENZER, A.Z. Exclusão includente e inclusão excludente: a nova forma de dualidade estrutural que objetiva as novas relações entre educação e trabalho. In: SAVIANI, D.; SANFELICE, J.L.; LOMBARDI, J.C. (Org.). *Capitalismo, trabalho e educação*. 3. ed. Campinas: Autores Associados, 2005. p. 77-96.

LEE, W. Propensity Score Matching and Variations on the Balancing Test. **Mimeo**. Melbourne Institute of Applied Economics and Social Research. 2006

LEVY, S., SCHADY, N. Latin America's Social Policy Challenge: Education, Social Insurance, Redistribution. **The Journal of Economic Perspectives**, v.27, n.2, p.193-218, 2013.

MARTINS, A. P. Pressupostos de Gramsci na educação profissional e tecnológica de nível médio. **Revista de Educação, Ciência e Tecnologia**, v. 1, n. 2, 2012.

MOEHLECKE, Sabrina. O Ensino Médio e as novas diretrizes curriculares nacionais: entre recorrências e novas inquietações. **Revista Brasileira de Educação**, v. 17, n. 49, Rio de Janeiro. Jan- abril, 2012. Disponível em: [http://www.scielo.br/scielo.php?pid=S1413-24782012000100003&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S1413-24782012000100003&script=sci_arttext) Acesso em: 10 julho 2015

NEUMAN, S; ZIDERMAN, A. Vocational Secondary Schools Can Be More Cost Effective than Academic Schools: The Case of Israel. **Comparative Education**, v. 25, n. 2, p. 151- 163, 1989.

NIBOM, A.A.R. O processo de implantação e expansão da rede de Escola Estadual de Educação Profissional no Estado do Ceará (2008-2014): um olhar investigativo. Monografia. MBA em Excelência em Gestão de Projetos e Processos. Centro Estadual de Educação Tecnológica Paula Souza, São Paulo, 2014. Mimeo.

RELATÓRIO DE GESTÃO “O PENSAR E O FAZER DA EDUCAÇÃO PROFISSIONAL NO CEARÁ – 2008 a 2014”. Secretaria da Educação. Coordenadoria de Educação Profissional. – 1. ed. – Fortaleza: Secretaria da Educação, 2014

ROSENBAUM, P.; RUBIN, R. The Central Role of the Propensity Score in Observational Studies for Causal Effects. **Biometrika**, v.70 n.1, p. 41-55. 1983

RUBIN, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. **Journal of Educational Psychology**, v.66 n.5.

1974

SECRETARIA DE EDUCAÇÃO DO CEARÁ - SEDUC. Disponível em: <http://www.seduc.ce.gov.br/index.php/educacao-profissional> Acesso em: 15 de Jun. 2015

SEVERNINI, E. R.; ORELLANO, V. I. F. O efeito do ensino profissionalizante sobre a probabilidade de inserção no mercado de trabalho e sobre a renda no período pré-Planfor. **Revista Economia**, 2010.

SILVA, V. H. O. Análise da Participação das Escolas Públicas Estaduais Cearenses no Exame Nacional do Ensino Médio (ENEM): 2009-2011. **IPECE Informe**, n. 54. Fortaleza, 2013. Disponível em: < [http://www.ipece.ce.gov.br/publicacoes/ipece-informe/Ipece\\_Informe\\_54\\_A26\\_fevereiro\\_2013.pdf](http://www.ipece.ce.gov.br/publicacoes/ipece-informe/Ipece_Informe_54_A26_fevereiro_2013.pdf) > Acesso em: 12 junho 2015

SMITH, J.; TODD, P. Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators? **Journal of Econometrics**, v.125, p. 305-353. 2005

TAVARES, P. A. **Três Ensaio em Economia da Educação**. São Paulo, 2014. Disponível em: < <http://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/11542/Tr%C3%AAs%20Ensaio%20em%20Economia%20da%20Educa%C3%A7%C3%A3o.pdf?sequence=1&isAllowed=y> > Acesso em: 12 junho 2015



## ANEXO

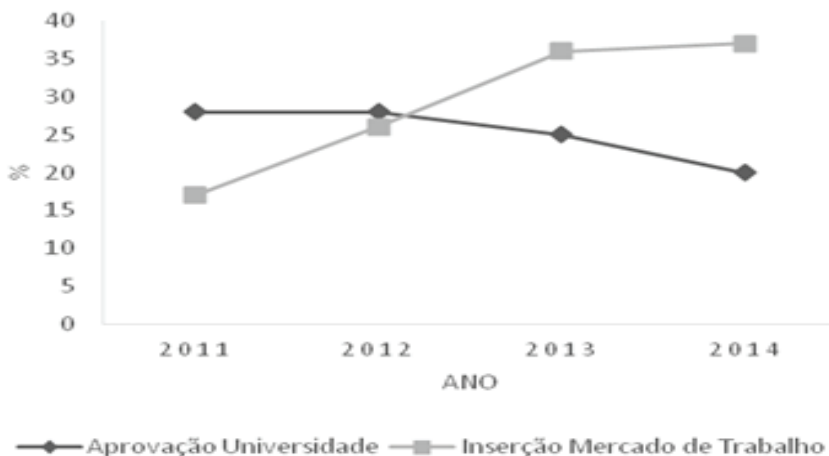
**TABELA A1 - Recursos Financeiros investidos na Educação Profissional do Ceará no período de 2008 a 2014**

Ano	RECURSOS INVESTIDOS* (R\$)		TOTAL (R\$)
	Federal	Estadual	
2008	0,00	2.734.025,15	2.734.025,15
2009	22.674.215,07	30.068.709,51	52.742.924,58
2010	64.507.454,63	159.923.487,47	224.430.942,10
2011	14.920.840,77	146.884.529,28	161.805.370,05
2012	87.538.689,16	135.748.389,70	223.287.078,86
2013	38.635.876,28	115.319.137,72	153.955.014,00
2014	68.096.884,07	149.044.771,41	217.141.655,48
Total	296.373.959,98	739.723.050,24	1.036.097.010,22

Nota(\*): Não incluído recursos de custeio.

Fonte: Sistema Integrado de Acompanhamento de Programas (SIAP/webmapp). Relatório de gestão (2014)

**GRÁFICO A1 - Comparação percentual entre a inserção de egressos das Escolas Estaduais de Educação Profissional no mercado de trabalho e aprovação na universidade, no Ceará, de 2011 a 2014**



Fonte: Secretaria da Educação do Ceará/Coordenadoria de Desenvolvimento da Educação Profissional. Relatório de gestão (2014).

**TABELA A2 - Teste de comparação de média dos grupos de EPT e Regular antes e depois do pareamento por nearest-neighbor, Radius e Kernel**

Variáveis	Sem Pareamento			Com Pareamento nearest-neighbor			Radius (0,01)			Kernel (0,01)		
	Media		p-valor	Media		p-valor	Media		p-valor	Media		p-valor
	Tratados	Controle		Tratados	Controle		Tratados	Controle		Tratados	Controle	
Características dos Alunos												
Idade	17,421	18,975	0,000	17,421	17,391	0,059	17,421	17,352	0,000	17,421	17,350	0,000
Sexo	0,4118	0,4308	0,002	0,4118	0,4226	0,181	0,4118	0,4100	0,825	0,4118	0,4099	0,818
Estado civil	0,9918	0,9497	0,000	0,9918	0,9922	0,781	0,9918	0,9938	0,154	0,9918	0,9938	0,144
Raça	0,2213	0,1918	0,000	0,2213	0,2338	0,080	0,2213	0,2215	0,977	0,2213	0,2215	0,974
Escolaridade_pai												
E_fund	0,6466	0,8033	0,000	0,6466	0,6541	0,342	0,6466	0,6622	0,047	0,6466	0,6617	0,057
E_sup	0,0561	0,0288	0,000	0,0561	0,0561	1,000	0,0561	0,0506	0,133	0,0561	0,0507	0,144
Escolaridade_mae												
E_fund	0,5302	0,7341	0,000	0,5302	0,5260	0,609	0,5302	0,5390	0,284	0,5302	0,5385	0,315
E_sup	0,1274	0,0586	0,000	0,1274	0,1229	0,411	0,1274	0,1150	0,021	0,1274	0,1153	0,025
Renda_familiar												
Ate2sm	0,9095	0,9402	0,000	0,9095	0,9096	0,977	0,9095	0,9147	0,258	0,9095	0,8146	0,271
CSE	0,4583	0,3958	0,000	0,4583	0,4621	0,095	0,4583	0,4542	0,080	0,4583	0,4543	0,092
Característica_escolas												
Infra	0,3035	0,2724	0,000	0,3035	0,2954	0,000	0,3035	0,2980	0,001	0,3035	0,2980	0,001
Ps_R2	0,090			0,002			0,002			0,002		
LR_chi2	4058,91		0,000	39,53		0,000	41,95		0,000	42,81		0,000

Fonte: Elaboração Própria

**TABELA A3 - Efeito médio do tratamento nas áreas de Conhecimento do ENEM e na inserção no mercado de trabalho por Radius, Kernel e Stratification**

	Radius (0,01)	Kernel (0,01)	Stratification
Notas_CN			
EEEEP	476,00	476,00	476,00
Regular	438,83	444,76	444,90
EMT	37,16	31,23	31,10
	(34,63)	(102,58)	(39,72)
Notas_CH			
EEEEP	524,93	524,93	542,93
Regular	473,70	481,93	500,11
EMT	51,23	43,00	42,82
	(59,30)	(73,56)	(54,58)
Notas_MT			
EEEEP	529,63	529,63	529,63
Regular	468,50	478,51	478,77
EMT	61,12	51,11	50,86
	(40,81)	(81,40)	(40,23)
Notas_LC			
EEEEP	502,35	502,35	502,35
Regular	453,25	462,94	463,14
EMT	49,10	39,41	39,21
	(57,3)	(48,29)	(69,52)
Notas_Red			
EEEEP	563,64	563,64	563,64
Regular	479,68	496,92	497,35
EMT	83,95	66,71	66,29
	(113,63)	(39,17)	(31,49)
Notas_MG			
EEEEP	519,31	519,31	519,31
Regular	462,79	473,01	473,25
EMT	56,51	46,29	46,06
	(63,08)	(109,24)	(48,20)
Trabalho			
EEEEP	0,016	0,016	0,016
Regular	0,115	0,121	0,122
EMT	-0,100	-0,105	-0,105
	(-44,56)	(-36,84)	(-58,43)

Nota: Estatística t entre parênteses.

Fonte: Elaboração própria

# A PERDA DE BEM ESTAR SOCIAL DA POPULAÇÃO DE JUAZEIRO DO NORTE COM O DESPERDÍCIO DE ÁGUA POTÁVEL

Joelice Alexandre de Oliveira\*

Wellington Ribeiro Justo\*\*

## Resumo

A exploração desenfreada dos recursos naturais tem gerado graves consequências para o planeta. Destaca-se a escassez de água potável, problema que já preocupa várias cidades e metrópoles em todo o mundo. O desperdício de água tratada, ocasionado pela distribuição ineficiente, é apontado como maior vilão da escassez. A alocação eficiente deste recurso natural eleva o bem-estar social. Nesse contexto, surge a necessidade de conhecer o impacto do desperdício de água potável no bem-estar social da população. O objetivo deste trabalho é mensurar os efeitos do desperdício de água tratada no bem-estar social da população de Juazeiro do Norte. Para atender esse objetivo, trabalhou-se com dados secundários fornecidos pela Companhia de Água e Esgoto do Ceará – CAGECE, unidade de Juazeiro do Norte, no período compreendido entre janeiro de 2006 e dezembro de 2013. A metodologia utilizada neste trabalho é a análise de Séries Temporais, através da abordagem de Box-Jenkins. De acordo com a metodologia, o modelo escolhido foi o AR (12) MA (1). Após efetuar os testes e realizar as análises, os resultados obtidos, atendendo as perspectivas iniciais, mostraram que a população de Juazeiro do Norte está perdendo bem-estar social em virtude do desperdício de água proporcionado pela CAGECE. Essa perda ultrapassou R\$ 6,2 milhões em 2014, a perspectiva é que ela cresça e aproxime-se de R\$ 6,3 milhões em 2015. Conclui-se, então, que a escassez de água potável é um problema de alta relevância e que o caso do desperdício precisa de maior atenção pelos agentes tomadores de decisão, essencialmente neste momento de escassez hídrica.

**Palavras-Chave:** Desperdício de água; Eficiência de Pareto; modelos ARIMA; abordagem de Box-Jenkins; Bem-Estar Social.

---

\* Economista (URCA). joeliceoliveira17@yahoo.com.br (88)99911 0746.

\*\* Professor Associado do Curso de Economia da URCA. Professor do PPGECON (UFPE). Doutor em Economia pelo PIMES (UFPE). justowr@yahoo.com.br (81) 98848 1898.

## Abstract

The unbridled exploration of natural resources has generated serious consequences for the planet. The shortage of drinking water stands out a problem that worries many towns and cities around the world. The treated water waste caused by inefficient distribution is named as greatest villain of shortage. The efficient allocation of this natural resource raises the social welfare. In this context, arises the need to know the impact of drinking water waste in the social welfare of the population. The objective of this study is to measure the effects of treated water waste in the welfare of Juazeiro population. To meet this goal, we worked with secondary data provided by the Ceará's Water and Sewerage Company - CAGECE, Juazeiro unit in the period between January 2006 and December 2013. The methodology used in this work is the analysis of Time Series, by Box-Jenkins approach. According to the methodology, the AR (12) MA (1) model was chosen. After performing the tests and accomplish the analyzes, the gotten results, met the initial perspectives, showed that the population of Juazeiro is losing welfare due to the waste of water provided by CAGECE. This loss exceeded R\$ 6.2 million in 2014, the prospect is it to grow and approach to R\$ 6.3 million in 2015. It follows, then, that the shortage of drinking water is a highly relevant issue and that the case of waste needs greater attention by decision maker agents, essentially this time of water scarcity.

**Key words:** Water waste; Pareto Efficiency; ARIMA models; Box-Jenkins approach; Social Welfare.

**JEL:**Q1, C5.

## 1 INTRODUÇÃO

Atualmente, quase todos os países têm se voltado para a preocupação com a preservação do planeta, o que se chama de Desenvolvimento Sustentável – uso consciente dos recursos naturais, preservação da fauna, da flora, bem como do meio ambiente como um todo. Conferências como a Rio+20<sup>1</sup> vêm acontecendo com a finalidade de discutir o problema e traçar metas, coletivas e individuais, para conciliar o crescimento econômico com a preservação do

---

1. A Rio + 20, Conferência das Nações Unidas sobre Desenvolvimento Sustentável, ocorreu entre os dias 13 a 22 de junho de 2012 na cidade do Rio de Janeiro. Um de seus maiores legados foi o compromisso por prefeitos das maiores cidades do mundo de reduzir as emissões de gases de efeito estufa em 12% até 2016, e em 1,3 bilhão de tonelada até 2030.

meio ambiente nas próximas décadas. Nesse contexto, muito se tem falado em economia sustentável e em uso consciente dos recursos naturais (ONU, 2014a).

Muitas empresas vêm dando atenção especial ao assunto e têm criado setores específicos para cuidar desta causa, a exemplo da Companhia de Água e Esgoto do Ceará – CAGECE, que tem demonstrado uma preocupação constante com o meio ambiente. Recentemente, a mesma empresa criou a Gerência de Meio Ambiente e implantou o Sistema de Gestão Ambiental (SGA), em três de suas principais Estações de Tratamento de Água (ETAs), visando compatibilizar suas atividades com a preservação ambiental dentro dos princípios do desenvolvimento sustentável (CAGECE, 2014).

Dentre os problemas gerados pela exploração infrene dos recursos naturais está a escassez de água doce que, por sua vez, está diretamente relacionada com o desperdício. Algumas cidades já sofrem as consequências advindas do mau uso da água, como é o caso de São Paulo, que ultimamente tem sofrido com o problema da falta de água potável. A falta de água na capital paulista é, na maior parte, gerada pela ineficiência da distribuição, que conta com uma tubulação de mais de trinta anos de uso para mais de 51% de sua totalidade, fato que aumenta os casos de vazamentos – maior vilão do desperdício (SANTOS, LOIOLA, 2014).

Segundo Santos e Loiola (2014), em 2013 a Companhia de Saneamento Básico do Estado de São Paulo (Sabesp) perdeu 31,2% da água produzida entre a estação de tratamento e as caixas d'água dos consumidores, o que representa uma perda de cerca de 950 bilhões de litros. O problema de escassez levou o governador de São Paulo a tomar medidas como estipular multas para o consumo excessivo de água e oferecer descontos para quem poupar na tentativa de mitigar o problema.

A água é um bem precioso e indispensável à sobrevivência do planeta. Por isso, faz-se necessário o seu uso consciente e a sua preservação, para que problemas de escassez não se alastrem pelo mundo e a sua falta venha causar maiores danos ao planeta como um todo (ONU, 2012b).

No contexto da preocupação com o meio ambiente e sustentabilidade, e considerando que, segundo Varian (2012), a alocação eficiente eleva o bem-estar social, assim como, segundo Santos (2014) a distribuição ineficiente produz escassez de água que traz graves consequências para a população. Daí surge a necessidade de conhecer qual é o impacto do desperdício de água potável para o bem-estar social de uma comunidade? A hipótese é de que o desperdício de

água potável causa um impacto negativo no bem estar social da população. O tema em debate é de grande relevância. Seu estudo justifica-se pelo fato que, o conhecimento sobre os efeitos do desperdício de água tratada e suas dimensões poderá ajudar no combate à escassez. Para conhecer melhor o assunto e descobrir os efeitos do desperdício de água sobre uma comunidade, elegeu-se a cidade de Juazeiro do Norte como área de estudo.

Além desta introdução o artigo tem mais quatro seções. A seção seguinte faz-se uma discussão sobre a sustentabilidade e Recursos hídricos, escassez e abastecimento de água em Juazeiro. A terceira seção traz a metodologia. A seção seguinte os resultados e por fim as conclusões.

## 2 SUSTENTABILIDADE E RECURSOS HÍDRICOS

A preocupação com o uso sustentável dos recursos hídricos ganhou peso em todo o mundo desde a Conferência Internacional de Água e Meio Ambiente<sup>2</sup>, acontecida em Dublin (Irlanda) nos dias 26 a 31 de janeiro de 1992, situação em que ficou evidente que a água é um recurso finito e vulnerável, essencial para manutenção da vida, do desenvolvimento e do meio ambiente. E, partindo do princípio que água sustenta a vida, ficou claro que a gestão dos recursos hídricos requer uma abordagem holística que integre o desenvolvimento econômico e social com a proteção dos ecossistemas naturais, e que sua gestão efetiva integra o uso do solo com os usos da água no âmbito da bacia de drenagem ou do aquífero subterrâneo (SANTILLI, 2001).

Segundo Santilli (2001), os recursos hídricos são indispensáveis para a manutenção da existência das espécies vivas e, portanto, merecem o devido cuidado e preservação. A necessidade de proteção das águas contra as diversas formas de poluição, contra o uso inadequado e o desperdício traduziu-se em normas legais que pretendem planejar, regular e controlar sua utilização, de acordo com padrões e critérios definidos por meio de uma Política Nacional de Recursos Hídricos (PNRH) e implementados por um Sistema Nacional de Gerenciamento de Recursos Hídricos (SNGRH). A instituição de uma Política Nacional de Recursos Hídricos e a criação de um Sistema Nacional de Gerenciamento de Recursos Hídricos, pela primeira vez na história do Brasil, é o que

---

2. A Conferência Internacional de Água e Meio Ambiente (ICWE), acontecida em Dublin, Irlanda, nos dias 26 a 31 de janeiro de 1992, contou com quinhentos participantes, incluindo especialistas designados pelo governo, de cem países e representantes de oitenta organismos internacionais, intergovernamentais e não governamentais (SANTILLI, 2001).

objetivou a Lei n. 9.433/97 (Lei das Águas). A qual afirma que a cobrança pelo uso dos recursos hídricos objetiva:

- reconhecer a água como um bem econômico e dar ao usuário uma indicação de seu valor real;
- incentivar a racionalização do uso da água;
- obter recursos financeiros para o financiamento dos programas e intervenções contemplados nos planos de recursos hídricos.

A Lei das Águas permite um preço para a água que é hoje captada de graça dos corpos d'água. Atualmente, o usuário paga apenas o custo do tratamento e da distribuição. Embora com um atraso de 30 anos em relação aos países mais avançados, como a França, a Lei das Águas brasileira incorpora os instrumentos mais adequados de gestão, tendo se beneficiado das experiências internacionais na administração dos recursos hídricos. Considera-se que a legislação existente é mais do que suficiente para o desenvolvimento da gestão. Porém, é de se esperar que leve um tempo para ser implementada (PEIXOTO FILHO e BONDAROVSKY, 2000).

Segundo a referida Lei, esse instrumento tem como objetivos assegurar o controle quantitativo e qualitativo da água e o efetivo exercício dos direitos de acesso a este recurso, disciplinando a sua utilização e compatibilizando demanda e disponibilidade hídrica (SILVA e MONTEIRO, 2010).

Em 17 de julho de 2000, a Lei n. 9.984 criou a Agência Nacional de Águas (ANA), entidade federal responsável pela implementação da Política Nacional de Recursos Hídricos, integrante do Sistema Nacional de Gerenciamento de Recursos Hídricos, estabelecendo regras para sua atuação, sua estrutura administrativa e suas fontes de recursos. Uma vez que está estabelecido na Constituição da República Federativa do Brasil, de 1988, (Art. 20, III e 26, I) que as águas de lagos, rios e as águas subterrâneas constituem bens ou da União ou dos Estados, e mesmo se admitindo a existência de águas particulares, estas teriam a sua utilização sempre condicionada e limitada pelo interesse público, como todos os bens ambientais (SANTILLI, 2001).

Os diversos usos da água – abastecimento humano, saciedade animal, irrigação, indústria, geração de energia elétrica, aquicultura, preservação ambiental, paisagismo, lazer, navegação, etc. – podem ser concorrentes, gerando conflitos entre setores usuários e impactos ambientais. Nesse sentido, gerir recursos hídricos é uma necessidade premente e que tem o objetivo de ajustar as



demandas econômicas, sociais e ambientais por água em níveis sustentáveis, de modo a permitir, sem conflitos, a convivência dos usos atuais e futuros da água (SILVA e MONTEIRO, 2010).

## 2.1 A Escassez de Água

Segundo MMA, MEC e IDEC (2005), a ameaça da falta de água, em níveis que podem até mesmo inviabilizar a nossa existência, pode parecer exagero, mas não é. Os efeitos na qualidade e na quantidade da água disponível, relacionados com o rápido crescimento da população mundial e com a concentração dessa população em megalópoles, já são evidentes em várias partes do mundo. Dados do Fundo das Nações Unidas para a Infância (UNICEF) e da Organização Mundial da Saúde (OMS) revelam que quase metade da população mundial (2,6 bilhões de pessoas) não conta com serviço de saneamento básico e que uma em cada seis pessoas (cerca de 1,1 bilhão de pessoas) ainda não possui sistema de abastecimento de água adequado. As projeções da Organização das Nações Unidas indicam que, se a tendência continuar, em 2050 mais de 45% da população mundial estará vivendo em países que não poderão garantir a cota diária mínima de 50 litros de água por pessoa. Com base nestes dados, em 2000, os 189 países membros da ONU assumiram como uma das metas de desenvolvimento do milênio reduzir à metade a quantidade de pessoas que não têm acesso à água potável e saneamento básico até 2015.

A escassez é uma realidade efetiva em algumas grandes cidades e metrópoles dos países em desenvolvimento, que padecem de forma crônica ou sazonal de déficits significativos no abastecimento de água potável. É o caso, por exemplo, da Grande São Paulo, como também das regiões metropolitanas como a de Fortaleza, de Recife e de Vitória. Algumas cidades de grande e médio porte do interior, também, vêm enfrentando há alguns anos problemas de descontinuidade no abastecimento de água. Trata-se de um problema que tem sido evitado nas metrópoles dos países desenvolvidos mediante planejamento adequado, interconexão de sistemas e mecanismos de financiamento estáveis para o setor (VARGAS, 1999).

[...] “a questão da escassez dos recursos hídricos não pode ser vista apenas em termos geofísicos e quantitativos, sendo uma questão eminentemente social relacionada também a padrões de desenvolvimento econômico, de demanda e de qualidade das águas” (VARGAS, 1999, p. 3).

Mesmo países que dispõem de recursos hídricos abundantes, como o Brasil, não estão livres da ameaça de uma crise. A disponibilidade varia muito de uma região para outra. Além disto, nossas reservas de água potável estão diminuindo. Entre as principais causas da diminuição da água potável estão o crescente aumento do consumo, o desperdício e a poluição das águas superficiais e subterrâneas por esgotos domésticos e resíduos tóxicos provenientes da indústria e da agricultura (MMA; MEC e IDEC, 2005).

“A água não é uma doação gratuita da natureza; ela tem um valor econômico: precisa-se saber que ela é, algumas vezes, rara e dispendiosa e que pode muito bem escassear em qualquer região do mundo” (MMA; MEC e IDEC, 2005, p. 40).

## 2.2 O Desperdício de Água

Quando se fala em desperdício sempre se evidencia os maus hábitos da população, mas ele também ocorre no sistema de abastecimento e na agricultura. O consumo humano de água em coisas básicas como saciar a sede, banhar-se, lavar a roupa e cozinhar é relativo e pode variar muito, dependendo da disponibilidade de acesso ao abastecimento e de aspectos culturais da população. Alguns estudos mostram que, por dia, uma pessoa no Brasil gasta de 50 litros a 200 litros de água. Então, com 200 litros/dia utilizando-se de forma racional vive-se confortavelmente. A quantidade de água economizada dentro de casa vai depender tanto do número de pessoas consumidoras, quanto das instalações hidráulicas de funcionamento e, principalmente de como ela é usada. Para se ter uma ideia dos maus hábitos com relação ao consumo, aproximadamente 75% da água consumida em casa é gasta no banheiro (CARVALHO JÚNIOR, 2014).

A população brasileira ainda mantém hábitos absurdos em relação ao desperdício de água. Quilômetros e quilômetros de calçadas seguem sendo lavados diariamente, até que a última folhinha caída da árvore seja empurrada por jatos de mangueira ao meio-fio. Nem sempre o consumidor tem o cuidado de impedir vazamento em aparelhos sanitários ou torneiras. Inúmeros postos de combustíveis oferecem duchas gratuitas como brinde para quem abastecer por lá, com água extraída dos lençóis freáticos para os quais não pagam tarifa de esgoto (ESTADÃO, 2014).

O Programa de Uso Racional de Água (PURA), desenvolvido na Grande

São Paulo pela SABESP em parceria com a USP, mostra que os desperdícios da água utilizada atingem níveis jamais imaginados. Os exemplos mais frequentes de desperdícios são: tomar banhos muito prolongados; lavar calçadas, pátios e carros com o jato da mangueira; usar bacias sanitárias que necessitam de 18 a 20 litros de água por descarga – quando, no mercado, já existem modelos modernos que necessitam de apenas 6 litros; utilização de equipamentos sanitários obsoletos – como, por exemplo, torneira de rosca e mictórios do tipo gamela – como descarga de água permanentemente aberta; utilização de água tratada em atividades que não exigem água potável, como para irrigar gramados esportivos públicos ou privados ou utilização de água potável em atividades industriais, tais como em torres de resfriamento (REBOUÇAS, 2003).

“As maiores vilãs do desperdício nas instalações prediais são as válvulas convencionais de descarga. Cada segundo que uma pessoa permanece com o dedo na descarga são dois litros de água desperdiçados. As válvulas de descarga, particularmente as mais antigas, são responsáveis por 40% de toda a água consumida dentro de uma residência. Entretanto, esse consumo é pouco se comparado com os 1.650 litros de água para produzir 1 kg de soja, 1.900 para 1 kg de arroz, 3.500 para 1 kg de aves e 15 mil para 1 kg de carne bovina. O mesmo ocorre com produtos industrializados. São 10 litros de água para 1 de gasolina, 95 para 1 kg de aço, 324 para 1 kg de papel e 600 litros para 1 kg de cana-de-açúcar voltada para a produção do etanol. Na agricultura, o desperdício de água também é muito grande: apenas 40% da água desviada é efetivamente utilizada na irrigação. O desperdício gira em torno de 60%, porque se aplica água em excesso, se aplica fora do período de necessidade da planta, em horários de maior evaporação do dia, pelo uso de técnicas de irrigação inadequadas ou, ainda, pela falta de manutenção desses sistemas de irrigação” (CARVALHO JÚNIOR, 2014, p. 1).

Também se perde muita água no sistema de abastecimento. Essa perda pode ser classificada em dois tipos: a física e a não física. A perda física é a água perdida em vazamentos, aquela que não chega ao consumidor. A perda não física é a água usada pelos consumidores, mas que não é medida pela empresa de abastecimento de água, como as ligações clandestinas e outros tipos de fraudes. A perda total nos sistemas públicos de abastecimento do Brasil equivale, em média, 40 a 45% da água produzida. Para evitar o desperdício, o distribuidor

deve procurar a otimização do rendimento da sua rede de distribuição, principalmente na detecção de vazamentos com instrumentos adequados, bem como definir a política de medição da água, concretamente, no que se refere a algumas das condições que determinam o nível da qualidade da medição (CARVALHO JÚNIOR, 2014).

Segundo MMA, MEC e IDEC (2005), o Brasil registra um elevado desperdício: de 20% a 60% da água tratada para consumo se perde na distribuição, dependendo das condições de conservação das redes de abastecimento. Além dessas perdas de água no caminho entre as estações de tratamento e o consumidor, o desperdício também é grande nas residências, envolvendo, por exemplo, o tempo necessário para se tomar um banho, a própria forma de se tomar banho, a utilização de descargas no vaso sanitário que consomem muita água, a lavagem da louça com água corrente, no uso da mangueira como vassoura na limpeza de calçadas, na lavagem de carros etc.

Para tentar reverter esse quadro, a conservação da água, riqueza natural mais preciosa do futuro, é necessário o desenvolvimento e o emprego de técnicas e instrumentos cada vez mais sensíveis e precisos para fazerem a medição do seu consumo e do seu uso (industrial, comercial, residencial e agrícola). Porém, a conservação da água depende, sobretudo, de ações educativas junto à comunidade, que deve ser esclarecida com relação à necessidade de redução dos desperdícios. Assim como também depende de uma série de leis e regulamentos que as autoridades devem implantar. Com essas medidas, é possível evitar os desperdícios de água, propiciando a sua utilização dentro do estritamente necessário para a sobrevivência humana e o desenvolvimento sustentável (CARVALHO JÚNIOR, 2014).

### **2.3 O Abastecimento de Água em Juazeiro do Norte**

A cidade de Juazeiro do Norte tem sua rede de abastecimento de água, que se sustenta em águas subterrâneas, proveniente de poços tubulares construídos pela Companhia de Água e Esgoto do Ceará e pela prefeitura local, ou até mesmo de cunho particular. A distribuição é de responsabilidade da CAGECE, uma vez que os poços construídos pela prefeitura possuem caráter de distribuição restrita, isto é, são utilizados apenas em chafarizes e em lavanderias públicas não chegando a ser utilizados para o consumo doméstico. Por sua vez, os poços de cunho particular são, na maioria, pertencentes a empresas e utilizados apenas para uso próprio. A CAGECE utiliza uma bateria de poços tubulares de diferentes profundidades e vazões para a captação de água subterrânea no

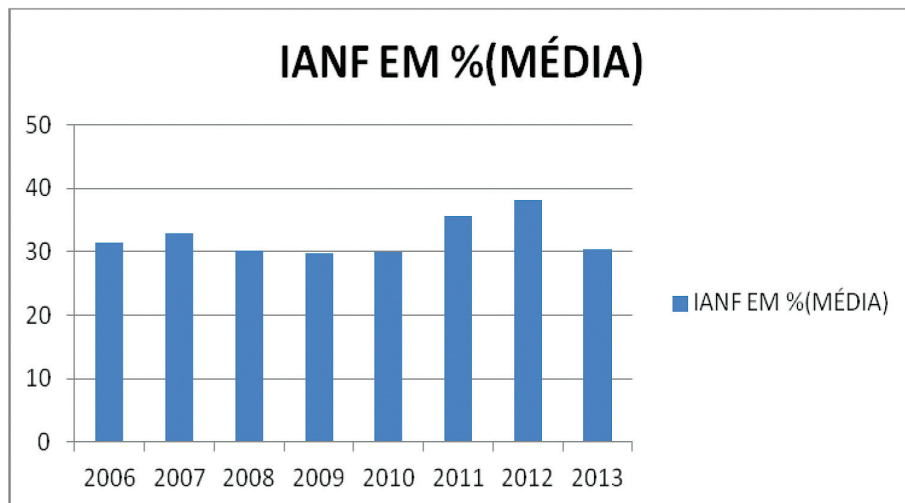
município (SILVA *et al*, 2012).

De acordo com os dados do Sistema de Informação da Atenção Básica (SIAB) do Ministério da Saúde, 92,9% da população do município de Juazeiro do Norte usufruem do benefício da rede de abastecimento de água da CAGECE em seus domicílios. Outros 5,9% usam água advinda de poços particulares e 1,1% contam com outros meios de abastecimento de água. Segundo os dados do SIAB, o número percentual de famílias com rede de abastecimento de água vem evoluindo nos últimos anos e os números municipais estão sempre acima dos nacionais (DEEPASK, 2014).

Segundo Lopes *et al* (2014), em agosto de 2000 todos os poços da cidade forneciam um volume mensal de 1.622.994 m<sup>3</sup> de água dos quais apenas 732.909 m<sup>3</sup> eram registrados pelos mediadores de águas instalados o que significava uma perda de 54,84% do volume distribuído. Perda essa, que poderia significar problemas na tubulação, vazamentos ou até mesmo ligações clandestinas feitas por usuários diversos, os famosos “jacarés”.

Segundo a CAGECE (2014), esses números vêm se modificando ao longo do tempo. Entre 2006 e 2013, período em estudo, o Índice de Água Não Faturada (IANF), em percentual médio anual, em Juazeiro do Norte oscilou entre 29,7 e 38,19%, como pode ser visto na Figura 4.

FIGURA 4: Índice de Água Não Faturada (IANF) em Juazeiro do Norte



Fonte: elaborado pela autora com base nos dados fornecidos pela CAGECE.

### 3 MATERIAL E MÉTODOS<sup>3</sup>

#### 3.1 Juazeiro do Norte: características e dados

Situada a 533 quilômetros da capital do estado, com uma população estimada de 263.704 habitantes, uma área de 248,832 km<sup>2</sup> e densidade demográfica de 1004,45 hab/km<sup>2</sup>, Juazeiro do Norte é a terceira maior cidade do Ceará e está localizada no extremo sul do estado, na Região Metropolitana do Cariri. Juazeiro do Norte é a cidade mais importante da região e tem um forte comércio que atende, também, cidades circunvizinhas. Possui ainda um grande número de indústrias, das quais cabe destaque para o setor calçadista – terceiro no *ranking* nacional – bem como para o de folheados, têxtil, alumínio etc. A cidade hoje atrai muitos investimentos de outras regiões, nacionais e internacionais, devido a seu forte crescimento nos últimos anos (IBGE, 2014).

FIGURA 3: Localização Geográfica de Juazeiro do Norte (CE)



Fonte: Guia JUANORTE (2014).

De acordo com o IPECE (2015), o município possui um clima tropical quente semiárido e tropical quente semiárido brando, tem uma pluviosidade

3. A fundamentação teórica do trabalho tem como base a teoria do Bem estar social. Segundo Pindyck e Rubinfeld (2013), a atuação dos consumidores, bem como dos produtores, pode gerar custos ou benefícios que não serão refletidos nos preços de mercado. São as chamadas externalidades. Quando existem externalidades, o preço de um bem não reflete necessariamente seu valor social. Pois, as mesmas não são refletidas nos preços de mercado, podendo, portanto, causar a ineficiência econômica. As externalidades podem surgir entre produtores e entre consumidores, ou ainda entre produtores e consumidores. Elas podem ser positivas, quando a ação de uma parte gera benefício para a outra, ou negativa, quando a ação de uma parte gera prejuízo para a outra.

média anual de aproximadamente 926 mm e seu período chuvoso está distribuído entre os meses de janeiro a maio.

A renda per capita média do município cresceu 98,98% nas últimas duas décadas, passando de R\$ 220,89 em 1991 para R\$ 439,53 em 2010. A taxa média anual de crescimento também se elevou nesse mesmo período passando de 3,14%, entre 1991 e 2000 para 4,18%, entre 2000 e 2010. A extrema pobreza passou por uma redução bastante considerável, de 28,54% em 1991, para 19,92% em 2000 e para 6,73% em 2010. Considerando o período que vai de 1990 a 2010, a desigualdade diminuiu: o Índice de Gini<sup>4</sup> passou de 0,59 em 1991 para 0,60 em 2000 e para 0,54 em 2010 (ATLAS, 2014).

O Índice de Desenvolvimento Humano Municipal (IDHM) de Juazeiro do Norte é 0,694, em 2010. O município está situado na faixa de Desenvolvimento Humano Médio (IDHM entre 0,6 e 0,699). Juazeiro do Norte teve um incremento no seu IDHM de 65,63% nas últimas duas décadas, acima da média de crescimento nacional (47%) e abaixo da média de crescimento estadual (68%). O hiato de desenvolvimento humano, ou seja, a distância entre o IDHM do município e o limite máximo do índice, que é 1, foi reduzido em 47,33% entre 1991 e 2010. Entre 1991 e 2000 e entre 2000 e 2010, a dimensão que mais cresceu em termos absolutos foi Educação (com crescimento de 0,165 e 0,252 respectivamente), seguida por Longevidade e por Renda (PNUD, 2013).

### 3.2 Fonte dos Dados

Essa pesquisa utilizou dados secundários fornecidos pela Companhia de Água e Esgoto do Ceará – CAGECE, unidade de Juazeiro do Norte. O estudo tem como base o volume de água distribuído, volume de água faturado e a estimativa de seu eventual desperdício na cidade de Juazeiro do Norte, no período compreendido entre janeiro de 2006 e dezembro de 2013, e os preços por faixa de consumo para 2014.

### 3.3 Os Modelos ARIMA e a Abordagem de Box-Jenkins

Os modelos ARIMA (Auto Regressive Integrated Moving Averages), que em português significa Modelos Auto-Regressivos Integrados de Médias

4. O Índice de Gini é um instrumento usado para medir o grau de concentração de renda. Ele aponta a diferença entre os rendimentos dos mais pobres e dos mais ricos. Numericamente, varia de 0 a 1, sendo que 0 representa a situação de total igualdade, ou seja, todos têm a mesma renda, e o valor 1 significa completa desigualdade de renda, ou seja, se uma só pessoa detém toda a renda do lugar (ATLAS, 2014).

Móveis, sistematizados por Box e Jenkins, são resultantes da combinação de três filtros denominados: componente auto-regressivo (AR), filtro de integração (I) e componente de médias móveis (MA). São modelos robustos do ponto de vista conceitual e estatístico, proporcionam previsões probabilísticas e são de fácil implementação, desde que se tenha os recursos computacionais adequados. Os mesmos representam uma generalização dos diversos métodos de análise de séries temporais (GOMES, 1989).

Segundo Gomes (1989), fórmula geral de um modelo ARIMA (p.d.q.) pode ser representada da seguinte forma:

$$W_t = \mu + \frac{\Theta(B)}{\Phi(B)} a_t \quad (1)$$

Onde:

t – índice do tempo;

$W_t$  – d'ésima diferença da variável de interesse  $Z_t$ ;

$\mu$  – “ponto de referência” do nível do processo;

$\Theta(B)$  – operador de “médias-móveis”:  $\Theta(B) = (1 - \Theta_1 B^1 - \Theta_2 B^2 - \dots - \Theta_q B^q)$ ;

$\Phi(B)$  – operador auto-regressivo:  $\Phi(B) = (1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p)$ ;

$B^p$  – operador de retrocesso:  $B^p Z_t = Z_{t-p}$ ;

$a_t$  – “ruído branco” ou erro aleatório.

Para Gomes (1989), de maneira mais clara, o modelo ARIMA (p.d.q.), ou modelo auto-regressivo-médias-móveis integrado de ordem (p.d.q.), acima estabelecido de forma sintética, pode ser expandido da seguinte maneira:

$$W_t = \theta_0 + \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2)$$

Onde:  $\theta_0 = \mu (1 - \phi_1 - \dots - \phi_p)$

Segundo Gomes (1989), de uma forma geral, os modelos ARIMA postulam que as séries temporais ( $z_t$ ) podem ser representadas por uma sequência de “choques” aleatórios ( $a_t$ ) submetidos a três operações de “filtragem” (médias-móveis, auto-regressiva e integração). Intuitivamente, pode-se afirmar que os modelos ARIMA representam as séries temporais como uma ponderação dos próprios valores e/ou erros passados da série. Um modelo ARIMA (p.d.q.) possui  $p+q+2$  parâmetros desconhecidos, que podem ser estimados a partir dos dados, a saber:



- a)  $\mu$ , i.e. o “ponto de referência” do nível do processo;
- b)  $p$ , parâmetros auto-regressivos  $\phi_1, \phi_2, \dots, \phi_p$ ;
- c)  $q$ , parâmetros médias-móveis  $\theta_1, \theta_2, \dots, \theta_q$ ;
- d)  $\sigma^2$ , a variância do ruído branco  $a_t$ .

O propósito da modelagem ARIMA é estabelecer uma relação entre o valor presente de uma série temporal e seus valores passados de modo que as previsões possam ser feitas somente com base nos valores que constam no histórico da própria variável (GOMES, 1989).

A modelagem Box-Jenkins (B-J) baseado em séries temporais, é indicada para fazer previsões futuras. Esta metodologia é considerada a mais popular das técnicas quantitativas de previsão, realizada a partir de uma série temporal univariada<sup>5</sup>. Ela é baseada apenas no passado da própria variável para fins de previsões, isto é, não utiliza a história de quaisquer outras variáveis. Ou seja, os dados falam por si mesmo, espera-se que a série apresente uma correlação seriada no tempo (MARTINS E JUSTO, 2014).

Segundo Martins e Justo (2014), uma série temporal é um conjunto de observações ordenadas no tempo. Sua representação é  $\{y_t \mid t \in T\}$ , em que  $y$  é a variável<sup>6</sup> da série e  $t$  (tempo) é o índice<sup>7</sup> pertencente ao conjunto  $T$ . A composição de uma série temporal é feita por quatro elementos: a tendência cujo objetivo é verificar o sentido de deslocamento da série e mostra redução, aumento ou estabilização dos dados; o ciclo que descreve o movimento ondulatório de tendência periódica; a sazonalidade que mostra o movimento ondulatório de curta duração, e segue um padrão temporal e está relacionada às influências climáticas, datas festivas etc.; e o erro que compreende a variabilidade intrínseca aos dados, que resulta de fatos excepcionais, como tragédias, mudanças conjunturais no país e crises mundiais, os quais não podem ser previstos, portanto esse elemento não pode ser modelado.

Segundo Martins e Justo (2014), a modelagem Box-Jenkins consiste em ajustar modelos, a uma dada série temporal. A proposta é que seja selecionado,

5. Segundo Lopes (2002), séries temporais univariada são aquelas que desprezam a possibilidade de existirem relações entre a variável de interesse e outras variáveis econométricas.

6. A variável pode ser discreta ou contínua, univariada ou multivariada (LOPES, 2002).

7. O índice pode ser discreto ou contínuo, unidimensional ou multidimensional, mas no geral é discreto, unidimensional e observado em intervalos iguais (GOMES, 1989).

em três estágios, o modelo mais apropriado para realizar a previsão do conjunto de dados disponíveis. A sequência desses estágios é a seguinte:

- primeira fase – transformação dos dados e identificação do modelo;
- segunda fase – estimação dos parâmetros e verificação se o modelo é ou não adequado;
- terceira fase – previsões (caso o modelo seja adequado)

O processo de escolha pode se repetir várias vezes. Uma forma de simplificar a previsão é identificar alguns modelos para estimação e verificação, encontrando-se um entre esses procede-se com a previsão (MARTINS e JUSTO, 2014).

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Identificação do Modelo

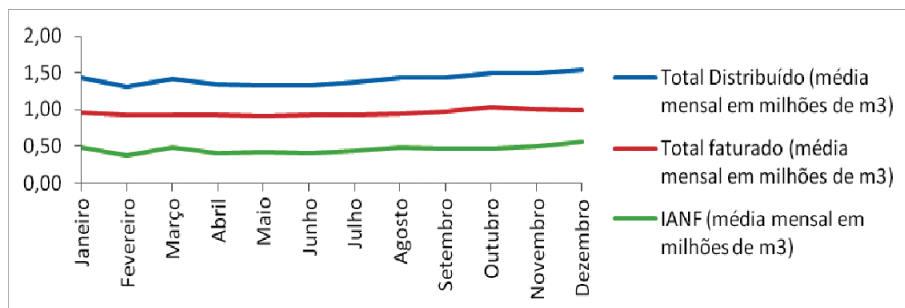
Na análise de modelos com séries temporais<sup>8</sup> é importante um exame visual das variáveis para identificar possíveis padrões de comportamento que podem indicar a necessidade de modelar a presença de determinados padrões como tendência, ciclos e sazonalidade e se há indícios que as séries sejam estacionárias. Embora o foco seja a perda de bem estar social em função do desperdício de água do sistema de distribuição no Município de Juazeiro do Norte, serão apresentadas inicialmente as séries do volume total distribuído, faturado e o Índice de Água Não Faturado (IANF). Desta forma, tem-se uma ideia mais apurada da representatividade da perda de água no sistema bem como correlações entre estas séries.

A Figura 5 mostra a média mensal do volume de água distribuído, faturado e o IANF em Juazeiro do Norte, compreendendo o período de janeiro de 2006 a dezembro de 2013. Observa-se que não há indícios de sazonalidade nestas séries.

8. Segundo Morretin e Tolo (2006) uma série temporal é um conjunto de observações ordenadas no tempo. Como os dados são coletados sequencialmente ao longo do tempo, espera-se que ela apresente correlação seriada no tempo.

Para Santana, Russo e Figueroa (2012) a análise de séries temporais objetiva investigar o mecanismo gerador da série temporal; fazer previsões de valores futuros da série; descrever o comportamento da série e procurar periodicidades relevantes nos dados.

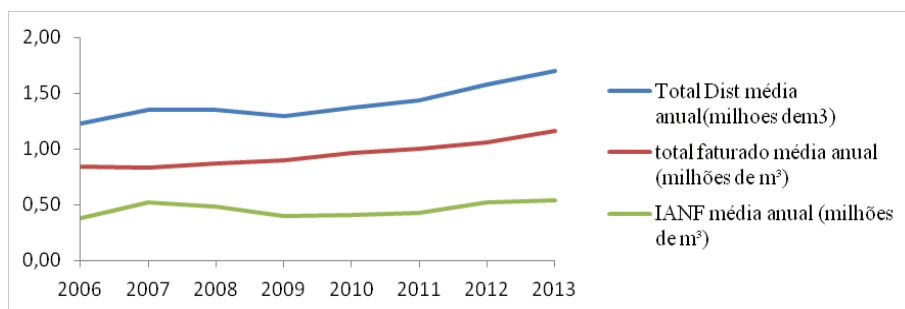
**FIGURA 5: Média mensal do volume de água distribuído, faturado e IANF em milhões de m<sup>3</sup>**



Fonte: Elaborado pelos autores com base nos dados da CAGECE.

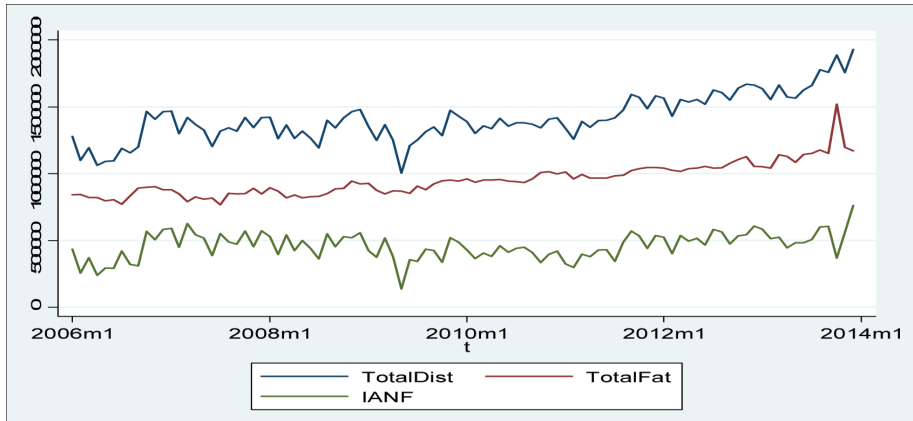
Na Figura 6 tem-se a média anual do volume de água distribuído, faturado e o IANF em Juazeiro do Norte de 2006 a 2013. Observa-se uma ligeira tendência no crescimento destas variáveis e ausência de indicações de presenças de ciclos. O crescimento é esperado em virtude da política de expansão da rede de distribuição de água e do crescimento da população ao longo do período.

**FIGURA 6: Média anual do volume de água distribuído, faturado e ianf em milhões de m<sup>3</sup>: 2006 a 2013**



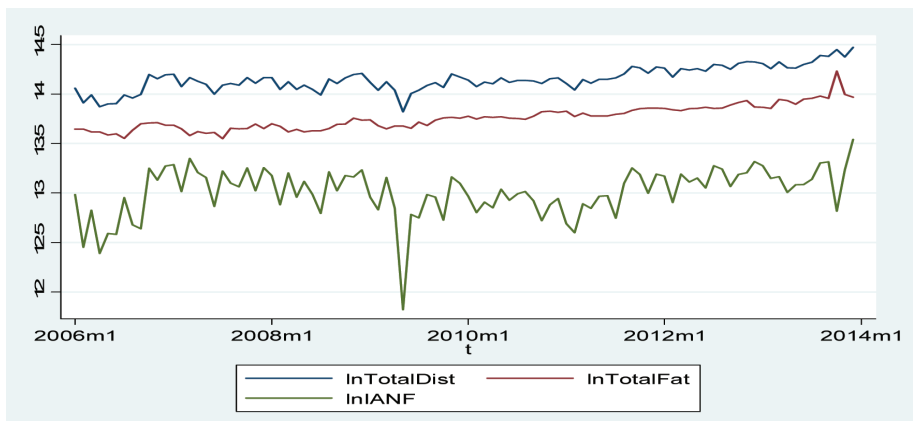
Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

A Figura 7 apresenta as séries temporais do volume de água distribuído, faturado pela CAGECE e o Índice IANF em Juazeiro do Norte no período de janeiro de 2006 a dezembro de 2013 em milhões m<sup>3</sup>. A análise visual sugere uma maior correlação entre o IANF e o volume de água distribuído. Há indícios de que as séries de volume de água distribuído e do IANF não sejam estacionárias. Contudo, serão realizados testes formais de identificação da estacionaridade.

FIGURA 7: **Volume de água distribuído, faturado e IANF em milhões de m<sup>3</sup>**

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

A utilização das séries em logaritmos pode, em muitas situações, tornar as séries estacionárias. Desta forma, a Figura 8 traz as séries do volume de água distribuído, faturado e IANF em logaritmo. Ressalta-se que para os testes de estacionaridade e para a seleção dos modelos de previsão será utilizada a série do IANF em logaritmo.

FIGURA 8: **Logaritmo do volume de água distribuído, faturado e o IANF em Juazeiro do Norte: Janeiro de 2006 a Dezembro de 2013**

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

Na análise de modelos de previsão é importante a realização de testes que permitam identificar se a série é estacionária. Contudo, há necessidade da

realização de testes que apontem o número de defasagens que devem ser incluídas no modelo. Segundo Baum (2013) se forem introduzidas muitas defasagens perdem-se muitos graus de liberdade. Por outro lado, se forem incluídas poucas defasagens pode ocasionar erro de especificação do modelo e autocorrelação dos resíduos. Desta forma, faz-se necessária a utilização de testes para identificar a quantidade ótima de defasagens no modelo. De acordo com o Quadro 1, a maioria dos testes aponta para a necessidade de incorporar três defasagens. Desta forma o modelo estimado incorpora três defasagens da série do logaritmo do IANF.

O Quadro 1 mostra o resultado do teste de identificação do número ótimo de defasagens antes da realização dos testes de estacionaridade. Os resultados apontam para a necessidade de incorporar três defasagens nas séries.

#### QUADRO 1: Identificado o número de defasagens para fazer o teste de raiz unitária

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	3,56519				.055374	-.055765	-.044702	-.028354
1	12,4042	17,678	1	0,000	.046698	-.226178	-.204052	-.171357*
2	13,4637	2,119	1	0,145	.046639	-.227471	-.194282	-.145239
3	16,7617	6,5959*	1	0,010	.044367*	-.277427*	-.233174*	-.167784
4	16,8871	.25088	1	0,616	.045221	-.258415	-.203099	-.121361

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

Depois de identificado o número de defasagens para fazer o teste da raiz unitária e o mesmo apontar para a utilização de 3 lags, realizou-se o teste de Dickey-Fuller Aumentado (ADF) para identificar se a série é estacionária em nível ou em diferença, e percebeu-se que a 5% a série logaritimizada não é estacionária como pode ser visto na tabela 1, tendo em vista que o valor absoluto calculado é menor que o valor crítico a 1%.

#### TABELA 1: Teste de Raiz Unitária Dickey-Fuller Aumentado

DF	Valor Crítico 1%	Valor Crítico 5%	Valor Crítico 10%	Z (t)
- 3,044	- 3,521	- 2,896	- 2,583	0,0310

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

De acordo com a Tabela 2 o teste de Phillips-Perron também apontou que a série do IANF em logaritmo também não é estacionária. Novamente, os absolutos valores calculados são menores que os valores críticos a 1% e a 5%. Daí partiu-se para a diferenciação da série e realização novamente dos respectivos testes de estacionaridade.

**TABELA 2: Teste de Raiz Unitária Phillips-Perron**

Phillips-Perron	Valor Crítico 1%	Valor Crítico 5%	Valor Crítico 10%	Z (t)
- 10.178 - 2.879	- 19.710 - 3.517	- 13.660 - 2.894	- 10.970 - 2.582	0.0900

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

Segundo Granger e Newbold (1977) os modelos Box-Jenkins são excelentes modelos de previsão de curto prazo. Análises com estes modelos mostram que as melhores previsões são obtidas com 5 a 10 anos com dados mensais. Contudo, a importância da série ser estacionária é a possibilidade de fixar parâmetros do modelo válidos para a utilização em previsões futuras a partir do passado.

Foi realizada a diferenciação da série do logaritmo do IANF e realizados os testes de estacionaridade novamente. Conforme a Tabela 3, observa-se que os testes apontam que a série diferenciada é estacionária. Assim, tem-se que houve a remoção da não estacionaridade após a diferenciação da série.

Na Tabela 3 pode-se observar que o teste de Dickey-Fuler e o de Phillips-Perron indicam que a série do IANF logaritmizada em primeira diferença é estacionária. Daí, parte-se para a etapa seguinte do modelo de previsão de Box-Jenkins.

**TABELA 3: Teste de Raiz Unitária Dickey-Fuler e Phillips-Perron para a Série em Logaritmo Diferenciada**

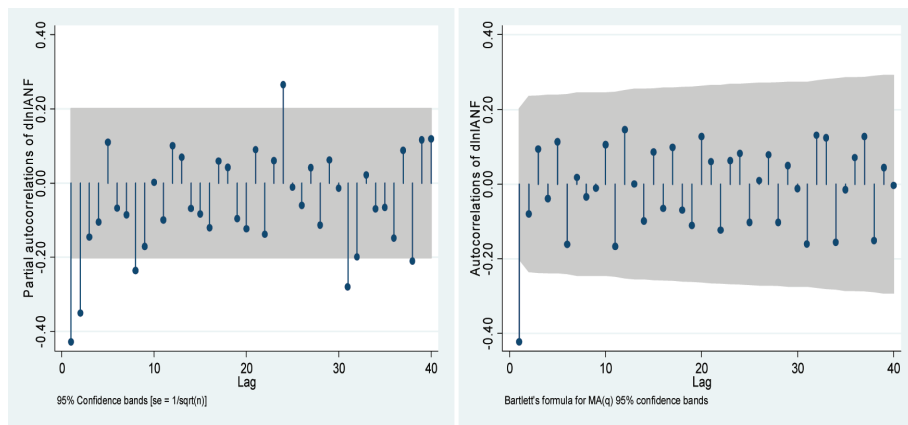
Phillips-Perron	Dickey-Fuler	V. Crítico 1%	V. Crítico 5%	V. Crítico 10%	Z (t)
- 18.802	- 15.466	- 3.518	- 2.895	- 2.582	0.0000

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

Uma vez identificada a estacionaridade da série, partiu-se para a identificação do modelo que melhor se ajusta à série do IANF. Inicialmente buscou-se analisar as funções de autocorrelação e autocorrelação parcial para identificar o

modelo adequado. A análise sugere que há um possível processo AR (1 2) e um MA (1) como pode ser visto na Figura 9. Desta forma, parte-se para a seleção dos modelos.

### FIGURA 9: Funções de Autocorrelação e Autocorreção Parciais da Série de IANF em Juazeiro do Norte



Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

A escolha do modelo deve ser parcimoniosa, ou seja, deve-se  $\min \{p, q\}$ , isto é, deve-se utilizar o menor conjunto de parâmetros possível para o ajustamento do modelo selecionado. Os parâmetros  $p$  e  $q$  representam o número de parâmetros relativos aos comprimentos de defasagem em que se observam valores significativos das autocorrelações e que correspondem a particularidades do sistema de geração das séries que devem ser explicadas pelo modelo (pois correspondem a um padrão de geração). O processo gerador dos dados da série é dito aleatório linear se o modelo ajustado  $Y_t$  pode ser descrito como uma combinação linear de valores defasados de  $Y_t$  e  $\epsilon_t$ .

Após a estimação de diversos modelos foram selecionados os modelos que são apresentados na Tabela 4. De acordo com os critérios de AIC (Akaike Information Criterion) e BIC (Bayesian information criterion) o modelo mais apropriado para a previsão do IANF é o modelo AR(1 2) MA (1) pois apresenta os menores valores do AIC e BIC e os coeficientes são significantes.

TABELA 4: **Características dos Modelos Propostos**

ARMA	AR(1)	AR(1 2)	MA(6)	Ar(1 2) MA(6)	AR(1 2) MA (1)
L.ar	-0.444*** (-6.07)	-0.604*** (-5.39)		-0.169 (-0.83)	-0.638*** (-5.21)
L2.ar		-0.362*** (-4.96)		-0.154 (-1.20)	-0.425*** (-5.29)
L6.ma			-0.148 (-1.14)		-0.261 (-1.70)
L.ma				-0.506** (-2.85)	
Sigma					
_cons	0.235*** (23.05)	0.221*** (21.13)	0.258*** (21.91)	0.217*** (20.38)	0.214*** (20.17)
N	95	95	95	95	95
AIC	0.905	-9.076	18.43	-9.679	-12.47
BIC	8.566	1.140	26.09	3.090	0.297

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

Nota: t statistics in parentheses \* p<0.05, \*\* p<0.01, \*\*\* p<0.001.

Um teste posterior à escolha do modelo para verificar a adequação do modelo selecionado é o teste de Portmanteau. Com os valores gerados pelo modelo, obtêm-se os resíduos. Se o ajuste feito pelo modelo à série em questão for adequado, a série dos resíduos gerados é um ruído branco. Conforme pode ser visto no Quadro 2, de fato, o modelo é adequado. Não se rejeita a hipótese nula de que os resíduos são ruído brancos.

#### QUADRO 2: **Teste de Portmanteau de Ajuste do Modelo selecionado**

Portmanteau test for white noise

$$\begin{aligned} \text{Portmanteau (Q) statistic} &= 44.5907 \\ \text{Prob} > \chi^2(40) &= 0.2848 \end{aligned}$$

Fonte: Elaborado pela autora com base nos dados da CAGECE (2014).

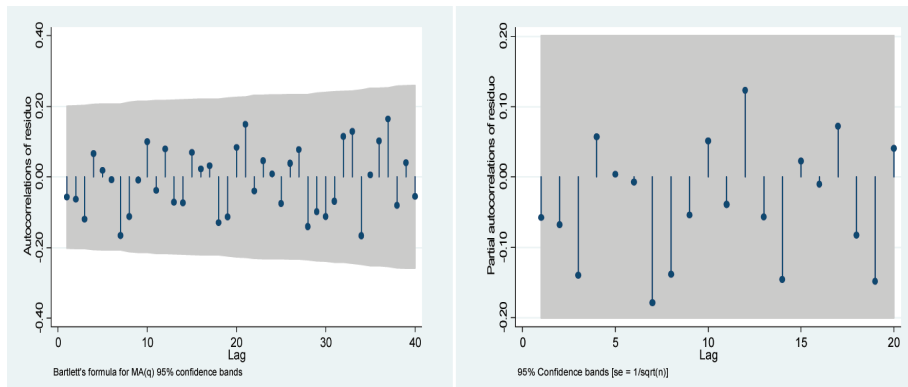
Nota: Ho: H0 = no serial correlation, if we reject it, then the errors are correlated and not white-noise.

Pelo resultado do teste de Portmanteau não se rejeita H0. Os erros são ruídos brancos. Assim, tem-se que o modelo selecionado é adequado.



Outra forma de testar se o modelo ajustado é adequado é analisar novamente a FAC e FACP após a estimação do modelo selecionado. Assim, analisando a Figura 10 observa-se, que, de fato, o modelo atende as exigências da metodologia, haja vista que os resíduos estão dentro do intervalo de confiança.

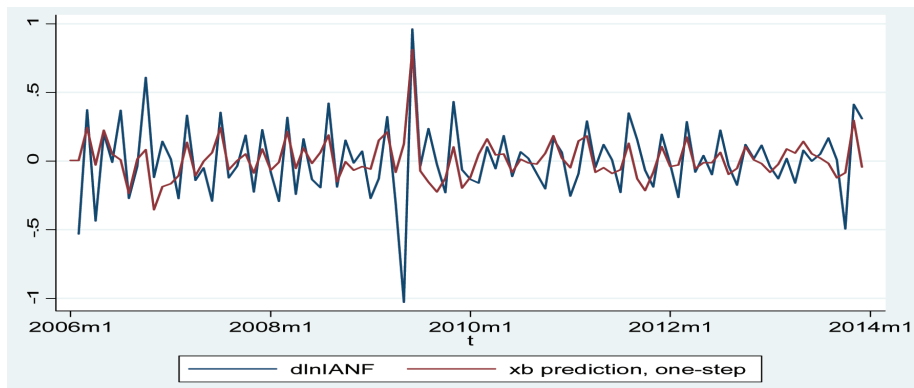
FIGURA 10: FAC e FACP após a estimação do modelo



Fonte: Elaborado pela autora com base nos dados da CAGECE (2014).

Uma vez identificado o modelo, segue-se para a etapa seguinte da metodologia de Box-Jenkins que é a previsão de acordo com o modelo selecionado. Em princípio, foi feita a previsão *ex-ant*. Conforme a Figura 11 percebe-se que o modelo selecionado é adequado para a série, pois, a previsão é bem próxima da série original indicando um bom ajuste.

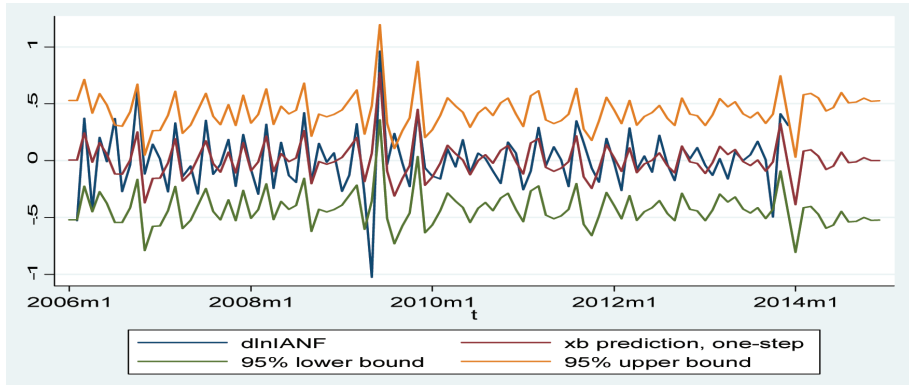
FIGURA 11: **Previsão *ex ant* com o modelo AR (1 2) MA (1) para o IANF em Juazeiro do Norte**



Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

Em seguida, partiu-se para a previsão *ex post*. Conforme se vê na Figura 12 a previsão para período posterior aos dados é bem ajustada corroborando com a previsão *ex-ant* de que o modelo selecionado é bem adequado aos dados.

**FIGURA 12: Previsão *ex post* com os intervalos de confiança com o modelo AR (1 2 ) MA (1) para o IANF em Juazeiro do Norte**



Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

#### 4.2 Estimação da Perda de Bem Estar Social com o Desperdício de Água em Juazeiro do Norte

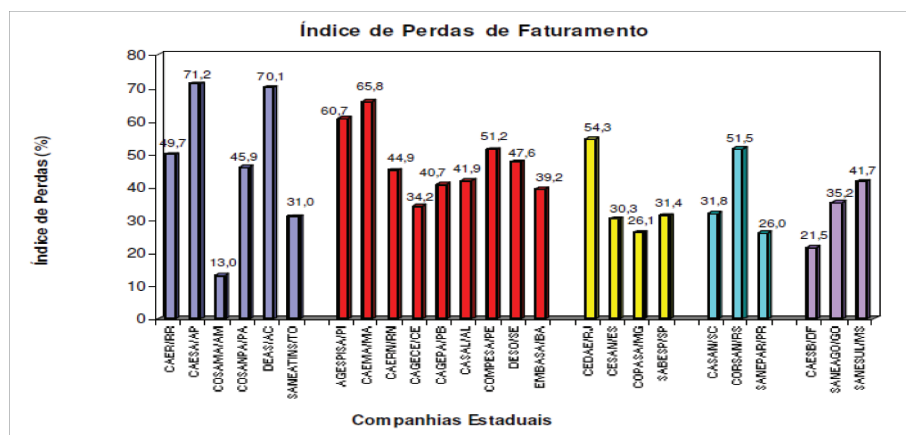
Como já apresentado anteriormente, o bem público e a presença de monopólio pode não permitir a eficiência no sentido de Pareto. Ou seja, o uso além da quantidade ótima por parte dos agentes envolvidos pode gerar ineficiência do ponto de vista social. Por outro lado, o preço cobrado pelo monopolista também pode gerar ineficiência. Daí a necessidade de regulação por parte do governo. O bem aqui em foco é a utilização da água tratada fornecida por uma empresa que atua como monopolista através de concessão pública. No entanto, a análise se dá através da estimação da perda de bem estar social, por intermédio do desperdício de água tratada pela CAGECE em Juazeiro do Norte.

Anteriormente, foi discutido que este desperdício tem várias fontes, tais como, ligações clandestinas, vazamentos pela tubulação, entre outros. Assim, a perda de bem estar será calculada com base no IANF, uma vez que a empresa internaliza os custos do desperdício e repassa para os consumidores. Ou seja, os consumidores que recebem água da referida empresa acabam por bancarem, nos preços cobrados, o custo do desperdício. Desta forma, caso não houvesse desperdício, uma vez que a empresa é regulada por agência reguladora

que tenta minimizar o efeito do monopólio no momento da estimação do preço cobrado, parte significativa da redução dos custos da empresa com o fim dos desperdícios seria repassada para os consumidores na forma de menores preços o que permitiria que elevasse a renda líquida destes elevando o seu bem estar. Aqui será levado a efeito que todo o custo com o desperdício seria repassado para os consumidores através da redução no preço da água fornecida.

De acordo com a seção anterior observou-se que em Juazeiro do Norte o padrão de desperdício está dentro dos padrões para o Nordeste brasileiro e abaixo dos padrões da região Norte, como pode ser visto na figura 13.

**FIGURA 13: Índice de perdas de faturamento para as concessionárias de água e esgoto para o ano de 2000**



Fonte: Elaborado por Neto (2003) a partir de SINS (2000).

Observa-se que para a CAGECE, os padrões de desperdícios são bastante similares ao que vêm ocorrendo em Juazeiro do Norte no período Analisado. Ainda pela Figura 13, tem-se que na região Norte estes desperdícios chegam a ultrapassar 70%.

A estimação da perda de bem-estar foi feita com base nos dados de preços apenas para o ano de 2014 que foram os dados de valores de tarifas cobradas pela CAGECE em Juazeiro do Norte de acordo com o tipo de consumidor e as faixas de consumo fornecidos pela empresa. Uma vez que estes preços fazem parte dos preços administrados, isto é, regulados pelo governo estadual que autoriza os reajustes anuais, parte-se da hipótese que estes preços estejam atualizados. Para previsões futuras da perda de bem estar foram utilizadas as previsões

de acordo com o modelo estimado e considerando estes preços fornecidos pela empresa. Assim, tem-se que os valores futuros estão em Reais de 2014.

Para se chegar a estimativa de valores consumidos de água e os valores desperdiçados por domicílio foram utilizados alguns procedimentos. Por exemplo, para o consumo per capita em litros por dia adotou-se os parâmetros propostos por Querido (2000), que podem ser vistos na Tabela 5.

**TABELA 5: Consumo de água de acordo com a classe Social**

Classe social	Percentual na População (%)	Consumo pcp (l/dia)
A	2,6	218
B	13,6	217
C	20,2	153
D	26,8	133
E	36,8	126

Fonte: Elaborado de acordo com Querido (2000).

Multiplicou-se a população urbana pelo percentual da população com água encanada em Juazeiro do Norte de acordo com dados do PNUD (2013) obtendo-se um total de 217.628 habitantes. Daí distribuiu-se essa população de acordo com as classes sociais conforme a tabela 5. Em seguida multiplicou-se a quantidade de habitantes pelo consumo diário e por 30 dias e obteve-se o consumo mensal em litros. Dividindo-se este valor por mil tem-se o volume de água consumida por classe social em m<sup>3</sup>. Este procedimento teve alta aderência, haja vista que a soma total correspondeu a 98,5% do volume médio mensal faturado pela CAGECE em Juazeiro do Norte. Assim, fez-se o ajuste do volume real distribuído proporcionalmente às classes sociais. Segundo o IPECE (2015) há em Juazeiro do Norte 66.731 domicílios urbanos. Manteve-se o mesmo padrão de distribuição conforme a tabela 5 para chegar as estimativas de consumo mensal de água em m<sup>3</sup> por domicílio. Assim, chegou-se aos valores conforme a tabela 6.

**TABELA 6: Consumo mensal e tarifa média por classe social em Juazeiro do Norte**

Classe social	Volume de água Ajustado (m3)	Número de Domicílios	Consumo Mensal (m3)	Tarifa média/ (m3) em R\$
A	36.472,59	1.735	21,02	32,45
B	189.904,55	9.075	20,93	29,24
C	198.874,70	13.480	14,75	12,69
D	229.362,90	17.884	12,83	7,82
E	298.369,96	24.557	12,15	6,20
Total	952.984,70	66.731		

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

Com esses dados e considerando que a razão IANF/Faturado é de 0,482, tem-se a estimativa de perda de bem estar para 2014 e, utilizando-se a previsão do modelo ajustado na seção anterior, obteve-se a projeção desta perda para o ano de 2015 como pode ser visto na Tabela 7.

**TABELA 7: Estimativa da perda de Bem estar por Classe social em Juazeiro do Norte em virtude do desperdício de Água**

Classe social	so- Valor Faturado (R\$) (A)	IANF/Faturado (B)	Perda de Bem Estar (R\$) (AxB) 2014	Perda de Bem Estar (R\$) 2015
A	1.183.535,46	0,482	570.464,09	576.274,73
B	5.552.809,11	0,482	2.676.453,99	2703.715,83
C	2.523.719,92	0,482	1.216.433,00	1.228.823,35
D	1.793.617,87	0,482	864.523,81	873.329,68
E	1.849.893,76	0,482	891.648,79	900.730,95
Total	12.903.576,11	0,482	6.219.523,69	6.282.874,55

Fonte: Elaborado pelos autores com base nos dados da CAGECE (2014).

Assim, a perda de bem estar social estimada com o desperdício de água em Juazeiro do Norte é de R\$ 6,2 milhões. A Tabela 7 traz a estimativa desta perda de bem estar por classe social.

Como foi dito anteriormente, esta seria a estimativa considerando que todo o valor que seria economizado com os desperdícios fosse repassado para os consumidores na forma de redução do preço das tarifas. Obviamente, que por se tratar de empresa monopolista ainda que regulada por agência pública, parte deste montante seria apropriado pela empresa.

## 5 CONSIDERAÇÕES FINAIS

A ameaça da falta de água em níveis que possam inviabilizar a existência humana não é exagero. Os efeitos do desperdício nas águas potáveis distribuídas já são evidentes em varias partes do mundo. A escassez de água é uma realidade efetiva em muitas cidades e metrópoles de países em desenvolvimento, que padecem de forma crônica ou sazonal de déficits significativos no abastecimento de água potável.

O desperdício de água potável precisa ser evitado o máximo possível para

que se possa afastar o fantasma da escassez. Para isso, as distribuidoras de água precisam buscar a otimização das suas redes de distribuição, principalmente, em se tratando de detectar vazamentos e definir políticas de medição de água. Pois, a água é um recurso natural indispensável à existência da vida, decisivo para o progresso econômico e social e essencial para o bem-estar humano, por isso não se pode colocá-lo em risco.

De acordo com os dados da pesquisa, mesmo sendo uma empresa bem conceituada, eleita destaque nacional por apresentar índices de produtividade elevados, menor tarifa em serviço por metro cúbico faturado, produção mais barata e menor índice de perda da Região Nordeste, receber vários prêmios e ser reconhecida nacionalmente como uma das melhores prestadoras de serviço público do país em inovação e qualidade, como toda empresa monopolista, a CAGECE ainda deixa a desejar no quesito eficiência econômica. E isso atinge diretamente o bem-estar social dos consumidores.

Segundo os dados fornecidos pela própria CAGECE, dos 952.912,5 m<sup>3</sup> de água (média mensal) que são distribuídos na cidade de Juazeiro do Norte, aproximadamente 32,4% são desperdiçados no caminho percorrido entre as estações de tratamento e os hidrômetros dos consumidores finais. As externalidades geradas com o desperdício de água são internalizadas nos preços e repassadas para os consumidores que acabam pagando mais do que consomem.

Utilizando-se a análise de Série Temporal, por meio da modelagem de Box-Jenkins, num processo criterioso, que compreende todas as etapas necessárias desde a transformação dos dados e identificação do modelo, a estimação dos parâmetros e verificação se o modelo é ou não adequado até chegar à previsão, constatou-se que os consumidores de água em Juazeiro do Norte continuarão perdendo bem-estar social em função do desperdício proporcionado pela distribuidora CAGECE.

A julgar pelos resultados obtidos, a perda de bem-estar social da comunidade de Juazeiro do Norte ocasionada pelo desperdício de água é bastante significativa. As classes mais afetadas pelo desperdício de água em Juazeiro do Norte em 2014 são as classes “B” e “C” que perderam respectivamente R\$ 2.676.453,99 e R\$ 1.216.433,00.

O valor total da perda de bem-estar social da população de Juazeiro do Norte em 2014 foi de R\$ 6.219.523,69. Valor esse, que se fosse repassado, ainda que parcialmente, aos consumidores poderia ser convertido em benefícios que poderiam fazer muita diferença, principalmente para a classe mais pobre

daquela população. A estimativa é que essa perda cresça ainda mais em 2015, chegando a atingir quase R\$ 6,3 milhões.

Num mundo em que tanto se tem falado de uso consciente dos recursos naturais, sustentabilidade, e em que a escassez de água já é uma realidade efetiva para tantas cidades e metrópoles, o caso do desperdício de água precisa de maior atenção, não só por parte da CAGECE e das demais distribuidoras de água, mas pelas autoridades competentes e por toda a população. Pois, o que é desperdiçado hoje, além de atingir o bolso do consumidor, poderá vir a faltar amanhã.

Por essa razão, faz-se necessário que os governantes adotem medidas preventivas no combate ao desperdício de água. Uma providência cabível seria investir mais em campanhas de conscientização que ajudasse a população a tomar conhecimento da magnitude que tem o problema de escassez de água potável no mundo. Pois, infelizmente nem todos têm conhecimento do problema e, muitos ainda acham que a escassez de água, em níveis que possam inviabilizar a existência humana, é exagero. Outra medida a ser considerada, seria uma maior fiscalização, também por parte do governo, das distribuidoras de água. A adesão de Taxas de juros diferenciadas nos empréstimos para as empresas com menor índice de desperdício poderiam tornar as distribuidoras mais eficientes.

As distribuidoras de água, por sua vez, poderiam melhorar as redes de distribuição que em média, no Brasil, estão além da sua vida útil. Para tanto, o investimento em novas tecnologias seria indispensável. A detecção de vazamento com hastes de escuta capazes de detectar até os vazamentos mais ocultos, tecnologia desenvolvida na Alemanha, poderia ser de grande utilidade nessa jornada. Essa técnica já se mostrou muito eficaz em muitas cidades e metrópoles dos países desenvolvidos, pois além de detectar os vazamentos ocultos ajuda na identificação de ligações clandestinas.

Por fim, cada indivíduo pode fazer sua parte. Pequenas atitudes como manter a torneira e o chuveiro fechados enquanto escova os dentes e se ensaboa no banho, lavar as calçadas com baldé ao invés da mangueira, utilizar a máquina de lavar louças ou roupas apenas quando estiver com a capacidade máxima etc. pode fazer a diferença. Se essas três instâncias – o governo, as distribuidoras e a população – agissem em conjunto, as chances de êxito seriam maiores e todos sairiam ganhando. Replicar esse trabalho, com dados de outras empresas concessionárias na região e outros municípios do estado, trará uma visão maior dessa problemática em nível regional e estadual e poderá ajudar na resolução do problema.

## REFERÊNCIAS BIBLIOGRÁFICAS

AKAIKE, H. Information Theory and an Extension of the Maximum Likelihood Principle, in Petrov, B. & Csaki, F. (eds). Second International Symposium on Information Theory, 267-281. 1973.

BAUM, C. VAR, SVAR and VECM models. **EC 823: Applied Econometrics**. Boston College, Spring 2013.

BILAS, Richard A. **Teoria Microeconômica: uma análise gráfica**/Richard A. Bilas; [Tradução de Paulo Neuhaus e Hélio Oliveira Portocarrero de Castro: revisão técnica de L. A. Madeira Coimbra]. – 12ª ed. – Rio de Janeiro: Forense universitária, 1991.

DICKEY, P.A.; FULLER, W.A. Distribution of estimators for autoregressive time series with a unit root. **Journal of the American Statistical Association**, Vol. 74, pp. 427-431, 1979.

CAGECE, Companhia de Água e Esgoto do Ceará. Unidade Juazeiro do Norte. 26 de maio de 2014.

CARVALHO JÚNIOR, Roberto de. **Gestão no Campo 2014. Site Eletrônico Permanente. Disponível em:** <<http://www.gestaonocampo.com.br/biblioteca/o-desperdicio-da-agua/>> Acessado em: 12 de setembro de 2014.

DEEPASK. Famílias com abastecimento de água: **Veja número de domicílios com rede pública, poço ou nascente por cidade do Brasil.**<<http://www.deepask.com/goes?page=Confira-o-abastecimento-de-agua-no-seu-municipio--rede-publica-poco-ou-nascente>> Acesso em Setembro de 2014.

2014

ESTADÃO. Desperdício de água. Economia, blog Celso Ming. <http://economia.estadao.com.br/blogs/celso-ming/desperdicio-de-agua/> Acesso em Setembro de 2014.

GOMES, Francisco Carlos. **Os Modelos ARIMA e a Abordagem de Box-Jenkins:**

Uma Aplicação na Previsão do IBOVESPA a Curtíssimo Prazo. Revista de Administração de Empresas: Paulo, Abr./Jun. 1989.

GRANGER, C. W. J.; NEWBOLD, P. **Forecasting economic time series**. Academic Press, New York, 1977.



Guia JUANORTE. **Site Eletrônico Permanente**. Disponível em: <<http://www.juanorte.com.br/guialocalizacao.html>> Acesso em: 15 de junho de 2014.

IBGE CIDADES. **Site Eletrônico Permanente**. Disponível em: <<http://cidades.ibge.gov.br/xtras/perfil.php?lang=&codmun=230730&search=ceara|juazeiro-do-norte|infograficos:-informacoes-completas>>. Acesso em: 11 de outubro de 2014.

IPECE, Instituto de Pesquisa e Estratégia Econômica do Ceará. **Perfil Básico Municipal**. Disponível em: <[http://www.ipece.ce.gov.br/publicacoes/perfil\\_basico/index\\_perfil\\_basico.htm](http://www.ipece.ce.gov.br/publicacoes/perfil_basico/index_perfil_basico.htm)>. Acesso em: 10 de março de 2015.

LOPES, Artur C. B. da Silva. **Modelização Univariada de Séries Temporais: uma Introdução**. Disponível em: <<https://www.repository.utl.pt/bitstream/10400.5/780/1/tapocap4.pdf>> Acesso em: 05 de junho de 2014.

LOPES, Cícero Roberto Martins *et al.* **As Águas Subterrâneas para o Abastecimento Público de Juazeiro do Norte, Região do Cariri, Ceará – Brasil**. Disponível em: <<http://aguassubterraneas.abas.org/asubterraneas/article/view/22855>> Acesso em: 04 de junho de 2014.

MARTINS, G. N.; JUSTO, W. R. **Previsão pelo Serviço de Fornecimento de Água no Semiárido Brasileiro: uma aplicação dos modelos Box-Jenkins**. R. Pol. Públ., São Luís, v. 18, n. 2, p. 631-646, jul./dez. 2014.

Ministério do Meio Ambiente (MMA); Ministério da Educação (MEC); Instituto Brasileiro de Defesa do Consumidor (IDEC). **Manual de Educação para o Consumo Sustentável**. Brasília, 2005.

MORRETIN, P. A. E TOLOI, C. M. **Análise de Séries Temporais**. Ed Edgard Blucher. São Paulo, 2006.

NETTO, M.A.C. **A Previsão Com Metodologia de Box-Jenkins**. IE:UFRJ, 2015.

NETO, M. F. L. **Avaliação de parâmetros intervenientes no consumo *per capita* de água: estudo para 96 municípios do estado de Minas Gerais**. Dissertação de Mestrado. Programa de Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos da Escola de Engenharia da Universidade Federal de Minas, Belo Horizonte: UFMG, 2003, 146p.

ONU (a). **Rio + 20 o Futuro que Queremos**. Disponível em: <http://www.onu.org.br/rio20/temas-agua/>>. Acesso em: 6 de junho de 2014.

ONU (b). **Manual de Educação para o Consumo Sustentável**. Disponível em: <[http://www.mma.gov.br/estruturas/sedr\\_proecotur/\\_publicacao/140\\_publicacao09062009025910.pdf](http://www.mma.gov.br/estruturas/sedr_proecotur/_publicacao/140_publicacao09062009025910.pdf)> Acesso em: 01 de junho de 2014.

PEIXINHO, Frederico Cláudio. **Gestão Sustentável dos Recursos Hídricos**, 2010.

PEIXOTO FILHO, Aser Cortines; BONDAROVSKY, Sandra Helena. Água, Bem Econômico e de Domínio Público. **Revista CEJ**, n. 12, Brasília, set/dez 2000 Brasília, 2000.

PEREIRA, Adriana Camargo; SIVA, Gibson Zucca; CARBONARI, Maria Elisa Ehrhardt. **Sustentabilidade, Responsabilidade Social e Meio Ambiente**. São Paulo: Saraiva, 2012.

PINDYCK, R.; RUBINFELD, D. **Microeconomia**. 8ª Edição. São Paulo: Pearson Education do Brasil, 2013.

PORTAL CAGECE. **Site Eletrônico Permanente**. Disponível em: <<http://www.cagece.com.br/>> Acesso em: 12 de maio de 2014.

PNUD. **Atlas do Desenvolvimento Humano no Brasil 2013**. PNUD, 2013.

QUERIDO, J.G. Caracterização da Cota per capita de Consumo de Água de Abastecimento Público em Função da Classe Social Consumidora. Porto Seguro. In: **Anais do IX Simpósio Luso-Brasileiro de Engenharia Sanitária e Ambiental**, Porto Seguro: ABES, 2000.

REBOUÇAS, Aldo da C. **Água no Brasil: abundância, desperdício e escassez**. Bahia Análises e Dados, 2003. Disponível em: <<file:///F:/abundabras%20-%20desperdicio%20-%20agua.pdf>> Acesso em: 13 de setembro de 2014.

SANTANA, L. V. R; RUSSO, S.L.; FIGUEROA, M.L. Aplicação da Metodologia Box-Jenkins na Previsão da Produção da Cana-de-Açúcar. In: **Anais do Congresso de Matemática Aplicada e Computacional (CMAC)**. Natal, 2012.

SANTILLI, Juliana, A Política Nacional de Recursos Hídricos (Lei 9.433/97) e sua implementação no Distrito Federa, in BENJAMIN, Antônio Herman V. e MILARÉ, Edis (coord)., **Revista de Direito Ambiental**, n. 24, ano 6 – São Paulo: Editora RT, out-dez/2001. SANTOS, Juliana; LOIOLA, Rita. **A Falta de Água em São Paulo Tem Solução?** Disponível em: <<http://veja.abril.com.br/noticia/ciencia/a-falta-de-agua-em-sao-paulo-tem-solucao>>. Acesso em: 7

de junho de 2014.

SCHWARZ, G. Estimating the Dimensions of a Model, **Annals of Statistics**, **6**, 444-464, 1978.

SISTEMA NACIONAL DE INFORMAÇÕES SOBRE SANEAMENTO – SNIS – **Diagnóstico dos Serviços de Água e Esgotos – 1999**. Brasília: Secretaria Especial de

Desenvolvimento Urbano da Presidência da República – SEDU/PR; Instituto de Pesquisa Econômica Aplicada – IPEA, v. 5, 2000.

SILVA, Carla Maria Salgado Vidal *et al.* **A Qualidade das Águas Subterâneas Usadas no Abastecimento do Município do Juazeiro do Norte**. Disponível em: <<http://aguassubterraneas.abas.org/assubterraneas/article/viewFile/23304/15398Ac>> Acesso em: 02 de junho de 2014.

SILVA, L. M. C.; MONTEIRO, R. A. Outorga de direito de uso de recursos hídricos: uma das possíveis abordagens. In: Machado, C. J. S. (Org.). **Gestão de águas doces**. Rio de Janeiro: **Interciência**. Cap. 5, 2004, p.135-178.

TOSTO, S. G.; PAIVA SOBRINHO, R.; PEREIRA, L. C.; MARQUES, J. F.; ROMEIRO, A. R.; ANDRADE, D. C. **Sustentabilidade Ambiental do Uso das Terras no Município de Araras, SP**. São Paulo, 2009.

VARGAS, Marcelo Coutinho. **O Gerenciamento Integrado dos Recursos Hídricos como Problema Socioambiental**. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1414-753X1999000200009](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1414-753X1999000200009)> Acessado em: 15 de setembro de 2014.

VARIAN, Hal R., **Microeconomia: uma abordagem moderna**. 8ª Edição. Rio de Janeiro: Elsevier, 2012.

YAFFEE, R.A. Forecast evaluation with Stata. **United Kingdom Stata Users Group Conference London School of Hygiene and Tropical Medicine**, 2010

YASSUDA, Eduardo Riomey. **Gestão de Recursos Hídricos: fundamentos e aspectos institucionais**. Revista de Administração Pública: Rio de Janeiro, Abr./Jun. 1993.

# ANÁLISE COMPARATIVA DA ESPECIALIZAÇÃO DA AGROPECUÁRIA NORDESTINA E CEARENSE A PARTIR DO ÍNDICE DE CONCENTRAÇÃO NORMALIZADO (ICN)

Maria Josiell Nascimento da Silva\*

Maria Daniele Cruz dos Santos\*\*

Otácio Pereira Gomes\*\*\*

Camila Pereira Brigido Rodrigues\*\*\*\*

## Resumo

O presente artigo visa apresentar as classes de atividade CNAE 2.0, classificadas como especializadas no setor agropecuário através do Índice de Concentração Normalizado – ICN, na macrorregião Nordeste do Brasil e fazer um comparativo com o Estado do Ceará, identificando neste as atividades especializadas em comum com a grande região. A pesquisa é fruto de investigações do grupo de pesquisa para o desenvolvimento da região centro-sul cearense, vinculado ao grupo CNPq Territorialidades Econômicas e Desenvolvimento Regional e Urbano. Pode-se identificar que das oito atividades, apenas dois setores são especializados em comum e pertencentes a mesma divisão: a aquicultura, tanto de água doce quanto salgada. Observa-se um crescimento exponencial da atividade nos últimos anos no Nordeste, alavancado em boa parte pelo Estado do Ceará, o que contribuiu significativamente para despontar seus resultados e a respectiva boa classificação no ranking.

**Palavras-Chave:** Setor agropecuário; ICN; Nordeste; Ceará.

\* Graduada em Ciências Econômicas pela Universidade Regional do Cariri – URCA/Campus Iguatu. Bolsista BAT/URCA integrante do grupo de pesquisa Estudos Regionais para o Desenvolvimento do Centro-sul cearense. E-mail: nascimentojosiell@hotmail.com. Telefone: (88) 99963-9388.

\*\* Professora do curso de Ciências Econômicas da Universidade Regional do Cariri – URCA, campus Iguatu. Especialista em Gestão Financeira e consultoria empresarial pela URCA. Graduada em Ciências Econômicas (URCA). E-mail: danielacruzconomia@gmail.com.

\*\*\* Mestre em Economia Rural pela Universidade Federal do Ceará (UFC). Graduado em Ciências Econômicas (URCA). E-mail: otaciopg@gmail.com.

\*\*\*\* Mestranda em Economia Rural pela Universidade Federal do Ceará (UFC). Graduada em Ciências Econômicas (URCA), Pós-graduada em Gestão Financeira e Consultoria Empresarial pela Universidade Regional do Cariri (URCA). E-mail: pbr.camila@gmail.com.

## Abstract

This article presents the CNAE 2.0 Activity classes, classified as specialized in the agricultural sector through the Normalized Concentration Index - ICN, on the macro-region Northeast of Brazil and make a comparison with the state of Ceará, identifying this specialized activities in common with the vast region. The research is the result of research group of research for the development of south-central region Ceará, linked to the group CNPq Territorialities Economic and Regional and Urban Development. It can identify which of the eight activities, only two sectors are specialized in common and belonging to the same division: aquaculture, both fresh as salt water. There has been an exponential growth of activity in recent years in the Northeast, driven in large part by the State of Ceará, which contributed significantly to blunt its results and its high rating in the rankings.

**Key words:** agricultural sector; ICN; Northeast; Ceará.

## 1 INTRODUÇÃO

Os estudos regionais estão ligados, entre outras coisas, à análise das desigualdades entre regiões, estados e nações. Algumas teorias da ciência regional, tais como a teoria dos polos de crescimento de Perroux (1977), a teoria dos lugares centrais de Christaller (1933) e a teoria das economias de aglomeração de Marshall (1982), abordam especificamente os fatores responsáveis por essas disparidades de desenvolvimento que são relacionadas com a forma como a humanidade organiza seu espaço, levando em conta fatores econômicos, sociais e naturais para que assim possam se reproduzir e desenvolver, gerando uma tendência locacional. Existem motivos específicos que levaram os agentes econômicos a escolherem um em vez de outro espaço.

Assim, o processo de crescimento econômico e suas etapas diferem de um local para outro, obtendo um grau maior ou menor de crescimento a partir da localização geográfica. A aproximação dos centros comerciais ou extrator de matéria prima determina a localização das indústrias. (ALMEIDA; ARAUJO; RODRIGUES, 2009).

Estudar a estrutura produtiva das regiões, visando a identificação de aglomerações particulares locais e, ao mesmo tempo, poder entendê-las, construindo propostas assertivas de redistribuição de atividades e setores, para a melhoria dos resultados individuais regionais, não é uma tarefa fácil. Promover um

planejamento regional com vistas a atender as expectativas locais de crescimento e aproveitar a mão-de-obra endógena (para implementar empreendimentos com os menores custos de produção), são algumas das premissas que estudiosos como Losch, Thünen, Weber e Isard se utilizaram para fundamentar suas teorias sobre localização das atividades socioeconômicas (FERREIRA, 1989).

A problemática que remeteu os autores desta pesquisa foi a de identificar as classes de atividades do setor agropecuário com aglomerações produtivas do emprego na região Nordeste e do Ceará no ano 2013, visando a identificação de tendência à concentração das classes de atividade regional e estadual. Apresentar-se-á quais os setores-chave propulsores do crescimento da agropecuária nordestina e cearense nesse ano.

O objetivo principal deste estudo é o de conhecer as classes de atividades do setor agropecuário cearense responsáveis pelo atual cenário de concentração das atividades produtivas, que são especializadas no Nordeste e em quais estados concentra a especialização de cada classe, para que se possa verificar a importância da agropecuária nordestina no cenário nacional.

## 2 REFERENCIAL TEÓRICO

A teoria dos polos de crescimento enfatiza o papel das forças internas de crescimento, caracterizando-se pelos recursos naturais, pelo mercado interno, pela tecnologia, pelo empresariado e pela própria estrutura industrial existente (PELINSKI; SILVA, 2005). Para Rippel e Lima (2009), a ideia de polarização parte da compreensão do espaço, um elemento ativo na dinâmica do sistema produtivo, tendo sua origem na atividade humana e constituindo-se de relações econômicas, tais como: produção, consumo, tributação, investimento, exportação, importação e migração.

Para Perroux (1977 *apud* PELINSKI e SILVA, 2005), o crescimento econômico não surge ao mesmo tempo e nem em todo o território, mas sim, por meio dos polos de crescimento. Segundo o autor, através de diversos canais de integração (como os meios de transporte e comunicação, através dos insumos industriais, do comércio de bens finais e das matérias primas) o crescimento acaba difundindo-se a partir desses polos para outras regiões, proporcionando um crescimento econômico generalizado. Porém, Benitez (1998), afirma que o crescimento e desenvolvimento somente serão conseguidos através de uma organização consciente do meio de propagação dos efeitos do polo, obtidos

através de planos de infraestrutura e comunicações comuns a várias regiões, elaborados pelo Estado.

A teoria dos lugares centrais de Christaller (1933) *apud* PIROLA (2012) sugere que as cidades formam uma hierarquia de lugares centrais e que os maiores abrigam um portfólio maior de atividades. A hierarquia espacial proposta pelo autor está relacionada com o dinamismo econômico. Lugares mais populosos tendem a possuir uma diversidade maior de atividades econômicas, atraindo maior parte da população, pondo o campo como um lugar pouco atrativo e daí a diferenciação no desenvolvimento.

Segundo Alves (2011), a Teoria dos Lugares Centrais mostra uma organização espacial da população de acordo com a importância e o dinamismo das atividades econômicas, principalmente o comércio e a indústria. A proximidade de centros industriais e comerciais faz com que a distribuição da população se dê em torno desses polos aglutinadores, ou seja, uma polarização ou redes desses centros urbanos.

A teoria das aglomerações de Alfred Marshall é explanada por Focchezatto (2010) da seguinte forma:

Para Marshall (1982), as economias de aglomeração são geralmente conhecidas como as economias de escala de uma localidade específica. O autor aponta as primeiras explicações para a atividade industrial apresentar economias de escala externas à firma e destaca três elementos pelos quais as vantagens aglomerativas manifestam-se: um mercado de trabalhadores com mão de obra qualificada; a disponibilidade de serviços e fornecedores de matéria-prima especializada; e a presença de *spillovers* de tecnologia e conhecimento. Esse conjunto de fontes ficou conhecido, posteriormente, como a “tríade marshalliana” (FOCCHIZZATTO, 2010).

A teoria dos polos de crescimento de Perroux (1977), dos lugares centrais de Christaller (1933), e a das economias de aglomerações de Marshall (1982) dizem, em suma, que o desenvolvimento surge a partir de um ponto específico, sendo esse ponto atrativo para os agentes que venham a se instalar naquele local. Esse ponto não é escolhido ao acaso, mas por possuir vantagens como, por exemplo, a facilidade em conseguir matéria prima, menor custo de transporte, formando assim aglomerados industriais ou comerciais, tornando o desenvolvimento polarizado. É o que eles chamam de economias de escala.

## 4 METODOLOGIA

As regiões selecionadas para estudo foram o Estado do Ceará e a macrorregião Nordeste do Brasil. O Nordeste é composto por nove estados, são eles: Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe.

Foram selecionados artigos de eventos publicados em anais eletrônicos (internet) para formação do referencial teórico, bem como livros consagrados em economia regional, como é o caso do Economia Regional de Haddad (1989).

Os dados do emprego formal foram conseguidos no site do Ministério do Trabalho e Emprego (MTE). Foram considerados apenas empregos ativos em 31/12.

Haddad (1989) sugere que se faça a desagregação dos setores para não incorrer em erros de interpretação por conta de valores muito altos devido ao agrupamento das classes. Então optou-se nesse trabalho por uma análise das classes de atividade da CNAE 2.0, por ser uma desagregação bem superior as demais classificações de atividade setorial. A escolha do período analisado foi pela disponibilidade de dados da CNAE 2.0 no SIDRA.

A seleção do emprego formal como variável de análise se deu pelo fácil acesso aos dados e também pelo fato de ser uma variável que não sofre alterações em termos absolutos, não precisando de conversão (como é o caso de Renda em determinadas metodologias e séries temporais).

Essa metodologia foi escolhida por se encaixar perfeitamente nos estudos que se pretende compor sobre exploração da concentração setorial por classe de atividade em 2013.

Nesse estudo, o cálculo do Índice de Concentração Normalizado (ICN) foi aplicado, visando à identificação da participação das classes das atividades e identificar quais possuem maior representatividade na economia do Nordeste e também do Ceará. O ICN é composto por três índices: O quociente Locacional (QL), O índice de Hirschmann-Herfindahl (IHH) e de Participação Relativa (PR). O cálculo do QL é dado pela equação (1):

$$QL = \frac{E_{ij} / E_j}{E_i / E_{..}} \quad (1)$$

Onde o  $E_{ij}$  representa o emprego no setor “i” na região “j”;  $E_j$  é o número do emprego total na região “j”;  $E_i$  É o emprego do setor “i” em todas as regiões



e;  $E_{..}$ , o emprego total nacional.

Na parte do numerador tem-se por dedução, a participação do setor na região estudada em relação ao emprego total da mesma, enquanto que no denominador verifica-se a representatividade percentual do setor total nacional. Monastério (2011) afirma que os valores de ICN maiores do que 1, significam que mais especializada é essa região em relação ao setor. Em contrário, os valores menores que um indicam que a região é importadora do bem produzido naquele setor, pois o mesmo tem uma menor representação na região do que no Estado.

Ele também sugere alguns cuidados na interpretação desse índice:

Se uma região possui uma tecnologia mais intensiva em trabalho do que outras regiões, seu QL pode ser enganoso, sugerindo exportações que não existem. Da mesma forma, caso exista uma diferença sensível no padrão de demanda local, o QL também poderá ser maior do que 1, e a região ser uma importadora do bem. Além disso, o QL é bastante sensível ao nível de análise e ao grau de detalhamento setorial. Em níveis de agregação maiores, o indicador tende a convergir para a unidade (MONASTÉRIO, 2011, p.318).

O (QL) visa identificar neste estudo, quais os subsetores econômicos que apresentam uma participação relativa superior à verificada na média da Região.

Utilizando da metodologia de análise proposta por Crocco *et al.* (2003) e Monastério (2011), se o valor do QL for superior a 1, a região Nordeste é, em termos relativos, significativamente especializado naquela atividade. Mas a análise dos critérios do QL deve ser utilizada com mais detalhamento, pois segundo Crocco *et al.* (2003), um Quociente Locacional maior que um (QL>1) indicaria apenas certa diferenciação produtiva da atividade; este conceito se aplica possivelmente por conta da dissimetria existente entre as atividades da região e ainda da alta representatividade que uma firma pode representar em uma região. É bastante importante não realizar afirmações concretas acerca do resultado do QL, sem antes mensurá-lo junto a outros índices, o que pode não apresentar resultados não significativos ou dar significância a dados que não expressam a realidade local.

Visando reduzir esses erros, Crocco *et al.* (2003) incluíram o IHH, o qual é extraído a partir da equação (2):

$$IHH = \left( \frac{E_{ij}}{E_{.j}} - \frac{E_{i.}}{E_{..}} \right)^2 \quad (2)$$

Para McCann (2001) *apud* Monastério (2011)

O IHH indica o quanto um setor está concentrado espacialmente. Ele se assemelha ao *coeficiente de localização*<sup>1</sup>, mas os afastamentos das regiões em relação à estrutura produtiva do país são elevados ao quadrado. Dessa forma, o IHH, que varia entre 0 e 2, é mais sensível a tais afastamentos do que o CL.

O resultado positivo deste índice indica que o Ceará está concentrando a produção do setor em análise e por isso ele terá maior poder de atração econômica devido ao seu nível de especialização. No trabalho de Crocco *et al.* (2003) este índice é utilizado sem elevação ao quadrado, mas no presente trabalho acredita-se que a elevação ao quadrado capta a maior sensibilidade ao afastamento entre as regiões como explica McCann (2001) *apud* Monastério (2011).

O PR destaca o grau de participação do setor da região estudada. Este índice tem variação entre zero e um, onde quanto mais próximo de um, maior a representatividade da atividade na estrutura do Estado. É adquirido através da equação (3).

$$PR = E_{ij} / E_j \quad (3)$$

Os indicadores acima relacionados são capazes de apresentar os dados necessários para a elaboração de um único indicador de concentração do subsetor de atividade econômica, o Índice de Concentração normalizado (ICN).

Crocco *et al.* (2003) destacam que:

Haja vista que cada um dos três índices utilizados como insumos do ICN podem ter distinta capacidade de representar as forças aglomerativas, principalmente quando se leva em conta os diversos setores industriais da economia, faz-se necessário calcular pesos específicos de cada um dos insumos em cada um dos setores produtivos (CROCCO *et al.* 2004, p.6).

A metodologia utilizada por Crocco *et al.* (2003) propõe uma combinação linear dos três índices, separada para cada subsetor da região estudada, que segue na equação (4):

$$ICN_j^i = \theta_1 QLn_j^i + \theta_2 PRn_j^i + \theta_3 HHn_j^i \quad (4)$$

1. OCL (coeficiente de localização) é um índice que mensura o quanto um setor está concentrado espacialmente (quanto mais próximo de 1). Para monastério (2011) o IHH tem uma maior precisão em relação ao CL. Grifonosso.

Os  $\emptyset$  são os pesos de cada um dos indicadores para cada índice correspondente. Este índice visa corrigir algumas falhas de análise que por ventura os demais índices em separado possam demonstrar (CROCCO *et al.*, 2003, p.7). Eles serão extraídos a partir da análise fatorial (A.F) a partir do método dos componentes principais (ACP).

Caso se tomasse como base o QL isolado, numa cidade cuja sua PR não seja significativa em relação ao país, tende-se a acreditar numa especialização inexistente, cujo aumento do índice se deu em detrimento de uma única fábrica numa cidade de pequeno porte, por exemplo.

Assim, o uso do ICN da forma como está sendo proposta pode ponderar tais distorções (CROCCO *et al.*, 2003, p. 7).

Antes, porém, da efetiva substituição dos respectivos pesos encontrados, foi realizada a normalização dos indicadores, para que não haja interferência da significância de um ou outro dado da amostra, tornando-os relativos e deixando-os com a mesma carga de importância na construção e análise do índice. Essa normalização se deu através do cálculo do desvio padrão e da média de cada indicador em cada setor. Ela é feita subtraindo-se do valor do indicador a média e o resultado é dividido pelo desvio padrão, como mostrado na equação 5. Os valores em negativo representam os setores que ficaram abaixo da média estadual. O cálculo da normalização é dado pela fórmula:

$$(5) \quad I_n = (I_i - I_{Média}) / I_{Desvio\_padrão}$$

Onde:

$I_n$  = Indicador normalizado

$I_i$  = Valor do Indicador no setor correspondente

$I_{Média}$  = Média do Indicador

$I_{Desvio\_padrão}$  = Desvio Padrão do Indicador

#### 4.1 A Análise Fatorial (A.F) e o Método dos Componentes Principais (ACP)

A análise fatorial se baseia na suposta existência de um número de fatores causais gerais, cuja presença dá origem às relações entre as variáveis observadas, de forma que, no total, o número de fatores seja consideravelmente inferior ao total de variáveis. Isso porque muitas relações entre as variáveis são, em grande medida, decorrentes do mesmo fator causal geral.

O modelo matemático (conforme FÁVERO *et al.*, 2009) da análise fato-

rial poderá ser representado de forma simplificada por:

$$Z_j = \sum a_{ji} F_i d_j u_j \quad (j = 1, 2, \dots, n); \quad (i = 1, 2, \dots, m) \quad (6)$$

Em que:

$Z_j$  = j-ésima variável padronizada;

$a_{ji}$  = é o coeficiente de saturação referente ao i-ésimo fator comum da j-ésima variável;

$F_i$  = é o i-ésimo fator comum;

$d_j$  = é o coeficiente de saturação referente ao j-ésimo fator específico da j-ésima variável;

$u_j$  = é o j-ésimo fator específico da j-ésima variável.

Assim, o objetivo da análise fatorial consiste em determinar um número menor de fatores que representem a estrutura das variáveis originais. Nesta etapa, é determinado o número de fatores comuns necessários para descrever adequadamente os dados, cabendo ao pesquisador a decisão de qual método de extração dos fatores e o número de fatores selecionados para representar a estrutura latente dos dados.

A fim de verificar a adequabilidade dos dados para a análise fatorial, serão utilizados o Índice Kaiser-Mayer-Olkin (KMO), o Teste de Esfericidade de Bartlett (BTS).

O Índice Kaiser-Mayer-Olkin (KMO), que varia de 0 a 1, serve para comparar as magnitudes dos coeficientes de correlações observados com as magnitudes dos coeficientes de correlações parciais. Portanto, o KMO é uma medida de homogeneidade das variáveis, que compara as correlações parciais observadas entre as variáveis, conforme a seguir:

$$KMO = \frac{\sum_i \sum_j r_{ij}^2}{\sum_i \sum_j r_{ij}^2 + \sum_i \sum_j a_j^2}$$

$r_{ij}$  é o coeficiente de correlação observado entre as variáveis  $i$  e  $j$   
 $a_j$  é o coeficiente de correlação observado entre as mesmas variáveis, que é, simultaneamente, uma estimativa das correlações entre os fatores. Os  $a_j$  deverão estar próximos de zero, pelo fato de os fatores serem ortogonais entre si.

A estatística do KMO (Kaiser-Mayer-Olkin) é a seguinte: quanto menor

o valor do respectivo teste, menor a relação entre as variáveis e os fatores, podendo o índice variar de 0 a 1. O índice menor do que (0,5) se caracteriza como inaceitável; o uso dessa técnica, caso contrário, com o índice próximo de 1, a utilização da técnica com os dados se torna bastante eficaz.

O Teste Bartlett de Esfericidade pode testar a hipótese nula de que a matriz de correlações é uma matriz identidade (o que inviabiliza a metodologia da análise fatorial proposta).

Caso a matriz de correlações seja matriz identidade, significa que as inter-relações das variáveis são iguais a zero e, portanto, a análise fatorial não deverá ser utilizada, sendo a hipótese  $H_0$  (a matriz de correlações é uma matriz identidade) e a hipótese  $H_a$  (a matriz de correlações não é uma matriz identidade). Caso  $H_0$  seja aceita a análise fatorial deve ser desconsiderada; se rejeitada, haverá indícios de que existam correlações entre as variáveis explicativas do processo.

A ACP ao tomar  $p$  variáveis  $X_1, X_2, \dots, X_p$ , encontra combinações lineares para elas, produzindo as variáveis  $Z_1, Z_2, \dots, Z_p$  sendo que

$$Z_i = \alpha_{i1}X_1 + \alpha_{i2}X_2 + \dots + \alpha_{ip}X_p \quad (7)$$

Variando tanto quanto possível, desde que condicionadas a

$$\alpha_{i1}^2 + \alpha_{i2}^2 + \dots + \alpha_{ip}^2 = 1 \quad (8)$$

As variâncias dos componentes principais são os autovalores dessa matriz, ao passo que os coeficientes  $a_{i1}, a_{i2}, \dots, a_{ip}$  são os seus autovetores associados. A matriz de variância é simétrica e tem a forma:

$$\lambda_1 + \lambda_2 + \dots + \lambda_p = C_{11} + C_{22} + \dots + C_{pp} \quad (9)$$

Em que:  $\lambda_i$  são os autovalores, ou variância, de cada um dos  $i$  componentes.

#### 4.1.1 Calculando os pesos via ACP

Uma característica dos autovalores é que a soma desses é igual à soma dos elementos da diagonal principal da matriz de covariância:

$$C = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1p} \\ C_{21} & C_{22} & \dots & C_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ C_{p1} & C_{p2} & \dots & C_{pp} \end{bmatrix}$$

Conforme explica Crocco *et al.* (2006) *apud* Piccicini, Finamore e Oli-

veira (2011), o procedimento para o cálculo dos pesos se inicia com os resultados extraídos conforme seguem na Tabela 1. Já a Tabela 2 mostra a matriz de coeficientes ou os autovetores da matriz de correlação. Na Tabela 3, apresenta os autovetores recalculados ou a participação relativa de cada índice nos componentes.

**TABELA 1 - Os autovalores da matriz de correlação ou variância explicada pelos componentes**

Componente	Variância explicada da componente	Variância explicada total
1	$\beta_1$	$\beta_1$
2	$\beta_2$	$\beta_1 + \beta_2$
3	$\beta_3$	$\beta_1 + \beta_2 + \beta_3$ (100%)

Fonte: Piccicini, Finamore e Oliveira (2011).

**Tabela 2 - Matriz de coeficientes ou autovetores da matriz de correlação**

Indicador	Componente 1	Componente 2	Componente 3
QL	$\alpha_{11}$	$\alpha_{12}$	$\alpha_{13}$
IHH	$\alpha_{21}$	$\alpha_{22}$	$\alpha_{23}$
PR	$\alpha_{31}$	$\alpha_{32}$	$\alpha_{33}$

Fonte: Piccicini, Finamore e Oliveira (2011).

A Tabela 3 é calculada com os resultados obtidos das Tabelas 1 e 2.

**TABELA 3 - Matriz de auto vetores recalculados ou participação relativa dos indicadores em cada componente**

Indicador	Componente 1	Componente 2	Componente 3
QL	$\alpha'_{11} \equiv \frac{\alpha_{11}}{C_1}$	$\alpha_{12} \equiv \frac{\alpha_{12}}{C_2}$	$\alpha_{13} \equiv \frac{\alpha_{13}}{C_3}$
IHH	$\alpha_{21} \equiv \frac{\alpha_{21}}{C_1}$	$\alpha_{22} \equiv \frac{\alpha_{22}}{C_2}$	$\alpha_{23} \equiv \frac{\alpha_{23}}{C_3}$
PR	$\alpha_{31} \equiv \frac{\alpha_{31}}{C_1}$	$\alpha_{32} \equiv \frac{\alpha_{32}}{C_2}$	$\alpha_{33} \equiv \frac{\alpha_{33}}{C_3}$

Fonte: Piccicini, Finamore e Oliveira (2011).

Os valores de  $\alpha'_{ij}$  da Tabela 3 correspondem aos pesos que cada uma das variáveis assume dentro de cada componente e os autovalores  $\beta_1$ ,  $\beta_2$  e  $\beta_3$

forneem a variância dos dados associada ao componente, o peso final de cada indicador insumo é o resultado da soma dos produtos dos  $\alpha'_{ij}$  pelo auto valor correspondente em cada componente. Apresentando-se nas equações 10, 11 e 12:

$$\Theta_1 = \alpha'_{11}\beta_1 + \alpha'_{12}\beta_2 + \alpha'_{13}\beta_3 \quad (10)$$

$$\Theta_2 = \alpha'_{21}\beta_1 + \alpha'_{22}\beta_2 + \alpha'_{23}\beta_3 \quad (11)$$

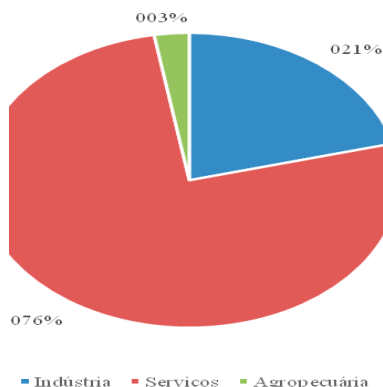
$$\Theta_3 = \alpha'_{31}\beta_1 + \alpha'_{32}\beta_2 + \alpha'_{33}\beta_3 \quad (12)$$

Onde respectivamente os pesos  $\Theta_1$ ,  $\Theta_2$  e  $\Theta_3$  representam o QL, IHH e PR, sendo que a soma dos três pesos deve ser igual a um. Outra observação que deve ser considerada é que o cálculo dos pesos não deve ser feito para a economia como um todo, mas individualmente para cada uma das atividades que se deseja estudar.

## 5 RESULTADOS

O total de empregos no Nordeste em 2013 foi de 8.926.710 e no Brasil, no mesmo ano, o número de empregos foi de 48.948.433, ou seja, a Região Nordeste detém 18% de todo o emprego nacional. Segundo estimativas do IBGE, a população residente no Brasil no ano de 2014 era de 202.768.563 pessoas e na região Nordeste era de 56.186.190, ou seja, 27,7% da população brasileira vive no Nordeste.

**GRÁFICO 1: Distribuição Percentual do Emprego por Grandes Setores no Nordeste 2013**

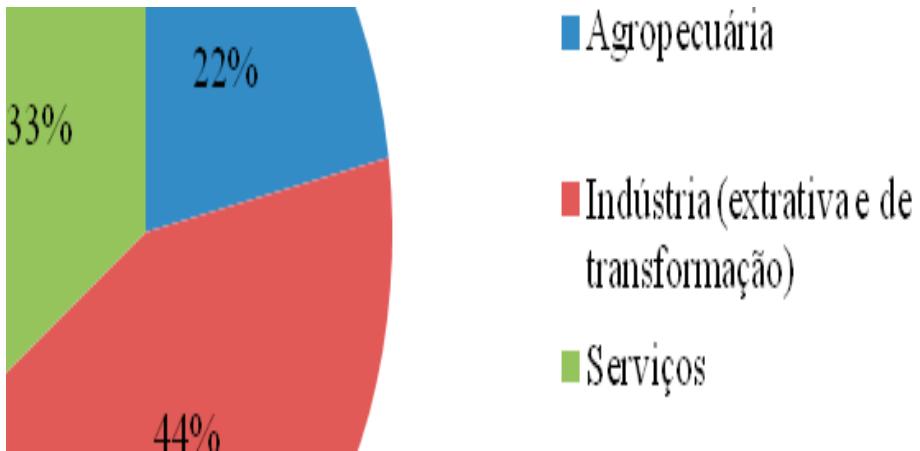


Fonte: elaboração dos autores a partir de dados do MTE.

Como pode ser observado acima, no gráfico 1, a distribuição do emprego

nos grandes setores da economia no Nordeste do Brasil se dá da seguinte maneira: vê-se que em 2013 o setor de serviços obteve a maior participação no emprego no Nordeste com 76,36% do total. Em seguida vem o setor industrial, com 20,97% e logo após é apresentado o setor agropecuário, com 2,67% do total de empregos. Ou seja, para cada cem pessoas empregadas na região Nordeste, 76 estão no setor de serviços, 21 no setor de industrial e 3 no setor agropecuário. Nota-se então a predominância do setor de serviços na composição do emprego regional nordestino.

**GRÁFICO 2: Distribuição percentual das classes de atividades especializadas por grandes setores no Nordeste-2013**



Fonte: SIDRA/IBGE.

No gráfico 2, onde mostra a distribuição da especialização das classes de atividades nordestinas se percebe que, ao contrário da distribuição emprego, onde o setor de serviços se mostra superior, na especialização é o setor industrial que tem o maior número de atividades especializadas, totalizando 44%. Isso corrobora com os estudos da teoria econômica regional, onde se descreve a respeito das aglomerações industriais, que se formam a partir de uma indústria motriz, criando uma tendência locacional. O setor de serviços é o segundo setor com o maior número de atividades especializadas, com 33%. Por último vem o setor agropecuário, com 22% das atividades especializadas no Nordeste. Apesar da agropecuária possuir o menor número de empregos na região e o menor número de atividades especializadas, estudos apontam que o este é um setor tático; o discurso internacional sobre desenvolvimento econômico situa a agricultura como setor estratégico para aferição de taxas maiores de crescimento econômico, acompanhadas de um melhor padrão de vida da população. Ressalte-se



que foi a partir dos incentivos à agricultura que os avanços tecnológicos viabilizaram a industrialização nos países centrais (FERREIRA; RAMOS; ROSA, 2006). Portanto, estudos ligados a este setor podem servir de apoio para a elaboração de políticas para a promoção do desenvolvimento regional e nacional.

**TABELA 1 - Relação das classes de atividades agropecuárias especializadas no Nordeste – 2013**

Ranking	Código e Atividade CNAE 2.0	ICN
1º.	01.14 Cultivo de fumo	4,41
2º.	03.21 Aquicultura em água salgada e salobra	4,23
3º.	01.35 Cultivo de cacau	4,07
4º.	01.32 Cultivo de uva	3,81
5º.	03.22 Aquicultura em água doce	1,58
6º.	01.12 Cultivo de algodão herbáceo e de outras fibras de lavoura temporária	1,33
7º.	01.33 Cultivo de frutas de lavoura permanente, exceto laranja e uva	1,28
8º.	01.53 Criação de caprinos e ovinos	1,26

Fonte: Tabela de elaboração dos autores com dados extraídos da metodologia. NOTA: números arredondados.

Apresenta-se então a tabela 1, que em 2013, a atividade do setor agropecuário nordestino obteve um ICN de maior expressividade no cultivo do fumo, que segundo Furtado (1991), na segunda metade do século XIX era o principal produto da economia baiana, destinado à exportação para o mercado europeu. No ano de 2004, a Bahia foi responsável por 74% das exportações de charutos e cigarros no Brasil, conforme o estudo feito por Oliveira (2006), o que mostra que essa atividade continua concentrada no estado da Bahia, desde o início do cultivo do fumo no Brasil até os dias atuais.

Este subsetor, no Nordeste, possui 1588 empregos formais e o Brasil possui 1675 empregos, ou seja, apenas a região Nordeste do país detém 94,81% de todo o emprego nacional nessa categoria.

Outro grande destaque na especialização agropecuária nordestina é o cultivo de cacau, com um ICN de 4,07. Este subsetor, no Nordeste, também está localizado em sua maioria, na Bahia, onde só no ano de 2002 respondia por 63% da produção de cacau do país, segundo o estudo feito por Cuenca e

Nazário (2004).

O cultivo de uva no Nordeste, em 2013, obteve um ICN de 3,81. A produção brasileira de uva em 2008 foi da ordem de 1,42 milhões de toneladas, destacando-se a região Sul como a maior produtora, participando com 936,7 mil toneladas (66,0%) da produção brasileira de uva, enquanto a participação da região Nordeste foi de 18,8% (267,1 mil toneladas), seguida da região Sudeste com 207,2 mil toneladas ou 14,6% (Instituto Brasileiro de Geografia e Estatística-IBGE, 2009). Muito embora a região Sul apresente-se como a maior produtora de uva do País, vale ressaltar que a uva produzida nessa região destina-se, principalmente, à produção de vinho, enquanto nas regiões Nordeste e Sudeste predominam a produção de uvas de mesa. Em 2008, a produção brasileira de uva destinada ao consumo *in natura* representou 691,2 mil toneladas, correspondendo a 49,4% da produção nacional, tendo a região Nordeste contribuindo com 37,7% (SILVA e CORREIA, 2004).

A área colhida de uva no Brasil, em 2008, segundo os dados do IBGE (2009), foi da ordem 81.286 ha, dos quais 11.552 ha estavam localizados na região Nordeste, que apreendia o equivalente a 14,2% da área total cultivada com videira no País. Nesse mesmo ano, enquanto as produtividades médias nas regiões Sul e Sudeste oscilavam em torno de 16 toneladas por hectare ao ano e 18 toneladas por hectare ao ano, respectivamente, na região Nordeste, a produtividade média alcançou 23,1 toneladas por hectare ao ano.

A produção de uva no Nordeste do Brasil concentra-se principalmente na região do Submédio do Vale do São Francisco, localizada nos sertões pernambucano e baiano. Favorecida pela potencialidade dos recursos naturais e pelos investimentos públicos e privados nos projetos de irrigação, esta cultura vem apresentando uma grande expansão ano após ano.

O cultivo do algodão herbáceo e de outras fibras de lavoura temporária possuía, em 2013, um ICN de 1,33. O estado do Ceará era, na safra de 1976/77, o principal estado do País em área cultivada de algodão, com 1,3 milhões de hectares, possuindo, portanto, a maior participação relativa dentre os estados, representando 32% da área plantada em todo o País. A área deste cultivo compreendia uma área de aproximadamente 10% do território do Estado, daí, percebe-se sua grande importância econômica e social para o Ceará, tendo participação relevante no produto local e na geração de ocupação para uma grande parcela da população rural, uma vez que seu cultivo baseava-se na agricultura familiar (CAVALCANTE *et al.*, 2007).

Em 1997, a produção dessa cultura no Ceará representava quase 2% da produção nacional passando a representar apenas 0,2% em 2005, reflexo da redução da produção ocorrida na maioria dos municípios cearenses. No primeiro ano, 91 municípios produziam algodão herbáceo no Estado, em 2005, esse número caiu para 61.

O Estado da Bahia caminhou no sentido contrário, expandindo a área cultivada e superando, em participação relativa, o Estado do Ceará. Os hectares plantados passaram de 55 mil, nos anos de 2000/01, para 247 mil na safra 2004/05, área quase 25 vezes maior que a área plantada de algodão no Estado do Ceará. Esse valor representa 21% da área cultivada em termos de Brasil, e 75% da Região Nordeste (CAVALCANTE *et al.*, 2007).

A fruticultura irrigada predomina nos perímetros irrigados, mas há também pomares dispersos à margem de açudes maiores e de rios perenizados, e próximos a outras fontes d'água, a exemplo de barragens subterrâneas, poços artesianos, entre outras.

O cultivo de frutas de lavoura permanente, exceto laranja e uva obteve um ICN de 1,28. No estudo feito por Gomes (2015), é destacado que a região Nordeste possui grande papel de destaque na atividade frutícola para o País. A fruticultura possui enorme potencial para a região na oferta de emprego e renda. As condições climáticas da sub-região semiárida nordestina, que representava um obstáculo para o cultivo de grãos e a pecuária, transformaram-se numa vantagem quando se trata da fruticultura irrigada. No Nordeste, são muitos os aglomerados de polos frutícolas, sendo o de maior referência em estudos nesse ramo o polo Petrolina-PE/Juazeiro-BA. A Região Nordeste, mediante a participação dos seus polos irrigados, é a principal região produtora e exportadora de frutas tropicais frescas do Brasil.

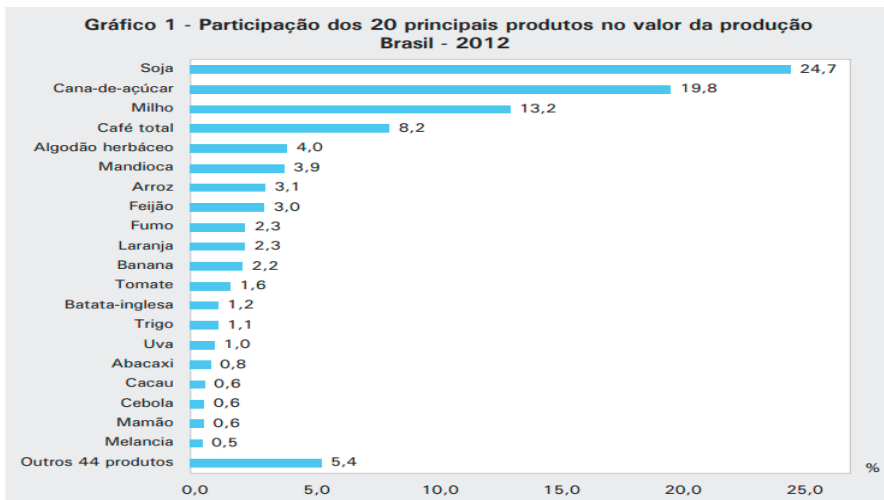
A criação de caprinos e ovinos possui um ICN de 1,26, tendo o menor ICN dentre as classes de atividades da agropecuária especializadas no Nordeste. Segundo a Embrapa (2002), a criação de caprinos e ovinos no Nordeste Brasileiro é praticada desde a colonização, principalmente pelo fato dessas espécies serem mais adaptadas às condições ambientais e climáticas desfavoráveis do que a maioria das outras espécies.

A região Nordeste está em mais de 80% coberta pela vegetação nativa da Caatinga. Este tipo de vegetação é utilizado como a principal fonte de alimentação para a maioria dos rebanhos. No entanto, durante a época seca o uso da vegetação como única fonte alimentar limita o potencial produtivo dos reba-

nhos. O mercado da carne ovina está em franca ascensão em todo o país. Os preços hoje praticados no âmbito da unidade produtiva representam bem mais do que o preço pago pela carne bovina nas mesmas condições.

No momento cerca de 50% da carne ovina consumida no Nordeste e Centro-Oeste são provenientes do Uruguai, da Argentina e da Nova Zelândia. Esta informação mostra uma possibilidade enorme de mercado a ser conquistado, principalmente porque no Brasil, especialmente no Nordeste, tem-se potencial para produzir carne de melhor qualidade do que àquela importada. A produção de carne proveniente de animais deslançados poderá perfeitamente atender à demanda interna e em futuro próximo adentrar no mercado internacional. A pele por sua vez, agrega valor ao produto, uma vez que forem adotadas regras básicas de manejo, este produto poderá representar ater 30% do preço final do animal (LEITE e SIMPLICIO, 2005).

**FIGURA 1: Participação dos 20 principais produtos no valor a produção Brasil-2012**



Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Agropecuária, Produção Agrícola Municipal 2012.

Fonte: IBGE (2012).

Fazendo um comparativo entre as culturas no total brasileiro (figura 1) e o Nordeste, dentre as vinte principais atividades que compõem a produção agrícola nacional, o Nordeste possui especialização em quatro delas. O algodão herbáceo, que possuem um ICN de 1,33 no Nordeste é o quarto produto de maior importância na agricultura nacional no ano de 2012. O cultivo do fumo

é a classe de atividade que possui o maior ICN agropecuário no Nordeste com um índice de 4,41 e é o nono produto de maior importância na produção agrícola do Brasil. o cultivo e uva é o décimo quinto principal produto a compor a produção agrícola nacional e é um dos principais produtos especializados no Nordeste, com um ICN de 3,81. O décimo sétimo principal produto a compor a produção agrícola do país é o Cacau, que na região Nordeste é um dos principais tipos de produção agrícola, com um ICN de 4,07.

Os produtos pecuários especializados no Nordeste são a aquicultura de água salgada e salobra e aquicultura de água doce, obtendo um ICN de 4,23 e 1,58, respectivamente. Além de serem as únicas classes de atividade pecuárias especializadas no Nordeste, são também as únicas classes de atividades agropecuárias especializadas que o nordeste tem em comum com o Ceará.

A seguir, será apresentada uma tabela que quantifica as atividades especializadas no setor agropecuário Cearense no ano de 2013.

**TABELA 2 - Relação das classes de atividades agropecuárias especializadas no Ceará em comum com o Nordeste – 2013**

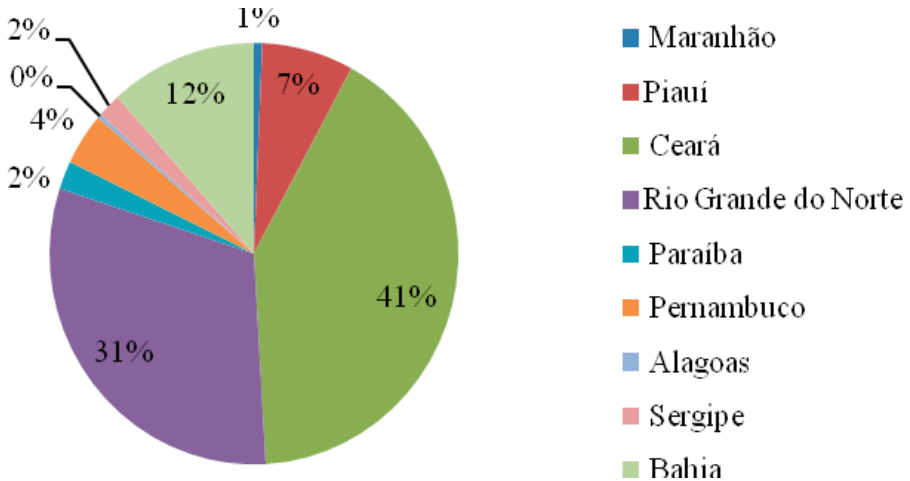
Ranking	Código e Atividade CNAE 2.0	ICN
1º.	Aquicultura em água salgada e salobra	4,23
2º.	Aquicultura em água doce	1,27

Tabela de elaboração dos autores com dados extraídos da metodologia.

NOTA: números arredondados.

Observa-se na tabela 2, um número extremamente reduzido de atividades especializadas no Ceará. O reflexo dessa redução pode ser em parte explicado pela redução da produção desse setor. De 2002 a 2012 o Valor adicionado bruto (VAB) a preços correntes do setor agropecuário reduziu a participação de 7,15% em 2002 do VAB total para 3,38% em 2012. A tabela 2 mostra o ICN das classes de atividades da agropecuária em comum que são especializadas no Nordeste e no Ceará. Como pode ser visto apenas a aquicultura de água salgada e salobra e a aquicultura de água doce, dentre as oito classes de atividades especializadas no nordeste são especializadas também no Ceará. A aquicultura de água salgada no estado do Ceará possui um ICN de 4,23 e a aquicultura de água doce possui um ICN de 1,27, um numero bem menor que a aquicultura em água salgada e salobra. A divisão do emprego na aquicultura entre os estados nordestinos pode ser observado no gráfico 2 a seguir.

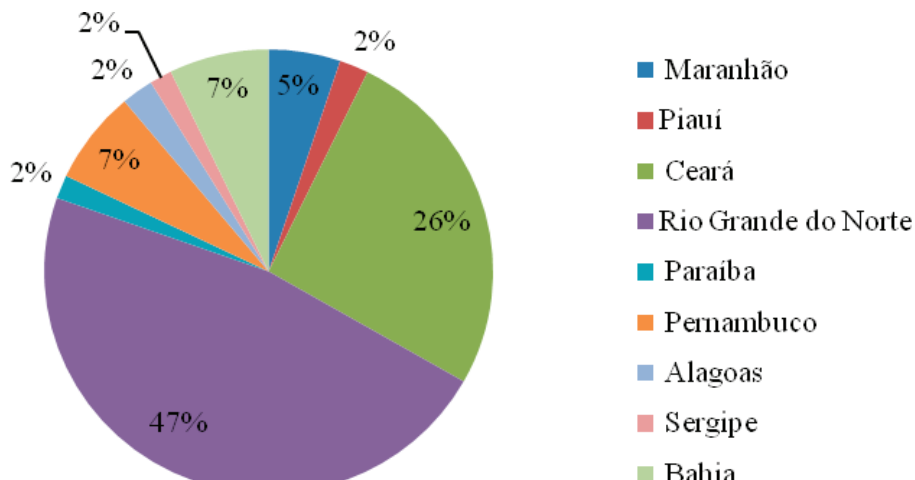
**GRÁFICO 2: Distribuição percentual do emprego da aquicultura em água salgada nos estados da região Nordeste -2013**



Fonte: elaboração própria a partir dos dados da pesquisa

O gráfico 2 apresenta o emprego da aquicultura em água salgada na região nordeste no ano de 2013, estava assim distribuído: o estado nordestino que concentra a maior parte dos vínculos empregatícios nessa classe de atividade é o Ceará, com 41% do total. Em seguida está o estado do Rio Grande do Norte, com 31%, a Bahia com 12%, o Piauí com 7%, o Pernambuco com 4%, a Paraíba e o Sergipe, ambos com 2%, o Maranhão com apenas 1% e Alagoas possuía um valor nulo. Se se considerar apenas os dois estados que concentram o maior número de empregos na aquicultura em água salgada, tem-se um total de 72%, ou seja, apenas o Ceará e o Rio Grande do Norte concentram mais da metade do emprego nessa classe de atividade na região Nordeste.

### GRÁFICO 3: Distribuição percentual do emprego da aquicultura em água doce nos estados da região Nordeste -2013



Fonte: elaboração própria a partir dos dados da pesquisa

No gráfico 3 acima, o emprego da aquicultura em água doce na região Nordeste no ano de 2013 estava assim distribuído: o estado nordestino que concentra a maior parte dos vínculos empregatícios nessa classe de atividade é o Rio Grande do Norte, com 47% do total. Em seguida está o estado do Ceará, com 26%, a Bahia e o Pernambuco, ambos com 7%, o Maranhão com 5%, Paraíba, Alagoas, Piauí e Sergipe com 2% cada. Considerando apenas o Ceará e o Rio Grande do Norte, juntos eles detêm 73% de todo o emprego na aquicultura de água doce do Nordeste do Brasil, 1% a mais que a concentração do emprego na aquicultura em água salgada e salobra desses dois estados no Nordeste.

O Ministério de Pesca e Aquicultura (MPA), em um levantamento preliminar de 2013, divulgou que o País produzia aproximadamente 2 milhões de toneladas de pescado, sendo 40% cultivados.

Segundo levantamento estatístico divulgado pelo MPA em 2010, essa atividade já apresentou significativo crescimento nos últimos anos, passando de 278 mil toneladas em 2003 para 415 mil em 2009, o que equivale a 35% de incremento em menos de uma década. Apenas a produção da piscicultura atingiu 60,2% de crescimento apenas entre 2007 e 2009. Isoladamente a produção de tilápia aumentou 105% em apenas sete anos (2003-2009). Em conjunto, a aquicultura cresceu 43,8%, entre 2007 e 2009, tornando a produção de pescado a que mais cresceu no mercado nacional de carnes no período. Todos estes

resultados demonstram a pronta resposta do setor às políticas de fomento.

FIGURA 2: **Distribuição das espécies aquícolas cultivadas no Brasil**



Fonte: MPA.

Atualmente, cada região brasileira vem se especializando em determinados tipos de pescado. Na Região Norte, predominam peixes como o tambaqui e o pirarucu. No Nordeste, a preferência é pela tilápia e pelo camarão-marinho. No Sudeste, a tilápia tem grande presença na aquicultura. No Sul, predominam as carpas, as tilápias, as ostras e os mexilhões. Já no Centro-Oeste os destaques são o tambaqui, o pacu e os pintados.

Segundo o SEBRAE (2009), O Ceará é o maior produtor de tilápia do Brasil, com uma produção média anual variando entre 22 e 24 mil toneladas. O setor de piscicultura é considerado o maior agronegócio do mundo. Participa com 16% da oferta mundial de proteína animal. Somando U 55 bilhões em exportações anuais e é um mercado duas vezes maior do que o complexo da soja, sete vezes maior do que o negócio da carne bovina e nove vezes em relação à carne de frango.



## CONSIDERAÇÕES FINAIS

O ICN neste trabalho, pode apresentar uma significativa participação da atividade de aquicultura no Nordeste brasileiro e seu resultado em boa parte é oriundo da atividade dentro do estado do Ceará. Isto posto, teve a intenção de avaliar a importância do setor agropecuário nordestino frente o cenário nacional e ver também qual a participação e importância do estado do Ceará para a agropecuária brasileira.

A análise feita neste trabalho mostra que apenas a atividade de aquicultura tanto em água salgada e salobra quanto a aquicultura em água doce, são atividades de especialização em comum no Nordeste e no Ceará. Esta classe se mantém especializada desde 2006, como explicita o trabalho de Santos, *et al.*, (2014). Na aquicultura em água salgada e salobra o cultivo predominante é o do camarão marinho e na aquicultura e água doce é o cultivo da tilápia, que pode ser encontrado em praticamente em todos os açudes da União controlados pelo Departamento Nacional de Obras Contra as Secas (DNOCS).

A Bahia é o estado que concentra a maior parte das atividades agropecuárias especializadas no Nordeste, inclusive no cultivo do algodão, onde o estado do Ceará deteve por longos anos a concentração dessa atividade, sendo este o principal produto da sua economia nas primeiras décadas do século XX, como também foi o responsável pelas primeiras indústrias no estado. O estado baiano juntamente com o de Pernambuco tem grande importância regional e nacional na fruticultura irrigada no agropolo do submédio São Francisco, com estaque para a produção de uva e de outras frutas tropicais.

Observa-se o excelente desempenho da atividade de aquicultura no Estado do Ceará, o que contribuiu significativamente com os resultados do Nordeste no setor. Este é um setor em plena expansão no Estado. Políticas públicas de criação de cursos técnicos já foram desenvolvidas pelo Governo do Estado, aperfeiçoando a mão de obra para melhoria dos processos tecnológicos e aumento da produtividade.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, N. A. S.; ARAÚJO, J. J. C. N.; RODRIGUES, F. M. **A Teoria dos Lugares Centrais e sua Aplicabilidade no Programa Zona Verde no Amazonas**. Santa Cruz do Sul: REDES, 2009. Disponível em: <<http://online.unisc.br/seer/index.php/redes/article/view/876/1457>> acesso em: 12/05/2015
- ALVES, F. N. **Notas Teórico- Metodológicas entre Geografia Econômica e Desenvolvimento Regional**. Santa Cruz do Sul: 2011. Disponível em: <<http://www.unifal-mg.edu.br/geres/files/unisc.pdf>> acesso em 20/06/2015
- BENITEZ, R. M. **O Capital Social Fixo Como Insumo do Desenvolvimento Regional**. Fortaleza: Revista Econômica do Nordeste, 1998.
- CAVALCANTE, A. L.; ALBUQUERQUE, D. P. L.; PAIVA, W. L.; MAGALHÃES, K. A. **Cultura do Algodão no Estado do Ceará**. Fortaleza: IPECE, 2007.
- CUENCA, M. A. G.; NAZÁRIO, C. C. **Importância Econômica e Evolução da Cultura do Cacau no Brasil e na Região dos Tabuleiros Costeiros da Bahia entre 1990 e 2002**. Aracaju: Embrapa Tabuleiros Costeiros, 2004. 25 p
- CROCCO, Marco Aurelio *et al.* **Metodologia de identificação de arranjos produtivos locais potenciais: uma nota técnica**. Belo Horizonte: UFMG/Ce-deplar, 2003. (Texto para discussão; 191)
- FÁVERO, L. P.; BELFIORE, P.; SILVA, F. L.; CHAN, B. L. **Análise de dados: modelagem multivariada para tomada de decisões**. Primeira edição. Rio de Janeiro: Elsevier, 2009, p.195-265.
- FERREIRA, Carlos Mauricio de C. **As teorias de localização e organização espacial da economia**. In: HADDAD, P. R. et. al. Economia Regional: Teorias e métodos de análise. (1989) BNB - ETENE p. 67-206
- FERREIRA, M. O.; RAMOS, L. M.; ROSA, A. L. T. **Crescimento da agropecuária cearense: comparação entre as produtividades parciais e total**. Brasília: Rev. Econ. Sociol. Rural vol.44 n.3. 2006.
- FOCHEZATTO, A. **Desenvolvimento Regional: notas sobre abordagens para novos paradigmas produtivos**. O Ambiente Regional, 2010. Disponível em: <<http://www.fee.rs.gov.br/3-decadas/downloads/volume1/5/adelar-fochezatto.pdf>> acesso em 18/06/2015.
- FURTADO, C. **Formação Econômica do Brasil**. 24. ed. São Paulo:

editora Nacional, 1991.

GOMES, O. P. **Perfil Socioeconômico e Tecnológico de Produtores de Fruticultura Irrigada na Região do Cariri, Ceará.** Fortaleza. 2015.

HADDAD, Paulo Roberto. **Medidas de localização e especialização.** In: \_\_\_\_\_, Paulo Roberto et. al. Economia Regional: Teorias e métodos de análise. (1989) BNB – ETENE p. 225-247.

INSTITUTO BRASIELIRO DE GEOGRAFIA E ESTATÍSTICA-IBGE. **Estimativas da População Residente no Brasil e Unidades da Federação com Data de Referência em 1º de Julho de 2014.** Disponível em: <[ftp://ftp.ibge.gov.br/Estimativas\\_de\\_Populacao/Estimativas\\_2014/estimativa\\_dou\\_2014.pdf](ftp://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2014/estimativa_dou_2014.pdf)> acesso em 08/09/2015.

INSTITUTO BRASIELIRO DE GEOGRAFIA E ESTATÍSTICA-IBGE. **Produção Agrícola Municipal:** culturas temporárias e permanentes. 2012. disponível em: < [ftp://ftp.ibge.gov.br/Producao\\_Agricola/Producao\\_Agricola\\_Municipal\\_\[anual\]/2012/pam2012.pdf](ftp://ftp.ibge.gov.br/Producao_Agricola/Producao_Agricola_Municipal_[anual]/2012/pam2012.pdf)> acesso em 16/09/2015.

LEITE, E. R.; SIMPLICIO, A. A. **Sistema de Produção de Caprinos e Ovinos de Corte para o Nordeste Brasileiro.** Embrapa Semi-Árido. 2005.

MONASTERIO, Leonardo. Indicadores de análise regional e espacial. In: CRUZ, Bruno de Oliveira *et al.* (org.) Economia regional e urbana: teorias e métodos com ênfase no Brasil. Brasília: Ipea, 2011.

PELINSKI, A.; SILVA, D. R. **A Influência da Indústria Motriz no Desenvolvimento Regional.** Cascavel: Unioeste, 2005. Disponível em: <<http://cac-phi.unioeste.br/projetos/gpps/midia/seminario2/trabalhos/economia/meco02.pdf>> acesso em 10/05/2015

PICCININI, Fabio Junior; FINAMORE, Eduardo B. M. C; OLIVEIRA, Guilherme de. **Identificação e mapeamento de aglomerações produtivas no Rio Grande do Sul: um enfoque na região da produção.** Disponível em: <http://webcache.googleusercontent.com/search?q=cache:7ZYJpMPdDWcJ:bell.unochapeco.edu.br/revistas/index.php/rce/article/download/1184/623+&cd=1&hl=pt-BR&ct=clnk&gl=br>. Acesso em: 10 Jan. 2015.

PIROLA, E. N. **Economia Urbano Regional:** uma resenha para contextualizar.

[S.l.: s.n.], 2012. Disponível em: <[http://www.pucrs.br/eventos/eeg/download/Mesa12/Economia\\_Urbano\\_Regional-Uma\\_resenha\\_para\\_Contextualizar\\_o\\_Debate.pdf](http://www.pucrs.br/eventos/eeg/download/Mesa12/Economia_Urbano_Regional-Uma_resenha_para_Contextualizar_o_Debate.pdf)> acesso em 22/06/2015.

OLIVEIRA, J. M. C de. **A Cultura do Fumo na Bahia**: refletindo sobre a Convenção-Quadro. Bahia Agrícola. Vol. 7, 2006. Disponível em: [http://www.seagri.ba.gov.br/sites/default/files/socioeconomia5\\_v7n2.pdf](http://www.seagri.ba.gov.br/sites/default/files/socioeconomia5_v7n2.pdf) acesso em 01/09/2015.

RIPPEL, R.; LIMA, J. F. **Polos de Crescimento Econômico**: notas sobre o caso do estado do Paraná. Santa Cruz do Sul: REDES, 2009.

SANTOS, D. C.; GOMES, O. P.; FEITOSA, A. M. B.; AMORIM, D. I. M. **Identificação de Aglomerações Produtivas no Estado do Ceará nos Anos 2006 e 2012 Via Índice de Concentração Normalizado (ICN)**. Crato-CE. 2014.

SILVA, P. C. G.; CORREIA, R. C. **Cultivo da Videira**. Embrapa Semi-Árido. 2004.

# ANÁLISE DO IMPACTO ECONÔMICO E SOCIAL DA ÁREA LIVRE DE MOSCA DAS FRUTAS NO ESTADO DO CEARÁ

Eliane Pinheiro de Sousa\*

Sílvia Helena Galvão de Miranda\*\*

## Resumo

O governo e os produtores de cucurbitáceas têm investido vultosos recursos para conservar a Área Livre de Mosca das Frutas (para *Anastrepha grandis*) no Ceará. Este trabalho busca avaliar a viabilidade de se manter essa ALP-MF, ou seja, se os benefícios econômicos e sociais estão compensando os custos incorridos por esses atores. Para tal, utilizou-se a técnica de análise benefício-custo, em que foram simulados cenários de manutenção da ALP-MF no estado do Ceará, considerando que se mantenha a parcela de comércio cearense de melão com os países quarentenários e de ampliação de suas vendas para os mercados internacionais, em um horizonte de quatro anos. Foram levantados dados de custos de produção e receitas com a comercialização de melão e melancia, referentes a 2014, por meio de entrevistas a produtores de cucurbitáceas. Adicionalmente, as despesas do governo estadual e outros dados sobre a manutenção da ALP-MF foram coletados em visitas e entrevistas a associações de produtores, a União do Agronegócio do Vale do Jaguaribe (UNIVALE) e Agência de Defesa Agropecuária do Estado do Ceará (ADAGRI). Os resultados simulados evidenciaram que os benefícios econômicos e sociais de manter a ALP-MF superam os gastos que o governo do Ceará e os produtores despendem com esta finalidade em todos os cenários analisados, sendo que o retorno é bem mais expressivo para o cenário que capta a ampliação do comércio cearense de melão para os Estados Unidos e a inserção no mercado do Japão.

**Palavras-Chave:** *Anastrepha grandis*, análise benefício-custo, defesa agropecuária, melão, melancia.

---

\* Pós-Doutora em Economia Aplicada pela Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo (ESALQ/USP). Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – Brasil. Doutora em Economia Aplicada pela Universidade Federal de Viçosa (UFV). Professora do Departamento de Economia da Universidade Regional do Cariri (URCA). E-mail: pinheiroeliane@hotmail.com.

\*\* Pós-Doutora na Pennsylvania State University. Doutora em Economia Aplicada pela Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo (ESALQ/USP). Professora do Departamento de Economia, Administração e Sociologia da ESALQ/USP. E-mail: shgdmira@usp.br.

## Abstract

The government and cucurbit producers have invested considerable resources to keep the Fruit-Fly Free Area for *Anastrepha grandis* in Ceará state. Therefore, this study aims to assess the feasibility of maintaining this Pest Free Area, i.e., if the economic and social benefits offset the costs incurred by these players. For that, the cost-benefit analysis technique has been applied, based on simulations for different scenarios involving the maintenance of the Pest Free Area in Ceará state, assuming that the market-share of melon Ceará exports to quarantined countries shall be kept, and alternatively - that its foreign sales will be extended -, along a timeline of four years. Data on production costs and melon and water melon sales have been collected, to 2014, by interviews to farmers. In addition, state government expenses and other information about the Pest Free Area maintenance have been obtained with farmers association, UNIVALE and ADAGRI. The simulation results showed that the economic and social benefits of maintaining the Pest Free Area outweigh the expenses faced by Ceará's government and by melon producers in all the scenarios. The economic return is much more expressive for the scenario which considers the expansion of Ceará's exports to the United States and the insertion in the Japanese market.

**Key words:** *Anastrepha grandis*, cost-benefit analysis, agriculture defense, melon, water melon.

## 1 INTRODUÇÃO

O Brasil é um dos principais países exportadores de frutas (tropicais e temperadas), mas ainda apresenta grande dificuldade em conquistar novos mercados, principalmente em função de questões fitossanitárias e de forma mais ampla para o atendimento dos inúmeros protocolos de qualidade e certificações requeridos pelos diversos países e, mais diretamente, empresas importadoras. Um dos maiores desafios fitossanitários é a presença de diversas espécies de moscas das frutas de importância quarentenária em seu território, fazendo com que os países importadores estabeleçam restrições fitossanitárias, que dificultam, encarecem e até impedem a entrada de frutas frescas. Dentre as moscas das frutas, destaca-se a mosca-sul-americana-das-cucurbitáceas, a *Anastrepha grandis*, que ataca as cucurbitáceas, entre as quais o melão e a melancia (AZEVEDO *et al.*, 2005).

Para viabilizar a exportação dessas frutas aos países quarentenários para essa praga<sup>1</sup>, os governos dos estados do Rio Grande do Norte e do Ceará, que se destacam como os maiores produtores e exportadores brasileiros de melão e de melancia, por meio dos seus órgãos estaduais de defesa vegetal, e sob a coordenação do Ministério de Agricultura, Pecuária e Abastecimento (MAPA), implementaram uma Área Livre de Mosca de Frutas (ALP-MF) para a *Anastrepha grandis*, respectivamente, a partir de 1985 e 1999.

Do ponto de vista de interesse econômico do Estado, é evidente o papel da produção e da exportação de melão e melancia para o Ceará. A contribuição desse setor não só ocorre pela ocupação de uma área agrícola que gera empregos e renda na região do Jaguaribe, mas também pela renda específica advinda das exportações. O Estado do Ceará foi, em 2014, o terceiro maior exportador de frutas do Brasil, ficando atrás somente da Bahia e de Pernambuco; o maior para o melão, sendo responsável por 59,75% e 56,41%, respectivamente, do valor e do volume exportados pelo Brasil; e o segundo maior para a melancia, sendo responsável por 45,10% e 38,71%, respectivamente, do valor e do volume exportados pelo Brasil.

Segundo Azevedo *et al.* (2005), é relevante manter uma área livre de pragas, visto que propicia benefícios socioeconômicos, como a conquista de novos mercados externos, a qualidade da produção de frutos, a geração de empregos diretos e indiretos e de divisas para o país. Ademais, a manutenção de uma área livre gera benefícios em termos ambientais, diminuindo os custos de produção, já que se reduz o uso de inseticidas.

Esses benefícios são verificados na ALP-MF no estado do Ceará, pois, conforme dados do Diário do Nordeste (2013), essa medida adotada pela Agência de Defesa Agropecuária do Estado do Ceará (ADAGRI) tem colaborado para a elevada qualidade da fruta cearense, possibilitando sua venda para os Estados Unidos. Graças à manutenção dessa área, a partir de 2014, conquistou-se também o mercado do Chile, que é um dos mais exigentes quanto ao aspecto fitossanitário. Dados do World Integrated Trade Solution (WITS, 2015) atestam que 96,4% e 24,31%, respectivamente, dos valores importados de melão e melancia, pelo Chile, em 2014, foram provenientes dessa ALP-MF no Brasil.

Em contrapartida, o estado do Ceará investiu mais de R\$ 1 milhão em

---

1 . A praga quarentenária refere-se a todo organismo de natureza animal e, ou vegetal, que, mesmo sob controle permanente, constitui ameaça à economia agrícola de um dado país ou região devido encontrar-se presente em outros países ou regiões. Esses organismos são, em geral, exóticos para esse país ou região e podem ser transportados de um local para outro pelo homem e seus meios de transporte por meio do trânsito de plantas, animais ou frutos e sementes contaminadas (BARBOSA; PARANHOS, 2007).

pesquisas laboratoriais, monitoramento, barreiras fitossanitárias e capacitação de técnicos agrícolas para impedir a entrada dessa praga (SOUZA, 2006).

Nesse sentido, em face do alto custo incorrido pelo estado do Ceará para manter a estrutura da ALP-MF e diante da restrição orçamentária que enfrenta os governos, estudos que tratem sobre essa temática revestem-se de importância. Em outros termos, é fundamental conhecer o custo atual de manutenção da ALP-MF para o governo do Ceará e para os produtores e o dano econômico evitado pelo controle da mosca das cucurbitáceas sul-americana, possibilitando uma análise da viabilidade econômica de manutenção dessa área. Além disso, também é relevante identificar seu retorno social, sob a perspectiva da alocação de recursos públicos. A esse respeito, conforme Miranda (2014), a quantificação de impactos ajuda a amparar as escolhas em termos de alocação de recursos.

Para Miranda *et al.* (2010), questões dessa natureza são relevantes, já que possibilitam antecipar os impactos, e, particularmente, as perdas de uma eventual entrada e disseminação de pragas ainda não presentes no Brasil, contribuindo com informações que podem ser utilizadas para o planejamento da defesa agropecuária. Ademais, podem subsidiar também as ações dos setores privados potencialmente afetados pelas pragas, como o de agrotóxicos, o de pesquisa agropecuária genética, e os próprios agentes produtivos e industriais. Vale destacar que a entrada de pragas exóticas pode afetar também direta e indiretamente o governo por meio da arrecadação tributária do setor envolvido, assim como requer maior alocação de verbas a ser destinadas às medidas de fiscalização e controle de infraestrutura laboratorial (MIRANDA, XIMENES, 2009).

Aliado a esses argumentos, é relevante destacar que recentemente foi lançado, no Brasil, o Programa Nacional de Mosca das Frutas pelo Ministério de Agricultura, Pecuária e Abastecimento. Esse programa faz parte do Plano de Defesa Agropecuária (PDA) e abarca as políticas já existentes, como a da Área Livre de *Anastrepha grandis*. A partir desse programa, espera-se que se consiga reduzir o impacto ocasionado pelas moscas das frutas, propiciando maior competitividade, aumento do número de postos de trabalho, acesso a novos mercados e maior *market share*, gerando, como consequência, benefícios para toda a cadeia produtiva (MAPA, 2015a).

Diante dessas considerações, este trabalho busca avaliar os benefícios econômicos e sociais de manter a Área Livre de *Anastrepha grandis* no Ceará. Especificamente, pretende-se determinar o dano econômico evitado pelo controle dessa mosca das cucurbitáceas sul-americana no Ceará; mensurar os custos de manutenção da ALP-MF da *Anastrepha grandis* para o governo do



Ceará e os custos de prevenção por parte dos produtores de melão e melancia; e identificar os impactos sociais ocasionados pela adoção dessa medida de defesa fitossanitária.

Estudos que avaliam os benefícios e os custos de monitoramento de pragas e doenças têm sido largamente debatidos na literatura econômica. No entanto, não se encontraram estudos científicos econômicos sobre a temática da Área Livre de Mosca de Frutas (ALP-MF) para a *Anastrepha grandis* no Ceará. Além de se tratar de um estudo pioneiro para este Estado, inova ao se procurar estimar os benefícios e os custos sociais e não centrar somente nos impactos econômicos dessa política. Portanto, este trabalho contribui neste sentido, buscando discutir o papel do Estado e do setor privado nas ações para controle dessa praga e em seus impactos macro e microeconômicos, buscando identificar não só aqueles aspectos positivos, mas também o seu ônus.

O artigo está organizado em quatro seções, além desta introdução. Na segunda, é realizada uma revisão bibliográfica sobre a Área Livre de Mosca de Frutas (ALP-MF) para *Anastrepha grandis* no Ceará, destacando a evolução dos postos de trabalho e das exportações de melão e melancia durante o período de funcionamento da ALP-MF no Ceará. Os referenciais teóricos e analíticos sobre Análise Benefício-Custo (ABC), os cenários e pressuposições adotados e as fontes dos dados são apresentados na terceira seção. Em seguida, são mostrados e discutidos os resultados relativos às perdas evitadas nas exportações e na massa salarial, aos custos com o programa de defesa agropecuária e à análise benefício-custo de manutenção da ALP-MF no Ceará. A última seção destina-se às considerações finais.

## **2 A IMPORTÂNCIA ECONÔMICA E SOCIAL DAS CUCURBITÁCEAS E DA ALP-MF PARA O CEARÁ**

Ao se analisar os programas de defesa fitossanitária, uma das grandes preocupações consideradas consiste no impacto sobre o comércio internacional, uma vez que a entrada e a disseminação de uma praga podem gerar redução das exportações ou proibição de comércio (MIRANDA, 2013). No caso da ocorrência de *Anastrepha grandis*, há proibição de comércio para os países quarentenários. Para viabilizar as exportações de cucurbitáceas (melão e melancia), do ponto de vista de conformidade fitossanitária, para tais países, tornou-se necessário implementar a Área Livre de Pragas (ALP).

A Área Livre compreende uma área geográfica, constituída por estados, municípios ou partes destes, em que se comprova cientificamente a ausência de uma praga específica. Para isso, precisa ser constantemente monitorada com uma rede de armadilhas com o intuito de identificar qualquer espécie da praga (AZEVEDO *et al.*, 2005). **O reconhecimento de área livre** é previsto no Acordo sobre Medidas Sanitárias e Fitossanitárias (SPS) da Organização Mundial do Comércio (OMC), que em seu artigo 6º, prevê a delimitação de áreas livres de pragas ou doenças e áreas de baixa prevalência de pragas ou doenças (WTO, 2015).

De acordo com o Ministério de Agricultura, Pecuária e Abastecimento (MAPA, 2015b), o interesse em se implantar essa área ocorreu a partir de restrições impostas pelos Estados Unidos à exportação brasileira de melões, decorrentes de diversas publicações que apontavam a presença dessa mosca no Brasil. Como a ocorrência da praga não foi confirmada por meio do acompanhamento no campo e corte de frutos, durante o período de 1985 a 1993, o Departamento de Agricultura dos Estados Unidos liberou as exportações brasileiras de cucurbitáceas produzidas na Área Livre de *Anastrepha grandis* no Rio Grande do Norte.

A partir de 1999, os municípios cearenses de Aracati, Icapuí, Itaiçaba, Jaguaruana, Russas, Quixeré e Limoeiro do Norte, localizados na região do Baixo Jaguaribe, vizinho aos municípios de Mossoró e Baraúna, no Rio Grande do Norte, passaram também a fazer parte da Área Livre. Esses municípios possuem uma área de cerca de 5.676 km<sup>2</sup> (MAPA, 2015b). Conforme dados do Instituto Brasileiro de Geografia e Estatística (IBGE, 2015), 85,08% e 32,47%, respectivamente, das quantidades produzidas de melão e melancia, no Ceará, em 1999, foram provenientes de tais municípios, sendo que, em 2014, praticamente toda produção de melão (99,96%) e 89,16% da produção de melancia se originaram dessa área, demonstrando a sua importância para o Estado. Este estudo foca na Área Livre de *Anastrepha grandis* no Ceará.

A relevância dos municípios que integram a ALP-MF no Ceará também é confirmada pela predominância dos valores da produção de melão e melancia obtidos por tais municípios. Com base nos dados do IBGE (2015), verifica-se que, dos 170.406 mil reais resultantes da produção de melão cearense, 170.363 mil reais foram provenientes dos municípios dessa Área Livre em 2014, perfazendo 99,97%. No tocante ao valor da produção de melancia, a participação resultante da ALP-MF foi de 92,06%, ou seja, dos 47.543 mil reais gerados pela produção de melancia no Ceará, 43.767 mil reais foram obtidos na ALP-

-MF. De posse dessa base de dados, também se pode inferir que a produção dessas cucurbitáceas foi responsável por 84,10% do valor total das lavouras temporárias gerado nessa ALP-MF. Em termos comparativos com o valor total das lavouras temporárias obtido pelo Ceará, em 2014, tem-se que 20,18% se originaram dos municípios que fazem parte da ALP-MF.

Segundo especialistas na *Anastrepha grandis*, essa praga tem pouco efeito sobre perda de produção de melão e melancia. Por essa razão, está se considerando que os impactos relevantes a serem avaliados nestes cenários é apenas o que é visto pelo setor privado e público como justificativa para manter a Área Livre.

Conforme descrito, a entrada e a disseminação de uma praga podem gerar impactos micro e macroeconômicos. Sob a ótica microeconômica, estes eventos podem refletir diretamente nos custos de produção e até gerar alterações na estrutura produtiva da propriedade e da região comprometidas. Quanto ao aspecto macroeconômico, as pragas exógenas podem impedir o comércio com mercados internacionais, prejudicando a geração de divisas (MIRANDA, ADAMI, 2015).

Além dos impactos econômicos, também são perceptíveis impactos no âmbito social, como perda de postos de emprego, não apenas os empregos diretos no campo, mas os empregos indiretos envolvidos em atividades como *packing house*, serviços de despacho e fiscalização, dentre outras. Todas as atividades que fazem parte da cadeia de produção são influenciadas, gerando uma movimentação da economia regional.

Em face da relevância do impacto sobre empregos, o Quadro 1 mostra, respectivamente, os postos de trabalho da categoria “Cultivo de outros produtos de lavoura temporária”, conforme a Classificação Nacional de Atividades Econômicas – CNAE 95 disponibilizados na base de dados da Relação Anual de Informações Sociais (RAIS) **do Ministério do Trabalho e Emprego (MTE, 2015)** para os municípios cearenses que fazem parte da ALP-MF. A escolha dessa categoria pode ser atribuída ao fato de 95,3% dos empregos formais dessa categoria ter sido alocada no cultivo do melão, em 2014.

**QUADRO 1 – Número de vínculos totais na categoria “Cultivo de outros produtos de lavoura temporária” nos municípios cearenses que fazem parte da ALP-MF, 1999-2014.**

Ano*	Aracati	Icapuí	Itaiçaba	Jaguarunana	Russas	Quixeré	Limoeiro do Norte	ALP-MF CE	CE	ALP-MF/CE (%)
1999	26	-	-	-	-	-	10	36	91	39,56
2000	19	-	-	355	-	236	9	619	819	75,58
2001	18	-	154	218	-	522	174	1.086	1.224	88,73
2005	52	68	-	27	138	106	-	391	3.925	11,87
2006	60	2.415	622	41	498	1.886	1	5.523	13.265	41,64
2007	622	2.518	331	62	1.252	1.204	-	5.989	15.022	39,87
2008	1.519	1.309	-	36	2.388	5.846	2.146	13.244	14.890	88,95
2009	1.562	1.792	-	2	1.390	5.113	2.103	11.962	13.170	90,83
2010	1.705	3.105	-	3	687	3.049	1.320	9.869	10.761	91,71
2011	1.960	2.801	-	12	917	1.872	280	7.842	8.654	90,62
2012	2.178	3.040	-	8	1.105	2.466	355	9.152	9.667	94,67
2013	2.300	3.845	-	11	1.610	2.665	407	10.838	11.273	96,14
2014	2.084	4.298	-	14	2.516	2.054	311	11.277	11.722	96,20
Média	1.085	1.937,77	85,15	60,69	961,62	2.078,38	547,38	6.756	8.757,92	72,80

Fonte: Elaborado pelas autoras com base nos dados da RAIS do Ministério do Trabalho e Emprego (MTE, 2015).

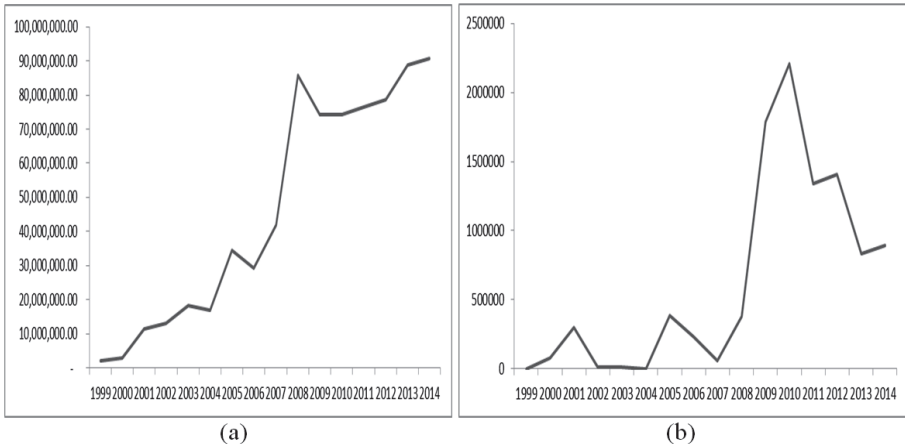
Nota: \* Não foram incluídos os anos de 2002, 2003 e 2004 devido à indisponibilidade de dados desagregados por municípios.

Como se observa, há uma grande mobilidade de mão-de-obra dessa atividade entre os municípios da ALP-MF e de um ano para outro. Verifica-se também que, embora o número de empregos formais nessa atividade oscile em alguns anos tanto na ALP-MF quanto no estado do Ceará, houve um acréscimo expressivo do número de empregos formais nos cultivos agrícolas durante o período de funcionamento da Área Livre de Mosca das Frutas, uma vez que os municípios da ALP-MF registraram uma média de crescimento anual de 255,4%, ao passo que o número de empregos formais nessa atividade no Ceará cresceu apenas 109,7% ao ano no decorrer do período analisado. Os dados indicam, ainda, que mais de 90% dos empregos gerados nessa atividade no estado do Ceará, a partir de 2009, foram provenientes dos municípios que fazem parte da ALP-MF, chegando a 96,2% em 2014, enquanto, em 1999, essa participação era inferior a 40%.

Dentre as frutas frescas comercializadas pelo Brasil no mercado externo, segundo o Ministério do Desenvolvimento, Indústria e Comércio Exterior (MDIC, 2015), o melão liderou, em termos de volume exportado, e destacou-se como a segunda fruta em termos de receita cambial, perfazendo um montante de US\$ 151,8 milhões. Assim, o melão respondeu por 25% e 33%, respectivamente, do valor e do volume exportados de frutas frescas brasileiras, em 2014. Desse valor exportado de melão pelo Brasil, US\$ 90,7 milhões foram provenientes do Ceará. Neste estado, o melão liderou o *ranking* tanto em quantidade quanto em receita exportada, sendo responsável por 87% e 85%, respectivamente, do valor e do volume exportados de frutas frescas cearenses.

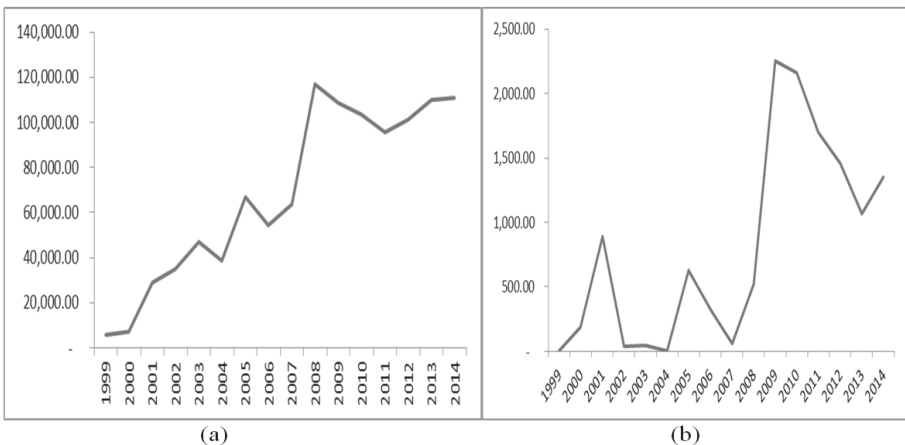
Tendo em vista que o principal impacto direto da Área Livre de *Anastrepha grandis* no Ceará ocorre no comércio externo, é relevante mostrar a evolução das exportações de melão durante esse período de funcionamento da ALP-MF no estado. As Figuras 1 e 2 ilustram, respectivamente, o valor e o volume exportado de melão. Conforme se observa, essas variáveis apresentam o mesmo comportamento ao longo do tempo tanto pela ótica das exportações totais quanto pelos movimentos oscilatórios evidenciados nas exportações para os países quarentenários. Apesar dessas oscilações, é perceptível que há uma tendência crescente em termos comparativos com o início da série temporal, quando foi implantada a ALP-MF.

**FIGURA 1 – Valor total exportado de melão pelo Ceará (a) e valor exportado de melão pelo Ceará aos países quarentenários (b), expressos em dólar, 1999-2014**



Fonte: Elaborado pelas autoras com base nos dados pesquisados no MDIC (2015).

**FIGURA 2 – Volume total exportado de melão pelo Ceará (a) e volume exportado de melão pelo Ceará aos países quarentenários (b), expressos em toneladas, 1999-2014.**



Fonte: Elaborado pelas autoras com base nos dados pesquisados no MDIC (2015).

Como se verifica nessa seção, os principais impactos da Área Livre de *Anastrepha grandis* no Ceará ocorre nos empregos e no comércio externo. Para mensurar esses impactos, utilizou-se a Análise Benefício-Custo (ABC), que será

apresentada na próxima seção. Conforme Miranda *et al.* (2010), a ABC é uma ferramenta que pode auxiliar o governo ou a iniciativa privada na tomada de decisão quanto à incidência e disseminação de uma praga exótica, por meio do estabelecimento de um programa de prevenção e controle. Essa técnica tem sido amplamente adotada em estudos internacionais e nacionais para avaliar políticas de defesa agropecuária. Mais recentemente, na literatura internacional, citam-se, os estudos de Amoabeng *et al.* (2014) e Enkerlin *et al.* (2015), este último também relacionado a moscas de frutas e, no Brasil, os estudos de Sanches *et al.* (2014), Adami e Miranda (2014) e Miranda *et al.* (2015), estes três últimos para frutas. Apesar da simplicidade matemática desse método, a dificuldade de sua aplicação, geralmente decorre da carência de dados sistematizados que possibilitem análises mais detalhadas e abrangentes (MIRANDA, ADAMI, 2015).

### 3 METODOLOGIA

#### 3.1 Referencial teórico

A análise benefício-custo (ABC), em seu sentido mais amplo, é o processo de classificação das opções políticas a partir de um ponto de vista econômico, tendo em conta tanto os benefícios da política quanto seus custos. Em termos econômicos, concentram-se nos benefícios líquidos da política (BOADWAY, 2006).

Em conformidade com SafetyNet (2009), a ABC apresenta quatro princípios fundamentais: soberania do consumidor; avaliação de bens de acordo com a disposição a pagar; “ótimo de Pareto” como critério de maximização do bem-estar; e neutralidade quanto à distribuição dos custos e dos benefícios. A soberania do consumidor é o princípio de que as escolhas feitas pelos consumidores em relação a como gastar os seus rendimentos são aceitos e são tratados como dados. No tocante ao segundo princípio, a disposição a pagar é considerada como uma medida dos benefícios utilizados na ABC. O “ótimo de Pareto” consiste no terceiro princípio da ABC e ocorre quando se melhora o bem-estar de pelo menos uma pessoa, sem reduzir o bem-estar de outra. Na prática, dificilmente uma medida tomada pelo governo terá apenas ganhadores sem perdedores. Assim, o critério geralmente aplicado em ABC é “Pareto potencial”. Este critério é satisfeito quando aqueles que ganham a partir de uma medida podem compensar os que perdem com isso, mantendo um benefício líquido. O quarto

princípio é que permanece neutro quanto à distribuição de custos e benefícios entre os grupos da população, desde que, em termos agregados, os benefícios excedam os custos. A ABC não se destina a encontrar a solução mais justa para um problema social, apenas a solução mais eficiente.

Conforme Zerbe e Bellas (2006), a operacionalização de uma ABC requer, em síntese, o cumprimento das seguintes etapas: esclarecer os custos e benefícios que serão incluídos, identificar as alternativas, definir hipóteses, listar os impactos de cada projeto alternativo, valorar tais impactos, lidar com os impactos não quantificados, determinar uma taxa de desconto para se obter o valor presente de todos os benefícios e custos futuros, identificar e explicar a incerteza, comparar os benefícios e os custos, e realizar uma análise de monitoramento após a conclusão do projeto. Nem todas as análises exigirão todas essas etapas. Um projeto de curta duração pode desconsiderar o desconto dos benefícios futuros e um projeto que vem sendo conduzido pode envolver pouco risco ou incerteza.

De acordo com Sanches *et al.* (2014) e Miranda *et al.* (2015), a ABC é indicada especialmente com o intuito de observar os fluxos financeiros da ação ou projeto e avaliar seus impactos econômicos. Esse instrumental é amplamente adotado para auxiliar na formulação de políticas públicas, já que identifica, quantifica e compara os gastos e os benefícios de alternativas de políticas, permitindo subsidiar a tomada de decisão. Nesse sentido, este método tem sido utilizado por países-membros da Organização para Cooperação e Desenvolvimento Econômico (OCDE) para análise de impacto regulatório, no escopo da implementação do instrumento de qualidade regulatória, a RIA – Regulatory Impact Analysis.

### 3.2 Referencial analítico

Conforme descrito, empregou-se o método analítico de ABC, buscando comparar os benefícios e os custos de se manter a Área Livre de Mosca de Frutas (ALP-MF) da *Anastrepha grandis* no estado do Ceará. Para tal, seguindo as etapas propostas por Zerbe e Bellas (2006) e os procedimentos adotados nos estudos desenvolvidos por Miranda *et al.* (2010, 2015), que aplicaram essa ferramenta analítica para as políticas de defesa agropecuária, a princípio, identificaram-se os benefícios valorados como perdas evitadas com a manutenção da ALP-MF e os custos envolvidos, tanto os incidentes sobre o governo quanto setor privado, na ação avaliada.



Feito isto, determinou-se o valor presente dos benefícios e dos custos projetados, conforme a expressão (1), para o período de tempo analisado. Neste estudo, os benefícios e os custos foram projetados para um horizonte temporal de quatro anos, referente a um mandato político, considerando valores de 2014 e tendo em vista que uma das motivações do estudo é evidenciar se há um retorno socioeconômico desta política de manutenção da ALP-MF.

$$VP = V / (1+i)^j \quad (1)$$

Na expressão (1),  $VP$  refere-se ao valor presente do montante  $V$ , que pode indicar os benefícios ( $B_j$ ) da manutenção da ALP-MF no Ceará ou os custos ( $C_j$ ) incorridos pelo governo do Ceará e pelo setor privado com essa manutenção;  $i$  corresponde à taxa de desconto; e  $j$ , ao horizonte temporal de projeção dos fluxos, medido em anos e definido neste trabalho para  $j = 0, 1, 2, 3$ . Se  $VP$  for positivo, conforme Boadway (2006), o projeto é socialmente rentável. Vale destacar que não está sendo considerado o valor inicial de investimento, visto que a área já está instalada há vários anos e o caso aqui é verificar apenas sua manutenção.

A taxa de desconto adotada foi a taxa de juros de longo prazo (TJLP), de dezembro de 2014. De acordo com Miranda *et al.* (2015), a TJLP é uma taxa de desconto em que o aspecto social pode ser mais bem evidenciado, já que é a referência para os financiamentos de investimentos de longo prazo do BNDES.

Os benefícios foram estimados pelas perdas evitadas, medidas pelas exportações de melão e melancia para os países quarentenários da *Anastrepha grandis* e pela massa salarial decorrente das atividades no mercado de trabalho que atendem tais exportações. Quanto aos custos, compreendem os gastos do governo do Ceará com a ALP-MF e as despesas diretas dos produtores de cucurbitáceas com essa área, que incluem o monitoramento em suas fazendas produtoras dessas cucurbitáceas, custo este que não é aplicável nas áreas de cucurbitáceas fora da ALP-MF.

Em seguida, projetaram-se os benefícios e os custos do programa para um horizonte de quatro anos e tais montantes foram trazidos a valores presentes. Somando os valores presentes dos benefícios e dos custos relativos à manutenção da ALP-MF, pode-se determinar a relação benefício-custo por meio da equação (2):

$$B/C = \sum_{j=0}^n B_j / (1+i)^j \Bigg/ \sum_{j=0}^n C_j / (1+i)^j \quad (2)$$

Na equação (2),  $B_j$  refere-se aos benefícios da manutenção da ALP-MF da *Anastrepha grandis* no Ceará no ano  $j$ ;  $C_j$  diz respeito aos custos

incorridos pelo governo do Ceará e pelo setor privado com a manutenção da ALP-MF no ano  $j$ ;  $i$ , taxa de desconto (TJLP); e  $n$ , vida útil da ação adotada, considerada, neste trabalho, como quatro anos.

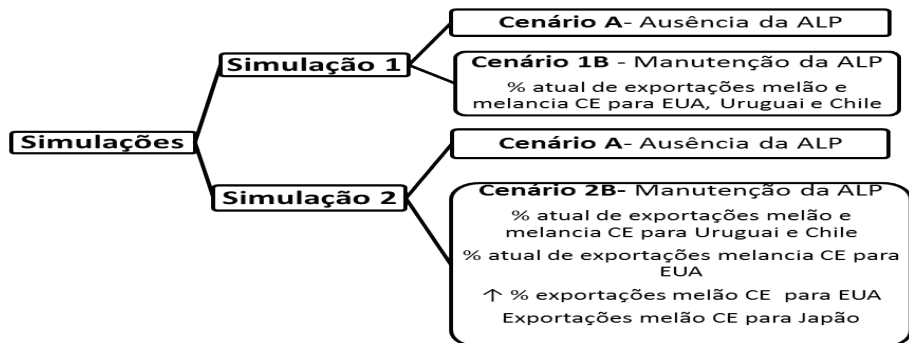
O programa deve ser mantido se os benefícios excederem os custos, ou seja, se a relação  $B/C$  for maior que a unidade (CONTADOR, 2014). Segundo Miranda *et al.* (2015), essa relação é um indicativo do retorno (benefício) para cada unidade monetária investida ou gasta (custo) no programa em análise.

### 3.3 Cenários e pressuposições adotados

Consideraram-se dois cenários, sendo que o cenário A considera a inexistência da ALP-MF da *Anastrepha grandis* no Ceará, que se refere ao cenário base para a identificação das perdas máximas, e o cenário B refere-se à manutenção desse programa nos moldes atuais. Os benefícios obtidos no cenário B são determinados pela diferença entre as perdas simuladas na ausência da ALP-MF (cenário A) e as perdas evitadas com o programa de defesa (cenário B).

Neste estudo, foram realizadas duas simulações, conforme ilustradas na Figura 3, sendo que a primeira (cenário 1B) buscou verificar se esse programa deve ser mantido nos próximos quatro anos, supondo que o Ceará continue exportando com a mesma participação de melão e melancia comercializada atualmente para os países quarentenários (Estados Unidos, Uruguai e Chile).

FIGURA 3 – Ilustração dos cenários simulados para análise da política de manutenção da Área Livre de Mosca de Frutas (ALP-MF) da *Anastrepha grandis*, CE.



Fonte: Elaborado pelas autoras.

A segunda simulação (cenário 2B) buscou investigar se esse programa deve ser mantido nos próximos quatro anos, admitindo manter as participações exportadas de melão para Uruguai e Chile, porém com aumento na participação das exportações cearenses de melão para os Estados Unidos e com a introdução de melão cearense também no mercado do Japão, grande importador dessa cucurbitácea. A construção desse cenário de simulação está fundamentada na pesquisa de campo realizada diretamente com os produtores de melão, que destacaram o interesse em ampliar as exportações dessa fruta para o mercado americano, como também conquistar o mercado do Japão.

Com base nos dados do World Integrated Trade Solution (WITS, 2015), o Brasil é o quarto maior exportador mundial de melão e responde por 11% do valor e do volume mundial exportado de melão. Nesta simulação, supõe-se que as exportações brasileiras de melão para os Estados Unidos alcancem essa participação, ou seja, que, do valor total importado de melão pelos Estados Unidos, 11% sejam atendidos pelo Brasil. Ademais, replica-se essa participação na simulação de importação de melão brasileiro pelo Japão. Vale ressaltar que se aplicaram tais percentuais considerando a participação relativa das exportações cearenses de melão. Conforme dados do MDIC (2015), 59,75% das exportações brasileiras de melão foram provenientes do estado do Ceará, em 2014, mas, do total de exportações cearenses, apenas 0,66% tiveram como destino os EUA.

No tocante às pressuposições adotadas neste estudo, é importante ressaltar que os impactos diretos na exportação foram computados para melão e melancia, que são as principais hospedeiras da *Anastrepha grandis* comercializadas para os países quarentenários. Embora a abóbora seja uma hospedeira preferencial desta praga, ela não foi considerada, visto que o estado do Ceará não exporta essa cucurbitácea. Pressupõe-se que se houver o abandono do programa da ALP-MF, apenas os países quarentenários para a mosca das cucurbitáceas reagirá com a interrupção de importações de melão e melancia vindas da região.

Apesar de especialistas na área terem apontado os parâmetros de perdas ocasionados pela *Anastrepha grandis* na produção de abóbora, não se incluiu na ABC, visto que a produção dessa cucurbitácea no Ceará ser inexpressiva, conforme dados do Censo Agropecuário 2006, o mais recente disponível para esta cultura. Segundo os especialistas, não há perda resultante do ataque dessa praga na produção de melão e melancia.

Em relação às perdas sociais, tendo em vista a dificuldade de obter dados e realizar simulações de realocação de culturas e outros impactos sociais ao lon-

go da cadeia produtiva, pressupõe-se que estas se limitam somente à perda de postos de emprego formais nas áreas destinadas à exportação. Esta perda, por sua vez, será estimada pelo valor correspondente de massa salarial perdida em decorrência da redução de postos de trabalho que estariam vinculados às áreas necessárias para exportação aos países quarentenários.

Por fim, admite-se que os custos foram considerados constantes ao longo do período analisado. Estes custos incluem gastos com os Postos de Vigilância Zoofitossanitária, com o monitoramento e com a capacitação de técnicos e confecção de material educativo sobre a ALP-MF. Existem cinco Postos de Vigilância Zoofitossanitária no Ceará, localizados nos municípios de Chorozinho, Limoeiro do Norte, São João do Jaguaribe, Fortim e Aracati e a manutenção de cada um deles requer recursos para cobrir os salários dos funcionários, as diárias para servidores realizarem atividades de Defesa Agropecuária junto a tais postos, além das despesas mensais com a infraestrutura, como telefone, internet, energia elétrica e água. Não foram levadas em consideração as despesas com a depreciação das instalações.

O monitoramento é realizado semanalmente em locais de risco, como pontos de fronteira, rodoviárias, centrais de recebimento e distribuição de frutas e hortaliças, mercados, entre outros. Para tal, são efetuadas despesas com armadilhas, atrativos, assim como diárias dos fiscais estaduais e agentes estaduais agropecuários e combustíveis. Também, neste caso, não foram incluídos os custos com a depreciação dos veículos para realizar a fiscalização.

Os agricultores se responsabilizam pelo monitoramento em suas propriedades produtoras de melão e melancia. Para isso, pagam uma taxa por hectare monitorado à União do Agronegócio do Vale do Jaguaribe (UNIVALE), que se incumbem de buscar o material coletado nas armadilhas dispostas nas propriedades que realizam o monitoramento e levar para análise laboratorial. Ademais, tais produtores precisam realocar mão-de-obra para cuidar da documentação do monitoramento (como em todo processo de certificação, um dos componentes mais importantes é o registro de documentação que permite a rastreabilidade dos processos), assim como a limpeza das armadilhas e a coleta do material contido em cada armadilha.

### **3.4 Fontes dos dados**

Neste estudo, foram utilizados dados de número de vínculos totais e remuneração média nominal, extraídos da Relação Anual de Informações Sociais (RAIS) do Ministério do Trabalho e Emprego (MTE, 2015) referentes ao ano

de 2014, para os municípios incluídos na ALP-MF. Também foram coletados dados da área plantada de melão do Instituto Brasileiro de Geografia e Estatística (IBGE, 2015) para o último ano disponível.

Para se valorar os benefícios (perdas evitadas), utilizaram-se também dados de exportações, obtidos da Secretaria de Comércio Exterior do Ministério do Desenvolvimento, Indústria e Comércio Exterior (MDIC, 2015). As estatísticas de comércio para o estado do Ceará e para o Brasil, para melão e melancia, foram obtidas para o período de 2012 a 2014 e as médias de valor em dólar (US\$) foram convertidas em reais pela taxa de câmbio de compra disponibilizada pelo IPEADATA, considerando a média dos três anos. Empregaram-se dados do World Integrated Trade Solution (WITS, 2015) para se estimar o percentual do valor e do volume mundial exportado de melão proveniente do Brasil, que foi utilizado na simulação do cenário 2B.

Quanto aos custos, os dados dos gastos do governo do Ceará foram obtidos diretamente com a Agência de Defesa Agropecuária do Estado do Ceará (ADAGRI) para 2014. Já os custos de produção dos produtores de cucurbitáceas foram levantados em pesquisa de campo, realizada em fevereiro e março de 2015 junto a uma amostra representativa de produtores. Também foi consultada a União do Agronegócio do Vale do Jaguaribe (UNIVALE), que informou a taxa paga pelos produtores por hectare monitorado. Esta taxa integra parte dos custos que os produtores certificados da ALP-MF têm para terem seu produto considerado livre da *Anastrepha grandis*.

#### 4 RESULTADOS E DISCUSSÃO

Existem diversos impactos econômicos, sociais e ambientais que podem estar relacionados à entrada e disseminação de pragas exóticas em um país. Segundo Miranda e Adami (2015), os principais impactos diretos resultantes dessas pragas referem-se à redução na produtividade; elevação nos custos de manejo para seu controle e comprometimento da sustentabilidade das culturas nas regiões afetadas; e comercialização, visto que pode ocasionar restrições às exportações de um dado produto.

Em termos de impactos indiretos, essas autoras, ao citarem Miranda e Ximenes (2009), destacam a concentração na atividade agrícola, visto que o aumento dos custos da cultura dificulta a permanência dos pequenos produtores agrícolas e os postos de trabalho perdidos. De acordo com essas autoras, as

pragas exóticas também podem ocasionar impactos diretos e indiretos para o governo, seja pela redução da arrecadação tributária do setor e região atingidos, como pela necessidade de maior alocação de gastos em defesa agropecuária. No que diz respeito aos impactos ambientais, tais autoras mencionam o aumento da degradação, em virtude do maior uso de defensivos agrícolas e do risco de interação com outras espécies vegetais ou animais. Esses impactos são evidenciados no caso da ocorrência de *Anastrepha grandis*, conforme a pesquisa de campo realizada com os produtores agrícolas.

Embora tanto a literatura quanto a pesquisa de campo realizada com os produtores de cucurbitáceas terem apontado vários impactos diretos e indiretos ocasionados, em geral, pelas pragas exóticas e, em particular, pela *Anastrepha grandis*, este estudo focou no impacto sobre as exportações de melão e melancia, visto que na ausência do reconhecimento de área livre da praga, tais exportações devem ser interrompidas para os países quarentenários para essa praga, e no impacto sobre os empregos formais perdidos na área destinada à exportação. Como já foi dito, a presença dessa mosca nos cultivos de melão e melancia não registra, em literatura, grandes perdas em produtividade, razão pela qual este tipo de prejuízo não foi considerado nos cenários examinados.

#### 4.1 Perdas evitadas

O foco dessa seção é estimar as perdas potenciais resultantes de uma eventual ocorrência da *Anastrepha grandis* no Ceará. Tais perdas são uma aproximação para os principais prejuízos evitados com a manutenção da Área Livre de Mosca das Frutas. De acordo com os produtores de cucurbitáceas entrevistados, essa ALP-MF propicia benefícios em termos econômicos como o acesso a novos mercados, além do potencial de manter a janela aberta para expandir a parcela no comércio cearense para os países quarentenários.

Nas entrevistas, os produtores também apontaram ganhos sociais provenientes da ALP-MF, como os empregos diretos na produção de melão e melancia e os empregos indiretos na produção de outras culturas frutícolas, manter a reserva patrimonial do país, potencializar o Estado para receber novas oportunidades, assim como assegurar a idoneidade e qualidade do produto comercializado. Ademais, destacaram a preservação do meio ambiente como um benefício ambiental.

Dentre tais benefícios, um dos mais evidenciados consiste em permitir o comércio cearense de cucurbitáceas com Estados Unidos, Chile e Uruguai. A conquista recente do mercado chileno é um exemplo do benefício direto pro-

movido pelo reconhecimento dessa ALP-MF. Destacadamente, os produtores demonstraram o interesse em ampliar a parcela nas vendas de melão para os Estados Unidos e conquistar o mercado japonês.

A Tabela 1 mostra os países de destino das exportações cearenses de melão e melancia e seus respectivos valores importados do Ceará, considerando as médias calculadas para o período de 2012 a 2014.

O principal destino do melão e da melancia cearenses é a União Europeia, que não impõe medidas restritivas à *Anastrepha grandis*. No entanto, apesar dos países quarentenários para essa praga terem respondido, em média, por menos de 1% do valor exportado dessas frutas pelo Ceará, no triênio 2012 a 2014, a ALP-MF garantiu para esse Estado uma receita de, aproximadamente, R\$ 2,6 milhões por ano, ou evitou uma perda desse montante, em termos de exportações que não ocorreriam na ausência do seu reconhecimento.

**TABELA 1 – Exportações de melão e melancia no Ceará por país de destino e perdas evitadas devido à Área Livre de Mosca das Frutas no estado do Ceará**

Países de destino	Exportações de melão	Exportações de melancia
	(R\$)	(R\$)
	Média – 2012 a 2014	Média – 2012 a 2014
União Europeia	178.013.094,75	17.038.420,25
Estados Unidos	2.006.321,30	8,62
Chile	193.319,45	93.259,71
Uruguai	53.381,76	16.214,33
Exportações totais	185.309.652,80	17.406.049,53
Exportações sem os países quarentenários	183.056.630,29	17.296.566,87
<b>Perda evitada (benefício)</b>	<b>2.253.022,51</b>	<b>109.482,67</b>

Fonte: Elaborado pelas autoras com base nos dados do MDIC (2015).

Os valores foram convertidos utilizando a taxa de câmbio de compra, média para o triênio 2012 a 2014.

Além desse montante de exportações que são viabilizadas pela existência da ALP-MF (perdas evitadas por esta política fitossanitária) para os países quarentenários, vale ressaltar também as perdas indiretas evitadas com a mão-de-obra empregada nessa área de exportação, estimada pela massa salarial correspondente à mão-de-obra empregada na área necessária para abastecer estes países quarentenários.

Para se mensurar essa perda evitada na cultura do melão, determinou-se, a priori, a área destinada à exportação de melão para os países quarentenários, obtida pela razão entre o volume exportado médio no período de 2012 a 2014

e a produtividade do melão, que varia entre 25 a 32 toneladas por hectare, conforme os produtores entrevistados.

A partir da área calculada para abastecer a exportação de melão para os países quarentenários (26,68 hectares, considerando uma produtividade média de 25 toneladas por hectare) e a mão-de-obra do melão por hectare (1,2 pessoas), foi possível obter uma estimativa do número de pessoas empregadas na exportação para os países quarentenários (32,04 pessoas).

Em seguida, determinou-se a massa salarial utilizada em uma safra de melão. No cenário de ausência da ALP-MF, conforme descrito, as exportações cearenses de cucurbitáceas serão banidas dos países quarentenários. Pode-se supor que essa mão-de-obra desempregada será realocada para outras atividades na próxima safra, como, por exemplo, o cultivo de frutas de lavoura permanente, exceto laranja e uva. Tendo em vista que a remuneração média é menor que a obtida no cultivo do melão, encontrou-se que o estado do Ceará obteve uma perda evitada de massa salarial para o cultivo do melão de aproximadamente R\$ 31,2 mil por ano, admitindo que a produtividade do melão nas áreas de exportação tenha sido 25 toneladas por hectare, em uma simulação conservadora.

Além de se estimar os benefícios, expressos em termos de perdas evitadas, torna-se necessário também levantar os custos envolvidos com o programa de defesa agropecuária, ou seja, no âmbito deste trabalho, os custos de manutenção da Área Livre de Mosca das Frutas, buscando aplicar a ABC. Para isso, deve-se inicialmente identificar os atores que estão incorrendo em custos para manter a ALP-MF. Após uma visita *in loco* para compreender o funcionamento da área, verificou-se que o governo do estado do Ceará e os produtores são os principais atores que estão arcando com as despesas para manter o status dessa área livre.

Considerando a soma das despesas com os postos de vigilância, monitoramento semanal, capacitação de técnicos e materiais educativos, o governo do Ceará tem arcado com um custo no valor de R\$ R\$ 745,4 mil por ano, segundo dados para 2014. No tocante às despesas diretas incorridas pelos produtores, resulta em um dispêndio de R\$ 108,9 mil por ano para toda a área monitorada, que abrange 3.009 hectares.

#### **4.2. Análise Benefício-Custo da manutenção da Área Livre de Mosca das Frutas**

Os resultados da primeira simulação (cenário A x 1B) para um período



de quatro anos, comparando o cenário com e sem a ALP-MF, e no qual se considera que com a manutenção da área livre, os volumes de melão importados do Ceará pelos países quarentenários permaneçam na proporção em que ocorrem atualmente, encontram-se na Tabela 2.

**TABELA 2 – Valores presentes totais dos benefícios (perdas evitadas) e dos custos da Área Livre de Mosca das Frutas no Ceará, para o melão e a melancia. Horizonte de simulação de quatro anos. Período base 2014. Cenário A e 1B**

Benefícios (perdas evitadas) e custos (R\$)		Cenários	
		A – Ausência da ALP-MF	1B- Manutenção da ALP-MF
Considerando os custos do governo do Ceará e dos produtores	Benefícios (exportações e empregos)	8.945.541,64	-
	Custos do programa	-	3.181.163,74
	Benefícios (perdas evitadas A – perdas B)		8.945.541,64
	Custo (custo A – custo 1B)		3.181.163,74
	Saldo líquido (benefícios - custos)	5.764.377,91	
Relação B/C		2,81	
Considerando apenas os custos do governo do Ceará	Benefícios (exportações e empregos)	8.945.541,64	-
	Custos do programa	-	2.775.421,71
	Benefícios (perdas evitadas A – perdas B)		8.945.541,64
	Custo (Custo A – Custo 1B)		2.775.421,71
	Saldo líquido (benefícios - custos)	6.170.119,93	
Relação B/C		3,22	
Considerando apenas os custos dos produtores	Benefícios (exportações e empregos)	8.945.541,64	-
	Custos do programa	-	405.742,03
	Benefícios (perdas evitadas A – perdas B)		8.945.541,64
	Custo (Custo A – Custo 1B)		405.742,03
	Saldo líquido (benefícios - custos)	8.539.799,61	
Relação B/C		22,05	

Fonte: Elaborado pelas autoras com base nos dados da pesquisa.

A partir desses resultados, nota-se que, se não houvesse a Área Livre de Mosca das Frutas no Ceará, as perdas de exportações e postos de trabalho para o estado do Ceará, trazidas a valor presente e acumuladas para os quatro anos alcançariam um montante de R\$ 8,9 milhões. Como existem diversos impactos diretos e indiretos que não foram computados, acredita-se que o prejuízo possa ser ainda maior.

Em relação os custos do governo e dos produtores para manter a ALP-MF, nesse mesmo horizonte de análise, adotando a taxa de desconto TJLP, obteve-se um valor presente total de quase R\$3,2 milhões. Comparando os cenários A e 1B, o valor máximo dos benefícios estimados da ALP-MF, em quatro

anos, é R\$5,7 milhões. Em termos da relação benefício-custo (B/C), verifica-se que, para cada R\$1,00 investido pelo governo estadual e pelos produtores na Área Livre, o retorno é R\$2,81, ou, dito de outra forma, evitam-se perdas de R\$2,81 para cada real investido pelo governo e pelos produtores com a ALP-MF, perdas estas relacionadas às exportações e empregos.

Replicando a análise, mas sob a ótica apenas dos custos incorridos pelo governo do Ceará, o valor máximo dos benefícios estimados da ALP-MF, em quatro anos, chega quase a R\$6,2 milhões quando se comparam os cenários A e 1B. Os resultados da relação B/C indicam que, para cada R\$1,00 investido pelo governo estadual na conservação da ALP-MF, o retorno é R\$3,22. Em contrapartida, essa relação B/C é muito maior quando se leva em conta somente os custos dos produtores, visto que, para cada R\$1,00 investido pelos produtores na ALP-MF, o retorno é de R\$22,05 (desde que o governo mantenha as ações necessárias para manter o reconhecimento da ALP-MF), gerando um benefício líquido de R\$8,5 milhões ao ano. Embora se tenha adotado um modelo determinístico, que não considera um intervalo de confiança para estas relações, tais resultados sinalizam a relevância econômica da ALP-MF para o estado.

Na segunda simulação, conforme descrito, projeta-se que no período de quatro anos ocorrerá uma expansão dos volumes exportados de melão para os Estados Unidos, de modo que estas alcancem a mesma participação que o Brasil ocupa hoje nas vendas mundiais de melão, qual seja de 11%. Adicionalmente, supõe-se que, conforme as expectativas dos produtores da ALP-MF, o mercado japonês passe a importar a fruta do Ceará, também em uma proporção de 11% do total de suas importações de melão. Ademais, considera-se que o comércio brasileiro de melão com o Uruguai e o Chile permaneça com a parcela atual; e não haja alteração em relação à melancia. Os resultados dessa simulação constam da Tabela 3.

**TABELA 3 – Valores presentes totais dos benefícios (perdas evitadas) e dos custos da Área Livre de Mosca das Frutas no Ceará, para o melão e a melancia. Horizonte de simulação de quatro anos. Período base 2014. Cenário A e 2B**

Benefícios (perdas evitadas) e custos (R\$)		Cenários	
		A – Ausência da ALP-MF	2B- Manutenção da ALP-MF
Considerando os custos do governo do Ceará e dos produtores	Perdas	173.765.292,63	-
	Custos do programa	-	3.181.163,74
	Benefícios (perdas evitadas A – perdas B)		170.584.128,89
	Custo (custo A – custo 2B)		3.181.163,74
	Saldo líquido (benefícios - custos)	170.584.128,89	
	Relação B/C		54,62
Considerando apenas os custos do governo do Ceará	Perdas	173.765.292,63	-
	Custos do programa	-	2.775.421,71
	Benefícios (perdas evitadas A – perdas B)		170.989.870,92
	Custo (Custo A – Custo 2B)		2.775.421,71
	Saldo líquido (benefícios - custos)	170.989.870,92	
	Relação B/C		62,61
Considerando apenas os custos dos produtores	Perdas	173.765.292,63	-
	Custos do programa	-	405.742,03
	Benefícios (perdas evitadas A – perdas B)		173.359.550,60
	Custo (Custo A – Custo 2B)		405.742,03
	Saldo líquido (benefícios - custos)	173.359.550,60	
	Relação B/C		428,27

Fonte: Elaborado pelas autoras com base nos dados da pesquisa.

Diante dessas pressuposições, percebe-se que os ganhos provenientes da ALP-MF seriam muito mais expressivos do que os estimados para o cenário atual, expresso pela simulação anterior. Avaliando do ponto de vista do dispendio total do governo e dos produtores com a manutenção da ALP-MF, o benefício líquido estimado de quase R\$ 170,6 milhões ao ano. Em termos da relação benefício-custo (B/C), pode-se dizer que, cada R\$1,00 investido pelos setores público e privado na ALP-MF para *Anastrepha grandis* tem um retorno de R\$54,62.

Ao se considerar apenas os custos do governo, ou seja, buscando-se responder à questão do retorno da política pública fitossanitária em relação a essa mosca de frutas, a relação B/C é ainda maior, alcançando 62,61. Desta forma, para cada R\$1,00 investido pelo governo do Ceará, o retorno será de R\$ 62,61, caso as expectativas dos produtores de que a ALP-MF permita expandir e abrir mercados para as frutas se concretize. Nota-se ainda que quando se leva em conta somente os custos dos produtores, a relação B/C é muito mais elevada, visto que cada R\$1,00 investido pelos produtores resultará em um retorno de

R\$ 428,27, considerando os quatro anos de projeção. Contudo, esta análise deve ser interpretada com cuidado, tendo em vista que este retorno só pode ser considerado, se o governo mantiver sua política de manutenção da ALP.

Esses resultados demonstram a persistente motivação do setor privado continuar investindo nessa área há tantos anos, pois os produtores entrevistados acreditam que podem ampliar suas vendas para o atrativo mercado americano. Ademais, esperam conquistar o mercado do Japão, uma vez que já estão enviando o melão para outros mercados também rigorosos em termos fitossanitários, como o americano e chileno. O Japão e os EUA absorvem, respectivamente, 1,83% e 16,10% do total de melão importado mundialmente.

## 5 CONSIDERAÇÕES FINAIS

Os programas de defesa fitossanitária certamente têm efeitos e implicações econômicos, sociais, ambientais e institucionais sobre os diversos atores envolvidos. Um dos impactos mais destacados na literatura econômica é o que tange ao comércio internacional, seja pelo potencial que as questões fitossanitárias têm de tornar-se barreira comercial, seja pela necessidade e importância econômica das políticas de prevenção da entrada de pragas nos territórios. É neste contexto que se insere o tema aqui examinado, da análise econômica da importância de se manter a Área Livre de *Anastrepha grandis*.

Para remover esse problema, o governo do Ceará implementou a Área Livre para essa Mosca das Frutas e tem incorrido em elevados custos para conservar essa área, o que tem levado a questionamentos sobre a necessidade e priorização de se manter estes gastos vis à vis outros programas e alocações orçamentárias na área fitossanitária. É fato que, no entanto, tais custos não recaem apenas sobre o setor público, uma vez que os produtores também arcam com custos e investimentos contínuos nessa área com a expectativa de conquistar novos mercados e serem recompensados pelo investimento realizado. Como já conquistaram o mercado americano e chileno, que são exigentes em termos fitossanitários, acreditam que a ALP-MF permitirá nos próximos anos, não só ampliar a parcela nesses países compradores, mas também acessar outros mercados, como o japonês.

Neste sentido, buscando avaliar a viabilidade econômica de se manter essa área considerando tais simulações, utilizou-se a ferramenta da ABC, cujos resultados indicaram que a relação benefício-custo foi maior que a unidade em todas as situações analisadas, sendo que, para cada R\$ 1,00 investido, obteve-se

um retorno que variou de R\$ 2,81, quando se consideram os custos agregados do governo e dos produtores e que se mantenha a parcela de comércio cearense com os países quarentenários, a R\$ 428,27, quando se leva em conta somente os custos privados e numa situação de ampliação das vendas cearenses de melão nos mercados internacionais. Portanto, estes resultados permitem constatar que a manutenção da área é vantajosa em termos econômicos e sociais para o governo do Ceará e para os produtores.

Além dos produtores e do governo, a política de defesa agropecuária também influencia outros agentes da cadeia produtiva, como as indústrias de insumos, cujos impactos também podem ser evidenciados. Em estudos posteriores, pode-se empregar simulações de Monte Carlo para construir intervalos de confiança sobre as relações B/C calculadas para variáveis, como preços, sendo possível verificar como as relações B/C se alteram quando tais variáveis se modificam.

## REFERÊNCIAS BIBLIOGRÁFICAS

ADAMI, A. C. O.; MIRANDA, S. H. G. Seguro sanitário para citricultura; oportunidades e desafios. **Revista de Política Agrícola**. Brasília, v. 23, n. 1, p. 77-90, 2014.

AMOABENG, B. W.; GURR, G. M.; GITAU, C. W.; STEVENSON, P. C. Cost: benefit analysis of botanical insecticide use in cabbage: implications for smallholder farmers in developing countries. **Crop Protection**, v. 57, p. 71-76, 2014.

AZEVEDO, F. R.; SOBRINHO, R. B.; OMETTO, A. C. F. Estratégias para o estabelecimento e manutenção de Áreas Livres e de Baixa Prevalência de Moscas das Frutas. **Circular Técnica n. 22**. Fortaleza: EMBRAPA, 2005. 14 p.

BARBOSA, F. R.; PARANHOS, B. A. J. **Pragas quarentenárias**. Brasília, DF: Embrapa Parque Estação Biológica, 2007. Disponível em:

<[http://www.agencia.cnptia.embrapa.br/Agencia22/AG01/arvore/AG01\\_116\\_24112005115225.html](http://www.agencia.cnptia.embrapa.br/Agencia22/AG01/arvore/AG01_116_24112005115225.html)>. Acesso em: 13 jul. 2015.

BOADWAY, R. Principles of Cost-Benefit Analysis. **Public Policy Review**, v. 2, n.1, p. 1-44, 2006.

CONTADOR, C. R. **Projetos Sociais**. 5 ed. São Paulo: Editora Atlas, 2014.

DIÁRIO DO NORDESTE. Vigilância – Exemplo para o mundo – Área Livre da Mosca da Fruta contribui para a alta qualidade dos produtos cearenses. **Jornal Diário do Nordeste**. Fortaleza, CE, fevereiro de 2013, p. 49-56.

ENKERLIN, W.; GUTIÉRREZ-RUELAS, J. M.; CORTES, A. V.; ROLDAN, E. C.; MIDGARDEN, D.; LIRA, E.; LÓPEZ, J. L. Z.; HENDRICH, J.; LIEDO, P.; ARRIAGA, F. J. T. Area freedom in Mexico from Mediterranean fruit fly (Diptera: Tephritidae): a review of over 30 years of a successful containment program using an integrated area-wide SIT approach. **Florida Entomologist**, v. 98, n. 2, p. 665-681, 2015.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. Banco de Dados Agregados - Sistema IBGE de Recuperação Automática – SIDRA. **Quantidade produzida e área plantada de melão**. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 09 dez. 2015.

MINISTÉRIO DE AGRICULTURA, PECUÁRIA E ABASTECIMENTO – MAPA. **Programa Nacional de Mosca das Frutas**. MAPA, 2015a (Painel apresentado no lançamento do Programa).

---

\_\_\_\_\_. **Áreas Livres de Pragas**. Disponível em: <<http://www.agricultura.gov.br/.../Relação%20de%20Pragas/pragasvegetais>>. Acesso em: 13 jul. 2015b.

MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO EXTERIOR – MDIC. **Sistema Aliceweb**. Disponível em: < <http://www.aliceweb2.mdic.gov.br> >. Acesso em: 13 jul. 2015.

MINISTÉRIO DO TRABALHO E EMPREGO – MTE. **Dados e Estatísticas da Relação Anual de Informações Sociais – RAIS**. Disponível em: <<http://portal.mte.gov.br/rais/estatisticas.htm>>. Acesso em: 13 jul. 2015.

MIRANDA, S. H. G. Mosca de Carambola e *Greening*: análise de benefício-custo dos programas de erradicação. In: **Ameaças Fitossanitárias – novas pragas colocam em risco a produção de alimentos no Brasil**. SBDA, p. 33-36, 2013.

MIRANDA, S. H. G. Combater ou conviver? Análise Benefício-Custo de programas de combate a pragas quarentenárias. In: **Workshop Ameaças Fitossanitárias**. Fortaleza, CE, 2014.

MIRANDA, S. H. G.; ADAMI, A. C. O. Métodos quantitativos na avaliação de risco de pragas. In: SUGAYAMA, R. L.; SILVA, M. L.; SILVA, S. X. B.; RI-

BEIRO, L. C.; RANGEL, L. E. P. **Defesa vegetal: fundamentos, ferramentas, políticas e perspectivas.** Belo Horizonte, MG: SBDA, p. 183-203, 2015.

MIRANDA, S. H. G.; NASCIMENTO, A. M.; XIMENES, V. P. Potenciais impactos socioeconômicos da expansão da mosca-da-carambola. In: VILELA, E. F.; ZUCCHI, R. A. **Pragas introduzidas no Brasil: insetos e ácaros.** Piracicaba, SP: FEALQ, p. 114-132, 2015.

MIRANDA, S. H. G.; NASCIMENTO, A. M.; XIMENES, V. P.; BASSANEZI, R. B. **Uma aplicação da Análise Benefício-Custo para políticas de defesa sanitária: alguns estudos de caso para o Brasil.** Brasília: CNPq, 2010 (Relatório de Pesquisa).

MIRANDA, S. H. G. & XIMENES, V. P. A importância econômica dos impactos da introdução de pragas exógenas no Brasil. **Agronegócios.** Curitiba: IBPEX, p. 95-106, 2009.

SAFETYNET. **Cost-benefit analysis.** Project co-financed by the European Commission, Directorate-General Transport and Energy, 2009. 24 p. Disponível em: < [ec.europa.eu/transport/.../cost\\_benefit\\_analysis.pdf](http://ec.europa.eu/transport/.../cost_benefit_analysis.pdf) >. Acesso em: 19 jan. 2014.

SANCHES, A. L. R.; MIRANDA, S. H. G.; BELASQUE JUNIOR, J.; BASSANEZI, R. B. Análise econômica da prevenção e controle do cancro cítrico no estado de São Paulo. **Revista de Economia e Sociologia Rural.** Piracicaba, SP, v. 52, n. 3, p. 549-566, 2014.

SOUZA, D. L. B. **Estudo das vantagens competitivas do melão no Ceará.** Fortaleza: Instituto Agropolos do Ceará, 2006.

WORLD INTEGRATED TRADE SOLUTION – WITS. **UN Comtrade Database.** Disponível em: <<http://wits.worldbank.org/>>. Acesso em: 07/07/2015.

WORLD TRADE ORGANIZATION – WTO. **Sanitary and Phytosanitary Agreement.** Disponível em: <<http://www.wto.org/>>. Acesso em: 26/10/2015.

ZERBE, R. O.; BELLAS, A. S. **A primer for benefit-cost analysis.** Cheltenham (United Kingdom), Northampton (Massachusetts): Edward Elgar, 2006.

# ANÁLISE DO PERFIL SOCIOECONÔMICO E TECNOLÓGICO DOS PRODUTORES DE FRUTICULTURA IRRIGADA NA REGIÃO SUL CEARENSE

Otácio Pereira Gomes\*

Kilmer Coelho Campos\*\*

Wellington Ribeiro Justo\*\*\*

Guillermo Gamarra Rojas\*\*\*\*

Maria Daniele Cruz dos Santos\*\*\*\*\*

## Resumo

O presente artigo visa apresentar as classes de atividade CNAE 2.0, classificadas como especializados no setor agropecuário através do Índice de Concentração Normalizado – ICN, na macrorregião Nordeste do Brasil e fazer um comparativo com o Estado do Ceará, identificando neste as atividades especializadas em comum com a grande região. A pesquisa é fruto de investigações do grupo de pesquisa para o desenvolvimento da região centro-sul cearense, vinculado ao grupo CNPq Territorialidades Econômicas e Desenvolvimento Regional e Urbano. Pode-se identificar que das oito atividades, apenas dois setores são especializados em comum e pertencentes a mesma divisão: a aquicultura, tanto de água doce quanto salgada. Observa-se um crescimento exponencial da atividade nos últimos anos no Nordeste, alavancado em boa parte pelo Estado do Ceará, o que contribuiu significativamente para despontar seus resultados e a respectiva boa classificação no ranking.

**Palavras-Chave:** Setor agropecuário; ICN; Nordeste; Ceará.

---

\* Prof. Msc. da Faculdade Vale do Salgado, Icó-CE. Endereço: Rua Maria Alves de Oliveira, nº 46, Bairro: Alto do Jucá, Iguatu, Ceará. CEP: 63.502.430, Fone: (88)99601-1930. Email: otaciopg@gmail.com.

\*\* Prof. Dr. Adjunto da do curso de mestrado em Economia Rural (MAER). Email: kilmer@ufc.br.

\*\*\* Prof. Dr. Adjunto da Universidade Regional do Cariri (URCA), Crato. Email: justowr@yahoo.com.br.

\*\*\*\* Prof. Dr. Adjunto da do curso de mestrado em Economia Rural (MAER). Email: ggamarra@terra.com.br.

\*\*\*\*\* Profª Especialista Temporária da Universidade Regional do Cariri – URCA, Campus Iguatu/CE. E-mail: danielecruzeconomia@gmail.com



## Abstract

It aims to analyze the profile of Ceará south horticulture and measure the technological level, identifying the factors that most influence. To prepare the technological index was used to factor analysis, and to measure the index of efficiency scores the envelopment analysis models were applied data (DEA) and quantile regression. They came from indicators of primary source obtained from the 86 growers of Ceará south. The results show that mostly fruit farmers (53.49%) have only to primary education, and only 16.28% can only sign his name. You can still see that 31.40% of producers have secondary education is low and that the number of illiterate farmers (3.49%) and they have 30-50 years of age (66.67%) and also that 75 , 68% of growers have over five years of experience in the activity. The results also show that 42.66% of fruit farmers in the region who do not participate unions have no access to any type of credit; 72.73% of these already, participating cooperatives or associations, hold some kind of credit. To technological contents, 61.63% of the growers have an average index and one (1.63%) showed high rate. For the less efficient growers group, it is observed that management index, credit and education played irrelevant role in explaining variations in technical efficiency level.

**Key words:** Fruit Growers; Technology; Technical Efficiency.

**JEL:**Q1, C5.

## 1 INTRODUÇÃO

A agricultura é um dos setores de maior importância da economia brasileira, dada sua elevada capacidade de geração de emprego e renda, na agricultura propriamente dita, assim como nos setores a ela relacionados, na produção e distribuição de insumos, bem como na compra, armazenamento, transformação e distribuição de seus produtos e subprodutos. Nesse conjunto, um dos setores cuja importância é crescente é o da fruticultura (SOUZA et al., 2009).

Esta atividade participa diretamente na economia do País por meio do valor das exportações e do mercado interno, podendo-se ainda salientar, a importância no caráter econômico e social, uma vez que se encontra em todos os estados brasileiros, sendo responsável por gerar milhões de empregos diretos, acolhendo um grande percentual da mão de obra agrícola.

De acordo com os documentos do Banco do Nordeste (2001, pág. 45),

“no Brasil, a produção por meio de cultivos irrigados é relativamente recente, tendo sua evolução ocorrida por meio de ações isoladas, dirigidas para alvos específicos, em termos setoriais e espaciais, sem a estrutura de políticas ou de programas nacionais”.

O Brasil possui pequenas áreas irrigadas, sendo que suas causas podem estar associadas a: pouca ênfase em políticas públicas de estímulo à agricultura irrigada; desarticulação dos órgãos envolvidos com a irrigação; desinformação quanto à disponibilidade de crédito para a irrigação e dificuldade de acesso, sobretudo em agentes privados; altos custos de energia; baixa organização dos produtores e das cadeias dos produtos oriundos da agricultura irrigada; falta de melhor qualificação técnica de parte dos agentes de comercialização de equipamentos de irrigação (...) e transferência de tecnologia e assistência técnica ineficiente, dada a complexidade dos sistemas de produção irrigados (PENSA, 2010).

Para a região Nordeste, a fruticultura possui enorme potencial na oferta de emprego e renda. As condições climáticas da sub-região semiárida nordestina, que representava um obstáculo para o cultivo de grãos e a pecuária, transformaram-se numa vantagem quando se trata da fruticultura irrigada. No Nordeste, são muitos os aglomerados de polos frutícolas, sendo o de maior referência em estudos nesse ramo o polo Petrolina-PE/Juazeiro-BA.

A fruticultura irrigada, principalmente, baseada em tecnologias modernas, se revela importante opção de investimento no setor agrícola, sendo capaz de ensejar produtos de maior valor agregado (SILVA; SILVA e KHAN, 2004).

Essa atividade, segundo Cardoso e Souza (2000) “apresenta, geralmente, rendimento superior a muitos outros produtos, possui substancial potencial para gerar empregos ao longo de sua cadeia produtiva e, dada sua diversidade, pode contribuir para minimizar outros problemas, tais como a sazonalidade de mão de obra”.

Diferentemente do semiárido nordestino em geral, a Região Sul do Ceará, detém notável potencial de recursos naturais, terras férteis, e recursos hídricos subterrâneos e é um dos climas mais favoráveis do Estado. Esses fatores não apenas servem para garantir bons resultados econômicos, mas também poderão contribuir para mudar a realidade das regiões, que vivem dessa prática agrícola, assim como um possível melhoramento de indicadores socioeconômicos.

Assim, a intensificação das políticas de agricultura irrigada, direcionadas, em especial, para a fruticultura irrigada na região sul cearense, e que buscam impulsionar essa atividade, são preponderantes, de forma a impedir a diminui-

ção da produção que ocorre paulatinamente a cada safra, em particular, nos períodos de estiagem prolongada, como os que aconteceram nos últimos dois anos, quando a queda da produção pode ter sido ainda mais acentuada.

Por essa e outras razões, justifica-se por meio de ações planejadas, verificar e analisar o desenvolvimento da fruticultura na região sul cearense, especificadamente nos Municípios de Crato, Missão Velha, Brejo Santo, Barbalha e Mauriti, pois estes demonstraram os melhores desempenhos relativos à produção de frutas na região.

Assim, o objetivo desse trabalho é identificar e analisar o perfil socioeconômico e tecnológico de produtores da fruticultura irrigada na região sul do Estado do Ceará.

## **2 FUNDAMENTAÇÃO TEÓRICA**

A agricultura brasileira sempre foi caracterizada como heterogênea, no que concerne à utilização de tecnologias, concentrando-se em poucos estabelecimentos agropecuários, e sendo diversos os motivos para seu emprego ou não no panorama agrícola do País. Essa desigualdade não está condicionada somente ao não acesso à tecnologia, mas também à dificuldade de acesso aos mercados, à capacidade de geração de renda, entre outras.

A adoção de modernas tecnologias causa efeitos positivos sobre a produção, elevando os níveis de produtividade, em todos os setores, sejam agrícolas ou não agrícolas, beneficiando a economia como um todo. O grande problema encontrado, no entanto, é que essa tecnologia não conseguiu se difundir de maneira uniforme entre todos os produtores nem com a mesma rapidez e intensidade que necessitaria para que pudesse promover elevações na produção agrícola.

Ganhos de produtividade refletem-se na redução dos preços dos produtos agrícolas, tornando-os, no longo prazo, mais acessíveis à população. Além disso, embora vários sejam os fatores que contribuem para elevar a competitividade de determinado setor ou atividade, os investimentos em tecnologia constituem, certamente, um de seus determinantes fundamentais (CARDO-SO, 2003).

Dada à importância da adoção de tecnologia, pode-se chegar à conclusão que ela é apenas parte de um processo de produção, sendo capaz de influenciar um conjunto de fatores específicos que podem desregular e até mesmo inviabi-

lizar a adoção por alguns grupos de produtores.

Questões relativas à eficiência constantemente são investigadas por tomadores de decisões, sobretudo quando se trata de ambientes competitivos e dinâmicos. A identificação do verdadeiro potencial de expansão da produção e as alterações na eficiência, no progresso tecnológico e na produtividade, ao longo do tempo, são condições necessárias para formulação de políticas econômicas coerentes com as necessidades dos setores analisados (GOMES; BAPTISTA, 2004; SOUSA; JUSTO; CAMPOS, 2013).

Esses efeitos proporcionados pela adoção de níveis tecnológicos mais sofisticados significam excelentes resultados, estimulando a competitividade entre os produtores na região, como também provocando a competitividade entre produtos nacionais e estrangeiros.

Na busca de maior competitividade, além da tecnologia, não podem ser negligenciados os aspectos econômicos envolvidos na produção. A produção econômica de qualquer cultura depende de uma série de fatores, que afetam seu desempenho e seu retorno financeiro. A variedade plantada, o espaçamento, o clima, o solo, os tratos culturais, o grau de incidência de pragas e doenças, os preços do produto e dos fatores de produção merecem especial atenção no planejamento da produção [...] (CARDOSO e SOUZA, 2000).

Por conta do intenso progresso tecnológico, ao inserir melhores práticas de cultivo ao campo, os produtores são induzidos a ter a maior compreensão das tecnologias para melhor manusearem os equipamentos. Com esse, mostra-se o quão é importante o papel da educação e dos gastos públicos em pesquisa, garantindo ganhos na produtividade agrícola.

Na promoção da modernização tecnológica da agricultura, Schultz (1965, pág. 43) enfatizou “a necessidade de melhorias na educação, de investimentos em pesquisa e nos serviços de assistência técnica. Esses investimentos, segundo o autor, têm papel primordial para tornar o uso de tecnologia mais acessível e próximo do agricultor”. E o mesmo autor complementa, ainda, expressando que, “apesar dos resultados favoráveis, a distância entre o conhecimento tecnológico disponibilizado nos serviços de pesquisa e de extensão rural e o seu emprego por parte da maioria dos produtores rurais é muito grande”.

Na realidade, o que se observa é a falta de elementos práticos e modernos, capazes de minimizar esse distanciamento e que venham efetivamente beneficiar o maior usuário, que é o produtor rural, do que apenas estudos associados ao problema sem nenhuma ação que possa lhe beneficiar. Assim, as

considerações até aqui procedidas, fica evidente a importância do uso de meios mais sofisticados, como é a tecnologia de produção, com sentido do produtor auferir renda relevante da implementação do recurso tecnológico, mantendo-o sempre competitivo no mercado onde atua.

### **3 METODOLOGIA**

#### **3.1 Área de estudo**

Localizada ao Sul do Ceará, a região do Cariri é composta por 32 municípios, que se limitam com os Estados de Pernambuco, Piauí e Paraíba, e pelo prolongamento da Chapada do Araripe. Os municípios da região do Cariri estão distribuídos em cinco microrregiões, assim estabelecidos. Sertão do Salgado, Serra do Caririáçu, Sertão Cariri, Chapada do Araripe e por fim o Cariri (REVISTA CARIRI INVEST, 1999). A pesquisa foi realizada em 5 municípios da região nos quais foram observados um avanço considerável na produção de frutas.

#### **3.2 Fonte dos dados e amostragem**

Para obtenção das informações empregadas neste experimento, foram utilizados dados primários, obtidos com aplicação de questionários junto aos produtores de frutas desses municípios, visando a obter uma caracterização socioeconômica e tecnológica desses produtores. O levantamento dos dados foi realizado em outubro a novembro e as informações correspondem ao período de 2014. A pesquisa foi realizada por amostragem não probabilística por conveniência, levando em conta os produtores que exploram a fruticultura irrigada na região sul cearense, especificamente os Municípios de Crato, Barbalha, Mauriti, Brejo Santo e Missão Velha, no Estado do Ceará, logo, foram entrevistados 86 fruticultores.

#### **3.3 Caracterização socioeconômica dos fruticultores**

A técnica da análise descritiva foi empregada no estudo proposto, com o objetivo de caracterizar o perfil dos produtores irrigantes de acordo com suas características sociais e econômicas. Foi utilizada a análise tabular, com exposição de tabelas de distribuição de frequências (absoluta e relativa) e medidas de tendência central.

A pesquisa descritiva tem como objetivo primordial a descrição das características de determinada população ou o estabelecimento de relação entre elas, que abordam as seguintes variáveis: (a) características do produtor (idade, gênero, experiência, escolaridade), (b) contato com outros produtores (associativismo, cooperativas) e (c) assistência técnica.

### 3.4 Identificação dos fatores representativos do desempenho tecnológico dos fruticultores

Para caracterizar o grau de adoção de tecnologia pelos produtores de frutas da região Sul do Ceará, foi empregada a análise fatorial.

O modelo matemático (conforme FÁVERO *et al.*, 2009) da análise fatorial poderá ser representado por:

$$Z_1 = a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + d_1u_1 \quad (1)$$

$$Z_2 = a_{21}F_1 + a_{22}F_2 + \dots + a_{2m}F_m + d_2u_2$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot$$

$$Z_n = a_{n1}F_1 + a_{n2}F_2 + \dots + a_{nm}F_m + d_nu_n$$

De forma simplificada, tem-se:

$$Z_j = \sum a_{ji}F_i + d_ju_j \quad (j=1,2,\dots,n); \quad (i=1,2,\dots,m)$$

Em que:

$Z_j$  = j-ésima variável padronizada;

$a_{ji}$  = é o coeficiente de saturação referente ao i-ésimo fator comum da j-ésima variável;

$F_i$  = é o i-ésimo fator comum;

$d_j$  = é o coeficiente de saturação referente ao j-ésimo fator específico da j-ésima variável;

$u_j$  = é o j-ésimo fator específico da j-ésima variável.

De acordo com a análise fatorial, cada fator é constituído por uma combinação linear das variáveis originais inseridas no estudo. A associação entre

fatores e variáveis se dá por meio das cargas fatoriais, podendo ser positivos ou negativos, mas nunca superiores a um. Esses coeficientes de saturação têm função similar aos coeficientes de regressão na análise de regressão (SIMPLICIO, 1985).

De acordo com Hair Jr. *et al.* (2005), a análise iniciar-se-á do exame da matriz de correlações para verificação da existência de valores significativos que justifiquem a utilização da técnica. Deve ser considerada também a normalidade dos dados para o uso da técnica (Teorema do Limite Central).

A fim de verificar a adequabilidade dos dados para a análise fatorial, serão utilizados o Índice Kaiser-Mayer-Olkin (KMO), o Teste de Esfericidade de Bartlett (BTS) e a Matriz Anti-Imagem.

O procedimento utilizado na pesquisa considerará a extração dos fatores iniciais por via da Análise dos Componentes Principais (ACP), que mostrará uma combinação linear das variáveis observadas, de maneira a maximizar a variância total explicada.

A escolha do número de fatores se dará por meio do critério da raiz latente (Critério de Kaiser) em que se escolhe o número de fatores a reter, em função dos valores próprios acima de 1 (*eigenvalues*), que mostraram a variância explicada por parte de cada fator, ou quanto cada fator conseguiu explicar da variância total (MINGOTI, 2005).

Com a finalidade de minimizar a dificuldade de interpretação dos fatores, utilizou-se o método de rotação ortogonal Varimax, que minimizará o número de variáveis com altas cargas em distintos fatores, permitindo a associação de uma variável a um só fator, mantendo a ortogonalidade entre eles. Neste estudo, para caracterizar o grau de desenvolvimento tecnológico dos fruticultores, foram consideradas as variáveis a seguir dispostas:

1) Práticas agrícolas; 2) Gestão rural; 3) Técnica de plantio; 4) Assistência técnica; 5) Anos de experiência; 6) Escolaridade; 7) Investimento; 8) Produção total; 9) Área total; 10) Renda Bruta e 11) Defensivos agrícolas.

### **3.5 Cálculo do Índice Tecnológico da Fruticultura Sul Cearense – ITFSC**

A análise fatorial permitiu criar um índice tecnológico dos fruticultores, com base nos escores fatoriais/fatores que mais contribuíram para o nível de tecnologia dos produtores. O Índice Tecnológico dos Fruticultores foi obtido da seguinte forma:

$$ITFSC_i = \sum_{j=1}^p \left[ \frac{\lambda_j}{\sum \lambda_j} \right] F_j^* \quad (2)$$

Em que:

$ITFSC_i$  = Índice Tecnológico do  $i$ -ésimo Fruticultor Sul Cearense

$j$  = é a  $j$ -ésima raiz característica ( $J= 5$  raízes)

$p$  = é o número de fatores extraídos na análise ( $F= 5$  fatores)

$F_{j_i}^*$  = é o  $j$ -ésimo escore fatorial do  $i$ -ésimo Fruticultor

$\sum \lambda_j$  = é o somatório das raízes características referentes aos  $p$  fatores extraídos

A participação relativa do fator  $j$  na explicação da variância total captada pelos  $p$  fatores extraídos e indicada por  $\frac{\lambda_j}{\sum \lambda_j}$ .

Para tornar todos os valores dos escores fatoriais  $F_{j_i}^*$  superiores ou iguais a zero, todos eles são colocados no primeiro quadrante (LEMOS, 2001), antes da elaboração do  $ITFSC_i$ , utilizando-se a expressão algébrica:

$$F_{j_i}^* = \frac{F_j - F_j^{\min}}{F_j^{\max} - F_j^{\min}} \quad (3)$$

Em que:

$F_j^{\min}$  é o menor escore observado para o  $j$ -ésimo fator, e  $F_j^{\max}$  é o maior escore observado para o  $j$ -ésimo fator.

### 3.6 Análise Envoltória de Dados (DEA)

De acordo com Gomes e Baptista (2004), as fronteiras podem ser estimadas por métodos paramétricos mediante procedimentos econométricos e por métodos não paramétricos. O método DEA fundamenta-se no trabalho proposto por Farrell (1957) e generalizado por Charnes, Cooper, Rhodes (1978), em que se incluíram múltiplos produtos e insumos.

As principais vantagens da utilização deste método consistem em permitir a obtenção das relações entre múltiplos produtos e insumos de modo menos complexo, identificar as ineficiências em cada insumo e produto e indicar as



unidades referenciais que servem como parâmetro de eficiência técnica para as unidades ineficientes (COOPER *et al.*, 2002).

O método DEA baseia-se numa amostra de insumos e produtos observados para distintas empresas ou unidades tomadoras de decisão (DMUs – *Decision Making Units*), em que se procura construir uma fronteira linear por partes e, utilizando-se de medidas radiais e de função de distância, analisa-se a eficiência das unidades de produção em relação à distância da fronteira elaborada com os *benchmarks* (os mais eficientes). Neste trabalho, a DMU corresponde ao produtor de frutas localizados no sul do Estado do Ceará.

Em sua versão inicial, o modelo DEA foi desenvolvido por Charnes, Cooper, Rhodes (1978) e ficou conhecido na literatura por modelo CCR em virtude das iniciais de seus nomes. Esse modelo pressupõe retornos constantes à escala e também é conhecido como CRS (*Constant Returns to Scale*).

Seguindo Coelli, Rao e Battese (1998), o modelo DEA com retornos constantes pode ser representado por:

$$\text{Min}_{\theta, \lambda} \theta, \text{ sujeito a: } -y_i + Y\lambda \geq 0, \quad \theta x_i - X\lambda \geq 0 \quad \text{e} \quad \lambda \geq 0 \quad (4)$$

Aqui,  $\theta$  é o escore de eficiência de uma dada DMU;  $y$  é o produto da DMU;  $x$  é o insumo;  $X$  é a matriz de insumos ( $n \times k$ );  $Y$  é a matriz de produtos ( $n \times m$ );  $\lambda$  é o vetor de constantes que multiplica a matriz de insumos e produtos.

De acordo com Ferreira e Gomes (2009), o pressuposto de retornos constantes de escala possibilita que se represente tal tecnologia empregada mediante uma isoquanta unitária. Sobre essa fronteira, o escore de eficiência é igual à unidade. Isso significa dizer que a DMU analisada é eficiente.

Em 1984, surgiu o modelo BCC, que teve essa denominação baseada nas iniciais de seus formuladores: Banker; Charnes e Cooper (1984). Esse modelo considera retornos variáveis, que podem assumir rendimentos crescentes ou decrescentes de escala na fronteira eficiente e também é conhecido como VRS (*Variable Returns to Scale*).

Nesse caso, conforme Coelli *et al.* (1998), o modelo DEA com retornos variáveis pode ser expresso por:

$$\text{Min}_{\theta, \lambda} \theta, \text{ sujeito a: } -y_i + Y\lambda \geq 0, \quad \theta x_i - X\lambda \geq 0, \quad N_1' \lambda = 1 \quad \text{e} \quad \lambda \geq 0 \quad (5)$$

em que  $N_1$  é um vetor ( $N \times 1$ ) de algarismos unitários.

Para uma DMU ser eficiente no modelo com retornos constantes, ela

deverá necessariamente ser eficiente no modelo com retornos variáveis, porém a recíproca não é verdadeira (COELLI *et al.*, 1998). Se o escore de eficiência técnica for distinto nos dois modelos, significa que a DMU considerada contém ineficiência de escala. Portanto, consoante Ferreira e Gomes (2009), a eficiência técnica global das unidades produtivas pode ser decomposta em duas formas de eficiência - a pura eficiência técnica e a eficiência de escala- sendo que a última corresponde ao quociente entre o escore obtido no modelo CCR e o encontrado no modelo BCC.

Neste estudo, foram determinados os escores de eficiência desses dois modelos, empregando-se a orientação insumo, que visa a reduzir os insumos sem modificar o nível dos produtos. A operacionalização de tais modelos foi realizada por meio do software *Efficiency Measurement System* (EMS) versão 1.3 (SCHEEL, 2000).

Ademais, é relevante destacar o fato de que, segundo Gomes e Baptista (2004), apenas uma observação discrepante na amostra influencia todas as medidas de eficiência. Desta forma, necessita-se verificar se existem *outliers* nos dados obtidos para não comprometer os resultados estimados e torná-los mais robustos. Souza e Stósic (2005) desenvolveram uma técnica de identificação de *outliers* e erros de medidas, baseados na associação do DEA com o método Jackstrap. Este método foi elaborado com base no teste Jacknife com o Bootstrap. O procedimento é elaborar uma medida de alavancagem que mensura a influência de cada DMU sobre as demais onde aquelas que apresentam maiores influências são descartadas da amostra por denotarem características que afetam as estimações do DEA.

Para identificar a influência de cada DMU, seguiram-se os procedimentos adotados por Araújo (2007), em que inicialmente se aplicou a técnica Jacknife, determinando os escores de eficiência por via do DEA para cada DMU contida no conjunto original de dados. Feito isto, removeu-se sucessivamente cada DMU, recalculando as eficiências. Como esse método é muito intensivo computacionalmente, utilizou-se a combinação desse procedimento Jacknife com a técnica de reamostragem Jackstrap, seguindo o procedimento adotado por Sousa, Justo e Campos (2013).

### 3.7 Regressão Quantílica

O método de regressão quantílica foi proposto inicialmente por Koenker e Bassett (1978). De acordo com esses autores, esse método contém vantagens em relação ao Modelo dos Mínimos Quadrados Ordinários (M.Q.O) pelo fato

de possibilitar a caracterização de toda distribuição condicional de uma variável resposta a partir de um conjunto de regressores; empregar todos os dados para estimar os coeficientes angulares dos quantis; não exercer forte interferência dos *outliers*, já que não se considera apenas o efeito médio do impacto de um regressor na distribuição condicional de um regressando, e os estimadores resultantes da Regressão Quantílica podem ser mais eficientes do que os obtidos pelo M.Q.O, visto que os erros não possuem distribuição normal.

Dadas essas vantagens, este método foi escolhido para fazer parte deste estudo, pois se está interessado não apenas em identificar os fatores que explicam a eficiência dos produtores de frutas, mas também em averiguar se as variáveis explicativas influenciam diferentemente a eficiência dos fruticultores com níveis de eficiência distintos.

Conforme Koenker e Basset (1978), a regressão quantílica  $\theta$  pode ser expressa com notação matemática:

$$\min_{\beta} \frac{1}{n} \sum_{i: y_i > x_i' \beta} \theta |y_i - x_i' \beta| + \sum_{i: y_i \leq x_i' \beta} (1 - \theta) |y_i - x_i' \beta| = \min_{\beta} \frac{1}{n} \sum_{i=1}^n \rho_{\theta}(\mu_{\theta_i}) \quad (6)$$

em que:  $\rho_{\theta}$  é a função *check* definida por:

$$\rho_{\theta}(u_{\theta_i}) = \begin{cases} \theta u_{\theta_i}, & u_{\theta_i} \geq 0 \\ (\theta - 1)u_{\theta_i}, & u_{\theta_i} < 0 \end{cases} \quad (7)$$

Os coeficientes da Regressão Quantílica podem ser interpretados por meio da derivada parcial do quantil condicional com respeito a um regressor particular. Em outros termos, tais coeficientes podem ser interpretados como uma variação marginal no  $\theta_h$  quantil condicional ocasionada por uma mudança no regressor (JUSTO, 2008).

Considerando que as variáveis explicativas não influenciam igualmente os diferentes níveis de eficiência técnica, foram estimadas regressões para os quantis 0,10 e 0,90, em função do número de observações, em que o  $\theta_h$  quantil condicional do nível de eficiência pode ser expresso por:

$$Q_{\theta}(y_i | X_i) = X_i' \beta_{\theta}, \quad \theta \in (0,1) \quad (8)$$

em que  $y_i$  corresponde aos escores de eficiência técnica obtidos no modelo DEA e  $X_i$  refere-se às variáveis capazes de explicar esses índices de eficiência.

A operacionalização desse modelo foi realizada por meio do *software* STATA 12.

Antes de definir as variáveis explicativas a serem consideradas no modelo de regressão quantílica, foi relevante identificar se o regressor será constituído pelos escores de eficiência do modelo com retornos constantes de escala (CRS) ou com retornos variáveis de escala (VRS). Para isso, foi empregada a Estatística não Paramétrica de Kolmogorov-Smirnov ( $T_{K-S}$ ), que, conforme Banker e Natarajan (2004), baseia-se na distância máxima entre as distribuições acumuladas dos escores de eficiência resultantes dos modelos CRS e VRS, podendo ser expresso por:

$$T_{K-S} = \max[F(\theta_{CRS}) - F(\theta_{VRS})] \quad (9)$$

Quando o valor dessa estatística  $T_{K-S}$  exceder seu valor crítico D, rejeita-se a hipótese nula de retornos constantes de escala. Isso significa dizer que se aceita a hipótese alternativa de retornos variáveis de escala. Foram realizados testes de igualdade dos coeficientes entre os quantis.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Perfil socioeconômico dos produtores de frutas da região do sul cearense

O estudo do perfil dos produtores exerce papel fundamental para o desenvolvimento da atividade da fruticultura. Com apoio nos resultados da Tabela 1 mostra-se a composição da idade dos fruticultores, por grau de instrução, tendo como resultado o fato de que, em sua maioria, os fruticultores possuem o ensino primário e secundário e os mesmos têm de 30 a 50 anos de idade. Destaque seja conferido à alta quantidade de fruticultores que possui ensino superior, o que se torna muito importante para a adoção de tecnologias voltadas à produção de frutas no que se refere à sua não utilização por aqueles que possuem baixo nível de escolaridade.

**TABELA 1 - Idade por grau de instrução dos fruticultores dos municípios pesquisados, 2014**

Idade (id)	Não lê nem escreve	%	Assina o nome	%	Lê e escreve	%
Id<30	0	0,00	0	0,00	0	0,00
30<id<50	2	66,67	6	42,86	1	33,33
Id>50	1	33,33	8	57,14	2	66,67
Total	3	100,00	14	100,00	3	100,00
Idade (id)	Primário	%	Secundário	%	Superior	%
Id<30	0	0,0	3	11,11	1	7,69
30<id<50	14	53,85	16	59,29	9	69,23
Id>50	12	46,15	8	29,60	3	23,08
Total	26	100,00	27	100,00	13	100,00

Fonte: Dados da Pesquisa (2014).

A Tabela 2 mostra o percentual associado à experiência dos fruticultores na atividade. Os resultados mostram que 75,58% dos fruticultores têm mais de cinco anos de experiência. Apesar da baixa idade desses fruticultores, sua experiência mostra que a atividade é praticada durante um bom período, e isso não se torna um empecilho, proporcionando bom desempenho na atividade.

**TABELA 2 - Experiência dos produtores do sul cearense com a fruticultura**

Experiência na fruticultura (anos)	Frequência	Percentual	Acumulado
Experiência > 2	1	1,16	1,16
2 < Experiência < 6	20	23,26	24,42
Experiência > 6	65	75,58	100
Total	86	100	

Fonte: Dados da Pesquisa (2014).

Os resultados da Tabela 3 mostram que cerca de 42% dos fruticultores não utilizam assistência técnica por agrônomo ou técnico agrícola.

**TABELA 3 - Percentual dos fruticultores que utilizam assistência técnica por agrônomo ou técnico agrícola**

Assistência técnica	Frequência	Percentual	Acumulado
Sim	50	58,14	58,14
Não	36	41,86	100
Total	86	100	

Fonte: Dados da Pesquisa (2014).

A Tabela 4 mostra participação dos fruticultores em cooperativas por acesso ao crédito agrícola na região sul cearense. Os resultados mostram que

42,66% dos fruticultores da Região não participam de cooperativas e não têm acesso a algum tipo de crédito. Já 72,73% destes que participam de cooperativas ou associações, detêm algum tipo de crédito.

**TABELA 4 - Participação dos fruticultores em cooperativas, por acesso ao crédito agrícola**

Sócio Cooperativa	Crédito Agrícola				Total
	Não	%	Sim	%	
Não	32	42,66	43	57,34	75
Sim	3	27,27	8	72,73	11
Total	35	40,7	51	59,3	86

Fonte: Dados da Pesquisa (2014).

#### **4.2 Identificação e análise do desempenho tecnológico dos fruticultores na região sul cearense**

Inicialmente, com o intuito de verificar a coesão dos dados coletados, foi calculado o teste Kaiser-Mayer-Olkin (KMO). Assim, considerando-se distribuição normal dos dados, que o KMO revelou valor de 0,674, portanto, indicando que os dados são consistentes. O Teste de Esfericidade de Bartlett indicou valor 291,658, sendo considerado elevado para garantir que a matriz de correlações não é uma matriz identidade, ao nível de significância 1%.

Com o uso da análise fatorial, pelo método dos componentes principais, foram obtidos cinco fatores característicos, com índices maiores do que 1, considerando o critério da raiz latente, conforme observado na Tabela 5.

Objetivando caracterizar ou representar um total de variáveis originais em um número menor possível de variáveis, a fim de explicar a tecnologia adotada pelos fruticultores, optou-se por trabalhar com os cinco fatores, considerando-se que o Fator 1 possui raiz 3,003, o Fator 2 tem raiz 1,919, o Fator 3, apresenta uma raiz característica de 1,268, o Fator 4, observou-se uma raiz de 1,087, e o Fator 5 possui raiz 1,028, ou seja, todos os fatores atendem a exigência da metodologia de apresentar raízes latentes maiores do que 1 e que, em conjunto, explicam 75,499% da variância total dos 11 indicadores de adoção de tecnologia pelos fruticultores.

**TABELA 5- Valores das raízes características e percentual de variância total explicada pela análise fatorial**

Fator	Raiz característica	Variância explicada pelo fator (%)	Variância acumulada (%)
1	3,003	24,058	24,058
2	1,919	16,431	40,489
3	1,268	12,418	52,907
4	1,087	11,375	64,282
5	1,028	11,217	75,499

Fonte: Resultados da Pesquisa (2014).

Conforme a Tabela 6, foram observadas as cargas fatoriais ou coeficientes de correlação após a rotação dos fatores de adoção de tecnologia e suas respectivas *comunalidades*. Admite-se que valores acima de 0,5 (em negrito) indicam intensiva associação entre a variável e o fator.

**TABELA 6 - Cargas fatoriais rotacionadas das variáveis de tecnologia obtidas na análise fatorial**

Variáveis	F1	F2	F3	F4	F5	Comunalidades
X1- Escolaridade	0,092	-0,146	0,122	0,841	-0,234	0,807
X2- Área total	0,913	0,019	-0,113	0,124	0,023	0,863
X3- Renda Bruta	0,934	0,075	0,054	-0,061	0,086	0,892
X4- Produção Total	0,895	0,152	0,154	-0,034	0,006	0,849
X5- Assistência Técnica	0,074	0,034	-0,008	-0,067	0,937	0,889
X6- Técnica de plantio	-0,002	0,523	0,385	0,201	0,306	0,555
X7- Gestão Rural	0,232	0,178	0,596	0,265	0,313	0,610
X8- Investimento	0,106	-0,349	0,061	-0,634	-0,255	0,603
X9- Experiência	-0,054	0,018	0,895	-0,076	-0,154	0,833
X10- Práticas Agrícolas	0,239	0,749	0,057	0,019	0,084	0,629
X11- Defensivos Agrícolas	0,007	0,876	0,020	-0,013	-0,075	0,774

Fonte: Resultados da Pesquisa (2014).

O primeiro fator (F1) está relacionado à tecnologia voltada para a produtividade, composta por: X2 (Área Total), X3 (Renda Bruta), X4 (Produção Total) sendo, portanto, *F1 representado pelo uso intensivo da tecnologia de Produtividade*.

Em relação às cargas do fator (F2), constatou-se que este possui correlação com as variáveis X6 (técnica de plantio), X10 (práticas agrícolas) e X11 (defensivos agrícolas), *indicando o uso intensivo de técnicas agrícolas*.

Analisando o fator (F3), percebeu-se que ele está intensivamente ligado

às variáveis X7 (Gestão Rural), X9 (Experiência), *representando o uso intensivo em Gestão.*

O fator F4 está relacionado com as variáveis referentes ao capital: X1 (Escolaridade) e X8 (Investimento), *representando assim o uso intensivo de Estoque de Capital.*

O fator F5 está relacionado com a variável referente à assistência: X5 (assistência técnica), *representando assim o uso intensivo de assistência técnica.*

#### 4.2 Classificação dos Produtores pelo Índice Tecnológico da Fruticultura Sul Cearense (ITFSC)

Por meio da análise fatorial, após a obtenção dos escores fatoriais extraídos pelo método de rotação *Varimax*, procedeu-se à elaboração do Índice Tecnológico para os 86 fruticultores pesquisados no período de 2014. Em seguida, foi feita a padronização do índice de forma que ele pudesse variar de zero a um. Quanto mais próximo de um, melhores são os níveis tecnológicos desenvolvidos pelo fruticultor.

De acordo com a Tabela 7, percebe-se que foram obtidas cinco classes para classificação de Índices Tecnológicos (ITFSC). Para os índices classificados como muito baixo, de 0 a 0,20, houve dois fruticultores com uma frequência relativa de 2,33%. Observa-se a quantidade de 30 fruticultores com frequência relativa de 34,88% considerados baixos; 53 apresentam índice considerado médio, com frequência relativa de 61,63%, e um fruticultor registrou índice considerado alto, com frequência relativa de 1,16%. Não foram encontrados fruticultores com índices muito altos para adoção de tecnologia.

**TABELA 7 - Índice tecnológico, número de produtores, segundo os grupos do índice de adoção de tecnologia dos fruticultores da região Sul do Ceará**

Grupos	ITFC	Número de produtores	Frequência relativa
1- Muito Baixo	0 - 0,20	2	2,33 %
2- Baixo	0,21-0,35	30	34,88 %
3- Médio	0,37-0,54	53	61,63 %
4- Alto	0,62-0,78	1	1,16 %
5-Muito Alto	0,79-1,00	0	0
Informações válidas	-	86	100,00

Fonte: Resultados da Pesquisa (2014).

As variáveis que mais impactaram na determinação do nível tecnoló-



gico do referido índice foram: X2 (área total), X3 (renda bruta), X4 (produção total), resultado justificado pela maior variância dos dados, num total de 24,058%, representadas pelo fator 1 sendo, *F1 representado pelo uso intensivo da tecnologia Produtividade*.

#### 4.3 Estimação dos escores de eficiência (DEA) e Regressão Quantílica para a identificação dos determinantes da eficiência

A distribuição amostral dos produtores de frutas em classes de eficiência técnica e de escala, sob orientação insumo, encontra-se na Tabela 8. Conforme indicação dos procedimentos de Jackknife e Jackstrap, ao sinalizarem o fato de que dentre as 86 DMUs analisadas, esse módulo foi o que mais exerceu influência dentro de um processo aleatório por meio de Jackstrap. Foi classificada como *outliers* e, portanto retirada da amostra. Assim, o espaço amostral a ser analisado neste estudo é constituído por 85 fruticultores.

Os dados mostram que, dos 85 fruticultores analisados, 70 deles - que corresponde a 82,35% - obtiveram medida de eficiência inferior a 0,6 sob a pressuposição de retornos constantes à escala. Verifica-se também que apenas sete produtores de frutas alcançaram a máxima eficiência.

TABELA 8 - Distribuições absolutas e relativas dos fruticultores da região sul cearense, segundo intervalos de medidas de eficiência técnica e de escala

Medidas de Eficiência	Eficiência Técnica				Eficiência de Escala	
	CRS		VRS		Fi	%
	fi	%	Fi	%		
0 + 0,2	26	30,59	7	8,24	7	8,24
0,2 + 0,4	28	32,94	12	14,12	15	17,65
0,4 + 0,6	16	18,82	14	16,47	24	28,24
0,6 + 0,8	8	9,41	13	15,29	22	25,88
0,8 + 1,0	0	0,00	6	7,06	10	11,76
1,0	7	8,24	33	38,82	7	8,24
Total	85	100,00	85	100,00	85	100,00

Fonte: Resultados da Pesquisa (2014).

As informações sobre os valores dos custos médios anuais excedentes referentes ao uso de insumos e à escala de produção adotada pelos produtores de frutas encontram-se na Tabela 9. Esses dados permitem identificar o nível de ineficiência ensejado por parte de cada um dos insumos utilizados por fruticultor em cada município analisado.

**TABELA 9 - Valores dos custos médios anuais excedentes dos fruticultores da região Sul do Ceará**

Custos médios anuais excedentes					
	Mão-de-obra	Energia	Investimentos	Outros Custos	Insumos
Amostra Total					
VRS	2.981,606	86,42051	3.255,688	238,9847	880,5913
Percentual	14,05	2,73	25,60	21,78	7,90
Amostra Total					
CRS	2.933,45	79,28	2.507,99	238,24	861,03
Percentual	13,82	2,50	19,72	21,72	7,73

Fonte: Resultados da Pesquisa (2014).

Os maiores gastos excessivos privilegiam os investimentos, cujo valor médio anual excedente atinge R\$ 3.255,68 para a amostra total pesquisada, correspondendo a um percentual de 25,60%.

Para identificar variáveis capazes de explicar os índices de eficiência encontrados, realizou-se uma estimação dos parâmetros por meio da regressão quantílica. De acordo com a metodologia o resultado do teste de Kolmogorov-Smirnov ( $T_{K-S}$ ) foi 0,6314, o que permitiu rejeitar a hipótese de retornos variáveis de escala.

Portanto, a regressão quantílica foi estimada considerando os escores de eficiência técnica obtidos no modelo de retornos constantes de escala. Conforme se verifica, existe diferenciação nos determinantes do nível de eficiência técnica entre os produtores de frutas menos eficientes, captados pelo 10º quantil, e os mais eficientes, traduzidos pelo 90º quantil. Para o grupo de fruticultores menos eficientes, observa-se que, apenas índice da gestão, crédito e escolaridade desempenharam papel irrelevante para explicar variações no nível de eficiência técnica.

Os valores dos parâmetros das variáveis - logaritmo da produção, assistência técnica e experiência - foram significativos ao nível de 1% e indicara sinais positivos, sinalizando que o nível de eficiência dos fruticultores pertencentes a esse grupo cresce com o aumento do logaritmo da produção e/ou da assistência técnica e/ou com a experiência.

TABELA 10 - Estimativa da eficiência técnica dos fruticultores da região Sul do Ceará, por meio do modelo de Regressão Quantílica, 2014

Variáveis explicativas	Quantis	
	0,10	0,90
Constante	-0.193* (-2.60)	0.791* (3.60)
Logaritmo da Produção	0.0247* (2.88)	0.0119** (2.13)
Índice de Gestão <sup>1</sup>	0.0371 (1.02)	0.495** (2.09)
		(continuação)
Assistência técnica (dummy)	0.0116* (2.53)	0.0114* (2.04)
Crédito (dummy)	-0.00714 (-0.26)	0.138* (3.53)
Experiência (dummy)	0.016* (3.43)	0.0742 * (3.33)
Escolaridade (dummy)	0.004 (1.02)	0.425* (2.31)

Fonte: Resultados da Pesquisa (2014)

Notas: \*\*\*, \*\* e \* indicam, respectivamente, significativos a 10%, 5% e 1% e os valores entre parênteses referem-se à estatística “t”. O teste de Wald para testar a hipótese de igualdade entre os coeficientes das duas regressões:  $F=42,62$  (prob=0,0000).

Quanto aos fatores explicativos da eficiência técnica para o grupo de produtores mais eficientes, constata-se que todas as variáveis são importantes para explicar a eficiência desse grupo. A *dummy* assistência técnica contribui positivamente para a melhoria de eficiência técnica dos fruticultores pertencentes ao grupo dos mais eficientes.

Os coeficientes das variáveis - logaritmo da produção e índice da gestão - foram significativos ao nível de 5% e denotaram sinais positivos, indicando que o nível de eficiência dos fruticultores pertencentes ao grupo dos mais eficientes cresce com o aumento do logaritmo da produção e/ou do índice da gestão.

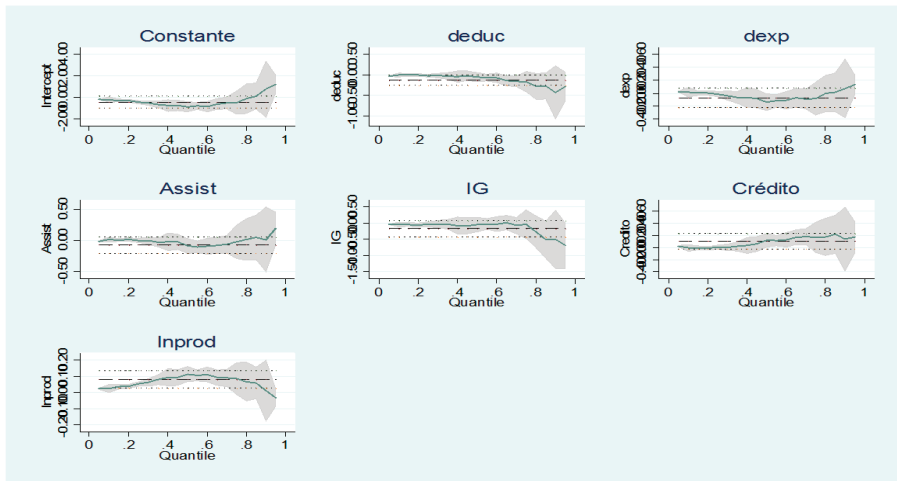
A Figura 1 ilustra o comportamento das variáveis explicativas da eficiência técnica dos produtores de frutas nos vários quantis. Como se observa, o efeito parcial de cada variável explicativa no nível de eficiência varia ao longo dos quantis, sendo que, as variáveis *dummies* assistência técnica, experiência

1. O índice de Gestão foi construído através do seguinte cálculo  $IG = \frac{\sum X_i}{N}$

$X_i$  = os valores das 10 variáveis de gestão que assumem valor 1 ou zero. Neste caso, quanto maior o valor de IG mais técnicas de gestão são adotadas pelo produtor.  $0 \leq IG \leq 1$ .  $IG=1$  indica que o produtor utiliza todas as técnicas de gestão. Se  $IG=0$  indica que o produtor não utiliza nenhuma das técnicas de gestão (aqui está chamando de técnicas de gestão cada uma das 10 variáveis).  $N=10$

e crédito demonstram um diferencial positivo maior para os produtores mais eficientes.

FIGURA 1 - Resultados das Regressões Quantílicas



Fonte: Resultados da Pesquisa (2014).

## 5 CONCLUSÕES

Os dados mostram que, dos 86 fruticultores analisados, 70 deles, o que corresponde a um percentual de 82,35%, obtiveram medida de eficiência inferior a 0,6 (considerada como média eficiência) sob a presunção de retornos constantes à escala. Verifica-se também que, desse total, apenas sete produtores de frutas alcançaram a máxima eficiência. Isso mostra que produtores ineficientes tem maiores dificuldades de desenvolver e gerir sua atividade.

Os resultados indicam, ainda, que os maiores gastos desses fruticultores estão relacionados aos investimentos, onerando assim o seu rendimento, cujo valor médio anual excedente atinge R\$ 3.255,68 para a amostra total pesquisada. Isso corresponde a um percentual de 25,60%. Ademais, analisando os níveis de eficiência desses produtores por meio do modelo de regressões quantílicas, pode-se inferir que os menores níveis de eficiência técnica podem ser explicados pela gestão, o crédito e a escolaridade. Quanto aos fatores explicativos da eficiência técnica para

o grupo de produtores mais eficientes, constata-se que todas as variáveis são importantes para explicar a eficiência desse grupo.

Como se observa, o efeito parcial de cada variável explicativa no nível de eficiência, as variáveis assistência técnica, experiência e crédito apresentam diferenciais positivos maiores para os produtores mais eficientes, ratificando a hipótese de que as práticas agrícolas, as formas de organização e gestão e a assistência técnica são importantes e contribuem intensivamente para maior desempenho tecnológico da fruticultura irrigada na região do Cariri, auxiliando os produtores na tomada de decisão. Por isso é necessário o aprimoramento das políticas públicas voltadas para a assistência técnica dos fruticultores, de forma que possa contribuir ainda mais no desenvolvimento da atividade.

## REFERÊNCIAS BIBLIOGRÁFICAS

ADECE – Agência de Desenvolvimento do Estado do Ceará. **Perfil da produção de frutas Brasil, Ceará.** Governado do Estado do Ceará. Conselho de Desenvolvimento Econômico, 2013.

BANCO DO NORDESTE. **A importância do agronegócio da irrigação para o desenvolvimento do Nordeste.** FRANÇA, Francisco Mavignier Cavalcante (coord.). Fortaleza: Banco do Nordeste, 2001.

BANKER, R.D., CHARNES, H., COOPER, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. **Management Science**, v. 30, n. 9, p. 1078-1092, 1984.

BANKER, R.D., NATARAJAN, R. Statistical tests based on DEA efficiency scores. In: COOPER, W. W.; SEIFORD, L. M.; ZHU, J. (Eds.). **Handbook on Data Envelopment Analysis.** Boston: Kluwer International Series, 2004. p. 299-321.

CARDOSO, C. E. L., SOUZA, J. S. Fruticultura tropical: perspectivas e tendências. **Revista Econômica do Nordeste**, Fortaleza – CE, vol. 31 nº 01 p. 84-95. 2000.

CHARNES, A., COOPER, W.W., LEWIN, A.Y., SEIFORD, L.M. **Data envelopment analysis: theory, methodology and application.** Dordrecht: Kluwer Academic, 1994. 513p.

CHARNES, A., COOPER, W.W., RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, v. 2, n. 6, p. 429-444, 1978.

COELLI, T.; RAO, D.S.P.; BATTESE, G. E. **An introduction to efficiency and productivity analysis**. Norwell: Kluwer Academic, 1998. 275p.

COOPER, W.W., SEIFORD, L. and TONE, K. **Data Envelopment Analysis: a comprehensive text with models, applications, references and DEA-Solver Software**. 3 ed. Norwell, Massachusetts: Kluwer Academic Publishers, 2002.

FÁVERO, L. P.; BELFIORE, P.; SILVA, F. L.; CHAN, B. L. **Análise de dados: modelagem multivariada para tomada de decisões**. Primeira edição. Rio de Janeiro: Elsevier, 2009, p.195-265;

FARREL, M.J. The measurement of productive efficiency. **Journal of the Royal Statistical Society**, Series A, part III, p. 253-290, 1957.

FERREIRA, C. M. C.; GOMES, A. P. **Introdução à Análise Envoltória de Dados: Teoria, modelos e aplicações**. Viçosa, MG: UFV, 2009. 389p. n.18, 2014.

GOMES, A. P.; BAPTISTA, A. J. M. S. Análise Envoltória de Dados. In: SANTOS, M. L., VIEIRA, W.C., (ed.) **Métodos Quantitativos em Economia**. Viçosa, MG: UFV, 2004. p. 121-160.

HAIR JÚNIOR, J. F.; ANDERSON, R. E. ; TATHAM, R. L.; BLACK, W. C. **Análise multivariada de dados**. 5ª ed. Porto Alegre: Bookman, 2005. p.89-126; 380- 419.

JUSTO, W. R. Políticas sociais e o papel nas disparidades regionais de renda no Brasil: evidências a partir de regressões quantílicas. In: Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos, 6, 2008. **Anais...** Aracaju, SE: ENABER, 2008. CDROM.

KOENKER, R.; BASSET, G. Regression quantiles. **Econometrica**, v. 46, p.33-50, 1978.

LEMOS, J. J. S. Indicadores de degradação no Nordeste Sub-úmido e Semi-árido. In: CONGRESSO BRASILEIRO DE ECONOMIA E SOCIOLOGIA RURAL, 34, 2001, Brasília, DF. **Anais...** Brasília, DF:

SOBER, 2001. p.1-10.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada:** uma abordagem aplicada. Belo Horizonte: UFMG, 2005.p. 99-138.

PENSA. **Desenvolvimento da cadeia produtiva de irrigação no Brasil.** São Paulo, 2010. 177p.

**Revista Cariri Invest.** Perfil Econômico da Agroindústria do Cariri. Fortaleza: Banco do Nordeste do Brasil, 1999.

SCHEEL, H. **EMS: Efficiency Measurement System User's Manual.** Version 1.3. 15/08/2000. Disponível em: <<http://www.wiso.uni-dortmund.de/lsg/or/scheel/ems/>>. Acesso em: 19/12/2014.

SCHULTZ, T. W. **A transformação da agricultura tradicional.** Rio de Janeiro: Zahar, 207 p. 1965.

SILVA, S.R., SILVA, L.M.R., KHAN, A.S. Fruticultura e a regionalização da produção agrícola no Estado do Ceará. XLII Congresso da Sociedade Brasileira de Economia, Administração e Sociologia Rural – SOBER. *Anais...* Cuiabá - MT: SOBER, 2004.

SIMPLÍCIO, T. A. **Caracterização socioeconômica do desenvolvimento do setor rural do Nordeste brasileiro.** Universidade Federal do Ceará – UFC, Fortaleza - CE, 1985. 99 p. (Dissertação de Mestrado).

SOUSA, E. P.; JUSTO, W. R.; CAMPOS, A. C. Eficiência Técnica da Fruticultura Irrigada no Ceará. **Revista Econômica do Nordeste**, Fortaleza, v. 44, n. 4, p.851-866, out-dez. 2013.

SOUSA, M. C. S.; STOSIC, B. Technical efficiency of the Brazilian municipalities: correcting non parametric frontier measurements for outliers. **Journal of Productivity Analysis**, v. 24, n. 2, p. 157-181, 2005.

SOUZA, P. M.; PAULO, R. M.; NERY, M. G.; MATA, H.T. Tecnologia de produção na fruticultura desenvolvida em municípios das regiões norte e noroeste do estado do rio de janeiro. XLVII Congresso da Sociedade Brasileira de Economia, Administração e Sociologia Rural – SOBER.

# AVALIAÇÃO DE FATORES SOCIOECONÔMICOS E COMPORTAMENTAIS NOS RESULTADOS DOS ALUNOS NA PROVA BRASIL DE 2011 NOS MUNICÍPIOS CEARENSES

Pedro Fernando Damasceno Rocha\*

Francisco José Silva Tabosa\*\*

Ricardo Brito Soares\*\*\*

Andrei Gomes Simonassi\*\*\*\*

## Resumo

O presente trabalho tem como objetivo identificar os fatores socioeconômicos e comportamentais dos alunos que mais influenciaram nos seus resultados na Prova Brasil de 2011. Para tanto, utiliza-se quatro modelos econométricos do tipo Logit para medir a chance destes alunos atingirem um ponto de corte nas disciplinas de língua portuguesa e matemática no 5º e 9º ano do ensino fundamental. Os resultados mostram a importância das variáveis relacionadas às famílias nos resultados dos alunos. Medidas simples como incentivar os filhos a estudar tiveram mais importância nas estimativas que variáveis relacionadas a posses de bens. Os alunos que afirmaram que os professores passam e corrigem tarefas, bem como os que fazem o dever de casa de forma regular têm probabilidade maior de atingir os resultados almejados, chegando a 16 pontos percentuais a mais de chance para matemática no 9º ano. Não ter reprovação anterior eleva as chances de forma significativa em ambas as séries e disciplinas o que indica a importância de um acompanhamento mais individualizado para as crianças com dificuldades de aprendizado logo nos primeiros anos de vida escolar. Por fim, cabe destacar a grande heterogeneidade dos resultados entre os municípios. Mesmo com PIB semelhantes alguns municípios têm notas médias muito distantes indicando a importância de políticas públicas na área de educação.

**Palavras-Chave:** Prova Brasil. Educação. Modelo Logit.

**JEL:** I21; I26; C25.

---

\* Mestre em economia pelo CAEN/UFC. Auditor da SEFAZ/CE. Rua Princesa Isabel, 1618/402 – Fortaleza/CE – CEP 60015061 – Tel (85) 999737771.

\*\* Economista. Dr. Professor do DEA/MAER/UFC. Campus do Pici, s/n, bloco 826 – Fortaleza/CE – CEP 60000-000 – Tel(85)33669716.

\*\*\* Economista. PhD. Professor DA/DA/UFC. Av. da Universidade, 2700 2º andar – Fortaleza/CE – CEP 60000-000 – Tel(85) 33667751.

\*\*\*\*Economista. PhD. Professor DA/DA/UFC. Av. da Universidade, 2700 2º andar – Fortaleza/CE – CEP 60000-000 – Tel(85) 33667751.



## Abstract

This study aims to identify the socioeconomic and behavioral factors that most influenced the students on their results of the “Prova Brasil 2011”. Therefore, we use four econometric models of the logit type to measure the chance of these students to achieve a cutoff in Portuguese-speaking subjects and mathematics on the 5th and 9th grade of elementary school. The results show the importance of the variables related to families on student outcomes. Simple measures such as encouraging children to study had more importance in estimates than variables related to goods of possessions. Students who reported that teachers give and correct tasks, as well as those who do regularly homework are more likely to achieve the desired results, reaching 16 percentage points more chance for math in grade 9. Having no previous failure increases the likelihood significantly in both series and disciplines which indicates the importance of a more personalized support for children with learning difficulties in the first years of school life. Finally, we highlight the great heterogeneity of results among municipalities. Even with similar GDP some municipalities have very distant middle notes indicating the importance of public policies in education.

**Key words:** Prova Brasil. Education. Logit Model.

## 1 INTRODUÇÃO

Educação pública é um direito social expresso no Artigo 6º da Constituição Federal de 1988 que tem por finalidade diminuir as desigualdades socioeconômicas e culturais e é atribuído a todos os entes federativos. Segundo Barreto (2009), do ponto de vista econômico, um aumento na escolaridade da população pode elevar a produtividade do trabalhador proporcionando um desenvolvimento econômico mais vigoroso e de longo prazo. Para Hanushek (2002), do ponto de vista social, a educação pode reduzir a concentração de renda, reduzir a criminalidade e, inclusive, melhorar a saúde da população.

Algumas políticas públicas na área de educação têm buscado universalizar o acesso à educação, por exemplo: vincular o recebimento de benefícios sociais à manutenção das crianças nas escolas. Mas não basta universalizar o ensino. Para que a sociedade tenha os benefícios econômicos e sociais esperados é preciso que a população receba uma educação de qualidade. Aliás, quando se atrai para a escola crianças de baixa ou baixíssima renda é esperado que o desempenho médio dos alunos seja reduzido por conta do background familiar

trazido por estas crianças e adolescentes (Colleman, 1966). Por conta disso, ganha cada vez mais espaço na sociedade temas como: alfabetização na idade certa, escolas de tempo integral, ingresso na escola pública através de creches ou pré-escolas que de certa forma procuram minorar a ausência do fundamental papel da família na educação dos seus filhos.

Desde 1990, uma importante medida da qualidade do ensino que vem sendo aplicada aos alunos da educação básica é o Sistema de Avaliação da Educação Básica (SAEB) promovida pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Seu objetivo é realizar um diagnóstico do sistema educacional brasileiro e de alguns fatores que possam interferir no desempenho do estudante, fornecendo um indicativo sobre a qualidade do ensino que é ofertado. Atualmente o SAEB é composto de duas avaliações: Avaliação Nacional de Educação Básica (ANEAB) e Avaliação Nacional do Rendimento Escolar (ANRESC), conhecida como Prova Brasil. Enquanto o ANEB é amostral e tem foco na gestão, a Prova Brasil é censitária e fornece informações sobre o ensino oferecido em diversos níveis de agregação. A Prova Brasil foi aplicada pela primeira vez em 2005 e é realizada a cada dois anos. Nas edições que se seguiram alguns pré-requisitos para realização dos testes foram alterados, principalmente no que diz respeito ao número mínimo de aluno por sala de aula e a zona rural.

Os resultados apresentados pelo INEP mostram que houve melhora no desempenho médio das notas dos alunos ao longo das avaliações. Mas mesmo assim, quando se cruza a informação desta média de desempenho com a escala de habilidades proposta pelo SAEB, vê-se que os alunos chegam ao final do ensino fundamental com um domínio sofrível da língua materna. Sabem ler e escrever mas não compreendem o que leem nem sabem se expressar por meio de um texto coeso e coerente.

Apesar do desempenho médio em matemática ser mais elevado que o obtido em língua portuguesa não é menos preocupante já que os retornos sociais e econômicos que se busca alcançar com a elevação da educação dependem do domínio da matéria no mesmo nível dos países desenvolvidos.

As notas médias totais dos alunos cearenses na Prova Brasil 2011 estão acima das dos estudantes nordestinos e um pouco abaixo das médias totais dos estudantes brasileiros, tanto em Língua Portuguesa como em matemática em todas as séries avaliadas. Quando se compara as notas médias destes alunos com as notas dos estados mais desenvolvidos a diferença se acentua ainda mais.

Mas para ter uma visão mais real do tamanho do problema, precisa-se comparar as médias de desempenho dos alunos brasileiros com os demais participantes do teste PISA (Programa Internacional de Avaliação de Alunos) promovido pela Organização para Cooperação e Desenvolvimento Econômico (OCDE). Este teste é aplicado, de forma amostral, aos estudantes na faixa de 15 anos de idade dos países membros da OCDE e a outros convidados, como é o caso do Brasil.

A Tabela 1 apresenta as notas médias obtidas no teste PISA, em 2012, na disciplina de matemática em diversos espaços geográficos para efeito de comparação. A maior nota obtida no teste PISA de 2012 foi a de Shangai-China, com 613 pontos. Mas optou-se por apresentar na tabela apenas países e democracias transparentes. A Coreia do Sul, pela tradição de boas notas em matemática, os Estados Unidos pela relevância econômica e espaço continental, o Chile pela proximidade geográfica. Além disso, inclui-se as notas médias do Brasil como um todo e de São Paulo, por ser o estado mais importante economicamente do país e, finalmente, o Ceará.

## 1 INTRODUÇÃO

Educação pública é um direito social expresso no Artigo 6º da Constituição Federal de 1988 que tem por finalidade diminuir as desigualdades socioeconômicas e culturais e é atribuído a todos os entes federativos. Segundo Barreto (2009), do ponto de vista econômico, um aumento na escolaridade da população pode elevar a produtividade do trabalhador proporcionando um desenvolvimento econômico mais vigoroso e de longo prazo. Para Hanushek (2002), do ponto de vista social, a educação pode reduzir a concentração de renda, reduzir a criminalidade e, inclusive, melhorar a saúde da população.

Algumas políticas públicas na área de educação têm buscado universalizar o acesso à educação, por exemplo: vincular o recebimento de benefícios sociais à manutenção das crianças nas escolas. Mas não basta universalizar o ensino. Para que a sociedade tenha os benefícios econômicos e sociais esperados é preciso que a população receba uma educação de qualidade. Aliás, quando se atrai para a escola crianças de baixa ou baixíssima renda é esperado que o desempenho médio dos alunos seja reduzido por conta do *background* familiar trazido por estas crianças e adolescentes (Coleman, 1966). Por conta disso, ganha cada vez mais espaço na sociedade temas como: alfabetização na idade certa, escolas de tempo integral, ingresso na escola pública através de creches ou

pré-escolas que de certa forma procuram minorar a ausência do fundamental papel da família na educação dos seus filhos.

Desde 1990, uma importante medida da qualidade do ensino que vem sendo aplicada aos alunos da educação básica é o Sistema de Avaliação da Educação Básica (SAEB) promovida pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Seu objetivo é realizar um diagnóstico do sistema educacional brasileiro e de alguns fatores que possam interferir no desempenho do estudante, fornecendo um indicativo sobre a qualidade do ensino que é ofertado. Atualmente o SAEB é composto de duas avaliações: Avaliação Nacional de Educação Básica (ANEAB) e Avaliação Nacional do Rendimento Escolar (ANRESC), conhecida como Prova Brasil. Enquanto o ANEB é amostral e tem foco na gestão, a Prova Brasil é censitária e fornece informações sobre o ensino oferecido em diversos níveis de agregação. A Prova Brasil foi aplicada pela primeira vez em 2005 e é realizada a cada dois anos. Nas edições que se seguiram alguns pré-requisitos para realização dos testes foram alterados, principalmente no que diz respeito ao número mínimo de aluno por sala de aula e a zona rural.

Os resultados apresentados pelo INEP mostram que houve melhora no desempenho médio das notas dos alunos ao longo das avaliações. Mas mesmo assim, quando se cruza a informação desta média de desempenho com a escala de habilidades proposta pelo SAEB, vê-se que os alunos chegam ao final do ensino fundamental com um domínio sofrível da língua materna. Sabem ler e escrever mas não compreendem o que leem nem sabem se expressar por meio de um texto coeso e coerente.

Apesar do desempenho médio em matemática ser mais elevado que o obtido em língua portuguesa não é menos preocupante já que os retornos sociais e econômicos que se busca alcançar com a elevação da educação dependem do domínio da matéria no mesmo nível dos países desenvolvidos.

As notas médias totais dos alunos cearenses na Prova Brasil 2011 estão acima das dos estudantes nordestinos e um pouco abaixo das médias totais dos estudantes brasileiros, tanto em Língua Portuguesa como em matemática em todas as séries avaliadas. Quando se compara as notas médias destes alunos com as notas dos estados mais desenvolvidos a diferença se acentua ainda mais.

Mas para ter uma visão mais real do tamanho do problema, precisa-se comparar as médias de desempenho dos alunos brasileiros com os demais participantes do teste PISA (Programa Internacional de Avaliação de Alunos) promovido pela Organização para Cooperação e Desenvolvimento Econômico

(OCDE). Este teste é aplicado, de forma amostral, aos estudantes na faixa de 15 anos de idade dos países membros da OCDE e a outros convidados, como é o caso do Brasil.

A Tabela 1 apresenta as notas médias obtidas no teste PISA, em 2012, na disciplina de matemática em diversos espaços geográficos para efeito de comparação. A maior nota obtida no teste PISA de 2012 foi a de Shangai-China, com 613 pontos. Mas optou-se por apresentar na tabela apenas países e democracias transparentes. A Coreia do Sul, pela tradição de boas notas em matemática, os Estados Unidos pela relevância econômica e espaço continental, o Chile pela proximidade geográfica. Além disso, inclui-se as notas médias do Brasil como um todo e de São Paulo, por ser o estado mais importante economicamente do país e, finalmente, o Ceará.

**TABELA 1 - Médias de matemática na edição do PISA de 2012 de acordo com os espaços geográficos selecionados.**

Espaço geográfico	Média PISA
Coreia do Sul	554,0
Estados Unidos	481,0
Chile	423,0
Brasil	391,0
São Paulo	403,6
Ceará	378,3

Fonte: OCDE. Elaboração dos autores.

Pela tabela acima, o Ceará está no Nível 1 da escala de proficiência do teste PISA, que varia de “abaixo de 1” até 6. Na verdade, nenhum estado brasileiro atingiu, em 2012, o limite inferior do Nível 2 (420,1 pontos) da escala de proficiência proposta por este teste. No ano de 2012 o Brasil, em matemática, ficou na posição 58 de um *ranking* de 65 participantes. Em leitura ocupou a posição 55 com uma nota média de 410 pontos.

Mesmo considerando que o teste PISA tem como objetivo medir a aquisição de competências (Emília Ferreiro, 2005) ao passo que a educação brasileira mira na perspectiva da aquisição de conteúdo, o resultado não deixa de ser um vexame. Sobretudo quando se sabe que o resultado é ainda pior. No Brasil, ao contrário dos países desenvolvidos, há um grande contingente de adolescentes fora da escola ou em situação de atraso escolar que não foram contabilizados. Pelos critérios da OCDE, alunos com 15 anos que não chegaram a 7ª Série.

É óbvio que a explicação dos resultados acima deve ser buscada de várias formas. Seja através das políticas públicas, desigualdades sociais, cultura, formação e valorização dos docentes, escolha técnica dos gestores, equipamentos pedagógicos adequados, estrutura física das escolas e muitos outros. Culpar apenas as escolas pelos resultados, além de improdutivo, seria irresponsável.

O presente trabalho focará no peso que os insumos relacionados com as características socioeconômicas e comportamentais dos alunos e do ambiente em que os mesmos estão inseridos tiveram nas suas respectivas notas na Prova Brasil do ano de 2011. Além disso, far-se-á uma breve análise descritiva dos dados com destaque nos resultados dos piores e melhores municípios buscando identificar possíveis causas de baixo e alto desempenho dos alunos das escolas públicas cearenses.

O trabalho está dividido em cinco partes, além desta introdução. Na próxima seção será feita uma revisão da literatura que focará nos estudos econômicos na área de educação. Na terceira parte será explicada a metodologia utilizada, base de dados, variáveis, espaço temporal e geográfico e modelos econométricos utilizados. Na quarta parte analisa-se os resultados encontrados juntamente com a discussão dos dados. Na quinta e última parte será feita a conclusão do trabalho.

## 2 REFERENCIAL TEÓRICO

O primeiro grande estudo que buscou relacionar o desempenho dos estudantes com um conjunto de insumos educacionais (características individuais e familiares dos alunos, características da escola e características da comunidade que frequenta a mesma escola) ficou conhecido como Coleman Report. Esse estudo, realizado em 1966 nos Estados Unidos e patrocinado pelo Governo Americano, foi conduzido por James S. Coleman e colaboradores, e tinha como hipótese inicial a idéia de que a diferença de desempenho dos alunos negros e brancos em testes padronizados era devido à disparidade na qualidade da escola frequentada pelos dois grupos étnicos.

No entanto, os resultados apresentados não se mostraram estatisticamente significativos para a hipótese levantada pelos pesquisadores. Ficou comprovado que os principais determinantes do desempenho dos alunos nas provas eram as características das famílias dos próprios alunos (*background* familiar).

É fácil imaginar o grande impacto que um resultado como este pode

ter nas políticas públicas de alocação de recursos para a educação. Seria mais produtivo criar condições econômicas favoráveis para que as famílias se desenvolvessem do que investir pesado em criação e reforma de escolas e contratação e aperfeiçoamento de professores e diretores. Muitas discussões se seguiram e vários estudos foram realizados com diferentes metodologias para tentar explicar ou contradizer os resultados obtidos por Coleman (Hanushek, 1986).

Nos últimos anos, muitos estudos têm conseguido, através de novas abordagens metodológicas e de posse de novos instrumentos de coleta de dados, produzir trabalhos que contradizem a visão pessimista do relatório Coleman que afirmava, como já dito anteriormente, que apenas os fatores individuais e familiares dos alunos eram importantes no desempenho destes em testes padronizados. Mas todos os estudos são unânimes em apontar a grande importância do *background* familiar no desempenho dos alunos.

Rivkin, Hanushek e Kain (2005), a partir de um conjunto de dados em painel, conseguiram identificar a importância dos professores para o aprendizado dos alunos. Este trabalho não identificou quais atributos dos professores constantes dos dados foram responsáveis pela elevação do desempenho dos alunos, mas isso não o invalida já que algumas características como didática e capacidade de motivação são de difíceis mensuração.

É razoável supor, por exemplo, que o grau de escolaridade do professor, sua didática, paciência e disciplina tenham importâncias diferentes ao longo da vida estudantil. Alunos do 5º ano do ensino fundamental (a primeira avaliação do SAEB é feita com este grupo) estão em um processo cognitivo muito mais verticalizado e expositivo do que os do 9º ano que já interagem e questionam os professores.

Argumentos semelhantes podem ser usados para a falta de evidência estatística de que salários maiores para os professores podem aumentar o desempenho dos alunos, pois estes podem estar associados muito mais ao tempo de serviço e o grau de escolaridade do professor do que a outros fatores. MENEZES FILHO (2007) afirma que a diferença salarial dos professores só faz diferença na rede privada de ensino e supõe que isto se deva ao fato de que, na iniciativa privada, os melhores professores obtêm os melhores salários, o que não se verifica na rede pública de ensino.

Aprofundando o tema sobre as características dos professores que podem influenciar o desempenho dos alunos, FERNANDES e FERRAZ (2014) utilizando dados da Secretaria de Educação de São Paulo, testaram o impacto de

duas variáveis críticas sobre o ensino na 8ª Série daquele estado: 1) O domínio da disciplina (os professores são avaliados por provas) e 2) práticas pedagógicas adotadas na sala de aula. Estas duas dimensões sobre a qualidade dos professores foram testadas utilizando uma função de produção educacional com valor adicionado e controlando de forma muito consistente, do ponto de vista estatístico, os demais fatores que poderiam influenciar o desempenho dos alunos.

Os resultados mostraram que “tanto o conhecimento da disciplina ministrada como a adoção frequente de práticas pedagógicas na sala de aula tem efeitos positivos e estatisticamente significantes sobre a aprendizagem dos alunos”. Contudo, a análise demonstrou que as práticas pedagógicas adotadas pelos professores tinham um impacto muito maior do que o conhecimento da matéria. E são práticas simples como passar e corrigir o dever de casa. Estas e outras variáveis estão presentes nos dados coletados do SAEB e serão exploradas nos modelos de análise propostos.

FELÍCIO E FERNANDES (2005) utilizando dados do SAEB 2001 propõem duas formas alternativas de avaliar o efeito escola no desempenho dos alunos e assim contornar os problemas clássicos que levam a subestimação dos insumos escolares. No primeiro modelo é feito a decomposição da desigualdade de notas em duas parcelas, uma explicada pelas características específicas das crianças e de suas famílias (intra-escolas), e outra que deve ser explicada pelas diferenças entre as escolas em que estudam tais crianças (entre escolas). A decomposição é feita utilizando o índice L de Theil, HOFFMAN (1998), que possibilita a separação das parcelas entre grupos e intragrupos. Chegaram à conclusão 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.

No segundo modelo proposto pelos mesmos autores é incluída uma variável dummy indicativa de escola, que permite estimar por efeitos fixos a importância de cada escola no aprendizado dos alunos e estabelecer um ranking da qualidade das escolas. “A partir desse ranking, foram feitas simulações em que se supõe que as crianças estudam em uma escola representativa da média das cinco melhores escolas. Tanto para as notas de Língua Portuguesa quanto de Matemática, as simulações utilizando todas as escolas apresentam um impacto de um desvio-padrão, o que ilustrativamente significa elevar o desempenho médio das 4ª série ao obtido pelas 7ª séries.”

Como bem lembrado por MENEZES FILHO (2007), é importante destacar que nenhuma política pública na área de educação pode fechar os olhos para as realidades individuais de cada família, de cada escola e de cada



município. O simples fato de uma criança estar na escola, mesmo com baixo rendimento escolar, é melhor do que ficar completamente à margem do sistema educacional sujeito à marginalidade ou ao uso de drogas.

Atualmente, muitas famílias espalhadas pelos rincões deste imenso país ou mesmo em comunidades carentes de grandes centros urbanos mantêm os filhos na escola apenas para receber benefícios sociais do Governo Federal ou para garantir alguma refeição de qualidade para os mesmos.

É óbvio que pessoas mergulhadas historicamente no caldeirão da ignorância não conseguem enxergar os benefícios da educação para seus filhos. E aí entra uma importante variável, de difícil mensuração e por isso quase sempre ausente nos diversos modelos econométricos, que é a motivação. Como uma pessoa vai incentivar outra a fazer determinada tarefa se para ela aquilo não faz o menor sentido? Juntamente com o esforço, dedicação e disciplina a motivação depende muito do ambiente familiar. No questionário socioeconômico do SAEB há variáveis de caráter motivacional que serão incluídas nos modelos econométricos com a intenção de captar os efeitos psicológico do incentivo dos pais nos resultados dos filhos.

### 3 METODOLOGIA

O presente estudo será realizado com o menor nível de agregação possível do SAEB, ou seja, os alunos. A partir destes microdados, pode-se expandir para as escolas, municípios e estado ao passo que o caminho contrário não é possível. Importante destacar que mesmo neste nível de agregação não é possível fazer a comparação entre alunos já que estes não fazem a mesma prova. Mas como o objetivo principal é a comparação do desempenho do aluno com suas próprias características socioeconômicas este obstáculo está superado.

Segundo a Lei 9.394/1996 que estabelece as diretrizes e bases da educação nacional, já com as recentes alterações promovidas pela Lei 12.796/2013, a educação básica é composta pela pré-escola, ensino fundamental e ensino médio. Nestas três fases cabe aos estados atuarem prioritariamente no nível médio e aos municípios prioritariamente no ensino fundamental. Mas todos os entes tributantes têm a faculdade de atuar nos diversos níveis desde que cumpram suas obrigações prioritárias.

Verificou-se neste trabalho que há 120 alunos no 9º ano do ensino fundamental que têm como dependência administrativa a União. Além disso há

1.658 alunos do 5º ano e 24.464 no 9º ano que têm vinculação com o Estado. Principalmente os alunos com dependência no nível Federal merecem um estudo à parte porque são alunos altamente selecionados e com um nível socioeconômico muito diferente dos demais. Nos modelos econométricos propostos para os alunos do 5º ano, incluir-se-á apenas alunos da rede municipal de ensino, ao passo que as estimativas para o 9º ano levarão em conta alunos da rede municipal e estadual. Neste caso, uma das variáveis explicativas será a dependência administrativa.

Serão utilizados os dados da Prova Brasil no ano de 2011 já que estes são os últimos dados disponíveis na forma de microdados para *download* no endereço eletrônico do INEP. Posteriormente, acrescentou-se as demais variáveis coletadas através dos questionários socioeconômicos aplicados aos alunos. Com a intenção de captar o “efeito escola”, fez-se um novo pareamento de banco de dados onde se acrescentou para cada registro de aluno a nota média obtida por sua escola naquele ano para cada disciplina. Optou-se por restringir o banco de dados para o estado do Ceará com o objetivo de facilitar a interpretação dos dados já que o estudo em conjunto de todos os estados, com todas as suas peculiaridades tanto orçamentárias como culturais dificultaria imensamente as análises dos dados.

O atual Plano Nacional de Educação, Lei 13.005/2014, estabelece em seu Art. 11, §1º, II, que deverão ser divulgados indicadores de avaliação do perfil do alunado, do corpo docente, infraestrutura da escola, entre outros. Certamente estes indicadores facilitarão a comparação entre as escolas já que conterão em seu bojo parte importante do *background* familiar que, como já explicado anteriormente, tem forte impacto no desempenho dos alunos.

Obviamente, que quando se opta por agrupar variáveis em indicadores ou índices se perde o valor o individual delas. Por conta disso, e porque estamos trabalhando apenas com os dados da Prova Brasil e no nível de agregação do indivíduo é que não será incluído, nestes modelos, índices e sim a forma primitiva das variáveis. Outros dados como salários dos pais, por exemplo, devem ser buscados no ENEM e na ANEB o que fugiria ao escopo do nosso trabalho.

Os resultados da avaliação da educação básica são apresentados em uma escala de desempenho capaz de descrever as competências e as habilidades que os alunos são capazes de demonstrar em cada nível. Esta metodologia conhecida como Teoria de Resposta ao Item-TRI permite apresentar em uma mesma métrica os resultados de desempenhos dos estudantes do 5º e do 9º ano do Ensino Fundamental e 3º ano do Ensino Médio.

Pela escala, pode-se verificar que percentual de alunos já construiu as competências e habilidades desejáveis para cada uma das séries avaliadas, quantos ainda estão em processo de construção, quantos estão abaixo do nível que seria desejável para a série e quantos estão acima do nível que seria esperado (INEP, 2011). Este é o principal motivo para se manter no estudo tanto os alunos do 5º ano como do 9º. Pretende-se verificar quais alunos do 5º ano do ensino fundamental já adquiriam habilidades semelhantes ao do 9º ano.

A partir dos acertos dos alunos na Prova Brasil as notas são padronizadas na escala SAEB tendo como referência a média de 250 e desvio-padrão de 50 para alunos do 9º ano (à época, 8ª Série) do Ensino Fundamental em 1997. Isso significa que um aluno do 5º ano com pontuação em matemática igual a 175 possui um desempenho nesta disciplina 1,5 desvios padrão abaixo da média de desempenho em matemática dos alunos do 9º ano em 1997 (FELÍCIO E FERNANDES, 2005).

Importante destacar que além da escala de 11 níveis de conhecimento proposta pelo INEP, vários autores propõem subdivisões pedagógicas desta escala. Inclusive com marcações diferentes para matemática e para língua portuguesa. Como o objetivo do presente trabalho não é classificar os alunos em níveis de conhecimento e sim estabelecer um ponto de corte, utilizaremos os pontos de referência de 250 para o 9º ano e 175 para o 5º ano, porque estão mais de acordo com a realidade dos alunos cearenses (próximos a média ou mediana).

A probabilidade de um aluno ter um desempenho considerado adequado pelo INEP é um caso de típico de modelo de escolha binária porque a variável resposta terá como resultado: sucesso ( $y=1$ : aluno atingiu o nível desejado) ou insucesso ( $y=0$ : aluno não atingiu o nível desejado). Assim, o modelo abaixo proposto tem como objetivo determinar a probabilidade de um aluno, levando em conta uma série de atributos (Tabela 2), atingir um nível de conhecimento considerado adequado para seu nível de escolaridade.

Segundo GUJARATI (2011), um modelo de probabilidade linear, estimada por mínimos quadrados ordinários, pode ser tendencioso. Ou seja, pode subestimar ou superestimar a inclinação da reta. Desta forma, optou-se por utilizar um modelo de resposta binária do tipo Logit que utiliza uma função de probabilidade acumulada que restringe os valores preditos no intervalo entre 0 e 1 (estimadores não tendenciosos). Em relação à escolha do Logit em detrimento do Probit, destaca-se que a escolha foi meramente por conveniência interpretativa uma vez que, segundo a literatura citada anteriormente, em gran-

des amostras, como é o presente estudo, não faria diferença estatística a escolha entre qualquer um destes modelos.

Tendo em vista que vários fatores de forma conjunta podem influenciar o desempenho dos alunos e possivelmente de forma diferente nas séries estudadas e disciplinas, utilizaremos 4 modelos econométricos do tipo Logit para medir a probabilidade dos alunos atingirem um desempenho aceitável. Aos alunos do 5º ano do ensino fundamental submeteremos o Modelo 1 (Língua Portuguesa) e o Modelo 2 (Matemática) tendo como ponto de corte a pontuação de 175. Os alunos do 9º ano serão submetidos ao Modelo 3 para língua portuguesa e ao Modelo 4 para matemática mas tendo como referência a pontuação de 250. Os modelos de regressão do tipo probabilístico terão a seguinte características:

$$P_i = E(Y = 1 / X_i) = \frac{1}{1 + e^{-Z_i}} \quad (1)$$

Onde:

$P_i$  = Probabilidade do aluno se encontrar no nível de desempenho desejado

$$Z_i = \sum_k \alpha_k DG_{ik} + \sum_k \alpha_k PB_{ik} + \sum_k \alpha_k EP_{ik} + \sum_k \alpha_k MT_{ik} + \sum_k \alpha_k HL_{ik} \\ + \sum_k \alpha_k TE_{ik} + \sum_k \alpha_k DE_{ik} + \sum_k \alpha_k DC_{ik} + \sum_k \alpha_k NE_{ik} + \sum_k \alpha_k DA_{ik} + X\beta$$

Sendo:

DG = Variáveis referentes ao Dados Gerais dos alunos;

PB = Variáveis das Posses de Bens das famílias dos alunos;

EP = Variáveis referentes à Escolaridade dos Pais;

MT = Variáveis referentes as Motivações que os pais/responsáveis Transmitem aos alunos;

HL = Variáveis referentes aos Hábitos de Leitura dos alunos;

TE = Variáveis sobre o Tempo disponível para os Estudos;

DE = Variáveis sobre os Dados Escolares progressos;

DC = Variáveis sobre Dever de Casa e outras práticas pedagógicas dos professores;

NE = Nota média da Escola onde o aluno estuda;

DA = Dependência Administrativa da escola.

Dentre as variáveis que constituem os grupos acima, disponíveis nos questionários aplicados pelo INEP, foram incluídas nas estimativas aquelas que tiveram alguma influência estatística ou algum peso teórico no desempenho dos alunos. A Tabela 2 abaixo discrimina as variáveis explicativas utilizadas nas estimativas, bem como suas médias.

**TABELA 2 - Descrição das variáveis explicativas, codificações e médias Prova Brasil, 2011, Ceará**

Variável	Codificação	Média
Sexo	1=Masculino; 0=Feminino	0,4972
Negro	1=Negro; 0=Demais	0,0742
Idade	Idade em anos	12,9848
Computador+internet	0=Não; 1=Sim	0,1664
Escolaridade mãe <4EF	0=demais;1=não completou a 4ª série	0,2243
Escolaridade mãe 4-8EF	0=demais;1=completou a 4ª mas não a 8ª série	0,2520
Escolaridade mãe 8-EM	0=demais;1=completou a 8ª mas não o EM	0,1940
Escolaridade mãe EM-ES	0=demais;1=completou o EM mas não ES	0,1684
Escolaridade mãe ES	0=demais;1=completou o Ensino Superior	0,0773
Pais incentivam a estudar	0=não; 1=sim	0,9654
Lê livros em geral as vezes	0=demais;1=Sim	0,4836
Lê livros em geral sempre	0=demais;1=sim	0,4021
Vai ao cinema as vezes	0=demais;1=sim	0,2460
Vai ao cinema sempre	0=demais;1=sim	0,1033
Vai música/dança as vezes	0=demais;1=sim	0,4193
Vai música/dança sempre	0=demais;1=sim	0,3015
Trabalha fora de casa	0=Não; 1= Sim	0,1775
Aprovações anteriores	1=Sempre; 0=1 ou + reprovações	0,6303
Faz dever de L. Portu. as vezes	0=demais;1=sim	0,2945
Faz dever de L. Portu. sempre	0=demais;1=sim	0,6630
Faz dever de Matem. as vezes	0=demais;1=sim	0,2996
Faz dever de Matem. sempre	0=demais;1=sim	0,6542
Corrige L.Portu. as vezes	0=demais;1=sim	0,1572
Corrige L.Portu. sempre	0=demais;1=sim	0,8184
Média Língua Portuguesa		206,52
Média Matemática		217,3
Dependência administrativa	1=Estado; 0=Município	0,1282

Fonte: dados fornecidos pelo INEP. Elaboração dos autores.

Salvo as próprias exclusões de dados propostas pelo INEP no que diz respeito ao “Indicador de Consistência”, “Indicador de Preenchimento” e “Indicador para Cálculo de Proficiência”, além dos já citados alunos com dependência administrativa federal (9º ano) e estadual (5º ano), optamos por manter nas análises todos os alunos, mesmo tendo consciência da presença de *outliers*. Para contornar os problemas intrínsecos a manutenção destes dados discrepantes utilizar-se-á um modelo robusto de regressão.

## 4 ANÁLISE DOS RESULTADOS

### 4.1 Análise descritiva

Os gráficos contidos nas Figuras 1 e 2 foram elaboradas utilizando os mesmos critérios para inclusão de dados nos modelos econométricos, o que pode levar a pequenas diferenças de resultados em consultas diretas ao endereço eletrônico do INEP. Calculou-se as médias das notas dos alunos por município para cada disciplina e série estudada. O município de Fortaleza não pertence nem aos 10 piores nem aos 10 melhores, foi inserido apenas para efeito de comparação.

Pode-se observar uma enorme heterogeneidade dos resultados entre os municípios. Há municípios com PIB pequeno, por exemplo, Ibaretama e Mucambo figurando em lados completamente opostos. Além disso, excluindo Sobral que merece comentários à parte, percebe-se a ausência das grandes cidades entre as melhores, tais como: Fortaleza, Caucaia, Maracanaú e Juazeiro do Norte. Provavelmente estas ausências se devem às falhas de governo e outras assimetrias que fazem com que municípios com maior PIB não consigam alocar de forma correta os recursos da educação.

É óbvio que municípios maiores têm outros problemas que direta ou indiretamente acabam atrapalhando os resultados na educação, por exemplo a favelização, concentração de renda e a violência. Mas os resultados acima mostram que Sobral passou por cima destes problemas e obteve resultados pujantes nas duas séries avaliadas.

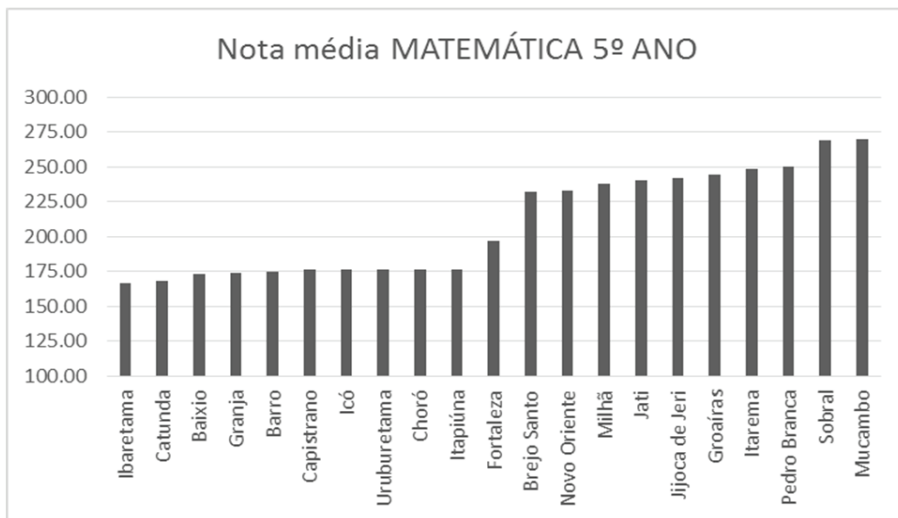
Em números absolutos Mucambo teve os melhores resultados. Média de 270,04 pontos em matemática no 5º ano (222 participantes) e 276,64 pontos no 9º ano (287 participantes). Mas Sobral, por ser um município polo, com todos os problemas intrínsecos aos grandes centros urbanos e com grande número de alunos realizando a Prova Brasil (2.533 participantes no 5º ano e 3.068

no 9º ano) foi o que mais chamou atenção. Sobral é um dos municípios onde o estado também atua no ensino fundamental. Se separarmos os 3.068 alunos do 9º ano, teremos: 2.558 alunos com média de 244,5 da rede estadual e 510 da rede municipal com 259,2 pontos.

Outro dado interessante das Figuras 1 e 2 é que na ponta de baixo há uma grande alternância entre as 10 piores notas. Ao passo que os 10 melhores parecem que conseguiram se desvencilhar do atraso na educação e avançam em ambas as séries. Pode-se observar também que nos municípios do topo da lista os alunos do 5º ano já adquiriram, em média, habilidades de resolver problemas de matemática semelhantes aos do 9º ano.

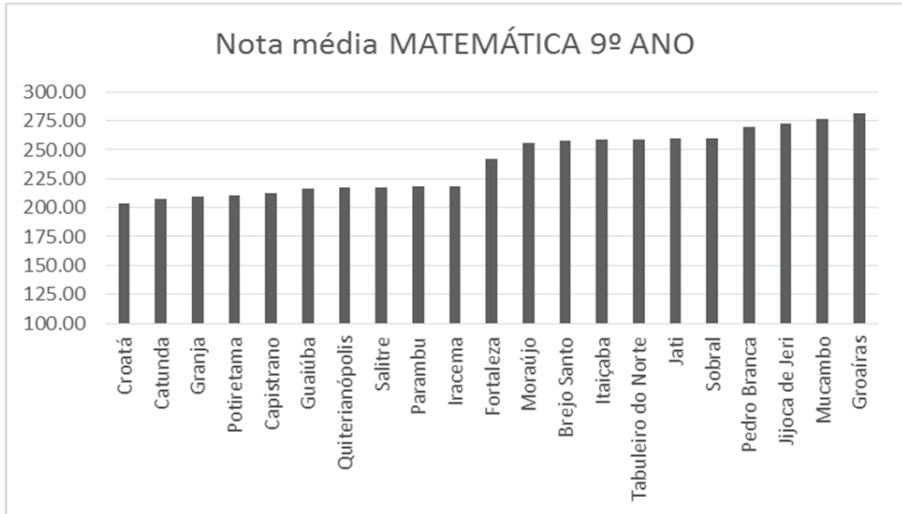
De forma individual, cabe destacar 55 alunos da rede municipal de ensino que ainda no 5º ano do ensino fundamental, fizeram mais de 350 pontos na prova de matemática. Sendo que destes, 6 eram de Caucaia e 6 de Fortaleza. E no 9º ano, 160 alunos (41 rede estadual e 119 rede municipal) obtiveram mais de 380 pontos, com destaque para Fortaleza com 36 participantes deste seletor grupo. Importante informar que estes alunos estão dispersos em diversas escolas destes municípios.

**FIGURA 1 - Distribuição das 10 melhores e 10 piores notas médias em matemática no 5º ano nos municípios cearenses, além de Fortaleza. Prova Brasil, 2011**



Fonte: Inep. Elaboração dos autores.

**FIGURA 2 - Distribuição das 10 melhores e 10 piores notas médias em matemática no 9º ano nos municípios cearenses, além de Fortaleza. Prova Brasil, 2011**



Fonte: Inep. Elaboração dos autores.

#### 4.2 Análise econométrica

Os resultados dos modelos multivariados de probabilidade de alcance de níveis de desempenho desejados na Prova Brasil no ano de 2011 estão expostos nas Tabelas 3 e 4. Ambas as tabelas estão divididas em colunas com as duas disciplinas aplicadas aos alunos: Língua Portuguesa e Matemática. A Tabela 3 resume os dados dos alunos do 5º ano do ensino fundamental das escolas públicas municipais e a Tabela 4 os dados dos alunos do 9º ano das redes públicas municipal e estadual.

Dentro de uma mesma tabela as únicas variáveis que diferem entre os dois modelos são as relacionadas ao “dever de casa”, “professor corrige o dever” e a “nota média da escola” pois cada um se refere à disciplina aplicada. Além disso, na Tabela 4 foi incluída a variável “dependência administrativa” para captar a diferença entre escolas municipais e estaduais. Após as estimativas de cada um dos 4 modelos, utilizando o *Software* Stata Intercooled na versão 12.0, calculou-se o efeito marginal, através do comando “mfx” para facilitar a interpretação dos resultados em termos de chances ou pontos percentuais.



**TABELA 3 - Modelos Logit de probabilidade para os alunos do 5º ano do ensino fundamental da rede pública municipal cearense de ensino, segundo a nota de corte de 175 pontos na Prova Brasil em 2011**

Variáveis	Modelo 1 (Língua Portuguesa 5º Ano)		Modelo 2 (Matemática 5º Ano)	
	Coeficiente	E. Marginal	Coeficiente	E. Marginal
Aluno Homem	-0,3631(14,93)*	-0,0826	0,3552(13,15)*	0,0571
Aluno Negro	-0,3094(6,86)*	-0,0730	-0,3841(8,22)*	-0,0678
Idade	-0,1127(9,82)*	-0,0257	-0,1000(8,25)*	-0,0160
Computador+internet	0,1173(3,49)*	0,0264	0,1039(2,78)*	0,0163
Escolaridade mãe <4EF	0,1395(2,89)*	0,0314	0,2974(5,96)*	0,0453
Escolaridade mãe 4-8EF	0,2616(5,53)*	0,0584	0,4039(8,26)*	0,0610
Escolaridade mãe 8-EM	0,2885(5,87)*	0,0640	0,3775(7,40)*	0,0566
Escolaridade mãe EM-ES	0,6522(12,36)*	0,1372	0,7116(12,63)*	0,0982
Escolaridade mãe ES	0,1646(2,94)*	0,0368	0,2514(4,23)*	0,0380
Pais incentivam estudar	0,7558(11,28)*	0,1838	0,6249(9,64)*	0,1175
Lê livros as vezes	0,3690(8,94)*	0,0832	0,3237(7,52)*	0,0511
Lê livros sempre	0,3816(9,25)*	0,0868	0,2817(6,53)*	0,0452
Vai ao cinema as vezes	-0,1588(5,47)*	-0,0366	-0,0615(1,94)***	-0,0099
Vai ao cinema sempre	-0,5202(12,98)*	-0,1243	-0,3828(9,02)*	-0,0670
Vai show as vezes	0,0041(0,13)	0,0009	0,0678(2,02)**	0,0108
Vai show sempre	-0,2451(7,74)*	-0,0564	-0,2511(7,34)*	-0,0411
Trabalha fora de casa	-0,6519(19,09)*	-0,1562	-0,4399(12,46)*	-0,0773
Aprovações anteriores	0,6996(24,43)*	0,1638	0,6093(19,42)*	0,1043
Faz dever as vezes	0,6202(8,53)*	0,1327	0,3562(4,38)*	0,0537
Faz dever sempre	0,9087(12,89)*	0,2159	0,5948(7,50)*	0,1045
Corrige dever as vezes	0,2741(3,05)*	0,0607	0,3810(4,01)*	0,0564
Corrige dever sempre	0,3837(4,43)*	0,0924	0,5121(5,59)*	0,0905
Média disciplina	0,0382(45,40)*	0,0087	0,0395(45,27)*	0,0063

Estatística robusta Z entre parênteses. \* = significante a 1%; \*\*=significante a 5%; \*\*\*=significante a 10%  
 Categorias base: sexo feminino, não negro, não possui computador com acesso à internet, mãe nunca estudou, não incentiva a estudar, nunca ou quase nunca lê livros, nunca ou quase nunca vai ao cinema, nunca ou quase nunca vai à apresentação de música ou de dança, não trabalha fora de casa, nunca foi reprovado, nunca ou quase nunca faz o dever de casa, professor nunca ou quase nunca corrige o dever de casa.

**TABELA 4 - Modelos Logit de probabilidade para os alunos do 9º ano do ensino fundamental da rede pública municipal e estadual cearense de ensino, segundo a nota de corte de 250 pontos na Prova Brasil de 2011**

Variáveis	Modelo 3 (Língua Portuguesa 9º Ano)		Modelo 4 Matemática 9º Ano	
	Coefficiente	E. Marginal	Coefficiente	E. Marginal
Aluno Homem	-0,3796(19,28)*	-0,0888	0,6031(30,57)*	0,1468
Aluno Negro	-0,0889(2,32)**	-0,0207	-0,1413(3,70)*	-0,0341
Idade	-0,1585(15,14)*	-0,0373	-0,1893(18,41)*	-0,0462
Computador+internet	0,2892(11,45)*	0,0693	0,2375(9,35)*	0,0585
Escolaridade mãe <4EF	0,0919(2,13)**	0,0217	0,1616(3,81)*	0,0396
Escolaridade mãe 4-8EF	0,3400(8,0)*	0,0814	0,3533(8,43)*	0,0870
Escolaridade mãe 8-EM	0,3449(7,83)*	0,0829	0,3918(9,03)*	0,0968
Escolaridade mãe EM-ES	0,5402(12,12)*	0,1308	0,5899(13,39)*	0,1458
Escolaridade mãe ES	0,5014(8,88)*	0,1225	0,6032(10,66)*	0,1496
Pais incentivam estudar	0,5822(7,24)*	0,1254	0,4320(5,68)*	0,1008
Lê livros as vezes	0,2093(6,14)*	0,0491	0,1445(4,38)*	0,0352
Lê livros sempre	0,4667(12,94)*	0,1111	0,1965(5,60)*	0,0481
Vai ao cinema as vezes	-0,1012(4,44)*	-0,0237	-0,1167(5,13)*	-0,0283
Vai ao cinema sempre	-0,2833(8,12)*	-0,0646	-0,3927(11,16)*	-0,0928
Vai show as vezes	-0,2559(11,44)*	-0,0600	-0,1531(6,88)*	-0,0373
Vai show sempre	-0,5465(20,66)*	-0,1236	-0,4535(17,24)*	-0,1082
Trabalha fora de casa	-0,2846(10,90)*	0,0654	-0,1041(4,12)*	-0,0252
Aprovações anteriores	0,6874(27,94)*	0,1563	0,6791(27,94)	0,1614
Faz dever as vezes	-0,1001(1,59)	-0,0235	0,1796(3,66)*	0,0440
Faz dever sempre	0,1120(1,78)**	0,0262	0,6786(13,98)*	0,1623
Corrige dever as vezes	-0,0290(0,32)	0,0068	0,1553(1,90)**	0,0382
Corrige dever sempre	0,1315(1,49)	0,0306	0,3067(3,92)*	0,0733
Média disciplina	0,0372(53,58)*	0,0087	0,0369(59,31)*	0,0090
Dependência adm	-0,0378(1,66)**	-0,0089	-0,0461(2,04)**	-0,0112

Estatística robusta Z entre parênteses. \* = significante a 1%; \*\*=significante a 5%; \*\*\*=significante a 10%

Categorias base: sexo feminino, não negro, não possui computador com acesso à internet, mãe nunca estudou, não incentiva a estudar, nunca ou quase nunca lê livros, nunca ou quase nunca vai ao cinema, nunca ou quase nunca vai à apresentação de música ou de dança, não trabalha fora de casa, nunca foi reprovado, nunca ou quase nunca faz o dever de casa, professor nunca ou quase nunca corrige o dever de casa, dependência administrativa municipal.

Como podemos observar nas tabelas acima, o gênero influencia o desempenho dos alunos nas disciplinas estudadas de forma inversa. Em língua portuguesa, ser do sexo masculino reduz em 8,2 pontos percentuais a chance deste obter a nota do ponto de corte no 5º ano, ao passo que em matemática eleva a chance em 5,7 pontos percentuais. E isso segue ao longo do ensino fundamental. No 9º ano, os meninos têm 8,9 pontos percentuais a menos de chance de atingirem o perfil desejado em leitura do que as meninas e em matemática a diferença se acentua ainda mais. Ser do sexo masculino eleva a chance de atingir a pontuação 250 em quase 15,0 pontos percentuais. Estas diferenças de gênero foram estatisticamente significantes no presente estudo e são corroboradas por estudos nacionais e internacionais.

Os alunos que se autodeclararam negros tiveram desempenho inferior aos demais em todas as disciplinas e em todas as séries. Sendo que esta diferença se mostrou mais acentuada, superior a 6,0 pontos percentuais, no 5º ano. Mas esta variável merece ser analisada com cautela já que se trata de uma auto percepção e que mais de 50% dos alunos se declararam pardos. Outra questão importante é que não se trata da cor da pele que leva ao menor desempenho em média destes alunos e sim a outras características familiares ou pessoais (observáveis ou não) tais como: baixa renda da família, baixo nível de escolaridade dos pais, discriminação, baixa autoestima e outros.

O aumento de um ano em relação a idade escolar correta provoca a redução de chance dos alunos obterem bons resultados nas provas. E isso se acentua a medida que eles avançam no ensino fundamental. Certamente em decorrência de repetências, abandonos e acúmulo de conteúdo. No 9º ano a chance do aluno ter um bom resultado em matemática, caso não esteja na idade correta, é reduzido em quase 5,0 pontos percentuais.

A variável escolhida para representar a “Posse de Bens/Habitação” foi a presença na residência do aluno de computador com acesso à internet. Este item, além de representar um bem com um custo fixo mensal associado, também pode interferir na aprendizagem do aluno já que permite, através da internet, acesso à informação. Além disso não se trata de um bem tão comum como tv em cores ou geladeira. Dos 203.849 alunos, apenas 33.386 (16,38%) acessam à internet de suas residências através de computadores. E como era esperado o sinal desta variável foi positivo e estatisticamente significativo nas duas séries e disciplinas, ou seja, alunos que dispõem deste bem tiveram notas melhores que os demais. E a importância deste item aumenta à medida que o aluno avança na vida escolar já que os efeitos marginais no 9º ano são muito

maiores, em ambas as disciplinas, que no 5º ano. Isso sugere que alunos com uma maior maturidade tiram um melhor proveito da internet.

Conforme visto, a importância do *background* familiar no desempenho do aluno é quase uma unanimidade na literatura. E dentro do amplo leque de variáveis, mensuráveis ou não, que compõem este grande grupo a escolaridade dos pais é uma das mais importantes. E mais especificamente ainda, a escolaridade da mãe. A mãe muitas vezes acumula sozinha, juntamente com outros afazeres domésticos, o papel de acompanhar as tarefas dos filhos.

Outro fato que motivou a escolha apenas da escolaridade da mãe para inclusão no modelo foi a enormidade de alunos que responderam que não sabem até que série seu pai estudou. São ao todo 77.140 alunos que não sabem esta informação. Somados aos dados em branco (“.”) com as marcações anuladas (“\*”) são 97.305 alunos onde não foi possível identificar a escolaridade do pai do aluno. No caso das mães, este número total ficou em 72.497.

A partir da variável referente à escolaridade da mãe, criaram-se *dummies* para medir o impacto que a quantidade de estudo da mãe tem no resultado do seu filho. Como era esperado a priori, quanto mais a mãe do aluno tenha estudado melhor o desempenho do filho. Isto fica muito evidente sobretudo nos anos finais do ensino fundamental. Filhos de mães que estejam cursando ou já tenham terminado faculdade têm quase 15 pontos percentuais a mais de chance de atingirem a nota de corte em matemática. O maior efeito marginal no 9º ano desta variável pode ser explicado em parte porque o conteúdo estudado é mais complexo e demanda mais ajuda para a compreensão.

Curiosamente, no 5º ano, o efeito marginal mais expressivo (13,7 p.p. para Português e 9,8 p.p. para matemática) foi para os alunos com mães que concluíram o ensino médio e ainda estão cursando a faculdade. Bem superior aos alunos com mães com nível superior completo (3,6 e 3,8, respectivamente). Este dado merece um estudo mais aprofundado, mas pode-se supor que as mães com grau de instrução superior tenham menos tempo para acompanhar as tarefas dos filhos uma vez que têm mais facilidade de obter bons empregos. Para esta faixa etária, é comum que a tarefa de casa seja a única forma de estudo fora da escola e este importante momento de aprendizado com os pais pode estar sendo delegado para terceiros. Através de um cruzamento simples de dados, pode-se perceber que 22,6% das mães com faculdade têm empregadas ou diaristas sendo que para as demais isto é uma realidade apenas para 8,9%.

O apoio psicológico dos pais nos estudos dos filhos tem um forte impacto nos seus resultados em todas as etapas estudadas. O fato dos pais incen-

tivarem os filhos a estudar eleva suas chances em língua portuguesa em 18,4 pontos percentuais no 5º ano e 12,5 no 9º ano. Já para matemática os acréscimos nas chances são, respectivamente, 11,7 e 10,1. Importante destacar como este aspecto motivacional teve um maior efeito marginal no 5º ano, sobretudo em língua portuguesa. O resultado sugere que matérias que necessitam uma maior dedicação, como matemática, e quando a quantidade e complexidade dos conteúdos aumentam, 9º ano, o incentivo se mantém importante mas deve ser complementado com atividades práticas tais como: dever de casa, leitura e disciplina.

O hábito de lê livros em geral, conforme esperado, se mostrou positivo e estatisticamente significante nas séries estudadas em ambas as disciplinas, com destaque para Língua Portuguesa. Este é um hábito que deve ser incentivado desde cedo e traz benefícios nas notas logo nos primeiros anos. Lê livros em qualquer quantidade no 5º ano eleva as notas de língua portuguesa em mais de 8,0 pontos percentuais e no 9º ano, o hábito de ler livre sempre ou quase sempre, eleva as chances dos alunos em mais 11,0 pontos percentuais nesta mesma disciplina.

De uma forma geral, os alunos que responderam que vão sempre ao cinema e vão sempre às apresentações musicais ou de dança tiveram uma redução no desempenho das notas da Prova Brasil. No 5º ano, a maior redução do efeito marginal se deu nos alunos que disseram que frequentam com maior frequência o cinema (12,4 p.p. português e 6,7 p.p. matemática) e no 9º ano o maior efeito marginal se deu nos alunos que assistem com frequência apresentações de música e dança (12,4 p.p. e 10,8 p.p., respectivamente). Essa alternância entre cinema e show nas duas séries pode ser explicada pelas diferenças de idades. Alunos do 5º ano têm uma média de 11,4 anos e os do 9º ano 14,9, natural, portanto, que os do 9º estejam mais expostos a apresentações de música e dança.

Importante neste momento, uma breve explicação sobre como alguns dados foram coletadas pelo INEP, inclusive os que foram analisados acima. É feita uma pergunta ao aluno e é dado a ele a opção de responder de 3 formas possíveis: A) Sempre ou quase sempre; B) De vez em quando; C) Nunca ou quase nunca. Como não há uma quantificação destas categorias (por exemplo: 2 x por semana) o que um aluno considera como muito ou pouco pode não ser para outro, neste caso, alunos com uma mesma frequência de hábito poderiam estar se classificando de forma distinta. Além disso há um claro entrelaçamento entre as três respostas. O “...quase sempre” contido na letra “A” pode ser confundido com o “De vez em quando” da letra “B” que, por sua vez, pode ser

confundido com o “...quase nunca” que está contido na letra “C”. Bom lembrar que se trata de um questionário muito extenso aplicado a alunos com uma baixa capacidade de discernimento, sobretudo no 5º ano.

Os alunos que responderam que trabalham fora de casa tiveram um desempenho inferior aos demais em todas as disciplinas e séries. E esta redução chega a 15,6 pontos percentuais em Língua Portuguesa no 5º ano. O maior efeito marginal nesta série, onde as crianças têm em média 11 anos, sugere, no mínimo dois efeitos: 1) o trabalho reduz o tempo disponível para os estudos; 2) trata-se de crianças de baixíssima renda. No 9º ano o fato do aluno trabalhar fora de casa continua sendo um fator que reduz o seu desempenho na prova de forma estatisticamente significativa, no entanto, com menos força (2,5 p.p. para matemática).

Na verdade, as diferenças de efeito das variáveis entre as diversas séries devem ser feitas não somente com o olhar estatístico, mas também de forma qualitativa. Alunos que chegaram ao 9º ano já são, de certa forma, vencedoras. Já foram peneiradas. Quer seja através de reprovações, desestímulo e desmotivação; quer seja porque em muitos municípios as escolas com ensino fundamental II ficam mais distante de suas residências.

Nunca ter sido reprovado eleva muito a chance dos alunos atingirem os pontos de cortes e não importa a série ou a disciplina. Um aluno do 5º ano que nunca tenha sido reprovado tem 16,4 pontos percentuais a mais de chance de atingir o ponto de corte em língua portuguesa que outro que tenha uma ou mais reprovação anterior. No 9º ano um aluno sem reprovações anteriores tem 16,1 pontos percentuais a mais de chances em matemática.

Pode-se concluir que o aluno que repete um ano não recupera o conhecimento da forma desejada. Ou seja, nos primeiros sinais de dificuldade do aluno o professor ou uma equipe pedagógica deveria fazer um acompanhamento mais individualizado para evitar acúmulo de dúvidas e conteúdo. Além disso é preciso pensar de forma imparcial sobre programas de progressão continuada. Diferente do ensino superior, onde cada componente da grade curricular é independente, no ensino fundamental se o aluno ficar reprovado em uma disciplina é obrigado a repetir o ano inteiro, todas as disciplinas, o que onera o custo da educação para os estados, municípios e famílias. E o pior, conforme os dados acima, ele não vai recupera o conhecimento perdido.

Uma das variáveis com maior impacto no resultado dos alunos é justamente uma das mais fáceis de aplicar, já que carece de pouco ou nenhum aporte

de recursos públicos, talvez, no máximo, treinamento dos profissionais de educação e conscientização dos pais sobre a importância do assunto. Fazer o dever de casa sempre ou quase sempre elevou a chance de sucesso dos alunos em todos os modelos testados, Tabelas 3 e 4, chegando a mais de 16,2 pontos percentuais de aumento de chances de obter os resultados desejados para os alunos do 9º ano na disciplina de matemática e a 21,6 em língua portuguesa no 5º ano.

Outrossim, fazer o dever de casa traz outro benefício. Aumenta a quantidade de horas estudadas. No Brasil, o tempo que os alunos passam na escola é muito pouco comparado com países no topo do ranking PISA. Em um celebrado livro de 2014, Amanda Ripley afirma que os estudantes Sul Coreanos, por exemplo, estudam mais de 14 horas por dia. Ou seja, quanto maior a quantidade de estudo maior será o retorno em aprendizado.

Outra variável relacionada com o dever de casa, mas mais voltada para as práticas pedagógicas dos professores, diz respeito à correção das tarefas de casa. O fato do professor corrigir sempre ou quase sempre o dever de casa foi estatisticamente significativo em ambas as disciplinas e séries avaliadas. Elevando a chance em 9,0 pontos percentuais tanto em língua portuguesa como em matemática no 5º ano. No 9º ano corrigir sempre as tarefas só foi estatisticamente significativo para matemática (efeito marginal de 7,3). Tais resultados demonstram a importância do acompanhamento e cobrança por parte dos profissionais de educação, sobretudo nos anos iniciais do ensino fundamental.

Para todos os modelos econométricos foi incluída a correspondente nota média obtida pela escola em cada disciplina e série. Todos os resultados para esta variável foram estatisticamente significantes mas com baixíssimo efeito marginal correspondente. Mesmo assim, esta variável tem forte influência no ajuste global dos modelos. Provavelmente por conta do que os pesquisadores chamam de “efeito dos pares”, ou seja, os resultados de alguns acabam contagiando os demais o que leva ao aumento (ou diminuição) da média como um todo.

Na Tabela 4 foi incluída a variável “dependência administrativa” com o intuito de captar possíveis diferenças nos resultados dos alunos oriundos de escolas municipais e estaduais. Apesar desta variável ter sido significativa ao nível de 90% de confiança o efeito marginal é baixíssimo e, de forma surpreendente, como sinal negativo indicando que o aluno da rede estadual tem em torno de 1,0 ponto percentual a menos de chance de atingir os pontos de cortes em língua portuguesa e matemática.

## 5 CONCLUSÃO

Este trabalho procurou analisar os efeitos que um conjunto de variáveis contidas nos questionários socioeconômicos aplicados pelo INEP produziram nos resultados dos alunos das escolas públicas cearenses na Prova Brasil de 2011. Feito o uso de quatro modelos econométrico do tipo Logit para medir a probabilidade de obter um desempenho desejável nos testes padronizados, mostrou-se a importância de algumas destas variáveis em detrimento de outras.

Em primeiro lugar, observou-se uma heterogeneidade muito grande entre as notas dos municípios, mesmo levando em consideração características econômicas e demográficas, o que indica que políticas públicas na área de educação comprometidas com os resultados se traduzem em benefícios diretos no desempenho dos alunos.

Os modelos econométricos evidenciaram a importância das variáveis relacionadas às características das famílias nos resultados dos alunos, com destaque para os aspectos motivacionais. O simples fato dos pais incentivarem os filhos a estudar, eleva suas chances de sucesso no 9º ano em 12,5 pontos percentuais em língua portuguesa e 10,1 pontos percentuais em matemática. E para o 5º ano os resultados se mostraram mais importantes ainda com 18,4 e 11,7 pontos percentuais de acréscimo nas chances de obter a nota de corte respectivamente para língua portuguesa e matemática.

Fazer a tarefa de casa de maneira regular trouxe benefícios em ambas as disciplinas e séries avaliadas. Os alunos do 9º ano na disciplina de matemática que responderam que fazem sempre as tarefas de casa, tiveram um incremento nas chances da ordem de 16,2 pontos percentuais. Mas ficou provado que a prática deve ser incentivada desde os primeiros anos de estudo já que os acréscimos nas chances são visíveis já no 5º ano com um acréscimo de 21,6 pontos percentuais em língua portuguesa e 10,5 pontos em matemática.

Somando-se a isso um professor vigilante, que corrige as tarefas de casa, tem-se uma receita barata e com sucesso garantido já que esta prática pedagógica também se mostrou estatisticamente significativa em ambas as disciplinas do 5º ano, acréscimo de 9,0 pontos percentuais e em matemática no 9º ano com 7,0 pontos percentuais de acréscimo.

Outra característica que elevou muito a chance de sucesso na Prova Brasil foi o fato do aluno nunca ter sido reprovado. 16,4 e 10,4 pontos percentuais no 5º ano e 15,6 e 16,1 pontos percentuais no 9º ano em língua portuguesa e matemática, respectivamente. O que mostra a importância de um acompa-



nhamento individualizado ou aulas de reforço para alunos com dificuldade de aprendizado logo no início da vida escolar.

A observação dos 4 modelos econométricos nas duas disciplinas e em duas fases distintas do ensino fundamental, demonstram a necessidade de um olhar mais direcionado e preciso para cada situação. Variáveis que se mostram estatisticamente significantes no 5º ano podem não fazer diferença, ou até ter o sentido inverso, no 9º ano. Houve também variações entre disciplinas na mesma série de estudo.

Por último, cabe destacar que apesar das notas médias dos alunos cearenses ainda estarem abaixo do desejável, os dados apontaram pequenas ilhas de auto desempenho. Alunos, escolas e municípios que apesar das adversidades conseguiram se superar e se sobressair obtendo notas e médias semelhantes às obtidas por alunos, escolas e municípios de regiões mais aquinhoadas mostrando que o esforço individual alinhado com políticas públicas corretas na área de educação podem causar uma revolução no ensino público municipal e estadual cearense.

## REFERENCIAS BIBLIOGRÁFICAS

BARROS, R. P.; MENDONÇA, R. Os Determinantes da Desigualdade no Brasil: A Economia Brasileira em Perspectiva. Rio de Janeiro, IPEA, 1996.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF, Senado, 1988.

COLEMAN, James S. et al. **Equality of educational opportunity**. Washington, 1966

FELÍCIO, F. e FERNANDES, R. **O Efeito da Qualidade da Escola sobre o Desempenho Escolar: uma Avaliação do Ensino Fundamental no Estado de São Paulo**, Anais do XXXIII Encontro Nacional de Economia, 2005.

FERNANDES, M. M., FERRAZ, C. **Conhecimento ou Práticas Pedagógicas. Medindo os Efeitos da Qualidade dos Professores no Desempenho dos Alunos**. PUC Rio, 2014.

GALTON, Francis. **Family likeness in stature**. Proceedings of Royal Society, Londres, 1886. V.40, p. 42-72.

GUJARATI, Damodar N.; PORTER, Dawn C. **Econometria Básica**. 5. ed. Porto Alegre: Mc Graw Hill, 2011.

- HANUSHEK, E. A.; LUQUE J. A. **Efficiency and Equity in Schools around the World**. Mimeo, 2002.
- HOFFMAN, Rodolfo. **Distribuição de Renda – Medidas de Desigualdade e Pobreza**. São Paulo: Edusp, 1998.
- LAVY, V. **What makes an effective teacher? Quasi-experimental evidence**. NBER Working Paper, n. 16885, National Bureau of Economic Research, Cambridge, MA, 2011.
- MANSO, BARRETO, FRANÇA. **Foi o Crescimento Econômico Pós-Plano Real a Favor dos Mais Pobres? Evidências Para as Regiões Brasileiras a Partir dos Ganhos de Produtividade no Mercado de Trabalho**. CAEN, 2008.
- MENEZES FILHO, N. **Os Determinantes do desempenho escolar do Brasil**. São Paulo: Instituto Futuro Brasil/IBMEC, 2007
- OLIVEIRA, CARVALHO. **A Educação e o Desenvolvimento Sócio - Econômico do Ceará no Período de 1996 E 2006**. IPECE, 2007.
- REVISTA NOVA ESCOLA. São Paulo. Out. 2008.
- RIPLEY, Amanda. **As crianças mais inteligentes do mundo**. Três Estrelas, 2014
- RIVKIN, Steven G.; HANUSHEK, Eric A.; e KAIN, John F. **Teachers, Schools, and Academic Achievement**. *Econometrica*. vol. 73, No.2, p417-58. 2005
- VALLE, R. C. **Teoria de Resposta ao Item**. Fundação Carlos Chagas, São Paulo, v.21. p. 7-92, 2000.

# COMPORTAMENTO RELIGIOSO DA POPULAÇÃO CEARENSE A PARTIR DE UMA PERSPECTIVA ECONÔMICA E SOCIAL

Eveline Carvalho\*

Luan Falcão\*\*

Guilherme Irffi\*\*\*

## Resumo

Este artigo analisa fatores que determinam a escolha da religião pela população do Ceará e investiga se o estado está se tornando menos católico e quais os fatores determinantes. Foi inicialmente realizada análise das adesões no Ceará em relação ao nordeste e ao país por religiões agregadas em católica, evangélica (tradicional e pentecostais), espiritualista, afro-brasileira, sem religião e outras religiões utilizando dados da POF 2009. Em seguida, utilizando dados do Censo Demográfico 2010, foram estimados modelos logit binários destacando a diferenças de comportamento das populações rural e urbana e modelo multinomial para o calculo das razões de chance para o Ceará e região nordeste. Os resultados corroboram parcialmente estudos anteriores quanto influência do nível de instrução na escolha por religiões não tradicionais e quanto à irrelevância da renda para a opção religiosa independente da situação do domicílio no estado comprovando que não é a seca e a pobreza que alimentam a tradição católica.

**Palavras-Chave:** Religião, comportamento, Ceará.

## Abstract

This article analyzes the factors that determine the choice of religion by the Ceará population and investigates whether the state is becoming less Catholic and what are the determining factors. Adhesions analysis was initially

---

\* Professora do Departamento de Teoria Econômica da UFC Avenida da Universidade, 2700, Benfica Fortaleza/Ceará CEP: 60020-181. Contatos: eveline@ufc.br (85) 996607400.

\*\* Mestrando CAEN-UFC Contato: luan\_aero@yahoo.com.br. (85)986496775.

\*\*\* Professor do Departamento de Economia Aplicada da UFC Avenida da Universidade, 2700, 2º Andar, Benfica Fortaleza/Ceará CEP: 60020-181. Contatos: guidirffi@gmail.com (85) 3366.7751.

performed in Ceará in relation to the northeast and to the country by religions aggregated into Catholic, Evangelical (Pentecostal and traditional), spiritual, african-Brazilian, without religion and other religions using POF data 2009. Then, using data Population Census 2010 were estimated binary logit models highlighting the differences in behavior of rural and urban populations and multinomial model for the calculation of odds ratios for Ceará and the Northeast. The results partially corroborate previous studies on the influence of the level of education in choosing a non-traditional religions and as to irrelevance of income for the preference for a religion regardless of the domicile situation in the state proving that it is not drought and poverty that feed the catholic tradition.

**Key words:** Key-words: religion, behavior, Ceará.

## 1 INTRODUÇÃO

*Enfim caiu a primeira chuva de dezembro. Dona Inácia, agarrada ao rosário, de mãos postas, suplicava a todos os santos que aquilo fosse um bom começo.* (Queiróz, 2004, p.134).

Percebe-se pelo texto extraído do livro *O Quinze* da escritora cearense Raquel de Queiróz uma confirmação de que a fé e a religiosidade sempre estiveram muito presentes no Ceará.

O Ceará é o segundo estado mais católico do Brasil, dos 8,4 milhões de habitantes do estado, 78% são católicos, o que corresponde a cerca de 6,6 milhões de pessoas de acordo com dados do Censo de 2010. Conforme a Pesquisa de Orçamentos Familiares-POF/IBGE, apenas 4,08% da população de Ceará não tem religião, o que significa que no estado mais de oito milhões de pessoas possuem alguma filiação religiosa.

O Ceará abriga várias manifestações que reforçam a religiosidade no estado. Um dos eventos religiosos mais antigos do Brasil é a festa de São Francisco das Chagas, que ocorre no mês de outubro na cidade de Canindé. São Francisco é o padroeiro do município que possui vários lugares de peregrinação e monumentos dedicados ao santo. Na cidade do Crato, geralmente em setembro, há homenagem à padroeira Nossa Senhora da Penha. Além das famosas romarias do “Padim Padre Cíço” em Juazeiro do Norte que reúnem anualmente cerca de

dois milhões de fiéis.

No entanto, apesar da tradição católica, a população evangélica do Ceará dobrou nos últimos dez anos e o número de adesão a denominações religiosas menos tradicionais vem aumentando.

Diante disso, esta pesquisa visa ampliar o debate em Economia da Religião, porém tendo como foco, o estado do Ceará e a região nordeste do país. Para isso, analisam-se as religiões na sociedade cearense na busca de determinar os fatores que contribuem para a escolha de uma religião.

Para analisar a trajetória das religiões no Ceará a partir da investigação das características dos fiéis, emprega-se a definição de religião descrita por Iannaccone (1998), que conceitua religião como sendo qualquer conjunto compartilhado de crenças, atividades e instituições tendo como premissa a fé em forças sobrenaturais.

Neste sentido, o objetivo consiste em estimar a probabilidade de escolha da religião com base nas características apontadas por Iannaccone (1998), considerando os microdados do Censo Demográfico do IBGE de 2010 para identificar os atributos geográficos, econômicos, educacionais, demográficos e sociais da população cearense e nordestina que influenciam na escolha da religião.

Para alcançar esses objetivos, optou-se por dividir o artigo em mais cinco seções. Na próxima seção, discorre-se sobre os aspectos teóricos relativos à adesão a uma religião. Em seguida é apresentada a fonte dos dados juntamente com as evidências empíricas. As estratégias de estimação e análise dos resultados compõem a quarta seção. Por fim, são feitas as considerações finais.

## **2 ASPECTOS TEÓRICOS RELATIVOS À ADESÃO A UMA RELIGIÃO**

Os principais teóricos da vertente relativa à escolha racional da religião consideram a religião como um mercado, em que a demanda se mantém relativamente estável, e as mudanças na oferta religiosa é que vão determinar as mudanças no comportamento religioso das pessoas (Mariano, 2008).

A primeira tentativa feita por economistas a respeito dos determinantes individuais de participação em atividades religiosas é o estudo de Azzi & Ehrenberg (1975) a partir da utilização de modelo de maximização da utilidade em domicílios.

Atualmente a maior parte dos chamados, economistas de religião seguem

os passos Becker (1976) que dá ênfase à importância da abordagem econômica ao analisar os objetivos do comportamento humano no estudo. De acordo com Becker, a competição por novos membros entre as muitas denominações é importante uma vez que uma competição sempre inclui competidores que tentam satisfazer as aspirações de seus “consumidores”.

Lehrer (2004) sintetiza o papel da religião em vários aspectos relacionando o comportamento econômico e demográfico dos indivíduos e das famílias nos Estados Unidos, utilizando abordagem teórica com base nas contribuições de Becker (1976), à economia da família. Em referido estudo, Lehrer constata que o envolvimento religioso tem efeitos benéficos à saúde e ao bem-estar.

Portanto os adeptos da teoria da escolha racional aplicam a teoria microeconômica e desse modo consideram a livre concorrência entre religiões e a diversidade de opções de religiões a serem escolhidas como benéfico para o bem estar dos fiéis e, portanto, o monopólio religioso, ou seja, a presença de uma igreja única levaria perda de bem-estar.

Outros estudos analisam como as religiões afetam o crescimento econômico como, por exemplo, Barro e McCleary (2003) que afirmam que as crenças religiosas influenciam o comportamento individual tendo como consequência o benefício do desempenho econômico.

No Brasil, os estudos sobre religião em geral focalizam a transformação que o Brasil vem apresentando com relação à mudança de crença religiosa, com ênfase no aumento no número de evangélicos como é o caso dos estudos de Almeida e Montero (2001); Anuatti-Neto e Narita (2004); Neri (2011); Alves, Barros e Cavenaghi (2012); Rosas e Muniz (2014).

### **3 EVIDÊNCIAS EMPÍRICAS**

Antes de analisar as características dos fiéis, é preciso apresentar as denominações religiosas, as quais foram agregadas em oito grandes segmentos, a saber: Sem religião, Católica, Evangélica, Evangélicos pentecostais, Espírita, Afro brasileiras, Orientais e Outras denominações; o Quadro 1, contempla as denominações e as suas respectivas agregações.

**QUADRO 1: Denominações Religiosas Agregadas**

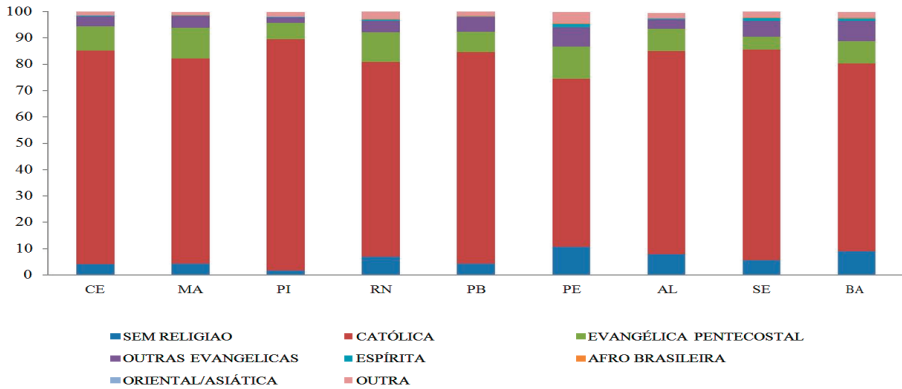
Denominação	Agregação
Sem religião	Sem religião, ateu, agnóstico
Católica	Apostólica romana, Apostólica Brasileira, Ortodoxa, Ortodoxa Cristã, Outras católicas
Evangélica	Luterana, Presbiteriana, Metodista, Batista, Congregacional, Adventista, Episcopal Anglicana, Menonita, Jesus Cristo dos Santos dos últimos Dias, Testemunha de Jeová, Outras Evangélicas
Evangélicos pentecostais	Assembleia de Deus, Congregação Cristã do Brasil, O Brasil para Cristo, Evangelho Quadrangular, Universal do Reino e Deus, Casa da Bênção, Casa de Oração, Deus é Amor, Neopentecostal Maranata, Comunidade Cristã, Nova Vida, Comunidade Evangélica, Avivamento Bíblico, Cadeia da Preçe, Igreja do Nazareno, Outras Evangélicas de Origem Pentecostal
Espírita	Kardecista, outras espíritas ou espiritualistas
Afro brasileiras	Umbanda, Cambomblé, outras religiões afro brasileiras
Orientais	Hinduísmo, Budismo, Outras religiões orientais asiáticas
Outras denominações	Judaísmo, Islamismo, Tradições esotéricas, Tradições indígenas, outras religiões

Fonte: Elaborado pelos autores.

A partir dos dados da Pesquisa de Orçamento Familiar-POF/IBGE 2009, analisa-se o comportamento religioso da população cearense em relação a outros estados do país.

Observe pelo Gráfico 1 que, o Piauí é o estado do Nordeste com o menor percentual de pessoas sem religião, enquanto o Ceará ocupa a sexta posição, entre os nove estados. Os cinco estados com o maior número de católicos estão localizados na região nordeste do país, Piauí (88%), Ceará (81,1%), Paraíba (80,3%), Sergipe (80,0%) e Maranhão (78,0%).

**GRÁFICO 1: Percentual das Denominações Religiosas Agregadas para os Estados do Nordeste.**

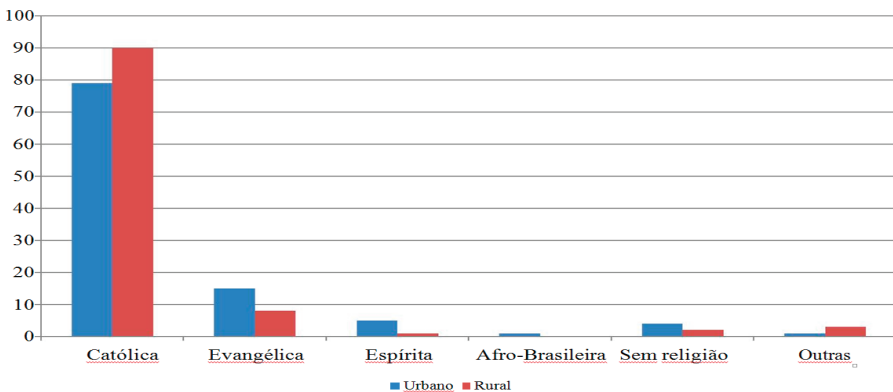


Fonte: Elaborado pelos autores com base nos microdados da POF-IBGE, 2009 extraídos pelo Centro de Pesquisa Social da Fundação Getúlio Vargas-CPS/FGV.

No entanto, analisando particularmente cada estado é possível observar diferenças significativas, nos percentuais de adesão a uma religião entre os habitantes de zonas urbanas e rurais.

No Ceará, 67,5% da população reside em área urbana, e o percentual de católicos na área rural é superior ao da urbana, já com relação a religiões evangélicas essa lógica se inverte, isto é, os evangélicos estão mais presentes em áreas urbanas. Constata-se também que o percentual dos sem religião e pertencentes a outras religiões é superior em áreas urbanas.

**GRÁFICO 2: Percentual de fiéis por Religião por situação do Domicílio.**



Fonte: Elaborado pelos autores com Base nos microdados do Censo 2010-IBGE.



Em relação ao rendimento domiciliar, em número de salários mínimos em 2010, também se verifica uma disparidade, haja vista que em média nas zonas urbanas as pessoas ganham 3,24 salários mínimos, enquanto que na área rural a renda média é de 1,43 salários mínimos. Os espíritas cearenses possuem o maior rendimento domiciliar e domiciliar per capita, seguido pelos adeptos do judaísmo. Note ainda que os evangélicos possuem, em média, os menores rendimentos.

**TABELA 1: Estatística Descritiva – Rendimento por religiões Agregadas no Ceará.**

		Religião						
		Católica	Evangélica	Espírita	Afro-brasileira	Judaísmo	Sem religião	Outras
Rendimento domiciliar	Média	3,22945	3,05985	9,38439	3,15923	3,95645	3,14569	3,84882
	Desvio padrão	6,91211	5,83168	14,41934	5,58225	5,19411	7,53264	6,58305
Rendimento domiciliar per capita	Média	0,87530	0,80030	2,82709	0,94043	1,56532	0,91708	1,08868
	Desvio padrão	2,30867	1,71590	4,81969	2,47119	2,41794	2,72198	2,51862

Fonte: elaboração dos autores com base nos dados do Censo 2010. Rendimentos em número de salários mínimos em julho de 2010.

A Tabela 2 apresenta o número e o percentual de fiéis por religião no Ceará por faixa de renda. Em relação aos católicos, observa-se que 83,6% possui renda domiciliar menor que dois salários mínimos porém esse percentual para católicos não é muito diferente do obtido para as demais faixas de renda cujo percentual médio é 79%.

Ainda de acordo com a tabela 2, o percentual de evangélicos é equivalente nas faixas extremas de renda, girando em torno de 11%. Porém o maior percentual de evangélicos está concentrado na faixa entre maior do que dois e menor do que dez salários mínimos onde o percentual de adeptos é em média 14,7%. Já a adesão pela religião espírita cresce com a faixa de renda e os sem religião representam 4,2% das pessoas com renda superior a vinte salários mínimos, se mantendo em média 3,7% nas demais faixas. No que se refere aos indivíduos com renda superior a quatro salários mínimos em média 1,7% pertence a outras denominações religiosas.

**TABELA 2: Ceará-Número e percentagem de fiéis por Religião e faixa de renda domiciliar**

Salários mínimos		Católica	Evangélica	Espírita	Afro-Brasileira	Sem Religião	Outras	Total
< 2	No.	397.765	56.334	677	283	17.364	3.588	476.011
	%	83,6	11,8	0,1	0,1	3,6	0,8	100,0
>2 e < 4	No.	153.800	27.397	631	184	7.167	2.466	191.645
	%	80,3	14,3	0,3	0,1	3,7	1,3	100,0
>4 e < 6	No.	43.127	8.216	457	22	1.971	830	54.623
	%	79,0	15,0	0,8	0,0	3,6	1,6	100,0
>6 e < 8	No.	16.630	3.039	281	33	830	351	21.164
	%	78,6	14,4	1,3	0,2	3,9	1,6	100,0
>8 e < 10	No.	8.608	1.682	193	11	399	206	11.099
	%	77,6	15,2	1,7	0,1	3,6	1,8	100,0
>10 e <20	No.	14.230	2.189	438	15	674	368	17.914
	%	79,4	12,2	2,4	0,1	3,8	2,1	100,0
>20	No.	7.530	1.036	327	8	394	155	9.450
	%	79,7	11,0	3,5	0,1	4,2	1,5	100,0

Fonte: elaboração dos autores a partir do Censo 2010 – IBGE

A tabela 3, que mostra o percentual por religião e faixa de renda nas zonas urbanas do Ceará, ratifica os achados da tabela 2 porém evidencia a elevada concentração de católicos na zona rural. De fato na área urbana o percentual de católicos é praticamente o mesmo para todas as faixas de renda em torno de 77 a 79%. Diferentemente dos dados para o total do estado, a tabela para o Ceará Urbano evidencia uma concentração de evangélicos para os níveis de renda inferiores a dez salários mínimos. As demais religiões apresentam concentração em termos percentuais semelhantes aos da tabela 2.

**TABELA 3: Ceará área Urbana - Número e percentagem de fiéis por Religião e faixa de renda domiciliar**

Salários mínimos		Católica	Evangélica	Espírita	Afro-Brasileira	Sem Religião	Outras	Total
< 2	No.	223.033	41.178	602	260	12.837	3.065	280.975
	%	79,4	14,7	0,2	0,1	4,5	1,1	100,0
>2 e < 4	No.	112.390	23.491	608	180	6.297	2.300	145.276
	%	77,4	16,2	0,4	0,1	4,3	1,6	100,0
>4 e < 6	No.	36.066	7.507	445	22	1.808	800	46.648
	%	77,3	16,1	1,0	0,0	3,9	1,7	100,0
>6 e < 8	No.	15.102	2.930	276	33	793	326	19.460
	%	77,6	15,1	1,4	0,2	4,1	1,7	100,0
>8 e < 10	No.	8.024	1.620	188	11	384	191	10.418
	%	77,0	15,6	1,8	0,1	3,7	1,8	100,0
>10 e <20	No.	13.586	2.123	433	15	653	358	17.168
	%	79,1	12,4	2,5	0,1	3,8	2,1	100,0
>20	No.	7.321	1.017	321	8	383	155	9.205
	%	79,5	11	3,5	0,1	4,2	1,7	100,0

Fonte: elaboração dos autores a partir do Censo 2010 – IBGE

Para o estado como um todo, que inclui áreas urbana e rural observa-se diferença de percentuais, porém não muito acentuadas, em decorrência da renda.

A tabela 4 mostra o percentual e número de respondentes por religião e nível de instrução para o estado do Ceará, sendo considerados os analfabetos, ou sem instrução, pessoas com ensino fundamental completo, com ensino médio completo e com curso superior completo.

Observa-se que não há uma grande variação percentual considerando os diferentes níveis de instrução para cada agrupamento de religião e mesmo para os sem religião assim como não se observou grande variação nos percentuais de adesão para cada faixa de renda, principalmente na zona urbana do estado.

**TABELA 4: Ceará-Número e percentual de fiéis por Religião e nível de instrução**

Nível de Instrução		Católica	Evangélica	Espírita	Afro-Brasileira	Sem Religião	Outras	Total
Sem instrução	N	470.479	67.548	949	338	20.305	4.474	564.093
	%	83,4	12,0	0,2	0,1	3,6	0,7	100,0
Fundamental Completo	N	96.616	16.464	363	103	4.610	1.454	119.610
	%	80,8	13,8	0,3	0,1	3,9	1,2	100,0
Médio Completo	N	104.290	19.736	1.065	113	4.966	2.106	132.276
	%	78,8	14,9	0,8	0,1	3,8	1,6	100,0
Superior Completo	N	21.383	3.138	775	28	1.023	435	26.782
	%	79,8	11,7	2,9	0,1	3,8	1,6	100,0

Fonte: elaboração dos autores a partir do Censo 2010 – IBGE.

Tais resultados ratificam parcialmente os achados de Iannaccone (1998) que conclui que a religião não é só aderida por pobres e sem educação uma vez que segundo o autor “...a taxa de crença religiosa, tende a não declinar com a renda e a maioria das taxas aumentam com a educação”.

#### 4 ESTRATÉGIAS DE ESTIMAÇÃO E ANÁLISE DOS RESULTADOS

Para analisar a influência de variáveis na escolha da religião, inicialmente foi utilizado logit binário em dois momentos. O logit binário é um modelo econométrico de escolha discreta, adequado para o caso uma vez que a variável dependente utilizada possui caráter qualitativo. Como modelo de escolha discreta, a teoria microeconômica da utilidade é utilizada para lidar com vários

tidos de escolha, correlacionadas e não correlacionadas.

O modelo logit binário trata do quanto é provável que uma observação pertença a um determinado grupo já que cada evento pode ter dois resultados. Para o presente estudo foram utilizados dois modelos logit binários com duas diferentes variáveis dependentes. No primeiro momento foi utilizada variável dependente ( $Y_{1i}$ ) binária para a escolha da religião, sendo codificada com valor 1 para católico e 0 para não católico. No segundo momento foi utilizada variável dependente ( $Y_{2i}$ ) binária para a escolha da religião, sendo codificada com valor 1 para evangélico e 0 para não evangélico. O modelo incluiu ainda, quatro variáveis binárias endógenas, para situação do domicílio, se urbano ou rural, se a pessoa é casada no religioso ou não, se tem mãe viva ou não e se passou pela experiência de perder um filho ou não.

A probabilidade de sucesso, ser católico, no modelo 1 e a probabilidade de sucesso, ser evangélico no modelo 2, podem modelados da seguinte forma, onde  $p$  representa a probabilidade de sucesso,  $X_i$  são as variáveis independentes, e  $\beta_i$  são os parâmetros estimados:

$$\ln \frac{p}{1-p} = \beta_2 X_2 + \dots + \beta_k X_k$$

ou

$$\ln \frac{p}{1-p} = \beta_1 X_i$$

Considerou-se:

$$Y_{1i/2} = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + U_i$$

Sendo:

$X_2$ , a situação de domicílio, sendo 1 atribuído a zona urbana e 0 para zona rural;

$X_3$ , a renda domiciliar per capita;

$X_4$ , o nível de instrução;

$X_5$ , o casamento religioso, sendo 1, sim e 0 não;  $X_6$ , ter mãe viva, sendo 1, sim e 0 não;

$X_7$ , a idade;

$X_8$ , se teve algum filho que faleceu, sendo 1, sim e 0, não.

Além da variável dependente, três variáveis explicativas são binárias nos dois modelos. A amostra obtida na base de dados do Censo 2010 do IBGE incluiu 1.588.572 observações da região nordeste do Brasil. A convergência foi encontrada após quatro iterações e a matriz de co-variância foi computada

utilizando derivadas de segunda ordem.

Para o cálculo das razões de chances foi utilizado logit multinomial para o Ceará (215.875 observações) e nordeste. O modelo multinomial logit foi utilizado como extensão dos modelos binários logit; contudo, sem a separação por área rural/urbana.

Tendo em vista que a variável  $X_8$ , se teve algum filho que faleceu, na amostra só existem mulheres uma vez que só é computado o número de filhos para indivíduos do sexo feminino.

A tabela 5 mostra os coeficientes e estatística z dos modelos considerando zonas urbanas e rurais conjuntamente, apenas zonas urbanas e apenas zonas rurais. O modelo 1 tem como dependente a variável binária 1 para católicos e o modelo 2 tem como dependente a variável binária 1 para evangélicos.

Os coeficientes estimados de modelos com variáveis discretas como o logit, possuem interpretação pouco intuitiva, contudo as direções dos sinais apontam para respostas dos possíveis determinantes de opção pela religião adotada e das possibilidades de optar por ser católico ou evangélico. Já a estatística z indica quantos desvios padrão o resultado está acima ou abaixo da média do grupo.

**TABELA 5: Nordeste: Estimativas dos Coeficientes dos modelos-católico/evangélico.**

Variáveis	Católico (modelo 1)			Evangélico (modelo 2)		
	Total	Urbana	Rural	Total	Urbana	Rural
Situação do Domicílio	-0.688996 (-129.72)	-	-	0.654613 (107.4)	-	-
Renda Domiciliar per capita	-3.60E-06 (-1.90)	-3.36E-07 (-0.16)	-0.000134 (-9.74)	-5.97E-05 (-15.54)	-6.83E-05 (-16.66)	6.96E-05 (5.47)
Nível de Instrução	-0.099308 (-38.76)	-0.080712 (-28.69)	-0.144244 (-22.47)	0.114965 (38.70)	0.083864 (25.66)	0.238848 (33.96)
Casamento Religioso	0.496851 (95.41)	0.383782 (62.82)	0.788996 (78.65)	-0.215123 (-37.25)	-0.089072 (-13.25)	-0.552190 (-48.80)
Mãe Viva	0.062575 (10.68)	0.056309 (8.22)	0.074707 (6.62)	-0.009214 (-1.39)	-0.003293 (-0.43)	-0.017603 (-1.35)
Idade	0.001931 (9.06)	0.004444 (17.67)	-0.003900 (-9.49)	0.004104 (16.99)	-0.001561 (5.52)	0.010942 (23.48)
Filho Falecido	0.068263 (9.59)	0.110389 (12.66)	-0.005073 (-0.41)	-0.080277 (10.09)	-0.107197 (-11.10)	-0.038317 (-2.73)

Fonte: Elaborado pelos autores a partir dos resultados estimados. Estatística z entre parênteses.

Nos modelos 1 e 2, os coeficientes relativos a variável situação do domicílio indicam que para o total da região, que inclui domicílios urbanos e rurais, a possibilidade de ser católico tende a cair (no modelo 1) e a possibilidade de ser

de evangélico tende a aumentar (modelo 2). No modelo 1 a possibilidade de ser católico tende a diminuir à medida que aumenta a renda domiciliar independente da situação do domicílio (se rural, urbano ou os dois). Já os coeficientes do modelo 2 mostram que a possibilidade de ser evangélico tende a diminuir com o aumento da renda nas zonas urbanas e por outro lado tende a crescer com o aumento da renda nas áreas rurais e a região como um todo.

Com relação ao nível de instrução a possibilidade de ser católico cai com o aumento do nível de instrução nas áreas urbanas, rural e no total do estado e de modo oposto aumenta a possibilidade de ser evangélico com a melhoria do nível de instrução nas áreas consideradas.

Quando o respondente afirma ter sido casado no religioso e ter mãe viva, aumentam as possibilidades de ser católico e não evangélico, independente da área onde habita, se em área urbana, rural e também no total. Já o avanço da idade parece não ter interferência na escolha da religião católica na área rural, mas aumenta a possibilidade de ser católico na área urbana e de ser evangélico independente do local onde habita. O fato de ter tido a experiência de ter filho falecido reduz a possibilidade de ser evangélico nas três áreas consideradas e aumenta a possibilidade de ser católico na área urbana

O quadro 2 resume as possibilidades de ser católico e evangélico com base nos coeficientes estimados das variáveis consideradas e a direção do sinal. Setas para cima indicam aumento na possibilidade de ser católico/evangélico e setas para baixo indicam redução de tais chances. Setas mais escuras indicam maiores possibilidades.

#### QUADRO 2: Possibilidade de ser católico/evangélico por variável – Nordeste

Variáveis	Possibilidade de ser Católico		Possibilidade de ser Evangélico	
	Urbano	Rural	Urbano	Rural
Quando a renda domiciliar per capita aumenta	↓	↓	↓	↑
Quando o nível de instrução aumenta	↓	↓	↑	↑
Se teve casamento no religioso	↑	↑	↓	↓
Se tem mãe viva	↑	↑	↓	↓
Quando a idade aumenta	↑	↓	↑	↑
Se teve filho que faleceu	↑	↓	↓	↓

Fonte: Elaboração dos autores com base nos resultados estimados.

De acordo com o quadro 2 quanto maior a renda domiciliar per capita

menores as possibilidades de um residente da área urbana ser católico. Esse resultado aponta para o avanço de outras religiões, não católicas ou evangélicas e para o avanço da secularização nas áreas urbanas conforme evidenciado na tabela 3, o que corrobora os achados de Carvalho e Irffi (2015).

Para as áreas rurais da região, quanto maior o nível de instrução, medido pelo número de anos de estudo, menor a possibilidade de ser católico e maior a possibilidade de ser evangélico o que sugere o maior sucesso dos missionários evangélicos junto aos residentes de áreas rurais com melhores níveis de instrução.

O fato de ter tido casamento religioso, aumenta a possibilidade de ser católico tanto nas áreas urbanas como nas áreas rurais e, por outro lado, diminui a possibilidade de ser evangélico e residir na zona rural. Esse resultado, provavelmente, deriva da tradição católica, especialmente, em áreas rurais da região.

Dentre os residentes de área urbana que passou pela experiência da perda de um filho, há maior possibilidade de ser católico, e menor possibilidade de ser evangélico provavelmente por ser essa a religião dominante e ser a área urbana o espaço geográfico onde ocorre o maior número de mortes prematuras da região.

Mas o objetivo de estimar a chance de ser católico ou evangélico com base nas variáveis seria obtido de forma mais coerente através das razões das chances ou *odds ratio* que é a razão da probabilidade de escolher uma religião em relação à probabilidade de escolher outra categoria de base e isso foi feito para o Ceará (tabela 6) e o Nordeste (tabela 7).

**TABELA 6: Estimativas das Razões de Chances de ser católico no Ceará**

religião	intercepto	situação	renda <i>per capita</i>	Idade	educação	casamento	Mãe viva	filho
evangélico	2.10157 1e-02	2.642925	1.000012 2	0.998157 7	1.1840409	9.125876e- 01	1.0161290 4	0.7755399
pentecostal	1.06259 7e-01	1.743227	0.999873 9	0.997442 5	0.9694917	8.995717e- 01	1.0221317 1	1.0160763
espírita	1.53139 5e-04	4.850798	1.000055 2	1.015605 5	2.3471266	4.506617e- 01	0.9718331 7	0.8328924
afro-brasil.	1.66399 6e-05	623.04961 1	1.000077 2	0.972405 5	0.8873768	6.425928e- 05	0.6835614 9	1.1908345
judaísmo	2.92823 0e-10	4.268011	1.000226 5	0.938044 3	25.999220 2	1.757114e+0 0	0.0421734 2	20.406280 3
sem religião	1.26757 3e-01	1.791415	1.000056 7	0.967710 9	0.8830078	2.758399e- 01	0.8793949 5	1.0433153
outras	1.81118 4e-03	4.941715	1.000003 4	1.005049 2	1.2584650	8.840747e- 01	1.0111123 1	0.7371901

Fonte: Elaborado pelos autores a partir dos resultados estimados.

**TABELA 7: Estimativas das Razões de Chances de ser católico no Nordeste**

religião	intercepto	situação	renda per capita	idade	educação	casamento	Mãe viva	filho
evangélico	2.127926e-02	2.7258733	0.9999901	1.0034091	1.320627	0.8773011	0.9674749	0.8059348
pentecostal	1.065441e-01	1.7579805	0.9998403	1.0007190	1.007086	0.8264893	0.9686530	0.9839493
espírita	5.657717e-05	9.0708604	1.0000280	1.0274322	2.692742	0.4560805	1.0100233	0.5989725
afro-brasil.	5.879757e-04	5.8098842	1.0000081	0.9967523	1.143843	0.1584153	0.7219105	0.9878961
judaísmo	1.849090e-04	0.8545696	1.0000256	0.9967254	1.552090	0.9180685	0.6711901	0.8111901
Sem religião	2.948885e-01	1.6702136	1.0000156	0.9697102	0.881993	0.2699131	0.8273972	1.0521591
outras	4.395013e-03	3.0801648	1.0000088	1.0076104	1.269591	0.6557391	0.9423064	0.7706825

Fonte: Elaborado pelos autores a partir dos resultados estimados.

De acordo com os resultados das razões de chances, pertencer à área urbana reduz as chances de ser católico tanto para o Ceará como para o Nordeste o que ratifica os achados do modelo logit binário.

O aumento na renda *per capita* não demonstrou definir a preferência por religião, mas a educação sim. De fato, tanto para o Ceará como para o nordeste a razão de chances para o aumento de uma unidade na variável educação é 1,18 e 1,32, respectivamente, para a escolha da religião evangélica em detrimento da religião católica. O aumento da escolaridade aumenta as chances de ser evangélico tradicional, espírita e outras religiões (incluindo o judaísmo), tanto para o Ceará como para a região nordeste.

Esses achados mais uma vez corroboram parcialmente os já mencionados achados de Ianacotti (1998) que em seu estudo conclui que a taxa de crença religiosa tende a não declinar com a renda e a maioria das taxas aumentam com a educação.

Para a população do Ceará, o casamento demonstra grande influência sobre a escolha de outras religiões que não a católica no Ceará, o que não era esperado. Principalmente porque para a região nordeste como um todo o fato de ter casamento formal privilegia a religião católica como seria esperado.

O avanço da idade aumenta as chances de ser católico em especial no Ceará o que aponta para o possível comprometimento futuro das tradições religiosas pela população mais jovem. Com relação à variável “ter mãe viva” para o nordeste aumenta as chances de ser católico. Já para o Ceará essa direção não



fica bem clara.

Ter a experiência de ter tido filho que faleceu prematuramente, aumenta a chance de ser católico e não evangélico tradicional, mas aumenta as chances de ser evangélico pentecostal principalmente no Ceará, o que indica que na ocorrência desse evento, perda de filho, a preferência é por se manter católico, mas em alguns casos há a opção por religiões menos tradicionais.

## CONSIDERAÇÕES FINAIS

O Ceará, assim como a região onde o estado se insere, tem experimentado mudanças na demanda por religião e expansão na oferta de novas denominações religiosas.

A presente pesquisa utilizou quase dois milhões de observações do Censo Geográfico do IBGE de 2010, tendo sido consideradas oito variáveis dependentes e oito denominações religiosas agregadas com o objetivo de analisar o comportamento religioso no Ceará por intermédio da identificação das variáveis responsáveis pelas escolhas por uma religião no Ceará e Nordeste do país.

A análise dos resultados com base nos coeficientes do modelo logit binário considerando inicialmente as populações rurais e urbanas separadamente e a razão de chances a partir de logit multinomial remetem a conclusões relevantes.

O Ceará, que tem 67,5% de sua população na zona urbana e 32,5 na zona rural, apresenta percentual de católicos na zona rural superior ao da zona urbana. De fato, evangélicos, sem religião e pertencentes a outras denominações religiosas estão mais presentes nas zonas urbanas do estado.

A religião católica domina em todas as faixas de renda do estado. De fato, na população com percentual de renda inferior a dois salários mínimos o percentual de católicos é de 83,6% e para as demais faixas de renda o percentual médio de católicos é de 79%.

Para a região nordeste os coeficientes dos modelos logit indicaram que a possibilidade de ser católico tende a diminuir e a possibilidade de ser evangélico tende a aumentar independente da localização do domicílio.

Com relação ao nível de renda na região, a possibilidade de ser católico tende a diminuir com o aumento no nível de renda na área urbana. Já a possibilidade de ser evangélico tende a aumentar com o aumento da renda nas áreas rurais da região.

Quando o nível de instrução aumenta nas áreas rurais diminui a possibilidade de ser católico e aumenta a possibilidade de ser evangélico. Ser casado no religioso e ter mãe viva aumentam a possibilidade de adesão à religião católica na região nordeste.

A idade não interfere na escolha de ser católico na área rural mas o avanço da idade aumenta a possibilidade de ser católico em áreas urbanas o que evidencia a força da fé católica. Por outro lado, ter filho falecido prematuramente, aumenta a possibilidade de ser católico e diminui a possibilidade de ser evangélico em áreas urbana.

Tais resultados foram parcialmente confirmados a partir das razões das chances uma vez que pertencer à área urbana aumenta as chances de não ser católico tanto no Ceará como no Nordeste.

O aumento na renda per capita não demonstrou alterar a preferência pela religião católica que permanece dominante independente do nível de renda. Mas o aumento no nível de escolaridade modifica essa tendência tanto no Ceará quanto no Nordeste. De fato, o aumento no nível de escolaridade reduz as chances de ser católico e aumenta as chances de ser evangélico tradicional, espírita ou pertencer a outras denominações religiosas tanto no estado quanto na região.

Esses achados com relação à baixa influência da renda e possibilidade de adesão a outras religiões não tradicionais ou sem religião corroboram parcialmente os achados de Ianacotti (1998) e Carvalho e Irffi (2015), respectivamente.

Os achados explicam em parte a razão do crescimento das religiões evangélicas no estado e na região e evidenciam o trabalho de oferta por parte dos missionários evangélicos. O crescimento dos evangélicos tem sido notório em toda América Latina que em 1900 representavam cerca de 1% da população e em 2010 representavam 17% dessa população.

A pesquisa evidencia que a cultura e tradições do estado estão mais arraigadas nas áreas rurais mas o nível de renda não influencia a escolha da religião e apesar do aumento no nível de instrução indicar preferência no estado pela religião evangélica tradicional, espírita e outras denominações, tanto o população do Ceará como da região nordeste privilegia a escolha da religião católica em detrimento de não ter religião, o que aponta para a força da religiosidade no estado.

Uma vez que independente do nível de renda, a tradição católica do es-

tado do Ceará a qual a população é historicamente ligada, se mantém, não mais em virtude da seca e dos problemas estruturais relativos a ela que persistem.

## REFERÊNCIAS BIBLIOGRÁFICAS

Almeida, R. D., & Monteiro, P. (2001). Trânsito religioso no Brasil. **São Paulo em perspectiva**, 15(3), 92-100.

Anuatti-Neto, F., & Narita, R. D. T. (2004). A influência da opção religiosa na acumulação de capital humano: um estudo exploratório. **Estudos Econômicos**, v. 34, n. 3, p. 451-486.

Alves, J. E. D., Barros, L. F. W., & Cavenaghi, S. (2012). A dinâmica das filiações religiosas no Brasil entre 2000 e 2010: diversificação e processo de mudança de hegemonia. **Revista de Estudos da Religião**, v. 12, n. 2, p. 145-174.

Azzi, C., & Ehrenberg, R. (1975). Household allocation of time and church attendance. **Journal of Political Economy**, v. 83, n. 1, p. 27-56.

Berman, E. & Laitin, D.D.(2008). Religion, Terrorism and public goods: testing the club model”. **Journal of Public Economics**, v. 92 (10-11), p. 1942-1967.

Barro, R.J., & McCleary, R. M. (2003). Religion and Economic Growth across countries. **American Sociological Review**. v. 68, n. 5, p. 760-781.

Becker, G.S. (1976). *The Economic Approachdo Human Behavior*. The University of Chicago Press. Chicago-IL.

Carvalho, C. & Irfi, G.(2015). Existe Secularização no Brasil? Análise a partir da Desfiliação, Descrença e Ausência de Prática Religiosa. In: **43º. Encontro da Associação Nacional dos centros de Pós-Graduação em Economia – ANPEC Nacional**, 2015, Florianópolis.

Coutinho, R. Z., & Golgher, A. B. (2014). The changing landscape of religious affiliation in Brazil 1980-2010: age, period and cohort perspectives. **Revista Brasileira de Estudos Populacionais**, v. 31, n.1, p. 73-98.

Iannaccone, L. R. (1998). Introduction to the Economics of Religion. **Journal of economic Literature**, v. 36, n. 3. p. 1465-1495.

Irfi, G.; Cruz, M. S.; Carvalho, E. B. S (2014). Determinantes da Frequência Religiosa da Mulher Brasileira. In: **III Encontro Pernambucano de Economia – ENPECON**, 2014, Recife. Políticas Para o Desenvolvimento Estadual.

- Lehrer, E.L. (2004). **Religion as a Determinant of Economic and Demographic Behavior in the United States**. Discussion Paper 1390. Institute for the Study of Labor (IZA).
- Mariano, R. (2008). Usos e limites da teoria da escolha racional da religião. **Tempo social**, 20(2), 41-66.
- McCleary, R. M., & Barro, R. J. (2006). Religion and economy. **Journal of Economic Perspectives**, v. 20, n. 2, p. 49-72.
- Moreira-Almeida, A., Pinsky, I., Zaleski, M., & Laranjeira, R. (2010). Envolvimento religioso e fatores sociodemográficos: resultados de um levantamento nacional no Brasil. **Revista de Psiquiatria Clínica**, v. 37, n. 1, p. 12-15.
- Neri, M. (2011). Novo Mapa das Religiões. Centro de Políticas Sociais-CPS. Fundação Getúlio Vargas-FGV. Rio de Janeiro.
- Oliveira, L. L. S. D., Cortes, R. X., & Balbinotto Neto, G. (2013). Quem vai à igreja? Um teste de regressão logística ordenada do modelo de Azzi-Ehrenberg para o Brasil. **Estudos Econômicos**, v. 43, n. 2, p. 363-396.
- Queiróz, R. de.(2004). **O Quinze**. 77<sup>a</sup>. ed. Rio de Janeiro: José Olímpio, São Paulo.
- Rosas, N., & Muniz, J. O. (2014). O hábito faz o monge? Frequência e auto-percepção religiosas no Brasil. **Mediações – Revista de Ciências Sociais**, v. 19, n. 1, p. 187-213.
- Smith, A. (1994). **The Wealth of Nations**. Modern Library Edition, New York-NY.

# DECOMPOSIÇÃO DOS DIFERENCIAIS DE RENDIMENTOS POR COR E GÊNERO NO MERCADO DE TRABALHO BRASILEIRO: UMA ANÁLISE POR NÍVEIS OCUPACIONAIS NAS REGIÕES METROPOLITANAS

Edward Martins Costa\*

Daniel Tomaz de Sousa\*\*

Daniel Barbosa Guimarães\*\*\*

## Resumo

Verificar o comportamento dos diferenciais de rendimentos entre as ocupações nas regiões metropolitanas (RMs) é o objetivo deste trabalho. Devido ao enorme número de ocupações, realiza-se um agrupamento ocupacional, reduzindo para quatro níveis de ocupação. Para captar o impacto das variáveis, obtidas a partir da PNAD 2012, em cada ocupação, são estimadas equações mincerianas e para verificar a existência ou não de discriminação utiliza-se a decomposição de Oaxaca e Blinder (1973). Os resultados da equação minceriana, mostram que as variáveis de capital humano tem um maior impacto no grupo dos dirigentes, em contraponto aos profissionais dos serviços com o menor impacto. Em relação ao diferencial percebido em cada RM e nível ocupacional os resultados indicam que mulheres assim como os indivíduos não brancos têm piores remunerações. A discriminação por gênero está mais presente no grupo dos técnicos e menos concentrada entre os dirigentes, apesar de estes possuírem diferenciais totais elevados. Por fim, a discriminação racial é menos concentrada e apresenta o mesmo padrão na maioria das RMs e ocupações.

**Palavras-Chave:** Setor agropecuário; ICN; Nordeste; Ceará.

## Abstract

Verify the behavior of income differentials between occupations in metropolitan regions (RMs) is the central focus of this research. Due to the huge

---

\* Doutor em Economia PIMES/UFPE. Prof. Mestrado Acadêmico em Economia Rural MAER/UFC. edwardcosta@ufc.br.

\*\* Mestrando em Economia CAEN/UFC. daniel25tomaz@gmail.com.

\*\*\* Doutor em Economia CAEN/UFC. Prof. Do Departamento de Administração FEAACS/UFC. barbozadan@hotmail.com.

number of defined and cataloged by IBGE occupations, an occupational grouping scheme was used, reducing to four levels of occupation, each requiring different levels of human capital. The analysis will be made for the defined occupational groups for each RM, by virtue of their importance to the economy. To capture the impact of the variables in each occupation, Mincer's equations are estimated and is used after the decomposition process of Oaxaca and Blinder (1973), the data used are from PNAD 2012. The results of the Mincer's equation, show that the human capital variables have a greater impact in the group of leaders, as opposed to professional services with minimal impact. In relation to the perceived differential in each RM and occupational level the results indicate that women and non-white individuals have worse salaries. The tendency of differential components was as expected, and the majority of the discriminatory overlapped that related to productive capacity. Gender discrimination is more present in the technical group and less concentrated among leaders, despite these having high total differential. Finally, racial discrimination is less concentrated and showed different patterns in most metropolitan areas and occupations.

**Key words:** Income Differentials. Occupational Categories. Discrimination.

**JEL:**JEL: J71. C01.

## 1 INTRODUÇÃO

Nos últimos anos, o mercado de trabalho brasileiro tem apresentado redução da taxa de desemprego<sup>1</sup>, que passou de 13% para 5,4%, no período de 2003 a 2013, respectivamente. Além do mais, verifica-se uma tendência de aumento do aperfeiçoamento por parte dos agentes econômicos (empresas, trabalhadores e o governo), tanto de técnicas de administração, por parte dos gestores/empresários quanto de aperfeiçoamento através da educação técnica/ou não, por parte dos trabalhadores. Contudo, apesar desse esforço para alcançar patamares adequados de eficiência, ainda são perceptíveis certos indícios de discriminação no mercado, fazendo com que pessoas com características semelhantes tenham diferentes retornos salariais, podendo gerar um efeito negativo sobre a eficiência da economia.

Matos e Machado (2006) analisam a evolução da discriminação por cor e sexo no período de 1987 a 2001, utilizando *cross section* empilhados obtidos

---

1. Ver tabela 1, na seção 2.

das PNADs dos anos em questão. Para a análise criaram grupos homogêneos<sup>2</sup>: homens brancos, homens negros, mulheres brancas e mulheres negras. Pode-se destacar que quando a comparação é feita entre homens e mulheres de mesma cor há redução do diferencial de rendimentos. Tanto mulheres brancas como negras possuem, em média, escolaridade mais elevada que homens brancos e negros, contando com a diferença de habilidade a seu favor. No entanto quando a comparação se dá no âmbito da cor, para indivíduos do mesmo sexo, os autores constatam que parte do diferencial se deve a fatores associados à discriminação. A outra parte significativa deve-se ao fato de negros apresentarem uma deficiência quanto aos atributos de produtividade, em especial os homens. Em termos de rendimentos a pior situação é das mulheres negras, por possuir menor escolaridade do que mulheres brancas estão mais sujeitas a segregação e discriminação ocupacional, acabando dessa forma ocupando postos de trabalho mais precários.

Além das diferenças observadas de forma ampla entre pessoas de distintas etnias e gênero, estas podem ter comportamentos variados quando analisados para cada ocupação, pois dependendo do setor de trabalho e atividade desenvolvida, a discriminação pode ter impacto e intensidade diferentes, o que justifica verificar essas tendências para as ocupações. Menezes (2013) afirma que cada categoria ocupacional exige níveis diferentes de capital humano, partindo desse ponto faz-se necessário a investigação desagregando as ocupações.

Menezes (2013) utilizando os dados da PNAD 2009 e da metodologia de Machado e Mata (2005), além da separação em níveis ocupacionais a partir do esquema EGP<sup>3</sup>, realizam uma análise das lacunas existentes de rendimentos entre gênero, raça e região de habitação (Nordeste ou em outra região). Seus resultados demonstram que, em alguns agrupamentos ocupacionais a defasagem é maior no topo da distribuição, indicando um *glass ceiling effect*, entretanto na maioria dos casos há ocorrência de um *stick floor effect*, isto é, a defasagem é maior na parte inferior da distribuição. Na decomposição por região esse efeito é encontrado para todas as ocupações indicando que a discriminação diminui à medida que se avança na distribuição. Outro resultado importante encontrado é que tanto os residentes na região nordeste e indivíduos não brancos possuem características produtivas (educação e experiência) piores, por consequência au-

2. Um indivíduo pertencente a um grupo homogêneo é aquele que apresenta as mesmas características que todos os demais para as seguintes variáveis: sexo, cor, faixa etária, relação na família, posição na ocupação, região e ramo da atividade.

3. Técnica iniciada por Erickson, Goldthorpe e Portocarrero (1979) e melhorada por Goldthorpe, Llewellyn e Payne (1987). Todavia, o esquema ficou mundialmente conhecido por as iniciais dos primeiros autores, por isso o nome do método é EGP.

mentando os diferenciais de renda. A presença de discriminação é assinalada em todas as ocupações e independe de qual decomposição realizada.

Guimarães (2013) realizou uma análise dos rendimentos por setor (dirigentes, serviços, técnicos e todas as ocupações) e ainda por níveis regionais (Brasil, Nordeste, Sul e Sudeste), seus resultados, apontaram que independentemente do setor ou região, as mulheres ganham em média menos que os homens, controlando por setor de ocupação e região a diferença salarial por sexo é de no mínimo 36,31% na média. A diferença é menor para o grupo de dirigentes; pessoas pertencentes ao grupo de técnicos e região Nordeste apresentam os maiores diferenças por gênero.

Outro fator relevante, e pouco considerado em análises de equações de rendimentos, é o fato de que a maior participação do Produto Interno Bruto (PIB) se concentra nas regiões metropolitanas (RMs). De acordo com estudos do Observatório das Metrôpoles<sup>4</sup>, os Grandes Espaços Urbanos Brasileiros Metropolitanos (GEUBs), concentraram 52,5% do PIB da economia do país em 2005 e em 2009, 47,47% do PIB segundo dados do IBGE. Ainda segundo o mesmo instituto, utilizando dados do Censo 2010, as regiões consideradas metropolitanas concentram cerca de 62% da capacidade tecnológica<sup>5</sup> do país, fator fundamental para o desenvolvimento e crescimento econômico além de representarem 36,3% da população brasileira<sup>6</sup>.

Diante do exposto, o objetivo deste trabalho é mensurar o comportamento do diferencial salarial para as RMs, além do poder explicativo de seus componentes, através de cada grupo ocupacional. Para tal, são usados dados da PNAD 2012. Devido ao enorme número de ocupações definidas pelo IBGE, as ocupações serão agrupadas em alguns blocos, será utilizado o mesmo esquema de agrupamento ocupacional adotado por Guimarães (2013) reduzindo para quatro níveis ocupacionais (todas as ocupações, dirigentes, serviços e técnicos)<sup>7</sup>, facilitando o trato estatístico e econométrico. Além dessa introdução, este trabalho está dividido em mais quatro seções, a próxima faz uma breve revisão de literatura sobre o tema em análise e alguns índices recentes do mercado de trabalho metropolitano. Após, são enumerados a base de dados adotada e a metodologia para decompor o diferencial salarial. Em seguida a discussão dos resultados e considerações finais.

4. Pesquisa disponível em <http://www.observatoriodasmetrolopes.ufrj.br/relat0092009b.pdf>.

5. A capacidade tecnológica é medida pelo número de patentes, artigos científicos, população com mais de 12 anos de estudos e pelo valor bruto da transformação industrial (VTI).

6. Esse percentual da população das metrôpoles leva em consideração características especiais consideradas pelo Observatório das Metrôpoles. Já levando em consideração as regiões metropolitanas oficiais esse percentual se eleva para 44,7%.

7. Posteriormente essa divisão será mais bem detalhada.



## 2 MERCADO DE TRABALHO: EVIDÊNCIAS PARA AS REGIÕES METROPOLITANAS

O estudo do mercado de trabalho é peculiar e, por vezes, dificultoso. De acordo com Ramos (2007) a grande questão envolvida nas análises está em definir o trabalho como uma mercadoria que é transacionada, tendo dessa forma preço definido pelos mecanismos de oferta e demanda. Contudo o processo de definição de seu preço não é simples. As características dos indivíduos são de extrema importância para que os postos de trabalho tenham remunerações adequadas de acordo com o nível de preparo de cada um.

Além do papel econômico, o trabalho, tem funções sociais, gerando bem-estar e seu estudo está voltado também a questões de pobreza, desigualdade e equidade. Em suma, são três os fatores que norteiam o desempenho do mercado de trabalho: i) instituições reguladoras; ii) condições macroeconômicas internas e externas e iii) quantidade e qualidade da força de trabalho, (RAMOS, 2007).

Especificamente para as regiões metropolitanas, Amadeo *et al.* (1994) afirmam que na década de 1990 as taxas de desemprego eram similares aos níveis da década de 1980, confirmando a hipótese de que o mercado de trabalho brasileiro estava enfrentando dificuldade de absorver mão-de-obra. Houve queda, também na década de 90, do emprego da indústria e um acentuado crescimento do setor de serviços e comércio.

De acordo com dados da Pesquisa Mensal do Emprego (PME)<sup>8</sup>, posteriormente ao lançamento do Plano Real, a taxa de desemprego sofreu uma queda significativa, atingindo em dezembro de 1994 um patamar de 3,4%. Já em 1995 essa taxa começou a apresentar um leve crescimento tendo permanecido em torno do patamar de 6% durante todo o ano de 1997. O ano de 1998 ficou marcado por um aumento da taxa de desemprego em cerca de 2 p.p. (pontos percentuais), valor ultrapassado em meses desse ano e de 1999, fato que não ocorria desde 1981. No ano 2000 houve uma tendência de declínio, contudo não teve seguimento mantido em 2001. Em 2002 o desemprego apresentou taxa média de 7,2%. É importante notar que vários desses períodos tiveram conjunturas econômicas distintas (RAMOS; BRITTO, 2004).

Curi e Menezes-Filho (2004), analisando a evolução do mercado de trabalho metropolitano para o período 1990-2000, demonstraram que houve redução da taxa de emprego formal em todos os níveis de educação, em especial para níveis educacionais elevados. Simetricamente, a proporção de pessoas em

8. Pesquisa mensal domiciliar que cobre seis regiões metropolitanas, realizada pelo IBGE.

empregos informais (sem carteira) aumentou sendo mais evidente para trabalhadores com mais de 11 anos de estudo.

No Brasil Metropolitano, a População em Idade Ativa (PIA) obteve crescimento de 2,36% no período 1992-1999 e 2% para o período 2001-2005, para o Brasil Não Metropolitano as taxas foram similares. Dentre as regiões definidas oficialmente pelo IBGE, a Norte obteve a maior taxa de crescimento (4,29% e 3,95%, respectivamente para os períodos 1992-1999 e 2001-2005) e a Sul teve as menores taxas, de acordo com dados das PNADs 1992, 1999, 2001 e 2005. (RAMOS, 2007)

Ainda segundo Ramos (2007), a participação no mercado de trabalho apresentou tendências diferentes na desagregação por gênero. A participação masculina teve um decréscimo de aproximadamente 4 p.p. no período 1992-2005 já a feminina teve um aumento de 7 p.p. no mesmo período. Verificando os diferenciais salariais devido ao gênero, este reduziu de 80,5% para 46,9%. Essa diminuição de fatores puramente discriminatórios tem efeitos benéficos para a promoção de igualdade dentro da economia.

Em estudo recente Menezes-Filho, Cabanas e Komatsu (2014) compilam as últimas tendências do mercado de trabalho no Brasil. A Tabela 1 abaixo sintetiza algumas das informações destacadas pelos pesquisadores.

**TABELA 1 – Tendências do Mercado de Trabalho nas Regiões Metropolitanas**

	2003	2008	2013
Taxa de Desemprego (%)	13	7,7	5,4
Empregados com Carteira assinada (%)	53	58	66
PIA 1 (em milhões)	37	40	43
PEA (em milhões)	22	23	25
Proporção (%) da População em Idade Ativa na População total	84	87	89
Total de Empregados (Milhões)	14	16	18
Empregados com Carteira Assinada (Milhões)	7	9	12

Fonte: Menezes-Filho, Cabanas e Komatsu (2014) a partir de dados da PME/IBGE e PNAD/IBGE.

Percebe-se que o desemprego nas RMs apresentou queda, de acordo com as informações da Tabela 1, passando de 13% em 2003 para 5,4% em 2013. Os empregos formais, com carteira assinada, tiveram tendência ascendente, saindo de 53% em 2003 para 66% em 2013. Empregos formais significam melhores remunerações e condições de trabalho e segundo Menezes-Filho, Cabanas e Komatsu (2014) isso gera um aumento de gastos com seguro-desemprego, por

exemplo. Segundo dados do Ministério do Trabalho e Emprego (MTE), entre 2002 e 2012 os gastos do governo com o seguro desemprego passaram de cerca de R\$ 11 bilhões para R\$ 28,4 bilhões (em valores de 2012; deflator: INPC<sup>9</sup>). Além disso, em 2013 a PIA representou 89% da população total das RMs e os trabalhadores com carteira assinada, em relação ao total de empregados, representavam 50% em 2003, 56,25% em 2008 e 66,67% em 2013.

### 3 METODOLOGIA

#### 3.1 Base de Dados

O presente trabalho utiliza dados da Pesquisa Nacional por Amostra de Domicílio (PNAD), a qual é fornecida com frequência anual, exceto nos anos censitários, pelo Instituto Brasileiro de Geografia e Estatística (IBGE). A PNAD utilizada será a de 2012, ano mais recente de sua publicação. As variáveis selecionadas foram escolhidas baseadas na teoria do capital humano, assegurando, portanto, os seus níveis de importância quanto à explicação dos retornos salariais. O quadro abaixo sintetiza as variáveis adotadas.

QUADRO 1 – Descrição das variáveis

Variável	Descrição
<b>Dependente</b>	
<i>lnW</i>	Rendimento do trabalho principal.
<b>Independentes</b>	
<i>Educ</i>	Anos de Educação.
<i>Exp</i>	Experiência* do Indivíduo.
<i>exp2</i>	Experiência ao quadrado.
<i>Sexo</i>	1 se o indivíduo for do sexo masculino, 0 caso contrário.
<i>cor</i>	1 se o indivíduo for branco, 0 se não branco.
<i>RM</i>	Conjunto de variáveis dummies que indicam qual região metropolitana está sendo analisada. Ao todo são 9 dummies, tem como grupo-base a RM de São Paulo.

Fonte: Dados da Pesquisa. Elaboração Própria.

\* Experiência=idade-anos de educação -6, onde se assume que o indivíduo ingressa no mercado de trabalho após a fase escolar, 6 é a idade em que a pessoa ingressa na escola no Brasil. A anexação dessa variável enfatiza a importância do treinamento no mercado de trabalho.

9. Índice Nacional de Preços ao Consumidor.

É importante salientar que devido a algumas restrições da pesquisa, diversos filtros de seleção foram utilizados para chegar a uma amostra mais homogênea possível: foram selecionados indivíduos pertencentes à faixa etária dos 15 aos 65 anos de idade, onde os extremos, inferior e superior, foram definidos levando em conta a real possibilidade de estes indivíduos estarem no mercado de trabalho das regiões metropolitanas, além da PIA ser composta de pessoas com 15 ou mais anos de idade e que após 65 anos boa parte da força de trabalho pode já estar aposentada.

Os indivíduos que não residiam em regiões metropolitanas também foram excluídos, assim como aqueles que não declararam os rendimentos. As regiões metropolitanas consideradas pela PNAD são as seguintes: Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre, além do Distrito Federal. Além disso, foram criadas as *dummies* indicativas de sexo e raça, como explicitadas no Quadro 01.

O agrupamento das categorias está disponível na PNAD 2012. Adotando o mesmo esquema de divisão ocupacional de Guimarães (2013) as restrições abaixo foram impostas: 1) exclusão dos militares, por seus salários não serem determinados pelas forças de mercado; 2) Exclusão dos profissionais das Ciências e das Artes; 3) Exclusão dos trabalhadores agrícolas, por não ser uma parcela significativa dos trabalhadores das regiões metropolitanas e 4) exclusão das ocupações mal definidas.

Desse modo os novos agrupamentos ocupacionais<sup>10</sup> são dirigentes, serviços e técnicos e todas as ocupações. De acordo com Guimarães (2013) essa divisão é bastante significativa e explicativa, no sentido de que engloba ocupações que exigem qualificações elevadas (dirigentes), médias (serviços) e baixas (técnicos), argumento que corrobora o de Menezes (2013) que cada estrato ocupacional necessita de níveis distintos de capital humano.

## 3.2 Modelagem econométrica

### 3.2.1 Modelo Minceriano e Decomposição de Oaxaca e Blinder (1973)

Os diferenciais salariais podem ser explicados por diversos fatores. Oaxaca (1973) encontrou evidências que parte do diferencial relativo ao gênero é devido à discriminação presente no mercado de trabalho. Blinder (1973) também mostra que o diferencial salarial entre brancos e negros, homens e mulheres, é devido à presença de discriminação. A presença de

10. Para um melhor detalhamento dos níveis ocupacionais, ver Guimarães (2013).

discriminação é constatada quando um grupo com elementos produtivos iguais recebe menos que outro apenas por características que não tem interferência na produtividade, como por exemplo, sexo, cor ou origem. Para análises de discriminação, por gênero e cor usa-se o processo de decomposição de Oaxaca e Blinder (1973)<sup>11</sup>. A primeira parte do procedimento é composta pela estimação de equações mincerianas, onde é possível fazer inferências sobre o retorno de cada variável explicativa sobre o logaritmo natural do salário ( $\ln W$ ), o qual é explicado por um vetor de variáveis ( $X_i$ ) de controle e de discriminação. As variáveis de controle são: educação, experiência e RMs, já as de discriminação são gênero e cor.  $\ln W = \beta_i X_i + u_i$  (1)

Para calcular o diferencial, por meio da decomposição de Oaxaca e Blinder (1973)<sup>12</sup>, são estimadas equações de salários para ambos os grupos em análise.

$$\ln W_A = \beta_A X'_A + u_A \quad (2)$$

$$\ln W_B = \beta_B X'_B + u_B \quad (3)$$

Onde  $W$ ,  $X$  e  $u$  são o rendimento, conjunto de variáveis explicativas e o termo de erro aleatório, respectivamente. Os subscritos A e B representam os dois grupos em comparação. Decompondo o diferencial (decomposição *two-fold*, ou em dois termos), obtém-se a seguinte equação:

$$\ln \bar{W}_A - \ln \bar{W}_B = \underbrace{(\bar{X}_A - \bar{X}_B) \hat{\beta}_B}_{\text{Efeito Dotação}} + \underbrace{\bar{X}_A (\hat{\beta}_A - \hat{\beta}_B)}_{\text{Efeito Discriminação}} \quad (4)$$

O termo  $(\bar{X}_A - \bar{X}_B) \hat{\beta}_B$  é a parte do diferencial atribuída às características produtivas dos indivíduos, denominado neste trabalho de “efeito dotação”. Aqui o coeficiente de B é mantido fixo, e avalia-se o diferencial em função das características observadas (Miro e Suliano, 2009). Enquanto  $\bar{X}_A (\hat{\beta}_A - \hat{\beta}_B)$  é a parte do diferencial que é atribuído à diferença nos retornos as características individuais, mas que não são produtivas, aqui denominado de “efeito discriminação”. Esse procedimento será realizado para o grupo que pode haver discriminação, brancos e não brancos e mulheres e homens. Sendo os subscritos A, o grupo de mulheres ( $m$ ) e não brancos ( $nb$ ) e B, o grupo de homens ( $h$ ) e brancos ( $b$ ). Portanto, serão obtidas as seguintes decomposições:

$$\ln \bar{W}_m - \ln \bar{W}_h = \underbrace{(\bar{X}_m - \bar{X}_h) \hat{\beta}_h}_{\text{Efeito Dotação}} + \underbrace{\bar{X}_m (\hat{\beta}_m - \hat{\beta}_h)}_{\text{Efeito Discriminação}} \quad (5)$$

11. Em geral usa-se o termo decomposição de Oaxaca e Blinder (1973), devido a maioria dos trabalhos sobre discriminação abordarem tanto gênero quanto cor.

12. Para mais detalhes sobre o processo de decomposição ver FORTIN, N.M., T. LEMIEUX, S. FIRPO (2011) e JANN (2008).

$$\ln \bar{W}_{nb} - \ln \bar{W}_b = \underbrace{(\bar{X}_{nb} - \bar{X}_b)\hat{\beta}_b}_{\text{Efeito Dotação}} + \underbrace{\bar{X}_{nb}(\hat{\beta}_{nb} - \hat{\beta}_b)}_{\text{Efeito Discriminação}} \quad (6)$$

### 3.2.2 Correção da Seletividade Amostral: Heckman (1979)

É comum em análises de equações de salários, a presença de viés de seletividade amostral. De acordo com Heckman (1979) o viés de seletividade amostral ocorre devido a dois motivos: pelas decisões de seleção dos pesquisadores e pela auto seleção de indivíduos. De forma sintética o procedimento é o seguinte:

1. Estimação de um modelo *Probit* para a participação no mercado de trabalho, a partir da qual se pode obter a razão inversa de *Mills* para cada observação;
2. Estimação por mínimos quadrados de uma equação de rendimentos (*à la* Mincer) que inclua como regressor adicional a razão inversa de *Mills*.

### 3.3 Estatísticas Descritivas

Após definir a metodologia adotada e a origem das variáveis utilizadas é pertinente realizar uma análise dos dados para fazer uma verificação prévia do diferencial de rendimentos, tanto por gênero quanto por raça. A Tabela 2 contém as estatísticas descritivas separadas por grupo ocupacional, para os dados da PNAD 2012. De acordo com essa tabela, verifica-se diferença salarial entre gênero em todos os grupos ocupacionais, onde os trabalhadores do gênero masculino são os que tem os melhores rendimentos em média, apesar das mulheres apresentarem, exceto para o grupo dos serviços, os melhores níveis educacionais.

Dentre os agrupamentos ocupacionais, observa-se que, em média, os dirigentes apresentam os melhores rendimentos, em especial o grupo dos homens, com R\$ 5013,34 enquanto as mulheres, no mesmo grupo, recebem R\$ 3478,07. Comportamento semelhante pode ser observado para brancos e não brancos, onde os brancos possuem, em média, rendimento mais elevado. O grupo ocupacional dos técnicos apresenta a menor diferença salarial entre os gêneros, já em relação a raça, a menor diferença é encontrada entre os profissionais dos serviços. Um ponto importante a notar é que apesar do grupo dos dirigentes apresentar as melhores remunerações também é neste bloco de

ocupações que ocorrem as maiores diferenças salariais tanto para gênero quanto para raça.

A educação em todas as ocupações é maior para os brancos do que para os não brancos, mesmo comportamento das mulheres, conforme detalhado anteriormente. O nível de experiência para as mulheres é mais elevado do que o dos homens apenas para profissionais dos serviços. Verificando a tendência da variável experiência para a raça, nota-se que em todos os agrupamentos ocupacionais os não brancos possuem os maiores níveis, talvez pelo fato de possuírem menores níveis de educação, dessa forma se inserindo no mercado de trabalho mais cedo. Para dirigentes o nível médio de experiência entre as raças é praticamente o mesmo. A idade não variou muito nos grupos ocupacionais e entre grupos comparáveis, homens *versus* mulheres e brancos *versus* não brancos.

TABELA 2 – Estatísticas Descritivas por Grupo Ocupacional

	Todas as ocupações				Dirigentes			
	Homens		Mulheres		Homens		Mulheres	
	Média	DP	Média	DP	Média	DP	Média	DP
<i>Renda</i>	1644.01	2544.07	1141.05	1678.39	5013.34	6625.30	3478.07	3878.68
<i>Educ</i>	8.90	3.75	9.47	3.62	11.88	3.36	12.53	2.94
<i>Idade</i>	37.38	12.31	37.45	11.98	41.60	11.25	39.50	10.78
<i>Experiência</i>	22.48	13.55	21.98	13.42	23.72	12.06	20.96	11.37
<i>Raça</i>	0.43	0.50	0.45	0.50	0.67	0.47	0.66	0.47
	Brancos		Não Brancos		Brancos		Não Brancos	
	Média	DP	Média	DP	Média	DP	Média	DP
	<i>Renda</i>	1787.59	2930.69	1136.66	1366.16	4938.14	6457.21	3322.91
<i>Educ</i>	9.81	3.65	8.63	3.67	12.52	3.08	11.37	3.32
<i>Idade</i>	37.68	12.31	37.19	12.05	41.26	11.06	39.74	11.15
<i>Experiência</i>	21.88	13.50	22.56	13.49	22.74	11.78	22.37	12.04
<i>Sexo</i>	0.55	0.50	0.57	0.49	0.60	0.49	0.59	0.49
	Serviços				Técnicos			
	Homens		Mulheres		Homens		Mulheres	
	Média	DP	Média	DP	Média	DP	Média	DP
<i>Renda</i>	1242.92	1413.39	914.67	869.46	1501.39	1816.99	1330.18	2338.74
<i>Educ</i>	9.08	3.60	9.04	3.60	8.39	3.73	10.26	3.29
<i>Idade</i>	35.90	12.41	37.15	12.07	37.92	12.20	38.01	11.90
<i>Experiência</i>	20.82	13.72	22.11	13.66	23.54	13.49	21.75	12.99
<i>Raça</i>	0.41	0.49	0.43	0.49	0.42	0.49	0.50	0.50
	Brancos		Não Brancos		Brancos		Não Brancos	
	Média	DP	Média	DP	Média	DP	Média	DP
	<i>Renda</i>	1228.78	1429.68	912.81	824.46	1746.53	2462.14	1251.49
<i>Educ</i>	9.61	3.53	8.66	3.59	9.37	3.65	8.32	3.70
<i>Idade</i>	36.62	12.45	36.68	12.06	38.25	12.22	37.71	12.08
<i>Experiência</i>	21.01	13.81	22.02	13.60	22.87	13.39	23.40	13.42
<i>Sexo</i>	0.39	0.49	0.40	0.49	0.76	0.43	0.82	0.39

Fonte: Dados da Pesquisa. Elaboração Própria.

Notas: DP= Desvio Padrão.

Ao analisar brevemente os resultados descritivos da amostra, nota-se que existem diferenças salariais tanto por gênero quanto por raça, e que tais dispa-

ridades podem estar associadas à presença de discriminação. No entanto, tais evidências não podem atestar a existência de uma possível discriminação entre os gêneros e raças nas RM's. Diante do exposto, a próxima seção apresentará os resultados obtidos pelas estimações com vistas verificar se este comportamento é devido à discriminação ou por fatos relacionados à produtividade dos indivíduos.

## 4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

### 4.1 Resultados das Estimações da Equação Minceriana

O modelo proposto por Mincer (1974) é base para diversas pesquisas onde se tenta captar o efeito de um conjunto de variáveis explicativas sobre a variação média dos rendimentos individuais. A Tabela 3 expõe os resultados das estimações, com e sem correção do viés de seletividade amostral, a partir dos dados da PNAD/2012, além de mensurar o impacto das RMs no rendimento, tendo como grupo base a RM de São Paulo.

Em todos os estratos ocupacionais a razão inversa de *Mills* (Mills) foi estatisticamente significativa, indicando que se o viés de seleção amostral fosse desconsiderado, as estimativas dos coeficientes não iriam retratar os reais efeitos dos coeficientes. Analisando as categorias ocupacionais como um todo, o coeficiente da variável educação estaria subestimado caso a seletividade amostral não fosse considerada. O retorno da educação no modelo sem correção é de 9,53% já no modelo corrigido a taxa sobe para 10,6%. Nos demais grupos ocupacionais houve superestimação dos modelos sem correção para a variável educação, já para a variável que denota experiência o comportamento foi o oposto. Independente se no modelo estimado foi aplicado o processo de correção de Heckman (1979) as variáveis de capital humano (*educ* e *exp*) apresentam os sinais esperados. Diante do exposto, as análises serão feitas apenas com base nas estimações que utilizaram a correção de Heckman. Os coeficientes da variável *exp2* foram relevantes, na maioria dos casos, e apresentaram o sinal esperado (negativo), esta variável tem como objetivo verificar até que ponto anos a mais de experiência acarretam em acréscimo nos salários. Aglomerando o efeito da variável *exp* e *exp2*, infere-se que o salário cresce conforme a experiência aumenta, contudo a taxas decrescentes. Sintetizando os demais resultados de ambos os modelos, as principais diferenças observadas foram nos coeficientes das variáveis discriminatórias de cor e sexo; para as variáveis que representam as RMs as



diferenças não foram muito elevadas.

**TABELA 3 – Equações Mincerianas por Categoria Ocupacional com e sem correção do viés de seletividade amostral**

	Todas		Dirigentes		Serviços		Técnicos	
	Sem Correção	Corrigida	Sem Correção	Corrigida	Sem Correção	Corrigida	Sem Correção	Corrigida
<i>educ</i>	0.0953 (0.000)	0.106 (0.000)	0.112 (0.000)	0.732 (0.003)	0.0775 (0.000)	0.0456 (0.000)	0.0819 (0.000)	-0.0286* (0.679)
<i>exp</i>	0.0360 (0.000)	0.0645 (0.000)	0.0408 (0.000)	0.135 (0.001)	0.0269 (0.000)	0.00461* (0.361)	0.0392 (0.000)	0.147 (0.028)
<i>exp2</i>	-0.00042 (0.000)	-0.00097 (0.000)	-0.000418 (0.000)	-0.0010 (0.007)	-0.00033 (0.000)	-0.000054* (0.445)	-0.000501 (0.000)	-0.00227 (0.041)
<i>raça</i>	0.156 (0.000)	0.130 (0.000)	0.179 (0.000)	1.382 (0.005)	0.116 (0.000)	-0.0076* (0.807)	0.117 (0.000)	0.239 (0.126)
<i>sexo</i>	0.399 (0.000)	0.594 (0.000)	0.311 (0.000)	0.893 (0.000)	0.312 (0.000)	-0.677 (0.000)	0.342 (0.000)	8.360 (0.076)
<i>Belém</i>	-0.353 (0.000)	-0.429 (0.000)	-0.249 (0.001)	-0.231* (0.187)	-0.321 (0.000)	-0.320 (0.000)	-0.346 (0.000)	-0.344 (0.124)
<i>Fortaleza</i>	-0.398 (0.000)	-0.417 (0.000)	-0.349 (0.000)	-0.339 (0.025)	-0.376 (0.000)	-0.374 (0.000)	-0.441 (0.000)	-0.438 (0.029)
<i>Recife</i>	-0.405 (0.000)	-0.484 (0.000)	-0.404 (0.000)	-0.392 (0.008)	-0.374 (0.000)	-0.372 (0.000)	-0.399 (0.000)	-0.400 (0.047)
<i>Salvador</i>	-0.344 (0.000)	-0.378 (0.000)	-0.202 (0.001)	-0.202* (0.161)	-0.357 (0.000)	-0.356 (0.000)	-0.348 (0.000)	-0.348 (0.088)
<i>Belo Horizonte</i>	-0.0879 (0.000)	-0.0545 (0.000)	-0.104 (0.021)	-0.0972* (0.443)	-0.114 (0.000)	-0.113 (0.000)	-0.0888 (0.000)	-0.0890* (0.633)
<i>Rio de Janeiro</i>	-0.159 (0.000)	-0.219 (0.000)	-0.169 (0.000)	-0.177* (0.187)	-0.122 (0.000)	-0.121 (0.000)	-0.160 (0.000)	-0.162* (0.380)
<i>Distrito Federal</i>	0.0834 (0.000)	0.0686 (0.000)	0.168 (0.006)	0.171* (0.214)	0.0642 (0.000)	0.0645 (0.035)	0.0858 (0.000)	0.0776* (0.748)
<i>Curitiba</i>	0.00020* (0.988)	0.0309 (0.044)	-0.124 (0.045)	-0.119* (0.415)	0.00503* (0.751)	0.00718* (0.835)	-0.00606* (0.768)	-0.0026* (0.991)
<i>Porto Alegre</i>	-0.122 (0.000)	-0.0959 (0.000)	-0.256 (0.000)	-0.244 (0.042)	-0.0887 (0.000)	-0.0875 (0.000)	-0.128 (0.000)	-0.121* (0.485)
<i>Mills</i>		0.560 (0.000)		4.479 (0.012)		1.693 (0.000)		11.21 (0.089)
<i>cte</i>	5.404 (0.000)	4.654 (0.000)	5.800 (0.000)	-12.14* (0.090)	5.637 (0.000)	5.613 (0.000)	5.599 (0.000)	-10.98 (0.260)

Fonte: Dados da Pesquisa. Elaboração Própria.

Nota: \* Não significante.

Entre parênteses encontram os p-valores. cte= constante. Resultados robustos a heterocedasticidade.

Para as regiões com coeficientes estatisticamente significantes, e considerando a característica discriminatória raça, independente do nível ocupacional, verifica-se que os não brancos recebem, em média, menos que os brancos, porém a magnitude da diferença é variável de acordo com o grupo que está sendo observado. Técnicos apresentam a maior diferença e profissionais dos serviços a menor.

A variável usada como indicativo de discriminação por gênero, *sexo*, apresenta o comportamento esperado, os homens ganham em média mais que

as mulheres e assim como para raça a intensidade da diferença é variável entre as ocupações, sendo mais acentuada no grupo dos técnicos. Apesar desse comportamento ser condizente com a literatura, não se pode afirmar que essa diferença salarial é decorrente somente de componentes discriminatórios, porque se analisa apenas o coeficiente da variável discriminatória e não o retorno das variáveis para cada grupo (por exemplo, homens e mulheres). Devido a essa peculiaridade, o próximo tópico traz um método que decompõe a diferença salarial, para que possa ser feito uma análise com a perspectiva da discriminação.

Em relação às *dummies* indicativas de RM de residência dos indivíduos, a interpretação se dá levando em conta o grupo base selecionado, a RM de São Paulo. Dentre todas as RMs estudadas, apenas o Distrito Federal apresenta melhores rendimentos que a RM-base em todas as categorias ocupacionais. O efeito do uso da RM de São Paulo como base, pode está subestimando as análises, já que esta possui maiores proporções territoriais e concentra além da força de trabalho natural, os migrantes. Outro aspecto fundamental, é que se está analisando apenas o fato da localização da RM e não de cada atributo de capital humano e discriminatórios separadamente, dessa forma ainda não é possível verificar em qual dessas áreas há maior presença de comportamento discriminatório.

#### 4.2 Resultados da Decomposição de Oaxaca e Blinder (1973)

Para decompor o diferencial de rendimentos em dois componentes, um devido à capacidade produtiva (efeito dotação) e outro decorrente da discriminação contra algum grupo da população (efeito discriminação) é utilizado o método de Oaxaca e Blinder (1973), descrito na seção 3. Antes de analisar os resultados das estimações, é interessante destacar os seguintes aspectos gerais de cada decomposição:

- a) Gênero: O grupo dos homens tem os melhores salários, consequentemente como a diferença salarial é feita a partir do hiato entre mulheres e homens<sup>13</sup>, presume-se que o diferencial total seja negativo. Espera-se que o efeito dotação (positivo) haja no sentido contrário do efeito discriminação (negativo), devido às mulheres possuir melhores níveis de educação, como já mencionado. Esse é um resultado geral para todos os casos da decomposição por gênero, tanto por ocupação quanto por RM, ou seja, quando o sinal

13. A diferença calculada é visível a partir da equação (5).

do efeito dotação for positivo significa que o grupo com menores salários possui dotações produtivas maiores que o grupo com maior salário.

- b) Raça: Os não brancos, neste caso, são os que possuem a menor remuneração. A diferença calculada a partir da Equação (6) é entre não brancos e brancos, portanto, o sinal do diferencial deve ser negativo. Em relação aos componentes, como as características produtivas em média são menores para os não brancos, tanto o efeito dotação quanto o discriminação, agem no mesmo sentido (negativo), ou seja, ambos levam ao aumento do diferencial de rendimentos.

Expostos os aspectos gerais da decomposição e processo de correção, a primeira análise se dará com as RMs agrupadas em cada ocupação, estes resultados são descritos na Tabela 4. Como na seção anterior foi confirmada a existência do problema de seletividade amostral, que pode ser corrigido pelo procedimento de Heckman (1979), dessa forma nas decomposições aqui expostas foi também aplicado o processo de correção.

Decompondo a diferença em relação ao gênero, verifica-se que o efeito discriminatório, exceto no grupo dos serviços, se sobrepõe ao dotação. No grupo dos serviços a maior parte do diferencial é explicada pelo efeito dotação, indicando uma deficiência nos atributos desta parte da amostra e o efeito discriminação acaba agindo como um redutor, este fato, pode ser explicado pelo fato de as mulheres serem maioria nas atividades que compõem este grupo.

Em relação à raça, apenas no grupo dos dirigentes e técnicos o efeito discriminação se sobrepõe ao dotação, explicando mais de 100% do diferencial. Já no grupo dos profissionais dos serviços e no agregado das ocupações, o efeito dotação age no sentido de aumentar o diferencial, assim como o efeito discriminação, indicando que os não brancos possuem um menor nível de capital humano.

**TABELA 4 – Decomposição por Raça e Sexo: RMs agrupadas e Categorias Ocupacionais com correção de Heckman (1979)**

	Decomposição		
	Efeitos	Sexo	Raça
Todas as ocupações	TOT	-0.345	-0.323
	DOT	0.262 (-76%)	-0.100 (30%)
	DISC	-0.607 (176%)	-0.223 (70%)
Dirigentes	TOT	-0.296	-0.342
	DOT	0.644 (-217%)	0.141 (-41%)
	DISC	-0.940 (317%)	-0.484 (141%)
Serviços	TOT	-0.285	-0.247
	DOT	-1.032 (362%)	-0.187 (75%)
	DISC	0.747 (-262%)	-0.0604 (25%)
Técnicos	TOT	-0.206	-0.252
	DOT	0.388 (-188%)	0.0671 (-26%)
	DISC	-0.594 (288%)	-0.319 (126%)

Fonte: Dados da pesquisa. Elaboração própria.

Notas: TOT= Diferencial Total, DOT= Efeito Dotação, DISC= Efeito Discriminação. Entre parênteses encontram-se a porcentagem de cada efeito da decomposição no diferencial total.

#### 4.2.1 Análise das Diferenças Salariais por RM: Comparação entre os Gêneros

De acordo com a tabela 5, o diferencial salarial total entre homens e mulheres, em todos os grupos ocupacionais e RMs, se mostrou significante estatisticamente, exceto para a RM de Salvador no grupo de dirigentes, podendo ser um indicativo de existência de igualdade salarial intergêneros. As diferenças salariais apresentaram valores semelhantes internos a cada grupo, tendo como extremo mínimo a RM de Belém para o grupo dos técnicos e valor máximo para a RM de Recife na classe dos dirigentes.

No agregado das ocupações, com relação à decomposição por gênero, os dois componentes que explicam o diferencial foram relevantes estatisticamente e apresentaram a tendência esperada, exceto o elemento dotação do DF. Apesar da não relevância do fator referente aos atributos dos indivíduos, o componente discriminatório consegue explicar totalmente o diferencial.

No grupo dos dirigentes, apesar de algumas RMs apresentarem fatores dotação sem poder explicativo, o outro componente explica na totalidade a disparidade salarial entre homens e mulheres. Nos casos, onde ambos os fatores tinham poder de explicação, a direção esperada dos sinais foi mantida.

**TABELA 5 – Decomposição de Oaxaca e Blinder (1973) por Sexo: com correção de Heckman (1979)**

		RM1	RM2	RM3	RM4	RM5	RM6	RM7	RM8	RM9	RM10
Todas as ocupações	TOT	-0.329	-0.395	-0.414	-0.275	-0.314	-0.353	-0.345	-0.337	-0.365	-0.352
	DOT	0.175 (-53%)	0.339 (-85%)	0.317 (-76%)	0.148*	0.276 (-87%)	0.104 (-30%)	0.337 (-97%)	0.164 (-48%)	0.175 (-48%)	0.146 (-41%)
	DISC	-0.504 (153%)	-0.734 (185%)	-0.732 (176%)	-0.422 (153%)	-0.590 (187%)	-0.457 (130%)	-0.682 (197%)	-0.501 (148%)	-0.540 (148%)	-0.498 (141%)
Dirigentes	TOT	-0.305	-0.434	-0.277	-0.259	-0.393	-0.359	-0.445	-0.256	-0.103*	-0.168
	DOT	0.154*	0.0396*	0.516 (-186%)	0.664 (-256%)	0.231*	0.230 (-64%)	0.247*	0.199 (-77%)	0.0838*	0.849 (-504%)
	DISC	-0.459 (150%)	-0.474 (109%)	-0.793 (286%)	-0.923 (356%)	-0.624 (158%)	-0.589 (164%)	-0.692 (155%)	-0.454 (177%)	-0.187*	-1.016 (604%)
Serviços	TOT	-0.340	-0.291	-0.325	-0.305	-0.303	-0.260	-0.301	-0.275	-0.347	-0.274
	DOT	-1.266 (372%)	-0.495 (170%)	-0.457 (140%)	-1.270 (416%)	-1.127 (371%)	-0.827 (318%)	-0.874 (290%)	-0.742 (269%)	-0.635 (182%)	-1.007 (367%)
	DISC	0.926 (-272%)	0.204 (-70%)	0.132 (-40%)	0.965 (-316%)	0.824 (-271%)	0.567 (-218%)	0.573 (-190%)	0.466 (-169%)	0.287*	0.733 (-267%)
Técnicos	TOT	-0.095	-0.282	-0.342	0.279	-0.248	-0.298	-0.141	-0.193	-0.150	-0.200
	DOT	0.286 (-301%)	0.211 (-74%)	0.841 (-245%)	1.27 (455%)	0.775 (-312%)	0.778 (-261%)	0.076 (-54%)	0.858 (-444%)	0.760 (-506%)	0.713 (-356%)
	DISC	-0.381 (401%)	-0.493 (174%)	-1.184 (345%)	-0.986 (-355%)	-1.024 (412%)	-1.076 (361%)	-0.218 (154%)	-1.05 (544%)	-0.910 (606%)	-0.913 (456%)

Fonte: Dados da pesquisa. Elaboração própria. Notas: (a) RM1=Belém, RM2= Belo Horizonte, RM3= Curitiba, RM4= Distrito Federal, RM5= Fortaleza RM6= Porto Alegre, RM7= Recife, RM8= Rio de Janeiro, RM9= Salvador e RM10= São Paulo; (b) **TOT**: Diferencial Total; **DOT**: Efeito Dotação; **DISC**: Efeito Discriminação. \*Efeito não significativo. Entre parênteses encontram-se a porcentagem de cada efeito da decomposição no diferencial total.

Nas RMs de Belém, Belo Horizonte, Fortaleza e Recife, ainda para dirigentes, apesar do elemento de discriminação conseguir explicar o diferencial, a parte explicada pelos atributos não foi significativa. Essa característica indica

fortemente que há fatores particulares de cada local que não foram captados, apesar de ter sido aplicado, à decomposição, o processo de correção de viés de seletividade amostral de Heckman (1979).

Para o grupo dos serviços, os resultados apresentam um comportamento diferente do caso geral, os componentes da diferença salarial seguem direções distintas das observadas na literatura em questão. O diferencial total favorece os homens, tendo estes os melhores rendimentos, contudo os componentes que explicam a diferença seguem o seguinte padrão: o efeito dotação age no sentido de aumentar o diferencial e o efeito discriminação (relacionado apenas a característica de ser mulher) age como um redutor da disparidade. Esse novo padrão pode residir do fato das atividades desenvolvidas pelas pessoas pertencentes a este grupo serem, predominantemente, exercidas por mulheres<sup>14</sup>, logo ter essa característica favorece a inserção no mercado de trabalho, assim a pouca qualificação seria, realmente, o fator decisivo das diferenças salariais e não a discriminação.

Para os técnicos, grupo que teoricamente, de acordo com o agrupamento ocupacional, necessitaria de menores níveis de educação e experiência, apresenta resultados que pouco diferem das classes analisadas anteriormente, exceto o grupo dos serviços por não seguirem o caso geral. Uma peculiaridade é a intensidade do fator discriminatório, uma vez que este é mais elevado que nas demais classes. Contudo o efeito dotação age no sentido oposto, como esperado, a fim de dirimir as diferenças salariais entre homens e mulheres.

Ainda em relação ao grupo dos técnicos, observa-se o único caso em que o diferencial total favoreceu o grupo das mulheres, obtendo assim um sinal positivo, isto ocorreu para o DF. Apesar de outros fatores inerentes a região em destaque, possivelmente, terem sido omitidos, este é um resultado animador no que diz respeito à inserção e valorização da mulher no mercado de trabalho.

#### **4.2.2 Análise das Diferenças Salariais por RM: Comparação entre as Raças**

A lacuna existente entre os rendimentos de brancos e não brancos, captado pelo diferencial total, é em sua totalidade, negativa, pois os brancos possuem melhores rendimentos. As maiores diferenças são encontradas para o grupo de dirigentes e as menores para os profissionais dos serviços e técnicos. A RM de Rio de Janeiro apresenta a maior disparidade salarial para o grupo dos dirigentes. É importante salientar que esta é uma hipótese, apenas, qualquer atividade do grupo dos serviços, pode ser exercida por homens ou mulheres de forma igual. Contudo, para afirmar com mais veemência esta hipótese deve fazer um estudo mais aprofundado das atividades pertencentes ao grupo, o que não faz parte do escopo deste trabalho.

tes, já a RM de Curitiba tem a menor diferença salarial inter-racial, característica observada para os técnicos.

Com relação à diferença entre brancos e não brancos para a RM de Belém os resultados não foram significativos para o grupo dos dirigentes, o que pode ter sido causado pela baixa representatividade dos não brancos nesta categoria ocupacional, ou pela singularidade das etnias da Região Norte. Isto posto, esses resultados podem ser um indicativo de que os grupos em comparação não possuem dotações que os façam ter melhores salários. Ainda nessa região, há uma parcela maior do grupo de indígenas que, neste trabalho, não foram analisados de forma separada, podendo ter impactos nos resultados obtidos. Uma possibilidade de melhoria desse resultado seria a comparação entre três grupos raciais, brancos, negros e indígenas, entretanto para o propósito desta pesquisa, isto não será realizado.

**TABELA 6 – Decomposição de Oaxaca e Blinder (1973) por Raça: com correção de Heckman (1979)**

		RM1	RM2	RM3	RM4	RM5	RM6	RM7	RM8	RM9	RM10
Todas as ocupações	TOT	-0.170	-0.251	-0.156	-0.369	-0.174	-0.238	-0.223	-0.286	-0.501	-0.270
	DOT	-0.048 (28%)	-0.171 (68%)	-0.064 (41%)	-0.174 (47%)	-0.069 (40%)	-0.105 (44%)	-0.103 (46%)	-0.143 (50%)	-0.199 (40%)	-0.133 (49%)
	DISC	-0.122 (72%)	-0.079 (32%)	-0.091 (59%)	-0.195 (53%)	-0.105 (60%)	-0.132 (56%)	-0.121 (54%)	-0.143 (50%)	-0.302 (60%)	-0.137 (51%)
Dirigentes	TOT	-0.079*	-0.462	-0.223	-0.144	-0.174	-0.144	-0.316	-0.588	-0.573	-0.371
	DOT	0.458*	0.190*	0.031 (-13%)	0.208 (-144%)	0.289*	0.132 (-91%)	0.473*	0.866 (-147%)	-0.148*	0.931 (-250%)
	DISC	-0.538*	-0.651 (140%)	-0.254 (113%)	-0.352 (244%)	-0.462 (265%)	-0.276 (191%)	-0.789*	-1.454 (247%)	-0.425*	-1.301 (350%)
Serviços	TOT	-0.122	-0.143	-0.105	-0.271	-0.132	-0.188	-0.208	-0.213	-0.303	-0.155
	DOT	-0.020 (16%)	-0.198 (138%)	-0.199 (189%)	-0.181 (66%)	-0.0295 (22%)	-0.197 (104%)	-0.0991 (47%)	-0.133 (63%)	-0.202 (66%)	-0.124 (80%)
	DISC	-0.102 (84%)	0.0553 (-38%)	0.0944 (-89%)	-0.090 (34%)	-0.102 (78%)	0.0091*	-0.109 (53%)	-0.0795 (37%)	-0.101 (34%)	-0.0313 (20%)
Técnicos	TOT	-0.196	-0.152	-0.0664	-0.289	-0.133	-0.185	-0.121	-0.195	-0.314	-0.238
	DOT	-0.535 (272%)	-0.249 (163%)	0.030*	-0.783 (270%)	-0.712 (535%)	0.0549*	-0.159 (131%)	-0.379 (194%)	-0.324 (103%)	-0.258 (108%)
	DISC	0.338 (-172%)	0.0977 (-63%)	-0.096*	0.494 (-170%)	0.579 (-435%)	-0.240 (130%)	0.038*	0.184 (-94%)	0.010 (-3%)	0.020*

Fonte: Dados da pesquisa. Elaboração própria. Notas: (a) RM1= Belém, RM2= Belo Horizonte, RM3= Curitiba, RM4= Distrito Federal, RM5= Fortaleza, RM6= Porto Alegre, RM7= Recife, RM8= Rio de Janeiro, RM9= Salvador e RM10= São Paulo; (b) TOT: Diferencial Total; DOT: Efeito Dotação; DISC: Efeito Discriminação.\*Efeito não significativo. Entre parênteses encontram-se a porcentagem de cada efeito da decomposição no diferencial total.

Verificando os componentes da diferença salarial, estes trazem performances distintas para cada ocupação. Como já denotado anteriormente, am-

bos os efeitos decompostos agem no sentido de aumentar o diferencial para o aglomerado das ocupações, profissionais dos serviços e técnicos. Para todas as ocupações, o efeito discriminação tem mais impacto do que o efeito relativo aos atributos produtivos dos indivíduos, exceto para a RM de BH e do RJ, onde, respectivamente, a discriminação representa apenas 32% e 50% da diferença salarial. A região metropolitana de Salvador expressou, dentro desse grupo específico, o maior diferencial (-0.501) sendo este explicado em 60% pelo fator discriminatório e 37% devido ao baixo nível de fatores produtivos dos não brancos.

Para o grupo dos dirigentes, as RMs de Recife e Salvador, apesar de possuírem efeitos totais relevantes, não apresentaram componentes que conseguissem explicar esta diferença. As RMs de BH e Fortaleza, não tiveram o componente dotação significativo estatisticamente, contudo somente o efeito discriminatório consegue explicar mais de 100% do diferencial. Características inerentes as RMs em questão ou variáveis controles que não foram anexadas aos modelos podem ser as causas da não relevância estatística das partes decompostas que não obtiveram poder explicativo.

Ademais, as RMs de Curitiba, DF, Porto Alegre, RJ e SP no grupo dos dirigentes além de terem os elementos que compõem a diferença salarial, significantes, não seguiram a tendência do caso geral para a decomposição racial. O efeito dotação age no sentido oposto ao discriminação, indicando que para estas RMs os não brancos possuem dotações superiores ao brancos, todavia como o efeito total favorece os brancos, a diferença é explicada em sua maioria pelo fator discriminatório.

Para os profissionais dos serviços, o percentual do efeito discriminação sofreu elevações notáveis apenas nas RMs de Belém, Fortaleza e Recife. Nas demais regiões deste grupo os valores ficaram abaixo de 50% para o componente discriminatório, estando o efeito dotação com valores mais elevados. BH e Curitiba não seguem a tendência geral, tendo o efeito discriminatório como um redutor da disparidade salarial.

O grupo dos técnicos, assim como a maioria das RMs do grupo dos dirigentes, seguem o padrão não convencional para este tipo de decomposição, contudo aqui o efeito discriminação é que possuem o papel de reduzir a diferença salarial e não o efeito dotação.

Assim, apesar de possuírem deficiência de capital humano os profissionais do grupo dos técnicos, tem a característica raça a seu favor. Isso pode ser



advindo do tipo de atividade desenvolvida no grupo, onde por questões culturais e pouca qualificação<sup>15</sup>, os negros são aceitos mais facilmente. Apenas na RM de Curitiba, Porto Alegre e Recife algum dos componentes não tiveram poder explicativo.

## 5 CONSIDERAÇÕES FINAIS

O presente estudo utilizou dados da PNAD 2012 com o intuito de analisar o padrão da discriminação no mercado de trabalho das regiões metropolitanas brasileiras, contribuindo desta forma para a literatura em questão. Devido à importância da análise desagregada em níveis ocupacionais (MENEZES, 2013; GUIMARÃES, 2013), foi verificado o comportamento para quatro grupos distintos, cada qual com uma exigência diferente de níveis de capital humano.

Para captar o diferencial salarial inter-racial e intergêneros, foram estimadas equações mincerianas para cada estrato ocupacional, aqui foi observado o impacto do vetor de variáveis adotadas para explicar os rendimentos. Em seguida, utilizou-se o método de Oaxaca e Blinder (1973) para decompor o diferencial de rendimentos, em dois componentes: um devido a capacidade produtiva (efeito dotação) e outro puramente discriminatório. A decomposição foi realizada para cada RM em cada nível ocupacional. Em ambos os processos foi aplicado o método de correção do viés de seletividade amostral de Heckman (1979).

Os resultados do modelo de Mincer (1974) apresentam os comportamentos esperados, educação e experiência agem positivamente sobre os rendimentos e sua intensidade varia de acordo com o grupo ocupacional analisado. As variáveis indicativas de raça e gênero permitem verificar que existe uma tendência de que mulheres e não brancos recebem, em média, menos que os homens e indivíduos de cor branca, nesta ordem, onde tal resultado é verificado em todas as categorias ocupacionais utilizadas.

Entre todas as regiões analisadas, apenas o DF (Todas as ocupações e profissionais dos serviços) e RM de Curitiba (Todas as ocupações) apresentam melhores rendimentos que a RM de São Paulo (SP), esta heterogeneidade de resultados justifica a necessidade de verificar as partes do hiato salarial entre RMs e em cada nível ocupacional.

---

15. Para confirmar este comportamento deveria ser feita uma análise a fundo das atividades que compõem o grupo ocupacional. Ademais, atualmente, independente do setor de atuação a qualificação é um fator decisivo para o ingresso e permanência no mercado de trabalho.

Em relação ao gênero, verifica-se, a partir dos resultados da decomposição de Oaxaca e Blinder (1973) para o conjunto das RMs, que o diferencial é explicado principalmente pelo efeito discriminação, já quando se analisa pela ótica da raça, tanto a parte explicada como a não explicada tem impacto sobre o diferencial de rendimentos, isso para todas as ocupações e serviços.

Na análise para cada RM, no que diz respeito ao gênero, na maioria dos casos o comportamento foi o esperado, exceto para o DF no grupo dos técnicos. Os profissionais dos serviços obtiveram resultados que diferem do caso geral, neste grupo o componente discriminatório age como um redutor da desigualdade salarial, assim o fator decisivo para explicar o diferencial passa a ser o efeito dotação.

O grupo dos técnicos é o que mais discrimina por gênero. A presença mais acentuada para os técnicos pode ser derivada do fato destas ocupações serem antes ocupadas, principalmente, por homens, dessa forma a inserção das mulheres neste grupo pode estar acompanhada com a presença mais forte de preconceito.

Em relação à análise por raça, a análise conjunta das RMs em todos os níveis ocupacionais e para sua desagregação no agregado das ocupações, os resultados verificados se comportam como esperado, os efeitos se reforçam para explicar o diferencial. Em sua maioria o fator relativo à discriminação tende a permanecer mais elevado do que o efeito dotação. Nas demais categorias, a grande maioria dos efeitos decompostos fugiu da tendência verificada na literatura: a) para dirigentes: o efeito dotação age como um redutor da desigualdade para as RMs de Curitiba, DF, Porto Alegre, RJ e SP; b) para serviços: BH e Curitiba apresentam o fator discriminação como um redutor da desigualdade e c) para técnicos: em todas RMs o efeito discriminação agiu como um redutor da diferença salarial entre brancos e não brancos, exceto para a RM de Porto Alegre.

Em alguns casos os efeitos aqui decompostos e até mesmo a diferença total foram inconclusivos, devido a não significância estatística para os dados em questão. Esses resultados podem ser gerados pelas individualidades de cada RM, por estarem em regiões com características distintas, ou pela não utilização de variáveis que podem influenciar o salário, todavia aqui desconsideradas, por não fazer parte do objetivo deste trabalho.

De uma forma geral, a discriminação por gênero está mais acentuada no grupo dos técnicos, e menos presente no grupo dos dirigentes, apesar destes

possuírem efeitos totais elevados. Por outro lado, os resultados para raça são menos concentrados entre os agrupamentos ocupacionais. Para as duas decomposições ocorreram casos que divergiram das tendências esperadas, o que antes não foi verificado no modelo sem correção de seletividade amostral.

Os resultados encontrados sugerem que mesmo com a implantação de políticas públicas ativas sobre o mercado de trabalho e acesso a educação, como por exemplo, a lei de cotas<sup>16</sup> de acesso à Universidade, como também, da redução da taxa de desemprego ocorrida na última década, estes ainda não conseguiram mudar o curso do mercado de trabalho brasileiro, no sentido de reduzir o preconceito intergênero e inter-racial. Assim, sugere-se que os fazedores de políticas públicas deveriam ter mais atenção para os fatores ocupacionais, que acabam por gerar maiores diferenças salariais.

## REFERÊNCIAS BIBLIOGRÁFICAS

AMADEO, E; BARROS, R.P.; CAMARGO, J.M.; GONZAGA, G.; MENDONÇA, R. A natureza e o funcionamento do mercado de trabalho brasileiro desde 1980. IPEA, **Série Seminários**, 1994.

BARTALOTTI, O. A.C. **Discriminação Salarial por Cor e Gênero Revisitada**: Uma abordagem de decomposição contrafactual utilizando regressões quantílicas. 2007. 87 f. Dissertação (Mestrado em Economia). Escola de Economia de São Paulo. São Paulo. 2007.

BLINDER, A. S. Wage Discrimination: Reduced Form and Structural Estimates. **Journal of Human Resources**, v.8, p.436-455. 1973.

CAMBOTA, J. N. **Discriminação Salarial por raça e gênero no mercado de trabalho das regiões nordeste e sudeste do Brasil**: uma aplicação de simulações contrafactualis e regressão quantílica. 42f. Dissertação (Mestrado). Universidade Federal do Ceará. Programa de Pós Graduação em Economia, CAEN, Fortaleza, 2005.

CARVALHO, A. E.; SILVA, D.; NERI, M. **Diferenciais de Salários por Raça e Gênero: Aplicação dos procedimentos de Oaxaca e Heckman em Pesquisas Amostrais, complexas**. Disponível em: <[Economics Working Papers 638 \(Ensaios Econômicos da EPGE\)](#)>, 2006. Acesso em 4 de abril de 2014.

---

16 Lei nº 12.711/2012. Perguntas frequentes sobre a lei de cotas disponível em <http://portal.mec.gov.br/cotas/perguntas-frequentes.html>

CAVALIERI, C. H.; FERNANDES, R. Diferenciais de salários por gênero e cor: uma comparação entre as regiões brasileiras **Revista de Economia Política**, v. 18, n. 1(69), p. 158-175, janeiro-março/1998.

CURI, A.Z.; MENEZES-FILHO, N.A. “Os Determinantes das Transições Ocupacionais no Mercado de trabalho brasileiro”. In: XXXII Encontro Nacional da Anpec. 2004. **Anais do XXXII Encontro Nacional da Anpec**, 2004.

DIEESE. **Anuário do Sistema Público de Emprego, Trabalho e Renda 2010/2011: mercado de trabalho**. 3. ed. / Departamento Intersindical de Estatística e Estudos Socioeconômicos. São Paulo: DIEESE, 2011. Disponível em < <http://portal.mte.gov.br/geral/publicacoes/> >. Acesso em 04 de maio de 2014.

FORTIN, N.M., T. LEMIEUX, S. FIRPO. Decomposition Methods in Economics, in O. Ashenfelter and D. Card, eds., **Handbook of Economics**, Amsterdam: North-Holland, Vol. IV.A: 1-102. 2011

GIUBERTI, A. C.; MENEZES-FILHO, N. Discriminação de rendimentos por gênero: uma comparação entre o Brasil e os Estados Unidos. **Economia Aplicada**. Ribeirão Preto, v. 9, n. 3, Setembro, 2005.

GUIMARÃES, D. B. **Avaliações de abordagens metodológicas da equação de rendimentos aplicadas em cenários selecionados no mercado de trabalho**. 117f. Tese (Doutorado) – Universidade Federal do Ceará. Programa de Pós Graduação em Economia, CAEN, Fortaleza, 2013.

JANN, B., The Oaxaca-Blinder Decomposition for Linear Regression Models, **Stata Journal**. Number 8: p. 435-479. 2008

LEITE, M. P.; SOUZA, S. M. Igualdade de gênero e raça no Brasil: uma discussão sobre a política pública de emprego. **Dados-Revista de Ciências Sociais**, v. 53, n. 1, p. 195-231, 2010.

LEME, M. C. S.; WAJNMAN, S. Tendências de coorte nos diferenciais de rendimentos por sexo. In: Henriques, R. (org.), **Desigualdade e pobreza no Brasil**. IPEA, 2000.

MACHADO, J. A. F.; MATA, J. Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. **Journal of Applied Econometrics**. N. 20, p. 445-465, 2005.

MATOS, R. S. & MACHADO, A. F., Diferencial de rendimentos por cor e

sexo no Brasil (1987 – 2001). **Econômica**, Rio de Janeiro, v. 8, n.1, junho, 2006.

MENEZES, F. L. S. **Decomposição dos diferenciais de rendimentos entre os trabalhadores brasileiros por quantis e categorias ocupacionais**. 42f. Dissertação (Mestrado) – Universidade Federal do Ceará, Programa de Pós Graduação em Economia, CAEN, Fortaleza, 2013.

MENEZES FILHO, N. A.; CABANAS, P. H. F.; KOMATSU, B. K. Tendências Recentes do Mercado de Trabalho Brasileiro. *In*spere - Centro de Políticas Públicas (CPP), **Policy Paper**, nº 10, janeiro, 2014.

MINCER, J. **Schooling, Experience and Earning**. New York: Columbia University Press, 1974. 152p.

MIRO, V. H.; SULIANO, D. C. Discriminação de rendimentos por gênero e raça a partir de realidades socioeconômicas distintas. In: Encontro de Economia do Ceará em Debate – IPECE. 2009. **Anais do Encontro de Economia do Ceará em Debate**, v. 5, 2009.

MORAES, E. L., **Relação Gênero e Raça na Política Pública de Qualificação Social e Profissional**. Brasília, TEM/ SPPE/DEQ. 2005.

OAXACA, R. L. Male-Female Wage Differentials in Urban Labor Markets. **International Economic Review**, 14, 693-709. 1973.

OBSERVATÓRIO DAS METRÓPOLES. **Identificação dos espaços metropolitanos e construção de tipologias**: relatório da atividade 1. S.I., 2004. Projeto Análise das Regiões Metropolitanas do Brasil. Disponível em: <[http://www.observatoriodasmetropoles.ufrj.br/produtos/produto\\_mc\\_1.pdf](http://www.observatoriodasmetropoles.ufrj.br/produtos/produto_mc_1.pdf)>. Acesso em 08 de agosto de 2014.

\_\_\_\_\_. **Índice de Bem-Estar Urbano**. Luiz Cesar de Queiroz Ribeiro, Marcelo Gomes Ribeiro (Org.) 1. ed. - Rio de Janeiro: Letra Capital, 2013. 264p. Disponível em < [http://www.observatoriodasmetropoles.net/images/abook\\_file/ibe\\_u\\_livro.pdf](http://www.observatoriodasmetropoles.net/images/abook_file/ibe_u_livro.pdf)>. Acesso em 01 de outubro de 2014.

RAMOS, L. O desempenho recente do mercado de trabalho brasileiro: tendência, fatos estilizados e padrões espaciais. **Texto para Discussão**, n. 1.255, IPEA, 2007.

RAMOS, L.; BRITTO, M. O funcionamento do mercado de trabalho metro-

politano brasileiro no período 1991-2002: tendências, fatos estilizados e mudanças estruturais. **Textos para discussão**, n. 1011, IPEA, 2004.

SACHIDA, A.; LOUREIRO, P. R. A.; MENDONÇA, M. J. C. Discriminação Salarial

e Local de Moradia: um estudo para o Distrito Federal. In: XXXII Encontro Nacional de

Economia Anpec. **Anais do XXXII Encontro Nacional de Economia Anpec**, João Pessoa, 2004.

SANTOS, R. V. **Desigualdade de rendimentos e discriminação por gênero no Brasil em 1999**. 88 f. Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul, Programa de Pós Graduação em Economia, Porto Alegre, 2005.

SARAIVA, A. F. M. G. **Aspectos Microeconômicos da formação profissional e medida da discriminação sexual no mercado de trabalho português: uma abordagem semiparamétrica**. 207 f. Tese (Doutorado) – Faculdade de Economia, Universidade do Porto, Programa de Doutorado em Economia, Porto, 2007.

SOUZA, P. F. L. **A importância da discriminação nas diferenças salariais: uma análise para o Brasil e suas regiões para os anos de 2002, 2006 e 2009**. 98 f. Dissertação (Mestrado) – Universidade Federal do Ceará, Programa de Pós Graduação em Economia, CAEN, Fortaleza, 2011.

VILELA, T.; ARAÚJO, E.; RIBEIRO, E. P. Análise do Diferencial de Renda do Trabalho em 2008 entre Diferentes Gerações de Trabalhadores no Brasil. *Revista Economia*, 2012.

# ESTIMANDO RETORNOS DA EDUCAÇÃO NO CEARÁ: UMA APLICAÇÃO DO MÉTODO DO LASSO PARA SELEÇÃO DE VARIÁVEIS INSTRUMENTAIS FRACAS

Rafael B. Barbosa\*

Daniel Barboza Guimarães\*\*

## Resumo

Este trabalho pretende estimar os retornos da educação para o Estado do Ceará utilizando um conjunto grande de instrumentos fracos. Para identificar os instrumentos relevantes será utilizado o método de seleção de variáveis do Lasso (Least absolute shrinkage and select operator). Para corrigir a possível presença de viés na estimação de primeiro estágio, será adotado o procedimento do pós-lasso, como proposto por Belloni et al (2010, 2012, 2014). Com isso, espera-se obter estimativas consistentes e que não dependem unicamente da teoria econômica para sua determinação. Dos resultados encontrados, verificou-se que os retornos da educação tiveram magnitude significativamente menor do que àqueles encontrados na literatura. Isso pode servir como indicativo de que os métodos anteriores de seleção ad-hoc de instrumentos podem gerar resultados viesados.

**Palavras-Chave:** Retornos da educação, Seleção de variáveis, Instrumento fracos, Ceará.

## Abstract

The aim of this paper is to estimate the educational returns to Ceará using many weak instruments. To identify relevant instruments it was used the method of variable selection of Lasso (Least absolute shrinkage and select operator). The bias in the first stage was corrected by post-lasso method proposed

\* Professor dos Cursos de Ciências Econômicas e Finanças da Universidade Federal do Ceará (UFC), Campus de Sobral. Email: rafaelbarrosbarbosa@gmail.com (85) 999-301-938.

\*\* Professor do Departamento de Administração da Universidade Federal do Ceará (UFC). Email: barbozadan@hotmail.com.

by Belloni et al (2010, 2012, 2014). Then, we hope that the educational returns estimates to be consistent and that do not have only dependence by economic theory for its determination. The results show that the magnitude of estimates is significantly lesser than that was estimated by literature. That could indicate that the ad-hoc selection methods of instruments can generate biased estimates.

**Key words:** Educational Returns, Variable selection, Weak Instruments, Ceará.

**JEL:**J24, I21, I25.

## 1 INTRODUÇÃO

O capital humano é um dos principais determinantes da taxa de crescimento e do nível de bem-estar de um país. De acordo com Resende e Wyllie (2006) as fontes de crescimento econômico têm sido crescentemente associadas a investimentos em capital humano e o fenômeno da decolagem (*catching up*) de alguns países asiáticos esteve amplamente relacionado a investimentos expressivos em educação.

O debate em torno do peso da educação como fator explicativo da desigualdade de renda existente no Brasil é intenso. Destaca-se a corrente que defende que a escolaridade é o principal fator explicativo causal da desigualdade, por gerar diferenças de produtividade entre os indivíduos que podem persistir por todo seu ciclo de vida (MENEZES-FILHO (2001)).

As evidências indicam que os retornos econômicos à educação no Brasil estão entre os mais elevados do mundo. No entanto, recentemente, vêm sendo registrado um declínio deste retorno. Esse processo em parte deve-se ao próprio processo de expansão educacional, que aumentou a oferta relativa de pessoas com ensino fundamental e médio (MENEZES-FILHO, (2001)). De fato, o papel da escolaridade vem exercendo forte influência desde o início da década de 1990 em consequência da persistente redução dos retornos médios à educação, como apontado por alguns autores como, por exemplo, Ferreira e Veloso (2006).

Diversos artigos, recentemente publicados na literatura econômica brasileira, são enfáticos ao atribuírem papel importante à educação na redução contínua e acelerada da desigualdade de renda no Brasil, apesar do enorme papel das transferências ao longo dos últimos anos<sup>1</sup>.

---

1. Soares (2006b), Hoffman (2006), Barros et al. (2006), Barros, Carvalho, Franco e Mendonça (2007), Ferreira et al. (2006).



De acordo com Carvalho, Neri e Silva (2006) o propósito desses trabalhos é avaliar a situação em que indivíduos com atributos produtivos semelhantes, têm salários tão diferenciados. Para que tais propósitos sejam alcançados, ou seja, para capturar os efeitos sobre as diferenças de rendimentos causadas por diferenças de escolaridade formal dos indivíduos, grande parte desses estudos desenvolve versões da equação de salário proposta por Mincer (1974).

No entanto, equações mincerianas são fortemente marcadas pela presença da endogeneidade, isto é, a presença de variáveis omitidas que são fortemente correlacionadas com a educação e que possuem um efeito indireto sobre a renda do indivíduo. Uma das principais estratégias para corrigir o problema da endogeneidade está na utilização de variáveis instrumentais (VI).

A prática comum na utilização de variáveis instrumentais é bastante limitada por dois motivos. Primeiro, as VI's devem ter forte correlação com a variável endógena e baixa correlação com o erro. VI's que possuem esta característica são chamadas de VI's fortes (ANGRIST e PISCHKE (2009)). Tais tipos de VI são bastante difíceis de serem encontradas em base de dados tradicionais (Pesquisas Amostrais e Census), sendo mais comum encontrar um conjunto de VI's fracas, isto é, variáveis que possuem baixa correlação relativa com as variáveis endógenas. Segundo, a escolha das VI é normalmente realizada por meio da intuição teórica dos pesquisadores. Ou seja, poucos critérios estatísticos são utilizados para determinar se uma variável serve ou não como VI.

Diante destes dois tipos de problemas, pesquisadores têm buscado alternativas empíricas para justificar a utilização de um grande conjunto de VI's fracas. Duas estratégias têm sido tomadas, a redução da dimensão e a seleção de variáveis. Bai e Ng (2009) e Kapetanios e Marcellino (2010) buscam reduzir a dimensão de um conjunto de VI's fracas por meio da estimação de fatores comuns. Groen e Kapetanios (2009), por outro lado, utilizam mínimos quadrados parciais para o mesmo fim.

Mais recentemente, autores como Belloni et al (2014) propuseram métodos baseados no esparçamento de dados (*Approximate Sparse Models*), com destaque para o método do Lasso (*Least absolute shrinkage and select operator*) e Post-Lasso. A ideia básica consiste em permitir que sejam escolhidas, no primeiro estágio, variáveis instrumentais que têm elevado poder de explicação da variável endógena.

A técnica do Lasso consiste na estimação por Mínimos Quadrados Ordinários aplicando uma penalidade normada do tipo L1. Ou seja, trata-se de um

Método de Mínimos Quadrados Ordinários Restrito que faz com que os coeficientes estimados de variáveis que não contribuam fortemente para explicar o comportamento da variável endógena tendam a zero.

O método do Lasso<sup>2</sup> foi originalmente desenvolvido para selecionar variáveis que possuam um elevado poder preditivo. Devido ao trade-off entre variância e viés (HASTIE et al (2012)), o Lasso tende a produzir estimativas viesadas. Para reduzir o peso deste viés, Belloni et al (2014) propuseram o Post-Lasso, que consiste em estimar o modelo em segundo estágio, utilizando apenas os instrumentos selecionados pelo Lasso, por mínimos quadrados ordinários.

Alguns estudos posteriores têm mostrado relativo sucesso da aplicação de tais técnicas. Belloni et al (2014) comparam os resultados apresentados pelo Lasso e o Post-Lasso aplicados a mesma base de dados de Angrist e Krueger (1991) e mostram que o trimestre de nascimento constitui um instrumento forte. Outras aplicações recentes foram Hansen *et al* (2015), Belloni *et al* (2012), Belloni *et al* (2013).

O objetivo deste trabalho consiste em estimar as equações mincerianas para o Estado do Ceará utilizando as técnicas de seleção de variáveis do Lasso e do Post-Lasso. Espera-se que tais métodos forneçam estimativas mais apropriadas dos retornos da educação, visto a dificuldade de selecionar, teoricamente, variáveis instrumentais fortes no contexto brasileiro.

Será utilizada a base de dados da Pesquisa Nacional por Amostragem Domiciliar (PNAD) realizada anualmente pelo IBGE. Para comparar temporalmente os resultados serão estimados os retornos educacionais para os anos de 2002 e 2012. Além disso, a amostra será estratificada entre regiões metropolitanas e não metropolitanas. Para situar o retorno educacional no contexto nordestino, o mesmo exercício será realizado para a região Nordeste e posteriormente comparado com o Ceará.

Portanto, este estudo realiza duas contribuições para a literatura sobre o retorno da educação. Primeiro, contribui para a discussão da seleção de variáveis instrumentais fracas, tendo o retorno da educação como foco. Diante da dificuldade de obtenção de instrumentos fortes, as técnicas do Lasso e do Post-Lasso mostram-se como uma alternativa factível e promissora. Não é do conhecimento dos autores nenhum trabalho que tenha aplicado tais técnicas

---

2. Tal técnica faz parte de um conjunto de métodos que buscam selecionar variáveis que reduzam uma função perda, chamados de métodos de redução (Shrinkage Methods). Dentre este ainda se incluem, elastic net, ridge regressions, entre outros.

no Brasil. Segundo, ajuda a compreender o papel da educação na elevação da renda os indivíduos no Nordeste e no Ceará. Destaca-se o papel da educação na superação da pobreza entre os indivíduos.

Além desta introdução este trabalho está subdividido nas seguintes seções. Na seção seguinte, será feita uma análise das evidências empíricas da literatura. Na seção três, será apresentada a metodologia do Lasso e do Post-Lasso e discutido a base de dados utilizada. A seção quatro exhibe os resultados e a última seção apresenta e discute as principais conclusões.

## 2 REVISÃO DA LITERATURA

O debate sobre o retorno da educação no salário dos indivíduos foi iniciado por Schultz (1961) e Ben-Porath (1967). No entanto, segundo esses autores, tais discussões só foram intensificadas a partir de 1974, quando Mincer formalizou teoricamente a derivação da equação de rendimentos, a qual utiliza apenas o nível educacional como representativo do capital humano. Nesse estudo, o autor constatou que um ano adicional de escolaridade gerava um aumento de 11,5% nos rendimentos.

Desde então, essa formulação passou a ser amplamente utilizada em estudos do mercado de trabalho, os quais procuram obter informações sobre os determinantes do salário dos trabalhadores, a partir de suas características individuais (escolaridade, experiência, cor/raça ou local de moradia) e informações sobre o mercado de trabalho (setor de atividade ou ocupação).

Ashenfelter e Krueger (1994) coletaram dados sobre salários e educação para 198 gêmeos univitelinos, residentes nos Estados Unidos, com diferentes níveis de escolaridade. Os resultados indicaram que mesmo entre os indivíduos nascidos na mesma família e com características genéticas idênticas, aqueles que estudaram mais recebem um salário maior, onde um ano a mais de estudo pode gerar um aumento salarial entre 12% a 16%. Rouse (1999), Arias e McMahon (2001), os quais afirmaram que os retornos da educação variam em torno de 10%, 13% e de 5% a 10%, respectivamente.

Utilizando informações referentes ao nível educacional e a experiência de trabalho, Boudarbat, Lemieux e Riddell (2008) buscaram analisar a evolução dos retornos do capital humano no Canadá sobre as diferenças salariais, no período de 1980 a 2006. Observaram que os retornos à educação permaneceram maiores para as mulheres. No entanto, para o período de 2000-2006 os

resultados mostram uma tendência decrescente das diferenças salariais medidos pelo retorno à educação.

Zepeda *et al* (2009) utilizaram variáveis demográficas e socioeconômicas, dentre elas, o nível educacional para identificar os fatores que explicam as alterações observadas na média dos rendimentos em três países da América Latina – Brasil, Chile e México. Os autores constataram que a única mudança mais significativa em todos os países e períodos, foi o impacto do nível educacional sobre a força de trabalho, indicando que a escolaridade é o fator mais importante para o aumento nos rendimentos.

Já Sapelli (2009) verificou as taxas de retornos da educação no Chile, onde afirmou que estas taxas variam de acordo com o nível educacional, divididos em três grupos, educação básica, média e superior, conforme comprovado com os resultados empíricos de 5,6%, 14,3% e 25%, respectivamente. Além disso, o trabalho compara estes resultados com os observados na Argentina, Colômbia, México e Brasil e conclui que apenas os retornos no ensino básico se equiparam com os do Chile.

Um dos primeiros autores a ressaltar a importância da educação como fator explicativo para a desigualdade brasileira foi Langoni (1973). De acordo com esse autor, o aumento da desigualdade no Brasil, entre 1960 e 1970, ocorreu devido ao aumento na demanda por trabalhadores qualificados associados à industrialização.

Na tentativa de capturar a influência do nepotismo ou de habilidades não observadas na relação entre educação e salários, Lam e Schoeni (1993) utilizaram informações referentes a educação do pai, da mãe, do sogro e da sogra, além da educação do próprio indivíduo, contidas na PNAD de 1982, como determinantes dos rendimentos dos indivíduos. O efeito da educação sobre os salários dos indivíduos cujos pais, mães, sogros e sogras têm o mesmo nível educacional permanece em torno de 11%.

Barros (1997, 2000) enfatizou que um dos principais problemas sociais no Brasil decorre do baixo nível e da má distribuição da educação na população brasileira. Pessoa e Barbosa Filho (2008) verificaram que as taxas de retorno do ensino primário (10%), ginásial (14%) e universitário (18%) mantiveram-se extremamente elevadas no Brasil. Os autores concluem que o investimento em educação no Brasil é extremamente atrativo, uma vez que oferece elevadas taxas de retorno. Mariano e Arrais (2011) verificaram que as taxas de retorno da educação variam de acordo com as regiões brasileiras, sendo a região Nordeste a

que apresentou o menor valor dentre elas (14,68%). Sendo assim, o diferencial entre os retornos da educação causada pela diferença regional chega a 26%.

Percebe-se que, mesmo em um período onde a desigualdade de renda no Brasil esteve entre as mais baixas nos últimos 30 anos aliada ainda a uma rápida expansão educacional, os resultados encontrados, apesar de não comparáveis com outros estudos semelhantes devido às diferentes variáveis explicativas e base de dados, mostram que o prêmio à escolaridade no Brasil ainda se mantém em patamares elevados (SULIANO; SIQUEIRA, 2012).

Portanto, das diversas variáveis que influenciam o salário de um indivíduo, o nível de escolaridade é um dos que tem maior importância (IOSCHPE, 2004). Além disso, conforme enumerado nos trabalhos supracitados, essa influência apresenta uma variabilidade considerável, ou seja, o impacto da educação no salário apresenta variações significativas quando são realizadas comparações entre os gêneros, raças, países e regiões geográficas, por exemplo.

Esses estudos, ao procurarem evidenciar a importância da educação na explicação dos diferenciais de renda, acabam incorrendo em uma série de problemas. O primeiro problema é o viés de seleção amostral, o qual, de acordo com Heckman (1979), pode surgir praticamente por dois motivos: pela auto-seleção de indivíduos ou unidades amostrais investigadas e pelas decisões de seleção da amostra pelos analistas.

Cameron e Trivedi (2005) indicam que a omissão e/ou a característica endógena de uma variável em uma equação, especificamente de rendimentos, poderá causar um problema grave de viés e inconsistência nos estimadores quando se aplica o usual método de mínimos quadrados ordinários (MQO), em virtude do grave problema de endogeneidade entre educação e salários que as equações mincerianas enfrentam (SULIANO; SIQUEIRA, 2012).

Na tentativa de solucionar o problema da endogeneidade entre educação e salários, diversos autores passaram a utilizar o método de Variáveis Instrumentais (VI). No entanto, Heckman e Urzua (2009) relatam uma série de problemas que podem emergir a partir de uma estimação por VI, como por exemplo, deficientes instrumentos provêm estimativas viesadas; estimativas de VI advêm de fortes restrições tomadas a priori sobre os dados; em modelos heterogêneos, diferentes instrumentos fornecem estimativas distintas; estimativas de VI, dependendo dos instrumentos utilizados e hipóteses feitas, gerarão diferentes estimativas do retorno da educação, as quais podem levar a incorretas interpretações. Mariano e Arraes (2011) e Suliano e Siqueira (2012) observam

que, que o modelo estimado pelo método de VI apresenta um Erro Quadrado Médio (EQM) maior que o modelo endógeno, indicando que, o método de VI apresenta um pior ajuste, o que prejudica a questão da previsão.

### 3 METODOLOGIA

Um dos problemas recorrentes na estimação de equações de salário mincerianas deve-se à presença de endogeneidade, causada pela existência de correlação entre as variáveis explicativas e o erro. No caso, a educação apresenta forte correlação com o erro devido à existência de variáveis omitidas como habilidade inata, herança familiar entre outras.

Variáveis instrumentais (VI) tem sido frequentemente aplicadas, com relativo sucesso na literatura (ANGRIST e KRUEGER (1991), CARD (1995), ANGRIST (1990), entre outros <sup>3</sup>), para tentar reduzir o efeito da endogeneidade. Todavia, a identificação de quais variáveis devem ser consideradas instrumentos fortes tem sido, na maioria dos casos, realizada por meio do uso da teoria econômica. Isto é, determinada variável será considerada uma VI ideal desde que possua forte correlação com a variável endógena e nenhuma correlação com o erro. A primeira condição é facilmente testada, porém, a segunda condição não é, uma vez que depende do erro, o qual é não observável.

Dessa forma, para selecionar uma VI os pesquisadores buscam encontrar justificativas teóricas que possibilitem inferir a ausência de correlação com a variável e o erro. Tal estratégia pode não necessariamente resultar em instrumentos capazes de corrigir o problema da endogeneidade, uma vez que tal procedimento pode acabar selecionando variáveis que servem apenas como VI's fracas e, portanto, podem elevar o viés devido à endogeneidade<sup>4</sup>.

Recentemente, algumas estratégias econométricas alternativas têm sido propostas para lidar com a pouca existência de variáveis instrumentais fortes, as quais buscam extrair de um conjunto de variáveis instrumentais fracas aquelas que podem ser utilizadas para estimar os modelos estruturais. Tais alternativas objetivam evitar a total dependência da escolha das VI's pela teoria econômica. O objetivo consiste em aliar a teoria econômica, necessária mesmo que em menor grau, a técnicas estatísticas de eficiência sobre o poder preditivo das candidatas a VI.

3. Uma recente literatura tem buscado analisar os retornos da educação diante de heterogeneidade provocadas por políticas públicas, mas não captadas pelas variáveis instrumentais. Destacam-se: Heckman et al (2006, 2008), Carneiro et al (2011), entre outros.

4. Ver Angrist and Pischke (2009).

Dois métodos têm ganhado destaque: Redução da dimensionalidade e seleção de variáveis. A redução da dimensionalidade consiste em extrair, de um conjunto de variáveis instrumentais fracas, variáveis latentes que resumam o comportamento da relação entre educação e seus instrumentos. Bai e Ng (2008) e Kapetanios e Marcellino (2010) utilizam componentes principais e Groen e Kapetanios (2009) mínimos quadrados parciais.

A redução da dimensionalidade, porém, não se aplica diretamente ao problema aqui estudado, pois, requer um painel de dados, em que são extraídos fatores comuns. Estes formam uma combinação multiplicativa entre os fatores temporais e a heterogeneidade dos indivíduos (fator loadings).

A seleção de variáveis, por sua vez, busca escolher, dentre um conjunto de variáveis instrumentais fracas, aquelas que têm maior poder de explicação sobre a variável endógena. Tais métodos forçam as variáveis instrumentais fracas, com pouco poder explicativo, a ter seus coeficientes estimados tendendo a zero.

Existem dois motivos para excluir determinadas variáveis. Primeiro, é que tais variáveis podem não ter significativo poder de explicação do comportamento da variável endógena. Nesse caso, a inclusão desta variável pode resultar em relativa perda de eficiência em comparação ao ganho em poder de explicação. Segundo, duas variáveis podem ter elevada correlação, de forma que, a presença de uma apenas sirva para explicar o comportamento da variável endógena.

Dentre os métodos de seleção de variáveis disponíveis, destaca-se a aplicação do Lasso e do Post-Lasso, como em Belloni et al (2012, 2014, 2013) e Hansen et al (2015). O método do Lasso foi originalmente desenvolvido por Tibshirani (1996). Este método busca selecionar variáveis que tenham elevado poder preditivo para a variável dependente. Belloni et al (2012) adotam este método para estimar equações estruturais, utilizando variáveis instrumentais fracas, as quais foram previamente selecionadas.

Considere a seguinte equação estrutural minceriana:

$$y_i = \alpha + \beta_1 y_{1i} + x' \delta + e_i, \quad i = 1, \dots, n \quad (1)$$

Em que:  $y_i$  é o log do salário do indivíduo  $i$ ;  $y_{1i}$  corresponde a medida de capital humano, no caso deste artigo, será utilizado os anos de estudo;  $x$  corresponde a um vetor de variáveis de controle de tamanho  $(k \times n)$  e  $e_i$  corresponde ao erro, que por hipótese tem as seguinte propriedade estatística:  $e_i \sim iid N(0,1)$ . A endogeneidade é caracterizada pela forte correlação entre  $y_{1i}$  e  $e_i$ .

Suponha que exista um conjunto  $Z$  de tamanho  $(pxn)$  de variáveis instrumentais fracas. Baseando-se na estratégia de Mínimos Quadrados Ordinários em Dois Estágios (MQO2E) a equação de primeiro estágio pode ser escrita como:

$$y_{1i} = Z' \gamma + x' \delta + u_i \quad (2)$$

Observe que se for  $p$  muito grande, então o método dos mínimos quadrados se torna inviável na estimação do primeiro estágio. Dessa forma, o objetivo do método do Lasso consiste em selecionar  $s$  variáveis de  $Z$  ( $s \ll p$ ), de modo a tornar a estimação da equação (2) factível. Para tanto, o Lasso estima os parâmetros por meio da otimização convexa restrita, a qual se dá da seguinte forma:

$$\beta_L = \arg \min \hat{Q}(\beta) + \frac{\lambda}{n} \|\beta\|_1 \quad (3)$$

Em que,  $\hat{Q}(\beta) = E[(y_{1i} - f' \beta)^2]$ ,  $f = (Z, x)$ ,  $\beta = (\gamma, \delta)$ ,  $\|\beta\|_1 = \sum_{j=1}^p |\beta_j|$ . Isto é, o Lasso busca minimizar a soma dos quadrados dos resíduos impondo uma restrição aos parâmetros. O termo  $\lambda$ , chamado de *tuning point*, funciona com uma penalidade, uma vez que corresponde ao peso que será dado a restrição. Observe que a equação (3) pode ser reescrita da seguinte forma:

$$\beta_L = \arg \min \sum_{i=1}^n (y_{1i} - f' \beta)^2$$

sujeito a  $\sum_{j=1}^p |\beta_j| \leq t$

Esta última forma deixa mais explícita a forma de otimização restrita proposta pelo método do Lasso. Um elemento fundamental da seleção de variáveis via Lasso consiste na escolha correta do *tuning point*. Tibshirani (1996) propôs a escolha de via correlação cruzada, isto é, o lasso é aplicado considerando um intervalo para o valor do . O escolhido é aquele que minimiza o erro quadrático médio.

Tal estratégia, entretanto, não é diretamente aplicável a problemas de estimação de modelos estruturais, por meio de variáveis instrumentais, pois o interesse não é necessariamente realizar previsão. Desse modo, Bickel, Rivot e Tsybakov (2009) propuseram o seguinte valor para  $\lambda$ :

$$\lambda = 2 \cdot c \sigma \sqrt{2n \log(2p/\zeta)}$$

Em que,  $c > 1$ ,  $1 - \zeta$  refere-se ao intervalo de confiança necessário para ficar próximo de 1, onde consiste no desvio padrão do erro. Tal derivação é motivada buscando obter *near-optimal* taxas de convergência dos estimadores após



a seleção. Todavia, tal escolha possui duas limitações. Primeiro, foi derivada supondo a homocedasticidade. Segundo, não é factível, tendo em vista que  $\sigma$  não é observável. Portanto, visando resolver esse problema, Belloni *et al* (2010) propuseram um  $\lambda$  factível e aplicável na presença de heteroscedasticidade. Este será utilizado para estimar (3) neste artigo.

O Lasso é um caso particular de um conjunto de métodos de seleção de variáveis, conhecidos como *shrinkage methods*. Todavia, o Lasso possui duas vantagens frente às outras técnicas de seleção de variáveis. Primeiro, ele seleciona variáveis ao forçar certos parâmetros irrelevantes a apresentarem valores tendendo à zero. Assim, supondo que existam duas variáveis que possuem alta correlação com a variável explicativa, então uma delas terá seu parâmetro estimado tendendo à zero. Outras técnicas como regressões do tipo ridge não produzem esta característica.

Segundo, a forma funcional da equação (3) é convexa. Portanto, é computacionalmente mais fácil a obtenção de ótimos, onde tais ótimos são globais. Estas duas propriedades fazem com que o método do Lasso seja bastante apropriado para selecionar quais variáveis instrumentais fracas são relevantes para explicar o comportamento da variável endógena, como em Belloni *et al* (2010).

Entretanto, o método do Lasso possui também limitações. De fato, o Lasso não lida bem com a presença de multicolinearidade (ZOU e HASTIE, 2005). Isso ocorre justamente pela forma de seleção de variáveis em que aquelas variáveis que possuem elevada correlação entre si têm seus coeficientes estimados forçados a tenderem a zero.

Mais grave, porém, é o fato de o Lasso produzir estimativas viesadas. O método do Lasso foi desenvolvido para selecionar variáveis de forma a aumentar o poder preditivo dos modelos. Isto é, as variáveis são selecionadas buscando minimizar a função perda quadrática. Todavia, à medida que a variância se reduz, o viés se eleva <sup>5</sup>.

Para reduzir o peso do viés, métodos alternativos têm sido propostos. Por exemplo, Zou e Hastie (2002) propuseram o *elastic net*, que busca suavizar a perda com o viés do Lasso sem, no entanto, perder a qualidade na seleção das variáveis. Belloni *et al* (2012, 2013) por sua vez, adotam o Post-Lasso, onde este método consiste em reestimar a equação

Dessa forma, primeiro realiza-se a seleção de variáveis, em que dentre

---

5. De fato, sendo  $y_0$  a variável dependente e  $\hat{f}(x_0)$  os previsores estimados então, o erro quadrático médio pode ser decomposto em  $E(y_0 - \hat{f}(x_0))^2 = Var(\hat{f}(x_0)) + [Viés(\hat{f}(x_0))]^2 + Var(\epsilon)$ , em que: consiste no erro de aproximar  $y_0$  por  $\hat{f}(x_0)$ . Para maiores detalhes ver, Hastie *et al* (2009).

as possíveis variáveis instrumentais são escolhidas ( $s \ll p$ ). Logo em seguida, a equação (2) é reestimada considerando apenas as variáveis selecionadas. Por fim, utiliza-se a estimativa do primeiro estágio para estimar a equação estrutural (1) de segundo estágio.

Belloni *et al* (2010) mostraram que esta técnica se adapta bem a problemas em que existem poucas (ou nenhuma) variável instrumental forte, porém, existem muitos instrumentos fracos. Este é justamente o caso da maior parte das bases de dados brasileiras de onde se pode extrair dados para a análise do retorno da educação.

Portanto, neste trabalho adota-se a técnica de estimação do Post-Lasso, a qual consiste em aplicar uma seleção *ex-ante* a estimação do modelo estrutural sobre um conjunto grande de variáveis instrumentais fracas.

É importante observar que não necessariamente as variáveis selecionadas precisam ter significado econômico. De fato, estas terão apenas forte poder de explicação sobre a variável endógena.

Todavia, como teoricamente se sabe que algumas variáveis são fundamentais para a estimação das equações mincerianas, estas variáveis não estarão sujeitas a seleção via Lasso e serão consideradas variáveis explicativas durante a estimação da equação estrutural.

Vale ressaltar que, não é de conhecimento dos autores da existência de nenhum outro trabalho que tenha utilizado técnicas de seleção de variáveis para estimação de equações de salário no contexto brasileiro. Assim, ademais a importância social e de planejamento de políticas públicas, este trabalho proporciona um avanço na estimação do retorno da educação no Brasil.

### 3.1 Base de Dados

Será utilizada a base de dados da Pesquisa Nacional por Amostragem Domiciliar (PNAD) realizada anualmente pelo IBGE. Foram escolhidos dois períodos, com uma diferença de 10 anos, para verificar o comportamento temporal dos retornos da educação. Assim, os anos selecionados foram 2002 e 2012.

As variáveis de controle escolhidas são: experiência (*exp*), experiência ao quadrado (*exp\_2*), sexo (feminino como referência), negros e pardos. Tais variáveis serão utilizadas como potenciais instrumentos, entretanto, mesmo que não sejam selecionadas pelo Lasso, figurarão na equação estrutural.

Algumas observações foram eliminadas para gerar resultados plausíveis

com a média da população e evitar a presença de *outliers*. Desse modo, foram consideradas pessoas acima dos 22 anos, cuja renda individual é inferior a R\$ 5000,00<sup>6</sup>. A variável dependente é o log do salário.

Inicialmente, quarenta e oito (48) variáveis foram selecionadas como potenciais instrumentos fracos. Entretanto, sabe-se que certas transformações geram novas variáveis instrumentais fracas (BELLONI *et al* (2012), CHEN (2007)). Assim, foram realizadas transformações polinomiais (elevar ao quadrado e ao cubo) e transformações por interação entre as VI's fracos. Dessa forma, ao final, foram utilizados oitenta (80) potenciais instrumentos fracos<sup>7</sup>.

A amostra foi estratificada entre região metropolitana e não metropolitana. O objetivo é identificar se existem diferenças no retorno da educação entre capital e interior. Essa possível diferença decorre da existência de estruturas de mercado de trabalho e oferta de serviços educacionais diferenciados entre tais regiões. Por fim, o mesmo exercício foi realizado para a região Nordeste com o objetivo de se verificar como se comportam os resultados encontrados para o Ceará quando comparados à região a qual ele pertence.

A Tabela 01 apresenta os valores médios e os desvios-padrões da renda pessoal e dos anos de estudo para os anos de 2002 e 2012. A amostra foi estratificada entre região metropolitana e não metropolitana.

**TABELA 1: Estatísticas Descritivas das Variáveis Renda Pessoal e Anos de Estudo.**

<u>Variáveis</u>	<u>Média</u>		<u>Desvio-Padrão</u>	
	<u>2002</u>	<u>2012</u>	<u>2002</u>	<u>2012</u>
<u>Amostra Completa</u>				
<u>Renda Pessoal</u>	352,30	718,30	338,75	385,79
<u>Anos de Estudo</u>	8,29	9,35	4,68	4,2
<u>Região Metropolitana</u>				
<u>Renda Pessoal</u>	372,01	753,54	347,04	379,28
<u>Anos de Estudo</u>	8,63	9,64	4,56	4,07
<u>Região Não Metropolitana</u>				
<u>Renda Pessoal</u>	177,60	463,94	174,59	333,67
<u>Anos de Estudo</u>	5,2	7,24	4,64	4,53

Fonte: Elaboração dos autores a partir dos dados da PNAD.

6. Valores atualizados em 2002.

7. Todas as rotinas utilizadas, sejam para estimar o post-lasso ou para criar e ajustar a base de dados encontram-se disponíveis desde que solicitadas no email: rafael\_barbosa@ufc.br.

Em relação aos anos de estudo, verifica-se que também houve um crescimento deste indicador, em termos médios, no período analisado. Além disso, constatou-se que o crescimento dos anos de estudo seguiu o padrão do crescimento verificado na renda, uma vez que este foi mais relevante na região não metropolitana (39%) do que na região metropolitana (12%). Vale salientar que, mesmo com essa disparidade significativa, em média, os residentes na região metropolitana apresentam mais anos de estudo. Diante dos resultados verificados na tabela 1, verifica-se que, em média, indivíduos mais escolarizados recebem maiores salários.

Apesar de analisar o retorno da educação entre as regiões metropolitanas e não metropolitana do Ceará, os resultados da amostra completa serão comparados aos resultados para a região Nordeste. O objetivo deste exercício adicional consiste em entender se o retorno educacional do Ceará está inserido no contexto regional.

#### 4 RESULTADOS

Esta seção se reserva a apresentação e análise dos resultados a partir das estimações da equação minceriana de salários, equação (1), com o propósito de se verificar o impacto dos anos de estudo nas remunerações dos indivíduos no Nordeste e no Ceará. Convém mencionar que foram feitas algumas divisões na amostra referente ao Estado do Ceará, visando homogeneizá-la de forma a dar uma maior robustez aos resultados. Desse modo, as estimativas supracitadas foram realizadas levando-se em consideração a Região Metropolitana e a Região Não Metropolitana. Os resultados podem ser verificados na tabela 2, apresentada a seguir.

Percebe-se que os modelos estão bem ajustados e que quase todos os coeficientes estimados apresentam sinais condizentes com a literatura que aborda a problemática dos retornos da educação.

Em todas as estimações verifica-se que indivíduos do sexo masculino, com um maior nível de escolaridade, brancos e mais experientes recebem, em média, maiores salários. Estes resultados já são esperados na literatura.

Em relação ao retorno à educação observa-se que tanto as regiões Nordeste como o Estado do Ceará apresentaram uma redução neste índice entre os anos de 2002 e 2012. Na região Nordeste, o retorno à educação apresentou uma pequena redução, uma vez que passou de 5,9%, em 2002, para 4,2%, em 2012.

Já o Estado do Ceará, quando comparado ao Nordeste, apresentou uma redução bastante significativa no período analisado. Em 2002, um ano a mais de estudo acarretava em um aumento, em média, de 11% no rendimento do indivíduo, ao passo que em 2012 este aumento passou a ser de apenas 4%. Portanto, atualmente, o Estado do Ceará apresenta um retorno à educação inferior ao da Região Nordeste.

Este resultado pode ser devido ao aumento do crescimento educacional verificado nos últimos anos no Estado. De fato, com a expansão da educação, o mercado de trabalho se torna mais concorrido, reduzindo com isso os retornos médios à educação.

No tocante as análises estratificadas, ou seja, região metropolitana e não metropolitana, verifica-se que nos dois casos manteve-se a trajetória de queda no período analisado. Na região não metropolitana tem-se um maior retorno à educação, 11% em 2002 e 7% em 2012, contra 6% em 2002 e 0,9% em 2012 na região metropolitana.

**TABELA 2: Resultados para as Equações Estruturais após a Seleção das Variáveis.**

Variáveis	2002				2012			
	Nordeste	Ceará			Nordeste	Ceará		
		Amostra Toda	RM	Não RM		Amostra Toda	RM	Não RM
<u>Educ</u>	0.059* (0.013)	0.1152* (0.0087)	0.0671* (0.0044)	0.1101* (0.0103)	0.042* (0.006)	0.0466* (0.0059)	0.0091* (0.0044)	0.0760* (0.0107)
<u>Exp</u>	-18.306 NaN	0.0772* (0.0069)	0.0862* (0.0082)	0.0471* (0.0068)	0.043* (0.004)	0.0319* (0.0098)	0.0302* (0.0082)	0.0410* (0.0160)
<u>Exp_2</u>	-0.001* (0.000)	-0.0009* (0.0001)	-0.0011* (0.0002)	-0.0005* (0.0001)	-0.001* (0.000)	-0.0004* (0.0002)	-0.0004* (0.0002)	-0.0005* (0.0003)
<u>Sexo</u>	0.224* (0.007)	0.2020* (0.0208)	0.1915* (0.0197)	0.2452* (0.0279)	0.323* (0.021)	0.2325* (0.0237)	0.2459* (0.0197)	0.2935* (0.0436)
<u>Negro</u>	-0.122* (0.013)	-0.1141* (0.0643)	-0.1872* (0.0718)	-0.1026* (0.0804)	-0.043* (0.017)	0.0045 (0.0378)	-0.0303* (0.0718)	-0.0180* (0.0799)
<u>Const</u>	-80252 (1113385)	3.1700* (0.0978)	3.4489* (0.1150)	3.3089* (0.1221)	4.965* (0.111)	5.1259* (0.1168)	5.5538* (0.1150)	4.4009* (0.2179)
<u>Pardo</u>	-0.111* (0.013)	-0.0918* (0.0129)	-0.1188* (0.0139)	-0.0834* (0.0294)	-0.090* (0.010)	-0.0899* (0.0230)	-0.0788* (0.0139)	-0.0576* (0.0534)
<u>R2</u>	0.985	0.9846	0.9882	0.9792	0.991	0.9912	0.9948	0.9855

Fonte: Elaboração dos autores a partir dos dados da PNAD de 2002 e 2012.

Nota: Erros- Padrões entre parênteses, calculados de acordo com Belloni *et al* (2014). Coeficientes com asteriscos indicam que as estimativas foram significativas, considerando nível de significância de 5%. NaN é uma abreviação para *Not a Number*. A inferência no lasso requer técnicas especiais e em alguns casos, este resultado pode aparecer. Para maiores detalhes, ver Belloni *et al* (2014).

Tal resultado pode ser explicado pelo fato da incipiente oferta de trabalhadores escolarizados nos municípios pertencentes à região não metropolitana cearense. Além disso, vale destacar a expressiva redução do retorno à escolaridade na região metropolitana de Fortaleza, onde tal fato pode ser explicado pelo expressivo crescimento das instituições de ensino em tais municípios.

Ao se contrastar os resultados encontrados neste artigo com os de outros autores, verifica-se que a magnitude dos retornos à educação se mostraram bem inferiores quando comparados com os obtidos por Mariano e Arraes (2011) e Suliano e Siqueira (2012), os quais também realizaram estimações da equação minceriana de salários para medir o retorno à educação no Estado do Ceará.

Este resultado pode decorrer da melhor escolha de variáveis instrumentais realizadas pela técnica do lasso e do post-lasso. De fato, Teixeira e Menezes-Filho (2012), propuseram nova abordagem para a estimação de retorno da educação utilizando a legislação brasileira como instrumento. Os resultados encontrados por esses autores caminham no mesmo sentido dos evidenciados aqui, onde a escolha apropriada de instrumentos reduz significativamente a magnitude dos retornos da educação.

Diante do exposto, verifica-se que os resultados corroboram com a literatura no sentido de que se verifica uma redução no prêmio à escolaridade ao longo dos últimos anos. Além disso, este artigo apresenta um benefício considerável para a literatura que se dedica ao estudo dos retornos da educação, uma vez que apresenta uma metodologia inovadora que busca corrigir o problema da falta de instrumentos fortes para corrigir a endogeneidade entre educação e salários.

## 5 CONCLUSÕES

Este artigo buscou reexaminar os retornos da educação no Ceará utilizando técnicas robustas de correção da endogeneidade. Utilizou-se os métodos de seleção do Lasso e técnicas de estimação pós seleção (Post-Lasso) para lidar com tal problema.

Os métodos do Lasso e do Post-Lasso estabelecem critérios estatísticos para selecionar, dentre um conjunto de instrumentos fracos, aqueles que possuem o maior potencial preditivo da variável endógena. Dessa forma, evita-se, assim, a escolha de instrumentos ad-hoc, por meio, exclusivamente, da teoria econômica, a qual pode estar sujeita a enganos.

Os resultados foram estimados para dois anos diferentes, 2002 e 2012. Além disso, estratificou-se a amostra, em cada um dos anos, entre região metropolitana e não metropolitana. Esta separação deve-se somente a análise em ambientes diferenciados de mercado de trabalho e oferta educacional.

Além disso, o mesmo exercício foi realizado para a região Nordeste. Dessa forma, foi possível identificar se os retornos da educação no Ceará resultam de um contexto regional mais amplo.

Dos resultados encontrados, verificou-se o mesmo que já havia sido observado na literatura. Primeiro, que os retornos educacionais vêm se reduzindo em magnitude ao longo do tempo. Este resultado foi robusto a amostra completa, estratificada do Ceará e ao Nordeste.

Segundo, a região metropolitana apresenta menor retorno a educação do que a região não metropolitana. Como possível explicação para essa evidência destaca-se a carência, ou a falta de concorrência de pessoas qualificadas na região não metropolitana, como observado na análise descritiva dos dados.

Terceiro, em comparação com o Nordeste, embora em termos de magnitude os retornos da educação aparentemente estejam próximos, indicando a inserção no contexto regional, o Ceará reduziu muito mais fortemente os retornos da educação ao longo do tempo.

Por fim, e mais importante, a magnitude dos retornos da educação estimados neste trabalho foi significativamente menor do que os encontrados em outros trabalhos semelhantes, como, por exemplos, os de Mariano e Arraes (2011) e Suliano e Siqueira (2012). Este resultado pode indicar que a utilização de instrumentos utilizando apenas a teoria econômica não é recomendável sob a pena de selecionar instrumentos fracos que elevam o viés causado pela endogeneidade.

## REFERÊNCIAS BIBLIOGRÁFICAS

ANGRIST, Joshua. Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Record. *American Economic Review*, June 1990.

ANGRIST, J. and PISCHKE, Jorn-Steffen. **Mostly Harmless Econometrics**. An empiricist's companion. Princeton University Press, New Jersey, 2009.

ANGRIST, Joshua and KRUEGER, Alan. Does Compulsory School Attendan-

- ce Affect Schooling and Earnings? **Quarterly Journal of Economics**. November, 1991.
- ARIAS, O; MCMAHON, W. Dynamic Rates of Return to Education in the U.S. **Economics of Education Review**. v.20, n.1, p. 121-138, 2001.
- ASHENFELTER, O.; KRUEGER, A. Estimates of the Economic Return to Schooling from a New Sample of Twins. **American Economic Review**. v. 84, n. 5, p. 1157-1173, 1994.
- BAI, Jushuan and NG, Serena. Selecting instrument variables in a data rich environment. **Journal of Time Series Econometrics**, 1, Nº 4, 2009
- BARROS, R.P. (1997). Os Determinantes da Desigualdade no Brasil, Seminário 22/97, IPE-USP.
- BARROS, R.P., HENRIQUES, R. e MENDONÇA, R.(2000), Education and Equitable Economic Development, **Economia**, Volume 1, No1.
- BARROS, R. P.; CARVALHO, M.; FRANCO, S.; MEDONÇA, R. Uma Análise das Principais Causas da Queda Recente na Desigualdade de Renda Brasileira. **Econômica**, Rio de Janeiro, v.8, n.1, p.117-147, jun. 2006.
- BARROS, R. P.; CARVALHO, M.; FRANCO, S.; MEDONÇA, R. A Queda Recente da Desigualdade de Renda no Brasil. Rio de Janeiro: IPEA, jan. 2007. (Texto para Discussão, 1258).
- BELLONI, A; HANSEN, C. and CHERNOZHUKOV, V. Inference Methods for High Dimensional Sparse Econometric Models. **Advances in Economics and Econometrics**, 10th World Congress of the Econometric Society, 2013.
- \_\_\_\_\_. High Dimensional Methods and Inference on Structural and Treatment Effects. **Journal of Economic Perspectives**, 28(2), pg. 29-50, 2014.
- \_\_\_\_\_. LASSO Methods for Gaussian Instrumental Variable Models. Working paper, **Duke University**, 2010.
- BELLONI, A; HANSEN, C.; CHEN, D.; and CHERNOZHUKOV, V. Sparse Models and Methods for Optimal Instruments with Application to Eminent Domain. **Econometrica**, 80(6), pg. 2369-2429, 2012.
- BEN-PORATH, Y. The Production of Human Capital and the Life Cycle of Earnings. **Journal of Political Economy**, v. 75, n.4, p. 352-365, 1967.



BICKEL, P. J.; RITOV, Y. and TSYBAKOV, A. B. Simultaneous analysis of Lasso and Dantzig selector, **Annals of Statistics**, 37(4), pg. 1705-1732, 2009.

BOUDARBAT, B.; LEMIEUX, T.; RIDELL, W. C. **The Evolution of the returns to Human Capital in Canada**, 1980-2006. (Working Paper n.1) University of British Columbia, 2008.

CARD, David. Earnings, Schooling and Ability Revisited. In Solomon Polachek, editor, **Research in Labor Economics**, vol. 14, 1995.

CARNEIRO, P.; HECKMAN, J. and VYTLACIL, E. Estimating Marginal Returns to Education. **American Economic Review**, 101(6), 2011.

CARVALHO, A. P.; NERI, M. C. ; SILVA, D. B. do N. . Diferenciais de Salários por Raça e Gênero: Aplicação dos procedimentos de Oaxaca e Heckman em Pesquisas Amostrais

Complexas. In: XV Encontro Nacional de Estudos Populacionais, 2006, Caxambu - MG. Anais do **XV Encontro Nacional de Estudos Populacionais**, 2006.

FERREIRA, E. A.; VELOSO, F. A. Intergeneration Mobility of Wages in **Brazil**. **Brazilian Review of Econometrics**, v. 26, p. 181-211, 2006.

FERREIRA, F. H. G.; LEITE, P. G.; LITCHFIELD, J.; ULYSSEA, G. A. Ascensão e Queda da Desigualdade de Renda no Brasil. **Econômica**, Rio de Janeiro, v.8, n.1, p.147-169, jun., 2006.

HASTIE, Trevor, TIBSHIRANI, Robert and FRIEDMAN, Jerome. **The Elements of Statistical Learnings: Data Mining, Inference and Prediction**. Springer Series in Statistics, Second Edition, 2012.

HANSEN, C.; CHERNOZHUKOV, V. and SPLINDER, M. Post-Selection and Post-Regularization inference in linear models with very many controls and instruments. **American Economic Review: Papers and Proceedings**, 105(5), pg. 486-490, 2015. HECKMAN, J.J. Sample Selection Bias as a Specification Error. *Econometrica*, v. 47, n. 1, p. 153-161, Jan. 1979.

HECKMAN, James; URZUA, S., and VYTLACIL, E. Understanding Instrumental Variables in Models with Essential Heterogeneity. **Review of Economics and Statistics**, 88(3). Pg 389–432, 2006.

\_\_\_\_\_. Instrumental Variables in Models with Multiple Outcomes: The General Unordered Case. **Les Annales d’Economie et de Statistique**, N°

91, pg 151-174, 2008.

\_\_\_\_\_. **Comparing IV with Structural Models: What Simple IV Can and Cannot Identify.** NBER Working Paper, n. 14.706, 33 p., 2009.

HOFFMANN, R. Transferências de Renda e a Redução da Desigualdade no Brasil e cinco Regiões entre 1997 e 2004. **Econômica**, Rio de Janeiro, v.8, n.1, p.55-81, jun., 2006.

IOSCHPE, Gustavo. A ignorância custa um mundo: o valor da educação no desenvolvimento do Brasil. São Paulo: Francis, 2004. 324p.

KAPETANIOS, George and MARCELLINO, M. Factor-GMM Estimation with Large Sets of Possibly Weak Instruments. **Computational Statistics and Data Analysis**, 54(11), pg. 2655–2675, 2010.

KAPETANIOS, George and GROEN, J. J. Parsimonious Estimation with Many Instruments. **Federal Reserve Bank of New York**. Staff Report nº 386, 2009.

LANGONI, C. G. **Distribuição da renda e desenvolvimento econômico do Brasil.** Rio de Janeiro: Expressão e Cultura, 1973.

LAM, D.; SCHOENI, R. Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil. **Journal of Political Economy**, v.101, n.4, p.710-739, aug., 1993.

MARIANO, F. Z. ; ARRAES, R. A . Endogeneidade da Educação na Previsão da Taxa de Retorno: Avaliação Metodológica e Aplicação para Regiões Brasileiras e Estados Seleccionados. In: **VII Encontro da Economia do Ceará em Debate**, 2011, Fortaleza.

MENEZES-FILHO, N. A. A Evolução da Educação no Brasil e seu Impacto no Mercado de Trabalho. Instituto Futuro Brasil, 2001a.

MENEZES-FILHO, N. A. Equações de Rendimentos: Questões Metodológicas. In: CORSEUIL, C. H. (EDITOR). Estrutura Salarial: aspectos conceituais e novos resultados para o Brasil. Disponível em: [http://www.ipea.gov.br/agencia/images/stories/PDFs/livros/capitulo2\\_equacoes.pdf](http://www.ipea.gov.br/agencia/images/stories/PDFs/livros/capitulo2_equacoes.pdf). Acesso em: 5 jun. 2002

MINCER, J. B. Schooling, experience and earnings. New York: NBER, 1974. 152 p. PESSÔA, S. A.; BARBOSA FILHO, Fernando de Holanda. Retorno da Educação no Brasil. **Pesquisa e Planejamento Econômico (Rio de Janeiro)**, v.

38, p. 97-125, 2008. RESENDE, M. ; WYLLIE, R. . Retornos para Educação no Brasil: Evidências Empíricas Adicionais. *Economia Aplicada*, FIPE-USP, v. 10, p. 349-365, 2006.

ROUSE, C. Further Estimates of the Economic Return to Schooling from a New Sample of Twins. **Economics of Education Review**. v.18, n.2, p. 149-157, 1999.

SACHSIDA, A., LOUREIRO, P.R.A., MENDONÇA, M.J.C. (2004), Um Estudo sobre Retornos em Escolaridade no Brasil, **Revista Brasileira de Economia**, 58, 249-265

SAPELLI, C. **Los Retornos a la Educación en Chile**: Estimaciones por Corte Transversal y por Cohortes. Documento de Trabajo, n. 349. ISSN: 0717-7593. 2009

SCHULTZ, T. W. Investment in Human Capital. **The American Economic Review**, vol 51, nº 1, Mar 1961, p. 1-17.

SOARES, S. S. D. Análise do Bem-Estar e Decomposição por Fatores da Queda na Desigualdade entre 1995 e 2004. **Econômica**, Rio de Janeiro, v.8, n.1, p.83-115, jun., 2006b.

SULIANO, D. C., SIQUEIRA, M. L. Retornos da educação no Brasil em âmbito regional considerando um ambiente de menor desigualdade. **Economia Aplicada**, vol.16, nº 1, Ribeirão Preto Jan./Mar. 2012.

TEIXEIRA, W. M., MENEZES-FILHO, N. A. Estimando o Retorno à Educação do Brasil Considerando a Legislação Educacional Brasileira como um Instrumento. **Revista de Economia Política**, vol. 32, nº 3 (128), pp. 479-496, julho-setembro/2012.

TIBSHIRANI, R. Regression shrinkage and selection via lasso. **Journal of Royal Statistics Society B**, v. 58 (1), pg. 267-288, 1996.

ZEPEDA et al. **Changes in Earnings in Brazil, Chile And Mexico**: Disentangling The Forces Behind Pro-Poor Change In Labour Markets. (Working Paper, n.51). International Policy Centre for Inclusive Growth (IPC - IG). Mar. 2009

ZOU, H. and HASTIE, T. Regularization and variable selection via elastic net. **Journal of Royal Statistics Society B**, v. 67 (2), pg. 301-320, 2005.

# GERAÇÃO CANGURU? FATORES ASSOCIADOS À PERMANÊNCIA DOS JOVENS CEARENSES NO AMBIENTE FAMILIAR DE ORIGEM

Juliane da Silva Ciríaco\*

Otoniel Rodrigues dos Anjos Júnior\*\*

Priscila Silva Rodrigues\*\*\*

Nárdia Costa Alves\*\*\*\*

## Resumo

Este artigo tem como objetivo principal analisar de forma empírica os principais determinantes da geração canguru no Ceará com base nos dados oriundos da Pesquisa Nacional por Amostra de Domicílios (PNAD), realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) para o ano de 2012. Para tal fim, utiliza-se o modelo *logit* multinível para identificar os principais fatores atuantes na probabilidade do indivíduo de 25 a 34 anos de habitar no ambiente familiar de origem, considerando a influência da estrutura familiar para homens e mulheres. Os resultados apontaram que, independente do gênero, ter mãe viva e ser estudante tende em aumentar a probabilidade do jovem ser canguru.

**Palavras-Chave:** Demografia, Geração Canguru, Família.

## Abstract

This paper is meant to examine empirically the main determinants of kangaroo generation in Ceará based on data from the Pesquisa Nacional por Amostra de Domicílios (PNAD) conducted by the Brazilian Institute of Geography and Statistics (IBGE) for the year 2012. To this end, it uses the *logit*

---

\* Mestre em Economia pelo PPGE-UFPB. E-mail: julianeciriac@hotmai.com.

\*\* Mestre em Economia pelo PPGE-UFPB. E-mail: pbdosanjos@hotmai.com.

\*\*\* Mestranda em Economia pelo CAEN-UFC. E-mail: priscilarodrigues65@yahoo.com.br.

\*\*\*\* Graduada em Filosofia pela UECE. E-mail: nardiacosta@gmail.com.

multilevel model to identify the main factors acting on the probability of an individual of 25 to 34 years old to live in the family environment of origin, considering the influence of the family structure for men and women. The results show that, independent of gender, to have the mother alive and being a student tends to increase the likelihood of being young kangaroo.

**Key words:** Demography, Kangaroo Generation, Family.

**JEL:**J01, J11, J12.

## 1 INTRODUÇÃO

No apogeu dos anos 70 quanto mais jovem os indivíduos saíam de casa maior seria a certificação de liberdade e independência atribuído a estes pela sociedade da época. No entanto, sabe-se que de lá pra cá muita coisa mudou, sobretudo, porque o desenvolvimento da sociedade contemporânea atrelado ao expressivo processo de transição demográfica tem ocasionado relevantes modificações nos arranjos familiares e, conseqüentemente, mudado a concepção dos agentes. Logo, tal processo fez com que aqueles lares ditos “tradicionais” compostos por uma grande quantidade de indivíduos desse lugar a famílias menores e mais envelhecidas.

Nota-se, essencialmente, que as mudanças ocorridas na estrutura das famílias brasileiras, principalmente, nas últimas cinco décadas foram reflexos direto da mudança do papel da mulher no seio familiar. Dado, principalmente, pela necessidade da mulher adentrar ao mercado de trabalho, proporcionando recorrentes transformações na educação, tamanho, relação e, acima de tudo, nas regras até então vigente na sociedade. Nos dias atuais a mulher assume alguns papéis até então pertencentes efetivamente aos homens e isto tem alavancado o desenvolvimento da sociedade, modernizando as ideias e permitido novas interpretações das velhas regras sociais.

O jovem dos tempos modernos possui uma percepção do mundo bem diferente daquela encontrada na juventude de outrora. Tal fato pode ter sido impulsionado pelas mudanças comportamentais ocorridas dentro dos próprios lares brasileiros. Nota-se, por exemplo, que se tornou bastante comum jovens de idade adulta residirem com seus pais mesmo após ter alcançado relativa estabilidade financeira. Percebe-se que tal comportamento tem proporcionado novas mudanças na ordenação familiar e, sendo assim, bastante retratada e recorrentemente discutida na mídia, principalmente nas revistas, jornais, noticiários

e até filmes. Para Cobo e Saboia (2010) a decisão de morar no ambiente parental pode estar baseada em justificativas e explicações diversas. A lista é grande e envolve desde questões de ordem financeiras (custo habitacional), psicológicas (comodismo e Síndrome do Peter Pan) e até mesmo sociodemográficas (queda da taxa de fecundidade, aumento do número de divórcios).

Ressalta-se, no entanto, que o campo da justificativa é longo, abrangente e acaba incidindo em uma série bastante diversificada de explicações plausíveis. Fato consensual, porém, é que tal comportamento encontra, geralmente, apoio dos pais. Esta geração de jovens adultos que prolongam a convivência com seus respectivos genitores foi denominado na literatura por “cangurus”.

Segundo os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD, 2012) o Ceará detém aproximadamente 25,75% de indivíduos entre 25 a 34 anos na situação canguru (filhos que residem com pai ou/ e mãe), superando inclusive média nacional que é 24,3%. Diante do exposto, e sabendo que tal aspecto continua pouco explorado, pode-se dizer que o presente artigo contribui atualmente para a literatura ao investigar em âmbito estadual os principais determinantes da geração canguru entre os jovens de 25 a 34 anos no Ceará, levando em consideração a influência da conjuntura familiar. Para isto, utiliza-se o modelo econométrico de ordem qualitativa, o *logit* multinível, considerando o diferencial de gênero, utilizando como fonte de dados a PNAD (Pesquisa Nacional por Amostra de Domicílios) para o ano de 2012.

Para a consecução do presente estudo, optou-se por dividir a pesquisa em quatro partes, além desta introdução. Inicialmente, apresentam-se os estudos mais relevantes na literatura associada a temática. A próxima contempla a descrição e tratamento do banco de dados. A quarta seção reporta-se os principais resultados encontrados, ressaltando e discutindo os aspectos relevantes, que culminam nas considerações finais.

## 2 REFERENCIAL TEÓRICO

Desde o início do século XXI vêm sendo realizados estudos brasileiros em alguns ramos do conhecimento como economia, sociologia e psicologia, que permeiam sobre as interconexões de elementos que podem influenciar na condição do jovem na conjuntura familiar. A grande problemática ao se estudar a possível saída/ ou permanência dos jovens adultos no lar parental consiste principalmente na falta de disponibilidade de dados apropriados para análise,

principalmente empírica, sendo ainda poucas as contribuições técnicas, principalmente em países em desenvolvimento, em que boa parte dos estudos estão dispostos na América do Norte e continente Europeu [DE VOS, 1989; FILGUEIRA, AMOROSO, 1997; MORAIS, RÊGO, 2011].

Para Gallagher (2013) o fenômeno do prolongamento da convivência familiar possui uma característica multidimensional, abrangendo questões de ordem pessoal, familiar e social, podendo ser fruto da ausência de segurança no campo profissional e afetivo. O autor ainda ressalta que evolução da sociedade trouxe implicações sobre o comportamento dos arranjos familiares, isso por que a família outrora marcada pela hierarquia entre os membros, atualmente é regida pelos ideais igualitários, possibilitando maior espaço para negociação e diálogo entre as gerações. Ademais, compreende-se que um apego inseguro na infância pode causar alguns temores em um jovem adulto frente a uma nova vida sem a presença dos pais, fazendo com que o mesmo prolongue a permanência no lar parental adiando sua transição para um arranjo totalmente independente (BOWLBY, 2002; ABREU, 2005; VIEIRA, RAVA, 2010).

De acordo com Nico (2012) existe certa influência das transições escolares sobre o processo de saída da casa dos pais atuando de duas formas. A primeira, está relacionada ao próprio destino da saída, atuando de forma mais direta, baixando a idade na primeira saída da casa dos pais, mas aumentando a da “última”, ao mesmo tempo que aumenta a probabilidade de regresso. Enquanto a segunda tem a ver com o adiamento da entrada no mercado de trabalho, tendo efeitos sobre a autonomia habitacional completa. Além do mais agrega-se ao fato de que os jovens adultos estão de maneira particular mais vulneráveis aos efeitos da recessão e ao desemprego (OLIVEIRA, CARVALHO, 2010; NICO, 2012).

Por fim, ressalta-se que diversos fatores têm contribuído para a permanência dos filhos no lar de origem, tais como: maior liberdade de expressão; a diminuição/ou neutralização dos conflitos geracionais, ambivalência dos pais no que concerne à saída dos filhos de casa; a permissão para o sexo na casa dos pais; custo habitacional; falta de maturidade; comodismo; casamento tardio; as transformações dos laços afetivos; menos exigências e expectativas no relacionamentos entre os pares; instabilidade no mercado de trabalho; insegurança e violência urbana (CAMARANO *et al.*, 2003; HENRIQUES *et al.*, 2004; FERREIRA *et al.* 2008; CARVALHO, 2009; COBO, SABOIA, 2010).

### 3 BASE DE DADOS E ESTRATÉGIA METODOLÓGICA

#### 3.1 Abordagem Hierárquica

Os modelos multiníveis, em linhas gerais, são utilizados para estudar os dados hierarquicamente organizados, em que as unidades de observação em um nível estão aninhadas em unidades de observação em um nível superior (SILVA, *et al.*, 2011). Tendo em vista que ao se agregar ou desagregar a base de dados podem-se cometer erros vinculados à falácia ecológica ou atomística (HOX, 2002)<sup>1</sup>.

Para se obter um maior entendimento sobre os fatores que afetam as chances dos indivíduos de estarem na condição “canguru”, considera-se a influência simultânea das características de ordem individual (nível 1) e familiar (nível 2). Na qual, a variável independente é dicotômica, onde se atribui o valor de “um” para os indivíduos entre 25 a 34 anos que possuem status de “filho” no domicílio e “zero” caso o indivíduo nesta faixa de idade possua *status* de “chefe” ou “conjugue” do responsável pelo lar. Logo, a probabilidade de ocorrência do evento é dada por  $\eta_{ij} = \log\left(\frac{\Phi_{ij}}{1-\Phi_{ij}}\right)$  e  $Pr(\text{canguru} = 1/\alpha) = \Phi_{ij}$ , é modelado por uma função de ligação *logit*, descrita como:

$$\eta_{ij} = \log\left(\frac{\Phi_{ij}}{1-\Phi_{ij}}\right) = \alpha_{0j} + \sum_q \alpha_q X_{qij} + \xi_{ij} \quad (\text{nível 1}) \quad (01)$$

$$\alpha_{0j} = \phi_{00} + \sum_s \phi_{0s} W_{sj} + r_{0j} \quad (\text{nível 2}) \quad (02)$$

$$\alpha_{qj} = \phi_{q0} \quad (\text{nível 2}) \quad (03)$$

O subscrito  $i$  refere-se ao indivíduo e  $j$  a família ao qual este faz parte. Logo,  $\alpha_{0j}$  representa o intercepto;  $X_{qij}$  é o valor associado as  $q$  variáveis incluídas no modelo;  $\alpha_q$  é o efeito parcial das variáveis nas chances de ocorrência da condição “canguru” e  $\xi_{ij}$  corresponde o termo de erro (do nível 1) com distribuição logística padronizada com média zero e variância igual a  $\sigma_\xi^2$ .

No segundo nível, expresso pela equação 2, assume-se que o intercepto do nível 1,  $\alpha_{0j}$ , varia aleatoriamente em todas as famílias, enquanto admite-se inclinação igual para todas as unidades de nível 2 (ver equação 3), mudando somente o intercepto de cada unidade  $j$ , que é representada por  $(\phi_{00} + r_{0j})$ . Logo, o intercepto está decomposto no valor médio global para todos os indivíduos  $(\phi_{00})$  e pelo componente aleatório associado ao segundo nível  $(r_{0j})$ .

1. A primeira consiste em interpretar dados agregados á nível individual, enquanto a segunda refere-se em fazer inferências em um nível mais alto baseado na análise de dados individuais.



Substituindo a Equação (2) e (3), na (1), tem-se a equação (4) exposta a seguir:

$$\log\left(\frac{\phi_{ij}}{1-\phi_{ij}}\right) = \phi_{00} + \sum_s \phi_{0s} W_{sj} + \sum_q \phi_{q} X_{qij} + \xi_{ij} + r_{0j} \quad (04)$$

$W_{sj}$  = valor associado as  $s$  variáveis associadas ao nível 2.

Com,  $r_{0j} \sim N(0, \sigma_{00}^2)$ .

Na abordagem hierárquica é conveniente, a estimação do modelo de baixo para cima, tendo como ponto de partida o Modelo Nulo. Usa-se o coeficiente de correlação intraclasse com o intuito de verificar se é justificável a inclusão do segundo nível. A métrica é representada por:

$$\rho_{logit} = \frac{\sigma_{00}^2}{\sigma_{00}^2 + \sigma_{\xi}^2}$$

$$0 \leq \rho_{logit} \leq 1$$

No qual,  $\sigma_{00}^2$  representa a variância residual do nível 2, que por suposição é normalmente distribuída, enquanto  $\sigma_{\xi}^2 = (\pi^2/3) = 3,29$  representa a variância residual do nível 1. O  $\rho_{logit}$  mensura a proporção da variância entre o grupo frente à variância total, ou seja, indica o quanto da variação observada na variável explicada pode ser atribuído às características no nível dos grupos, variando entre 0 e 1. Quanto mais próximo de zero, significa que os grupos entre si considerados são mais homogêneos, logo a comportamento do indivíduo independe do grupo que este frequenta, enquanto, quanto mais o  $\rho_{logit}$  se aproxima de 1 indica que variabilidade no padrão comportamental dos indivíduos nos grupos se deve basicamente às diferenças existentes entre esses grupos.

## Construção do Modelo Multinível

O passo inicial para análise da trajetória de construção de um modelo Hierárquico consiste na especificação mais simples, chamado de Modelo Nulo ou Vazio, em que observa-se apenas as variações no comportamento dos indivíduos (nível 1), dadas pelas diferenças atribuídas a meio familiar (nível 2), expresso por:

$$\log[\phi_{ij}/(1 - \phi_{ij})] = \alpha_{0j} + \xi_{ij} \quad (\text{Indivíduo}) \quad (05)$$

$$\alpha_{0j} = \phi_{00} + r_{0j} \quad (\text{Família}) \quad (06)$$

Introduzindo as equações 6 na 5, obtém-se o 1º modelo expresso por:

✓ **Modelo Vazio ou Nulo**

$$\log[\phi_{ij}/(1 - \phi_{ij})] = \phi_{00} + r_{0j} + \xi_{ij} \quad (07)$$

Em seguida, procura-se estimar a existência de variações no comportamento dos indivíduos causados pelas diferenças existentes entre as famílias, ao incluir as variáveis relacionadas às características individuais. Dessa forma, o modelo é especificado como:

$$\log[\Phi_{ij}/(1 - \Phi_{ij})] = \alpha_{0j} + \alpha_{1j}Indivíduo_{ij} + \xi_{ij} \quad (\text{Indivíduo}) \quad (08)$$

$$\alpha_{0j} = \phi_{00} + r_{0j} \quad (\text{Família}) \quad (09)$$

$$\alpha_{kj} = \phi_{q0} \quad \text{tal que: } q = 1 \quad (\text{Família}) \quad (10)$$

Substituindo as equações (9) e (10) em (8), obtêm-se o 2º modelo:

✓ **Modelo 1**

$$\log[\Phi_{ij}/(1 - \Phi_{ij})] = \phi_{00} + \phi_{10}Indivíduo_{ij} + r_{0j} + \xi_{ij} \quad (11)$$

Finalizando, na última especificação, além das variáveis expostas anteriormente acrescenta-se a variável que representam a *Influência Familiar* ilustradas abaixo:

$$\log[\Phi_{ij}/(1 - \Phi_{ij})] = \alpha_{0j} + \alpha_{1j}Indivíduo_{ij} + \xi_{ij} \quad (\text{Indivíduo}) \quad (12)$$

$$\alpha_{0j} = \phi_{00} + \phi_{01}Influência\_Familiar_{ij} + r_{0j} \quad (\text{Família}) \quad (13)$$

$$\alpha_{kj} = \phi_{q0} \quad \text{tal que: } q = 1 \quad (\text{Família}) \quad (14)$$

Inserindo as equações 13 e 14 na equação 12, tem-se o modelo 3 expresso pela equação abaixo:

✓ **Modelo 2**

$$\log[\Phi_{ij}/(1 - \Phi_{ij})] = \phi_{00} + \phi_{01}Influência\_Familiar_{ij} + \phi_{10}Indivíduo_{ij} + r_{0j} + \xi_{ij} \quad (15)$$

As variáveis do modelo final representam o conjuntura familiar do “jovem adulto”, as quais buscam captar a influência de determinados fatores não observados que podem repercutir sobre o comportamento do mesmo. Adicionalmente, com o objetivo de entender a importância direta das características relacionadas a conjuntura familiar, usa-se a métrica proposta por Raudenbush e Bryk (2002). Este método visa verificar o quanto a introdução das variáveis do segundo nível ajudam a explicar a variabilidade relacionado ao intercepto do modelo não condicional. Ressalta-se que, quanto menor o componente de variância, mais elevado é o poder explicativo das variáveis atribuíveis ao segundo nível, obtidas pela seguinte expressão:

$$\text{Variância Explicada} = \left( \frac{\sigma_{00}^2(\text{não condicional}) - \sigma_{00}^2(\text{condicional})}{\sigma_{00}^2(\text{não condicional})} \right) * 100$$

Onde

$\sigma_{\alpha_0}^2$  = Estimativa do componente da variância do intercepto do modelo não condicional<sup>2</sup> e modelo condicional (o qual inclui-se as variáveis de nível 2).

### 3.3 Descrição Do Banco De Dados

Para a realização deste trabalho, utilizou-se os microdados da PNAD de 2012, abrangendo todo o território cearense, disponíveis de forma anual no portal do Instituto Brasileiro de Geografia e Estatística (IBGE). A escolha desse banco de dados consiste na gama de informações disponíveis sobre a população residente no país.

O estudo é compreendido por dois níveis, onde o primeiro refere-se a característica individual do jovem, ao passo que o segundo representa a estrutura de agrupamento relacionado a Família do mesmo. Com o intuito de atender os objetivos dessa pesquisa, foram aplicados alguns filtros como, por exemplo, excluir aqueles jovens que simultaneamente não se enquadravam em nenhuma das categorias analisadas; ou seja, aqueles indivíduos que não eram “cangurus” e não eram “independentes”. Além disso, foram considerados somente os jovens de 25 a 34 anos de idade, excluindo-se aqueles indivíduos com rendimento familiar *per capita* igual ou inferior a R\$ 96, ou seja, os lares 10% mais pobres do estado<sup>3</sup>. Destaca-se ainda que foram considerados como canguru, somente aqueles indivíduos que tinham posição familiar de dependente na família, o qual nesta pesquisa é designado pelos indivíduos que detinham *status* de filho na residência; e no caso do jovem independente dos pais, foram considerados aqueles indivíduos que tinham posição domiciliar de chefe da família e cônjuge do indivíduo considerado como responsável pelo lar, os demais foram excluídos da amostra.

Após todos esses filtros e exclusão das observações faltantes (*missings values*) restaram aproximadamente um total de 2.458 observações para o primeiro nível, enquanto o segundo é constituído de 1.927 observações, conjuntamente para os homens e mulheres.

Por sua vez, as variáveis explicativas são exemplificadas no Tabela 1:

2. Esse modelo é denominado de não condicional, pois permite mensurar a variabilidade não condicional do segundo nível.

3. Tal filtro foi realizado porque acredita-se que a condição canguru é menos evidente em famílias de baixa renda. Ademais, ressalta-se que alguns fatores referentes ao lar de origem daqueles jovens que haviam saído de casa para constituir seu próprio domicílio não puderam ser incorporados junto a estratégia empírica devido a sua não disponibilidade (ex: escolaridade do pai e mãe, nível de riqueza dos pais).

**TABELA 1- Ceará: Descrição das variáveis utilizadas no modelo econométrico**

Variáveis	Banco	Descrição
Idade	Nível 1 (Características Individuais)	Aferida em anos de vida, adiciona-se também o termo quadrático
Branca		1 para cor Branca e 0 caso contrário
Estuda		1 se estuda e 0 caso contrário
Anos de Estudo		Correspondente à série, ao nível ou ao grau do jovem. A equivalência é feita de maneira que cada série concluída com aprovação é considerada como 1 ano de estudo
Metro		1 para o jovem que mora na região metropolitana e 0 caso contrário
Trabalha		1 se trabalha e 0 caso contrário
Mãe viva	.Nível 2 (Características do Lar dos Jovens Independentes e Cangurus)	1 para mãe viva e 0 caso contrário
Ter Crianças		Número de crianças com idade inferior a 5 anos
Tamanho		Membros do Lar

Fonte: Elaboração dos autores.

Antes da exposição dos resultados econométricos, contempla-se na Tabela 2 as estatísticas descritivas das variáveis utilizadas neste artigo.

Observa-se, inicialmente para as mulheres, partindo-se de uma amostra de 1.240 indivíduos e 1.221 famílias, que a idade média dos jovens equivale a 29 anos de idade, sendo as pessoas por volta de 33% consideradas de raça branca, com aproximadamente 9,5 anos de estudo, em média, os lares são compostos por 3 membros, possuem um menor número de crianças (0,37), com mais de 92% das famílias com mãe viva. No referente ao status ocupacional e localização, 10% estudam e pouco mais de 62% trabalham, localizados em sua maioria na região metropolitana.

Para os Homens, a amostra foi composta por 1.218 indivíduos e 1.161 famílias, com características similares ao público feminino anteriormente analisada: média de 29 anos de idade, 8,6 anos de estudo, com 31,3% de cor branca, compostos por lares com 3 integrantes, menor média de crianças em

casa (0,33), mãe presente em aproximadamente 94 % dos domicílios, sendo que pouco mais 63% residentes da região metropolitana. Ademais, chama-se a atenção que 87% dos Homens trabalham, ao passo que 7% estudam.

Antes da exposição dos resultados econométricos, contempla-se na Tabela 2 as estatísticas descritivas das variáveis utilizadas neste artigo.

Observa-se, inicialmente para as mulheres, partindo-se de uma amostra de 1.240 indivíduos e 1.221 famílias, que a idade média dos jovens equivale a 29 anos de idade, sendo as pessoas por volta de 33% consideradas de raça branca, com aproximadamente 9,5 anos de estudo, em média, os lares são compostos por 3 membros, possuem um menor número de crianças (0,37), com mais de 92% das famílias com mãe viva. No referente ao status ocupacional e localização, 10% estudam e pouco mais de 62% trabalham, localizados em sua maioria na região metropolitana.

Para os Homens, a amostra foi composta por 1.218 indivíduos e 1.161 famílias, com características similares ao público feminino anteriormente analisada: média de 29 anos de idade, 8,6 anos de estudo, com 31,3% de cor branca, compostos por lares com 3 integrantes, menor média de crianças em casa (0,33), mãe presente em aproximadamente 94 % dos domicílios, sendo que pouco mais 63% residentes da região metropolitana. Ademais, chama-se a atenção que 87% dos Homens trabalham, ao passo que 7% estudam.

**TABELA 2- Ceará: Característica da Amostra, 2012**

Variáveis/ Sexo	Mulher				Homem			
	Média	Desvio	Mínimo	Máximo	Média	Desvio	Mínimo	Máximo
Mãe viva	0,927	0,260	0	1	0,936	0,245	0	1
Tamanho	3,504	1,334	1	12	3,534	1,477	1	12
Metro	0,661	0,473	0	1	0,637	0,481	0	1
Ter Crianças	0,377	0,568	0	3	0,332	0,551	0	3
Branca	0,333	0,471	0	1	0,313	0,464	0	1
Anos de Estudo	9,557	3,780	0	15	8,601	4,109	0	15
Idade	29,39	2,805	25	34	29,26	2,864	25	34
Idade <sup>2</sup>	871,8	165,1	625	1.156	864,5	168,4	625	1.156
Estuda	0,101	0,301	0	1	0,0704	0,256	0	1
Trabalha	0,626	0,484	0	1	0,876	0,329	0	1

Fonte: Elaboração dos autores a partir de dados da PNAD.

## 4 RESULTADOS ECONOMÉTRICOS

Dado à hierarquia dos dados, a utilização da modelagem Multinível se faz necessária, uma vez que possibilita a determinação dos efeitos diretos (nível 1) e dos efeitos indiretos (nível 2). A primeira etapa da abordagem consiste na análise do modelo Vazio, em que nenhuma variável é inserida, com intuito de constatar aleatoriedade dos coeficientes, exposto na Tabela 3. Adicionalmente para testar a significância estatística da variância foi realizado o teste de razão verossimilhança<sup>4</sup>, que possui como hipótese nula a variância igual a zero. Constatou-se que, no modelo nulo, independente do sexo, a variância contextual é significativamente diferente de zero, por sua vez, é aceitável que as Famílias apresentem valores diferenciados para a probabilidade dos “jovens adultos” serem canguru<sup>5</sup>. De forma adicional, o coeficiente de correlação intraclasse evidenciou que 70% e 80% da variação na probabilidade, respectivamente, do indivíduo do sexo feminino e masculino morar com os pais são atribuídas às características a nível familiar. O teste de razão verossimilhança bem como o valor do  $P_{logit}$  torna justificável o emprego da abordagem hierárquica<sup>6</sup>, ao invés de um modelo de regressão clássico.

Como para os seis modelos analisados (na totalidade considerando homens e mulheres) as variâncias contextuais são estatisticamente diferentes de zero, rejeita-se a hipótese nula de intercepto com efeito aleatório, logo as probabilidades de ser canguru diferem de acordo com a família no qual o mesmo está inserido para todas as especificações. Dessa forma, o próximo passo consiste em verificar a importância direta das variáveis medidas no nível 2, usa-se o percentual da variância explicada, comparando-se a *Modelo 2* (condicional) com o *Modelo 1* (não-condicional<sup>7</sup>). Chega-se à conclusão de que a inclusão das variáveis que caracterizam a “*Influência Familiar*” explicam conjuntamente para as mulheres 10,7% da variância do intercepto, enquanto para os homens a mesma representa 6,8%. Ou seja, na devida ordem, 10,7% e 6,8% da variação na probabilidade da mulher e do homem estar na condição canguru são explicados pelos diferenças atribuídas a presença materna, existência de criança

4 Optou-se por escolher esse teste devido ao procedimento de estimação adotado, máxima verossimilhança.

5 O teste da razão da máxima verossimilhança com um grau de liberdade para as Mulheres foi 28,42, com um significativo valor de probabilidade (<0,0001), enquanto para os Homens 73,98, com um significativo valor de probabilidade (<0,0001), indicando a possível existência de estrutura de agrupamento.

6 Recomenda-se o uso da modelagem Multinível sempre que o  $P_{logit}$  for maior que 0,01 (HOPE; SHANNON, 2005).

7 Utiliza-se esta nomenclatura devido à inclusão de apenas as variáveis explicativas associadas às características individuais da amostra.

e tamanho da família observados entre os lares cearenses.

Em linhas gerais, boa parte das variáveis mostram-se significativas, e com sinal esperado na maioria dos casos observados. Os resultados do último modelo evidenciam que jovens mais escolarizadas e que estão estudando possuem maior probabilidade de continuar morando na casa parental, o que já era esperado, dado que um dos principais suportes oferecidos pelos pais para a permanência juvenil no domicílio é possibilitar a obtenção de maior nível de escolaridade e qualificação, corroborando os argumentos de autores como Moraes e Rêgo (2011) e Cobo e Saboia (2010).

A existência materna mostrou significância estatística e sinal positivo na probabilidade do jovem adulto residir com a família, sendo este efeito superior para as mulheres. As evidências apontam que, para as jovens, possuir mãe viva aumenta em 26 vezes as chances de permanecer na casa dos pais em relação aquelas que não tinham a figura presente, destarte que para os homens estas chances representam pouco mais de 5 vezes. O que já era esperado, pois indivíduos que vivem em lares com ausência da figura maternal, tendem de certa forma acelerar a transição para um arranjo independente. Sugere-se que isso ocorre por que a não existência materna representa uma perda fundamental de recursos disponíveis, uma vez que na maioria dos casos é a mãe que tende a alocar seu tempo para a realização de afazeres domésticos (*ex*: comida e roupa lavada), trabalho e educação dos filhos. Carvalho (2009) corrobora com tais argumentos, ressaltando ainda que ambientes onde não existe a convivência maternal presente está devidamente susceptível a possíveis conflitos familiares, principalmente quando a ausência da mesma é substituída por uma madrasta.

Estar trabalhando aumenta as chances de morar com os genitores para a mulher, enquanto para o homem aumenta a probabilidade de sair da casa e formar um novo domicílio. Ressalva-se, que para certificar, se tal comportamento ocorre ou não, é essencial uma análise mais aprofundada do banco de dados, uma vez que que tal diferenças entre os sexos pode estar influenciado pelo nível de renda pessoal e tipo de inserção no mercado de trabalho.

Sugere-se que, independente do gênero, um dos principais fatores para a saída da casa dos pais está associado a maternidade ou casamento, tendo em vista que a maior existência de crianças reduz as chances de residir com os genitores. Ao passo, no referente ao tamanho da família, aqueles que residem em famílias numerosas, de forma geral, possuem chances maiores de residir com pai ou/e mãe.

TABELA 3 - Ceará: Determinantes da Geração Canguru, 2012

Componente Fixo	Mulher				Homem			
	Modelo Nulo	Modelo 1	Modelo 2	Razão das chances	Modelo Nulo	Modelo 1	Modelo 2	Razão das chances
Branca	0,336 (0,296)	0,289 (0,298)	1,335 (0,398)	0,229 (0,315)	0,202 (0,340)	1,223 (0,415)		
Trabalha	0,724** (0,311)	0,581* (0,322)	1,788* (0,576)	-4,553*** (0,628)	-4,051*** (0,745)	0,017*** (0,013)		
Estuda	1,408*** (0,464)	0,882* (0,464)	2,415* (1,119)	1,934*** (0,603)	1,474*** (0,677)	4,367*** (2,958)		
Anos de Estudo	0,145*** (0,0431)	0,223*** (0,0545)	1,249*** (0,068)	0,0727* (0,040)	0,150*** (0,048)	1,161*** (0,055)		
Idade	-1,869* (1,133)	-1,445 (1,167)	0,235 (0,275)	-0,724 (1,190)	0,164 (1,297)	1,178 (1,528)		
Idade <sup>2</sup>	0,0267 (0,0192)	0,0189 (0,0197)	1,019 (0,020)	0,00821 (0,022)	-0,00751 (0,022)	0,992 (0,021)		
Metro	-0,205 (0,293)	-0,244 (0,298)	0,734 (0,233)	-0,323 (0,318)	-0,355 (0,344)	0,700 (0,241)		
Ter Crianças			-2,628*** (0,518)	0,072*** (0,037)	-4,838*** (0,749)	0,007*** (0,006)		
Tamanho			0,994*** (0,195)	2,702*** (0,526)	1,674*** (0,260)	5,333*** (1,387)		
Mãe Viva			3,271*** (0,977)	26,324*** (25,727)	1,751** (0,749)	5,759*** (4,315)		
Constante	-2,040*** (0,259)	27,41* (16,46)	15,57 (16,86)	-	-1,432*** (0,203)	16,10 (17,33)	-3,478 (18,90)	-
Componente Aleatório								
$\sigma^2_{\epsilon}$	7,94***	8,70***	7,77***	13,20***	10,69***	9,96***	6,83	
Variação Explicada			10,7					
Indivíduo	1.240	1.240	1.240	1.240	1.218	1.218	1.218	1.218
Família	1.221	1.221	1.221	1.221	1.161	1.161	1.161	1.161

Fonte: Elaboração dos autores a partir dos dados da PNAD. Nota: Desvio Padrão em Parênteses. \*\*\* p < 0,01, \*\* p < 0,05, \* p < 0,1.



No referente as características relacionadas a raça, idade e de localização geográfica do jovem, estas não apresentaram significância estatisticamente, sugerindo que as mesmas não explicam este efeito.

É importante perceber que, apesar das características individuais e familiares explicarem em parte as chances de estar na condição canguru, ainda existe uma parcela relevante da proporção da variância ( $\sigma_{\text{oo}}^2$ ) que permanece ainda não explicada, dada pelo intercepto. Tal resultado, pode estar condicionado a fatores não observados nas famílias, como aspectos culturais, religiosos, e de insegurança tanto no lado profissional como afetivo.

Por fim, ressalta-se que os resultados encontrados sobre às características individuais, familiares ratificam a análise descritiva preliminar elaborada anteriormente e também corroboram em boa parte com aspectos expostos na literatura nacional e internacional [Morais e Rêgo (2011), Cobo e Saboia (2010), Carvalho (2009), Ferreira *et al.* (2008), Henriques *et al.* (2004), De Vos (1989)].

## 5 CONSIDERAÇÕES FINAIS

A proposta inicial deste estudo foi investigar o efeito da conjuntura familiar na determinação da condição canguru, considerando a estrutura hierárquica dos dados, afim de não recorrer a falácia ecológica ou atomística evidenciada por Hox (2002).

Este trabalho buscou complementar as pesquisas de Cobo e Saboia (2010), Morais e Rêgo (2011) e Carvalho (2009) ao buscar analisar o perfil dos adultos que permaneceram ou deixaram o lar de origem, considerando determinadas características dos lares. Em relação aos expostos dos demais trabalhos, este avança ao visar aprofundar, de forma empírica, o entendimento sobre a temática, tentando captar a influência simultânea das características de ordem individual (nível 1) e familiar (nível 2), que podem repercutir sobre o comportamento do indivíduo.

Os achados indicam uma forte influência da estrutura familiar, ou seja, acredita-se que existem fatores não observados nos lares que vão afetar o comportamento do jovem adulto. Os resultados mostram que 80% e 70 % da variância nas chances de estar na condição canguru, para homens e mulheres, respectivamente, são atribuídas as características a nível familiar, mostrando a importância do uso de Modelos de Regressão Logística Multinível ou Hierár-

quico, afim de evitar estimativas ineficientes e/ou enviesadas. Comprovou-se, a partir dos resultados, que os fatores que figuram como os mais impactantes na decisão de permanecer ou não no ninho estão relacionadas ao fato do jovem estar estudando, possuir mãe viva e tamanho da família.

A dificuldade ao se estudar os determinantes da geração canguru, consiste na falta de dados apropriados para as análises, visto que o banco de dados em questão não a permite acompanhar o indivíduo ao longo do tempo, bem como conhecer a fundo os principais motivos que levaram o indivíduo a sair ou não do lar de origem. Desta forma, sugere-se para futuras pesquisas verificar o tipo de arranjo domiciliar constituído, fazendo-se uma distinção entre tipos de famílias, ou seja, constituídas por: laços matrimoniais, da coabitação, unipessoais, com ou sem parentes. Tal subdivisão irá permitir um maior detalhamento sobre o destino e possíveis repercussões sobre a vida do mesmo, permitindo o maior entendimento sobre a conjuntura familiar em diversas óticas.

Por fim, esta pesquisa não visa encerrar as discussões do problema e nem apresentar a solução final acerca da questão. Sendo assim, busca-se encontrar não a fórmula que soluciona os seculares problemas advindos da inter-relação humana no seio familiar, e sim, enumerar a menor quantidade de respostas plausíveis e significantes para, com isso, auxiliar a elaboração de políticas públicas mais ajustadas. Neste contexto, tais iniciativas seriam elaboradas por meio da utilização de informações menos obscuras acerca do comportamento da dinâmica da população jovem dos tempos modernos.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABREU, C. N. de. *Teoria do Apego: Fundamentos, Pesquisas e Implicações Clínicas*. São Paulo: Casa do Psicólogo, 2005.
- BOWLBY, J. *Apego e perda: separação – angústia e raiva*. 3. ed. São Paulo: Martins Fontes, 2002.
- CAMARANO, A. A.; PAZINATO, M. T.; KANSO, S.; VIANA, C. A transição para a vida adulta: novos ou velhos desafios? *Boletim de Mercado - conjuntura e análise*, Rio de Janeiro, v. 21, p. 53-66, 2003.
- CARVALHO, R. L. *Casa, comida e roupa lavada: fatores associados à saída do jovem brasileiro do domicílio de origem*. Dissertação (Mestrado em Demografia). Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.
- COBO, B.; SABOIA, A. A geração canguru no Brasil. In: *XVII Encontro da Associação Brasileira de Estudos Populacionais, Caxambu*. Anais do XVII Encontro Nacional de Estudos Populacionais, ABEP, 2010.
- DE VOS, S. Leaving the parental home: patterns in six Latin American countries. *Journal of Marriage and the Family*, n. 51, p. 615-626, Aug. 1989.
- FERREIRA, P.A, REZENDE, D.C., LOURENÇO, C.D.S. Geração Canguru: Algumas Tendências que Orientam o Consumo Jovem e Modificam o Ciclo de Vida Familiar. In: *III Encontro de Marketing da ANPAD*, Curitiba, 2008.
- FILGUEIRA, C.; AMOROSO, G. *Condiciones habitacionales de la juventud: elementos para el diseño de una política de vivienda*. CEPAL, Oficina de Montevideo, 1997.
- GALLAGHER, I. M. Geração canguru entre o conforto e o desamparo. Dissertação (Mestrado em Psicologia) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.
- HENRIQUES, C. R.; JABLONSKI, B.; FERES-CARNEIRO, T. A “geração canguru”: algumas questões sobre o prolongamento da convivência familiar. *Revista Psico*, v. 35, 2, p. 195-205, 2004.
- HOPE, A.D.; SHANNON, E.D. *A comparison of two procedures to fit multi-level data: PROC*  
*GLM versus PROC MIXED*. Pennsylvania, 2005.
- HOX, J. J. *Multilevel Analysis: Techniques and Applications*. 4. ed. Routledge Academic, 2002.

MORAIS, M. da P.; RÊGO, P. A. Coabitação familiar e formação de novos domicílios nas áreas urbanas brasileiras. *Boletim regional, urbano e ambiental*, IPEA, 2011.

NICO, [M. L. Género e saída de casa dos pais. Os percursos de autonomia habitacional por diferentes camadas analíticas](#), *Sociologias e-Working Papers*, ISCS, 2012.

OLIVEIRA, L.; CARVALHO, H. *Regulação e Mercado de Trabalho*. Lisboa, 2010.

RAUDENBUSH, S. W.; BRYK, A. S. *Hierarchical linear models: applications and data analysis methods*. 2.ed. Londres, Nova Deli: Sage, 2002.

SILVA, A. R. A.; GONÇALVES, E.; FREGUGLIA, R. S. Condicionantes regionais e individuais da mobilidade de trabalhadores da indústria de transformação do Estado de São Paulo. *In: ENCONTRO NACIONAL DE ECONOMIA*, 2011, Foz do Iguaçu. XXXIX ENCONTRO NACIONAL DE ECONOMIA, 2011.

VIEIRA, A. C. S.; RAVA, P. G. S. Ninho cheio: uma nova etapa do ciclo vital familiar? *Barbarói*, 33,118-134, 2010.

# IMPACTOS DAS RENDAS NÃO-AGRÍCOLAS SOBRE OS INDICADORES DE POBREZA FOSTER-GREER-THORBECKE (FGT) PARA AS FAMÍLIAS RURAIS DO ESTADO DE CEARÁ

Alan Francisco Carvalho Pereira\*

Wellington Ribeiro Justo\*\*

João Ricardo Ferreira de Lima\*\*\*

## Resumo

A elevada incidência da pobreza e os grandes desequilíbrios nas condições socioeconômicas dos indivíduos residentes no meio rural dos estados da região Nordeste, são objeto de estudo quando se busca entender as diferentes dinâmicas geográficas do desenvolvimento econômico brasileiro. A partir dos anos 90, mesmo com a evolução positiva no desempenho econômico do país e com as grandes transformações econômicas ocorridas nas duas últimas décadas, observou-se que o progresso na redução dos problemas sociais no Nordeste rural foi bastante inferior aos das demais regiões. O objetivo do trabalho é mostrar os efeitos das rendas Não-Agrícolas sobre os indicadores de pobreza *Foster-Greer-Thorbecke (FGT)* para as famílias rurais do estado de Ceará utilizando a metodologia *Propensity Score Matching (PSM)*. De acordo com os resultados observados, o impacto médio das rendas não agrícolas sobre as exclusivamente agrícolas é de R\$ 190,28 e R\$ 195,41 pelos dois métodos de pareamento utilizados; a proporção de pobres cai mais de 25 pontos percentuais; o hiato da pobreza é reduzido em média 33 pontos percentuais e a severidade da pobreza diminui mais de 28 pontos percentuais.

**Palavras-Chave:** Setor agropecuário; ICN; Nordeste; Ceará.

---

\* Mestrando em economia PPGECON/UFPE. Endereço: Rua Iracema Araújo Freire, 170 – Maurício de Nassau Caruaru – PE. CEP: 55.014-450. e-mail: alanpereira1993@hotmail.com.

\*\* Professor associado URCA. Professor do PPGECON/UFPE. Doutor em Economia pelo PIMES-UFPE. e-mail: justowr@yahoo.com.br.

\*\*\* Pesquisador A da Embrapa Semiárido. Doutor em Economia Aplicada pela UFV. e-mail: joao.ricardo@embrapa.br.

## Abstract

The high incidence of poverty and the great imbalances in socioeconomic conditions of individuals residing in rural areas of the states of the Northeast, are the subject of study when it seeks to understand the different geographical dynamics of Brazilian economic development. From the 90s, even the positive developments in the economic performance of the country and the great economic transformations that have occurred in the last two decades, it was observed that the progress in reducing social problems in the rural Northeast were much lower than the other regions. The objective is to show the effects of non-agricultural incomes of the *Foster-Greer-Thorbecke poverty indices (FGT)* for rural households in the state of Ceará using the propensity score matching methodology (PSM). According to the observed results, the average impact of non-farm incomes on agricultural exclusively to R\$ 190.28 and R\$ 195.41 for the two pairing methods used; the proportion of poor fell more than 25 percentage points; the poverty gap is reduced on average by 33 percentage points and the severity of poverty decreases more than 28 percentage points.

**Key words:** Poverty. Non-agricultural incomes. Ceará. Propensity Score.

**JEL:** C8, C10, D63, I30, R58.

## 1 INTRODUÇÃO

A elevada incidência da pobreza e os grandes desequilíbrios nas condições socioeconômicas dos indivíduos residentes no meio rural da região Nordeste, são objeto de estudo quando se busca entender as diferentes dinâmicas geográficas do desenvolvimento econômico brasileiro (NASCIMENTO, 2008). O reconhecimento da caracterização de uma diferenciação regional no crescimento econômico se tornou tema principal nas discussões e estudos na década de 70, onde se observou que as altas taxas de expansão da capacidade produtiva não eram suficientes para resolver o aumento das deficiências na estrutura distributiva de renda (ROCHA, 2003).

Apesar do alto nível de crescimento do produto nacional, ocorreu no período de 1967 a 1979 um agravamento nas condições de vida de parte da população, principalmente no meio rural, e piora em muitos indicadores sociais como os de incidência de pobreza, analfabetismo e desemprego (MARIANO; NEDER, 2006).

A partir dos anos 90, mesmo com a evolução positiva no desempenho econômico do país e com as grandes transformações econômicas ocorridas nas duas últimas décadas, observou-se que os progressos na redução dos problemas sociais no Nordeste foram bastante inferiores aos das demais regiões, pois essas desigualdades continuaram em evidência na região devido à permanência de uma elevada população em situação de pobreza e também quando é observada a qualidade de vida dos pobres nordestinos em comparação com os pobres das demais áreas do Brasil (ROCHA, 2003).

A intensidade da representação do nordestino na pobreza nacional foi descrita por Rocha (2003), onde a autora mostra nos seus resultados que ao longo da década de 90, mesmo com a melhoria dos indicadores sociais trazidas pelo Plano Real, a participação da pobreza nordestina aumenta no contexto nacional de 41,58% para 42,03%. Este resultado se mostra mais intenso quando analisado com um recorte para o meio rural da região, que tem sua participação na pobreza rural do país elevada de 65,06% para 68,89%. Essa dinâmica negativa da evolução da pobreza rural do Nordeste foi compensada apenas em parte pela diminuição da pobreza nas famílias rurais residentes no Sul e no Centro-Oeste.

Entre as principais causas que trazem como consequência as discrepâncias na qualidade de vida da população pobre do Nordeste, segundo as análises do Grupo de Trabalho para o Desenvolvimento do Nordeste (GTDN, 1967), destacam-se o processo histórico de desenvolvimento do país que engloba desde a escassez de terras em condições adequadas para a produção; passando por uma política industrial que favoreceu o desenvolvimento econômico e urbanização das outras regiões (Sul, Sudeste e Centro-Oeste); e finalmente, devido à falta de compreensão dos problemas decorrentes das desigualdades de renda entre as regiões no planejamento econômico em que se deu o desenvolvimento nacional.

Ainda segundo o GTDN (1967), o resultado dessa falta de atenção com a região em questão, no que se refere à rota de investimentos e industrialização, aliado com a concentração de renda das monoculturas predominantes na região e a baixa produtividade agrícola local, deixam esta como um retrato da baixa qualidade de vida da população pobre brasileira.

Em relação à divisão espacial da pobreza e às inferiores condições de vida da população pobre do Nordeste, alguns autores como Graziano da Silva (1997), Lima (2008) e Schneider (2009), argumentam que esses problemas são mais graves no meio rural da região, pois as famílias rurais desse recorte geográfico estão envolvidas em um conjunto de atividades agrícolas voltadas

principalmente à subsistência que, na maioria dos casos, não agrega valor aos rendimentos e que demandam baixa qualificação técnica. Nascimento (2008) enfatiza outros possíveis determinantes da elevada pobreza no meio rural nordestino, entre eles: concentração de terras, influências negativas da seca (problema climático) na produção agrícola e pecuária, e as dificuldades de acesso ao crédito por parte dos pequenos produtores agrícolas da região.

A pouca diversificação nas fontes de renda associada com a forte dependência das famílias agrícolas para com as transferências governamentais, na forma de aposentadorias, pensões e programas sociais, formam uma barreira para o desenvolvimento social e produtivo do meio rural nordestino com uma falta de perspectiva para mudança, tendo como alternativa mais viável, os incentivos ao desenvolvimento de atividades produtivas não ligadas a agricultura como um meio de reversão desse quadro negativo de bem-estar social dessas famílias (LIMA, 2008; SCHNEIDER e FIALHO, 2000). Graziano da Silva (1997), Schneider (2009), e agências internacionais de pesquisa e financiamento (OIT, 2005; WORD BANK, 2005) também enfatizam a importância das atividades não-agrícolas e a combinação dessas com as demais atividades agrícolas (pluriatividade), como meio para a superação da pobreza e desigualdade social no campo, bem como para o desenvolvimento do meio rural principalmente em regiões com maior incidência de pobreza rural.

Desse modo o objetivo desse trabalho é verificar se as rendas não-agrícolas têm efeito de diminuição dos indicadores de pobreza para as famílias rurais cearenses e mensurar o tamanho dessa diminuição. A questão que se formula é: *Quais os impactos das rendas não-agrícolas sobre a pobreza rural no estado do Ceará?* Este trabalho está dividido em 5 partes: Esta introdução; revisão de literatura; metodologia; resultados e principais conclusões sobre as discussões levantadas.

## 2 REVISÃO DE LITERATURA

### 2.1 Rendas Não-agrícolas e Pluriatividade: Conceitos Fundamentais

O conceito de rendas não-agrícolas diz respeito aos rendimentos advindos de atividades não ligadas à agricultura<sup>1</sup>. A incidência dessas atividades não-

<sup>1</sup> Para o IBGE, as atividades agrícolas são aquelas relacionadas a agricultura, silvicultura, pecuária, extração vegetal, pesca e piscicultura. Deve-se ter em vista também que as atividades ligadas a todo beneficiamento de produtos agropecuários realizado em estabelecimentos rurais rudimentares e não licenciados é considerado "indústria rural" dentro das atividades agropecuárias (LIMA, 2008).



-agrícolas é resultado da estrutura de desenvolvimento das áreas rurais, bem como do processo de integralização e distribuição, no espaço, das indústrias e serviços que antes só eram oferecidos nos grandes centros urbanos. Essas novas dinâmicas aliadas também ao desenvolvimento tecnológico da agricultura mudaram a configuração da oferta de mão de obra para outros setores produtivos que não necessariamente estão ligados de maneira direta à produção agrícola (SCHNEIDER, 2009; VEIGA, 2000).

No meio rural de algumas regiões brasileiras e em muitos países do mundo, esse conjunto de atividades não-agrícolas ganharam importância nos estudos sociais a partir da década 70. Os argumentos quanto à relevância da introdução dessas atividades no meio rural por parte de membros das famílias agrícolas, discutem que as atividades não-agrícolas têm papel fundamental na manutenção das relações sociais no campo e incentivos ao próprio desenvolvimento e estruturação da agricultura. Torna-se fundamental observar que a busca pela inserção das atividades não-agrícolas faz parte de estratégias das famílias rurais objetivando superar as dificuldades financeiras e riscos elevados da produção agrícola, bem como acesso a novos mercados de trabalho e produtos refletindo maiores oportunidades e melhoria das condições de vida (KAGEYAMA, 1998).

As características fundamentais das atividades não-agrícolas que as tornam uma alternativa viável de estratégia para o desenvolvimento social e diminuição dos riscos por parte das famílias rurais são: maiores níveis e estabilidades nas rendas se comparados com a maior parte dos segmentos exclusivamente agrícolas, a exigência de capacitações não elevadas e fortalecimento das condições e perspectivas da ascensão social por parte dessas famílias rurais (DEL GROSSI, 1999).

Após o entendimento das rendas advindas de atividades não-agrícolas como meio de diversificação e aumento na renda, deve se levar em consideração a combinação dessas atividades não-agrícolas com as rendas agrícolas por parte das famílias rurais. Assim, tem-se a pluriatividade<sup>2</sup> que aborda uma situação em que pelo menos um membro da família exerce atividade exclusivamente agrícola e os demais empregados em outros setores.

Na literatura internacional, alguns autores como Fuller (1990), Carneiro (1995) e Gasson (1988) têm uma visão da pluriatividade como uma etapa posterior do desenvolvimento capitalista no meio rural. Além disso, esses

2. A pluriatividade foi utilizada até meados da década de 80, na literatura internacional, relacionada aos termos *part-time farming* e *pluriactivié* que designavam um fenômeno crescente de diversificação das fontes de renda e introdução de novas atividades no meio rural que começou a se tornar evidente a partir da década de 1970 (NASCIMENTO, 2005).

autores argumentam que a combinação de atividades agrícolas e não-agrícolas é uma ação organizada pelos agentes visando a convergência de desequilíbrios econômicos e sociais do desenvolvimento agrário em ascensão a partir da década de 50.

No Brasil, além do objetivo principal de servir como estratégia para superação das vulnerabilidades sociais e financeiras como citado anteriormente, a pluriatividade altera as relações demográficas e produtivas entre o campo e a cidade pois muda de maneira estrutural o mercado de trabalho (MATTEI, 1999). Um resumo das principais mudanças positivas resultantes da combinação das atividades agrícolas e não-agrícolas e sua contribuição para o desenvolvimento do rural brasileiro, segundo Graziano da Silva (1997) e Schneider (2009), são:

- a) A elevação da renda da família, pois essas atividades não-agrícolas, em sua essência, são diretamente relacionadas com o comércio, e agroindústrias que dão maior valor agregado aos salários;
- b) Maior contribuição para geração de emprego no espaço rural;
- c) Estabilização dos rendimentos, em contraste com os efeitos sazonais na qual está sujeita a agricultura;
- d) Redução das migrações do campo para a cidade, além do estímulo aos mercados e desenvolvimento dos territórios rurais locais.

## 2.2 Pobreza Rural e Suas Características

Os temas relativos à pobreza ganharam uma maior relevância nos estudos relacionados ao desenvolvimento econômico a partir do final do século XX, quando, em algumas partes da Europa e em outros países desenvolvidos, observaram-se que as questões relativas à pobreza e desigualdades sociais em nichos da população não estavam sendo equacionadas mesmo levando em consideração os efeitos do crescimento econômico. Assim, mesmo em países ricos ou nos países em desenvolvimento, a pobreza ganhou uma importância central na formulação de políticas; na elaboração de estudos sobre a sociedade e nos debates acadêmicos (CODES, 2008).

A incidência da pobreza em uma determinada sociedade pode ser relacionada com o não atendimento (ou atendimento de forma inadequada) de algumas necessidades básicas de uma parcela dos indivíduos que compõem esta

sociedade. Para que seja possível a investigação e caracterização desse fenômeno, faz-se necessário a definição dos tipos de necessidades que se busca investigar como fator condicionante da pobreza (ROCHA, 2006).

As definições do fenômeno que se tornaram bastante generalizadas ao longo do período de desenvolvimento de estudos referidos a sua incidência foram aquelas relacionadas aos conceitos: pobreza absoluta (ou extrema pobreza) e pobreza relativa. O conceito de pobreza absoluta está relacionado ao não atendimento de necessidades ligadas ao mínimo vital de um indivíduo em questão, este conceito vincula-se diretamente a ideia de sobrevivência física. O conceito relativo de pobreza é discutido para a sociedade em que o mínimo vital já é atendido, ou seja, quanto ao bem-estar da população, o único objetivo a ser alcançado é a diminuição da desigualdade entre os indivíduos (ROCHA, 2006). Outros autores como Sen (2010), estudam a incidência do fenômeno pobreza levando em consideração as capacidades que uma pessoa possui (seja pelo aspecto produtivo; de interação social; de acesso à educação, etc.). Sendo assim, o conceito de pobreza está ligado à privação dessas capacidades dos indivíduos, ou seja, são os níveis de privação de liberdades dos seres que determinam os estágios de desenvolvimento social de uma população.

A utilização do conceito de pobreza como privação de liberdades (ou capacidades) dos indivíduos, tem a característica de abranger as mais diversas necessidades básicas de sobrevivência desses agentes, já que leva em consideração as múltiplas facetas que compõem o desenvolvimento da pobreza em uma sociedade. Esse enfoque multidimensional traz consigo uma maior eficiência no combate ao problema, pois tem a vantagem de proporcionar um maior entendimento das diversas características e pontos que precisam ser atacados para equacionar as questões relativas à pobreza (BARROS *et al.*, 2006).

O meio rural Brasileiro tem algumas características relevantes no que diz respeito às condições de vida das famílias. O estudo separado dessa área em relação ao meio urbano é resultado de um entendimento diferenciado quanto ao conceito de pobreza e necessidades básicas relacionadas à observação da estrutura de acesso à terra e recursos naturais; e disponibilidade de serviços públicos básicos de fácil acesso para a população de cada espaço geográfico.

A incidência da pobreza no meio rural está diretamente relacionada à ocorrência de outros problemas, como o aumento do desemprego nos setores urbano-industriais; a pressão demográfica sobre as grandes cidades; a violência urbana, etc. Todos esses fatores quando somados ao processo de mecanização e revolução produtiva na agricultura que ascenderam à posição relativa das re-

giões Sul, Sudeste e Centro-Oeste em detrimento do setor agrícola nordestino, fizeram com que surgissem parcelas da população rural dependentes de uma agricultura não desenvolvida e que se tornaram o retrato da pobreza rural brasileira. Sendo assim, o atraso no desenvolvimento da agricultura nordestina, e a característica de subsistência dessa agricultura, é uma das principais causas da reprodução das desigualdades sociais e aumento das disparidades nos indicadores de qualidade de vida para as famílias com fontes de rendimento exclusivamente ligadas a agricultura (NASCIMENTO e CARDOSO, 2007).

As desigualdades nos indicadores sociais e deficiências no desenvolvimento estrutural do campo são resultados, basicamente, das marcas históricas que englobam o processo de colonização do país e de sua essência exploratória, na qual, a colônia brasileira servia apenas como uma base para atendimento das necessidades da metrópole portuguesa. Dentre as principais consequências trazidas por esse processo histórico, as que mais se destacam como retratos da discrepância nas condições de vida no meio rural são: a incidência da pobreza; a concentração de renda; e os altos níveis de marginalização social, medidos por indicadores como taxa de analfabetismo, desocupação e anos de estudo (HOFFMANN e NEY, 2007).

### 3 METODOLOGIA

#### Impactos das Rendas Não-agrícolas

Para atender ao objetivo principal do trabalho que é medir o impacto (tamanho da diferença) entre as rendas das famílias não-agrícolas sobre as exclusivamente agrícolas, ou seja, que todos os seus integrantes têm fontes de rendas advindas apenas da agricultura, lança-se um exemplo de uma família  $i$  como sendo essa exclusivamente agrícola e tendo renda per capita familiar igual a  $R_i$ . Sendo essa mesma família  $i$  agora não-agrícola, o rendimento familiar per capita da unidade familiar em questão é dada pela seguinte expressão dada por Becker e Ichino (2002):

$$R_i = DR_i + (1 - D)R_i \quad (1)$$

Onde,

$D$  é uma variável *dummy* que assume valores 1 para famílias do tipo não-agrícolas e 0 para as exclusivamente agrícolas.

Para medir o tamanho da diferença nos rendimentos que as rendas não-agrícolas proporcionam é necessário medir a diferença total dos rendimentos

dessa família quando ela passa a ser não-agrícola dado pela fórmula seguinte:

$$\Delta R_i = R_{iD=1} - R_{iD=0} \quad (2)$$

Expandindo a expressão para a população total  $n$ , em uma análise mais geral, introduz-se o impacto médio ( $\delta R_i$ ) das rendas não-agrícolas sobre as rendas não-agrícolas dada pela equação:

$$\delta R_i = E(R_{iD=1} - R_{iD=0} | D = 1) \quad i = 1, 2 \dots n \quad (3)$$

Sendo que  $E(\cdot | D = 1)$  é o valor esperado condicional da diferença dos rendimentos de uma família  $i$  dado que a família passou a ser não-agrícola. Como não é possível observar a família  $i$  sendo não-agrícola e exclusivamente agrícola no mesmo instante de tempo, uma medida aproximada seria realizar a comparação entre dois grupos: um grupo formado por famílias não-agrícolas, grupo tratado, e outro formado por famílias exclusivamente agrícolas, grupo de controle, dado da seguinte forma (DUARTE *et al.*, 2009):

$$\begin{aligned} & E(R_{iD=1} | D = 1) - E(R_{iD=0} | D = 0) = \\ E(R_{iD=1} | D = 1) - E(R_{iD=0} | D = 1) + E(R_{iD=0} | D = 1) - E(R_{iD=0} | D = 0) &= (4) \\ &= \delta' R_i + E(R_{iD=0} | D = 1) - E(R_{iD=0} | D = 0) \end{aligned}$$

Nesse caso o impacto médio entre os grupos ( $\delta' R_i$ ) tem o viés  $E(R_{iD=0} | D = 1) - E(R_{iD=0} | D = 0)$  que representam, basicamente, as diferenças dos rendimentos causados por características antes da introdução das rendas não-agrícolas: grau de escolaridade dos membros da família, se os indivíduos possuem ou não aposentadoria e etc. Sendo assim, deve se procurar uma medida de análise que minimize esse viés que pode ser dada pela metodologia *Propensity Score Matching*.

### 3. 2 Propensity Score Matching (PSM)

A metodologia *Propensity Score Matching*, ou mais resumidamente PSM, foi criada por Rosenbaum e Rubin (1983) objetivando reduzir o viés do cálculo do efeito médio do impacto de uma variável de interesse de um grupo tratado sobre um grupo de controle. A definição do *Propensity Score*, chamado então de  $p(X)$ , é entendida como a probabilidade condicional de participar do grupo tratado dado um vetor de características observáveis  $X$ , mais formalmente dado pela expressão:

$$p(X_i) = Pr(D_i = 1 | X_i) = E(D_i | X) \quad (5)$$

Após a estimação de  $p(X)$  torna-se factível o cálculo do impacto médio sobre uma população de participar ou não do grupo tratado. Assim surge o conceito de *Average effect of Treatment on the Treated (ATT)*, ou em português, efeito médio do tratamento sobre os tratados (considerado aqui por  $\tau$ ). Segundo Becker e Ichino (2002) aplicado ao contexto do trabalho é dado por:

$$\begin{aligned}\tau &= E(R_{i,D=1} - R_{i,D=0} | D_i = 1) = \\ &E[E\{R_{i,D=1} - R_{i,D=0} | D_i = 1, p(X_i)\}] = \quad (6) \\ &= E[E\{R_{i,D=1} | D_i = 1, p(X_i)\} - E\{R_{i,D=0} | D_i = 0, p(X_i)\} | D_i = 1]\end{aligned}$$

Para o cálculo de  $\tau$  deve se admitir duas hipóteses sobre  $p(X_i)$ :

**Hipótese 1** – Balanceamento das variáveis observadas antes do tratamento, ou seja, que a seleção da amostra seja independente das características observadas condicionais às probabilidades de ter ou não renda não-agrícola, de maneira formal:

$$D \perp X_i | p(X_i) \quad (7)$$

**Hipótese 2** – As rendas dos dois grupos (tratamento e controle) independem das variáveis observáveis utilizadas dado o *propensity score*:

$$R_{i,D=1}, R_{i,D=0} \perp X_i | p(X_i) \quad (8)$$

O cálculo do ATT para este trabalho foi feito considerando duas metodologias: método de *Kernel* e método *Nearest Neighbor* (ou vizinho mais próximo). No método *Nearest Neighbor (NN)* calcula-se a média da diferença da variável de interesse entre uma observação do grupo tratado com uma equivalente do grupo de controle levando em consideração o  $p(X_i)$  mais próximo. No método de *Kernel (K)* a comparação é feita com todas as observações tratadas e a média ponderada das observações no grupo de controle. Os pesos que entram no cálculo da média são dados pelo inverso da diferença de  $p(X_i)$  para  $i$  no grupo de tratamento e um  $\hat{i}$  correspondente no grupo de controle.

Para o presente trabalho o a aplicação do PSM se dará com o seguinte roteiro:

- Primeiro: calcular os indicadores de pobreza FGT para as rendas *per capita* das famílias exclusivamente agrícolas e não agrícolas individualmente considerando as características individuais de cada grupo;
- Segundo: estimar o  $p(X_i)$  com base em um modelo *logit* e depois reali-

zar o pareamento considerando os métodos de *Kernel* e *Nearest Neighbor*;

- Terceiro: Calcular o ATT levando em consideração a variável de interesse renda familiar per capita com base nos dois métodos referidos e, posteriormente, gerar duas variáveis, sendo cada uma um novo vetor de renda agrícola tratada correspondente a ampliação do rendimento resultante do ATT de cada método;

- E quarto: calcular os novos indicadores FGT para as famílias exclusivamente agrícolas considerando os novos vetores de rendas tratadas dada pela soma do rendimento anterior com cada ATT calculado na etapa anterior e observar o impacto dos rendimentos não agrícolas sobre os indicadores de pobreza rural.

### 3.3 Características Observáveis e Modelo *Logit*

Para a estimação do PSM, por meio do modelo *logit*, foram consideradas as variáveis dadas, segundo Lima (2008), como as mais relevantes que influenciam a probabilidade de uma família ser ou não do tipo não-agrícola. A variável dependente do modelo *logit* é uma *dummy*, assumindo o valor 1 se a família for não-agrícola e 0, caso contrário. As variáveis explicativas do modelo, incluídas no vetor de características  $X_i$  são: idade média da PEA restrita; idade média da PEA restrita ao quadrado<sup>3</sup>; média de anos de estudo; número de componentes da família; duas *dummies* para posição na ocupação, definidas como 1 para conta-própria e 0, caso contrário, além de 1 para empregados e 0, caso contrário; uma variável *dummy* para local de moradia, assumindo valor de 1 para o rural mais distante (agropecuário) e 0, caso contrário. Também foi considerada a razão de dependência formada dividindo o número de membros da família considerados dependentes (com idade menor que 10 anos ou maior que 65 anos) pela idade da PEA ativa; uma variável *dummy* indicando se a família possui ou não renda do não-trabalho, como aposentadorias, pensões ou transferência de renda do governo, por exemplo.

Para a análise dos determinantes das famílias serem ou não não-agrícolas foi utilizado o modelo de repressão logística ou simplesmente *logit*. Esse modelo é baseado na função de distribuição logística acumulada especificada segundo Pindick e Rubinfeld (2004), da seguinte forma:

---

<sup>3</sup> PEA restrita se refere às pessoas com 10 anos ou mais que trabalharam mais de 15 horas na semana de referência, excluídas as que se dedicam apenas ao autoconsumo. A variável idade da PEA restrita ao quadrado é inserida levando em consideração o efeito do ciclo de vida sobre as escolhas dos indivíduos, ver Lima (2008).

$$P_i = \frac{1}{1 + e^{-(\beta_0 + \beta_i X_i)}} \quad (9)$$

Na equação,  $e$  é a base dos logaritmos naturais e tem valor aproximado de 2,7182.  $P_i$  é a probabilidade de um indivíduo fazer uma escolha  $i$  dado um vetor de características (ou variáveis)  $X_i$ . Ainda segundo Pindick e Rubinfeld (2004) para que seja estimado um modelo de regressão de acordo com essa função de distribuição logística é necessário que ambos os membros da equação anterior sejam multiplicados por  $1 + e^{-(\beta_0 + \beta_i X_i)}$ , para que se obtenha:

$$(1 + e^{-(\beta_0 + \beta_i X_i)})P_i = 1 \quad (10)$$

Dividindo-se ambos os lados por  $P_i$  e subtraindo-lhes uma unidade, a expressão anterior será:

$$e^{-(\beta_0 + \beta_i X_i)} = \frac{1}{P_i} - 1 = \frac{1}{e^{-(\beta_0 + \beta_i X_i)}} = \frac{P_i}{1 - P_i} \quad (11)$$

Finalizando, toma-se o logaritmo natural de ambos os lados da expressão anterior e o modelo de *logit*, como é conhecido, pode ser estimado segundo Gujarati (2009) e Greene (2008) pela seguinte equação:

$$\tilde{L} = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_i X_i + \hat{\varepsilon}_i \quad (12)$$

Onde

$P_i$  é a probabilidade de um evento ocorrer assumindo os valores de 0 e 1;

$\hat{\beta}_i$  é o coeficiente angular estimado relacionado à  $i$ -ésima variável;

$X_i$  é o vetor de características do  $i$ -ésimo indivíduo (família);

$\hat{\varepsilon}_i$  é termo de erro estimado.

No contexto do presente estudo, a regressão estimada permitiu analisar o efeito de variáveis selecionadas no aumento ou diminuição da probabilidade de uma família rural do estado de Ceará ser do tipo não-agrícola. Assim, para este trabalho o modelo *logit* estimado pode ser descrito da seguinte maneira:

$$\begin{aligned} \tilde{L} &= \ln\left(\frac{P_i}{1 - P_i}\right) = \\ &= \beta_0 + \beta_1 IDPEARES + \beta_2 IDPEARES^2 + \beta_3 ANOEST + \\ &+ \beta_4 NUMCOMPFAM + \beta_5 RNT + \beta_6 CONTAPRO + \beta_7 EMPREGADOS + \\ &+ \beta_8 LOCALMORA + \beta_9 RAZAODEP + \varepsilon \end{aligned} \quad (13)$$



De acordo com as variáveis dependentes definidas anteriormente.

### 3.4 Parâmetros e Índices de Pobreza Foster-Greer-Thorbecke (FGT)

Para um estudo que vise a análise conceitual e das características da pobreza, faz-se necessário o estabelecimento de um parâmetro que não seja afetado pelo tempo em relação às condições de vida dos indivíduos. Essa medida é a linha de pobreza, a qual permite uma sociedade específica definir como pobres todos aqueles indivíduos com indicadores abaixo desse parâmetro (BARROS *et al.*, 2000).

A orientação no que se refere às medidas de pobreza é que se leve em consideração as informações nutricionais no caso de formulação de linhas de indigência, e informações de cesta de consumo combinadas com coeficientes que representem as necessidades de moradia, saúde, transporte, etc. Porém, devido à complexidade e, muitas vezes, ausência dessas informações, as linhas de pobreza são adotadas com certa arbitrariedade, pois são referentes a informações de rendimento (HOFFMANN, 1998).

Como argumento a favor do uso de múltiplos de renda como linhas de pobreza Hoffmann (1998) e Rocha (2006) citam o alto grau de monetarização da economia brasileira, onde a grande maioria das necessidades dos indivíduos é atendida pela via das trocas mercantis de forma indireta por meio dos rendimentos. Assim, para este trabalho a incidência de pobreza foi calculada utilizando-se uma linha de pobreza igual à metade do salário mínimo de cada ano estudado.

A análise da incidência da pobreza é feita aqui por meio dos índices *Foster-Greer-Thorbecke* (FGT) que mostram os vários aspectos do fenômeno. O indicador de pobreza FGT é avaliado por três índices (FOSTER, *et al.*, 1984):

P0 – proporção dos pobres que mede o tamanho do número de pobres em relação à população total.

P1 – hiato da pobreza que mede a intensidade da pobreza como um déficit de renda.

P2 – severidade da pobreza que mostra quão desigual é a distribuição de renda entre os mais pobres da população estudada, dando uma maior relevância a esses últimos.

Segundo Lima *et al.* (2011), os índices FGT são descritos pelas seguintes equações:

$$P0 = \frac{q}{n} \quad (14)$$

$$P1 = \frac{1}{n} \sum_{i=1}^q \frac{z - y_i}{z} \quad (15)$$

$$P2 = \frac{1}{n} \sum_{i=1}^q \left( \frac{z - y_i}{z} \right)^2 \quad (16)$$

Na qual,

$q$  é o número de pessoas ou famílias abaixo da linha da pobreza;

$n$  é o tamanho da população analisada;

$z$  é a linha de pobreza<sup>4</sup>;

$y_i$  é a renda per capita do  $i$ -ésimo elemento da população (renda familiar per capita).

Para Lima (2008), os dois últimos indicadores (hiato da pobreza e severidade da pobreza) são mais relevantes do que a proporção de pobres (P0), pois levam em consideração os mais pobres entre os pobres dando um peso maior a esses na análise de incidência da pobreza.

A unidade consumidora levada em consideração neste trabalho é a família (conjunto de indivíduos ligados por traços de parentesco) que representa a união das diversas necessidades dos indivíduos que a compõe e todo o esforço para otimização e atendimento dessas necessidades (HOFFMANN, 1998; KAGEYAMA, 1998). A renda per capita equivale ao rendimento familiar per capita (soma de todos os rendimentos da família divididos pela quantidade de pessoas da mesma família residentes no mesmo domicílio).

### 3.5 Fonte de Dados

A base de dados para este trabalho é a Pesquisa Nacional por Amostra de Domicílios (Pnad) do IBGE para o ano de 2013 já com as novas ponderações. A Pnad possui um desenho de amostragem complexa com estratificação, conglomerado, probabilidades desiguais de seleção e ajustes de pesos amostrais. Segundo Lima (2008), tratar os dados da Pnad como *IID* (independentes e igualmente distribuídos) significa desconsiderar os efeitos de seu plano amostral e com isso construir os intervalos de confiança e estimar os níveis de signi-

4. A linha de pobreza utilizada para é meio salário mínimo equivalente a R\$ 339,00 para 2013.

ficância erroneamente.

A estratificação da amostra da Pnad é feita dividindo o país em 36 estratos, dos quais 18 unidades são estados que formam um estrato único e independente<sup>5</sup>, nas outras nove unidades da federação são definidos mais dois estratos em cada uma: um contando com todos os nove municípios da região metropolitana com sede na capital<sup>6</sup> e o outro com os demais municípios, auto representativos e não auto representativos formando-se assim, mais 18 unidades de amostragem. Os municípios auto representativos (conhecidos como área censitária 2) são aqueles não pertencentes às regiões metropolitanas, mas que tem sua presença na amostra garantida, devido ao seu tamanho populacional maior que o estabelecido para os estratos. Os municípios não auto representativos (área censitária 3), são os demais municípios que tem população inferior ao estabelecida pela estratificação e não têm sua participação garantida na amostra ficando dependentes da seleção de municípios na realização da amostra (são selecionados 2 municípios não auto representativos para comporem o extrato da área censitária 3).

O *software* utilizado na pesquisa empírica é o Stata 12.1 comercializado pela Stacorp, College Station, Texas 77845 USA.

#### 4 RESULTADOS E DISCUSSÃO

Iniciando os resultados apresentando indicadores relativos às condições de vida das famílias exclusivamente agrícolas e não agrícolas no meio rural de Ceará, observa-se pela Tabela 1 que a renda per capita média das famílias rurais como um todo é de R\$ 342,02; para as famílias exclusivamente agrícolas a média é de R\$ 259,24; já para as não-agrícolas a média sobe para R\$ 480,45, acima aproximadamente 40,35% da média para o estado como um todo e 85,35% maior que as rendas das famílias exclusivamente agrícolas. Mostrando assim que as famílias não-agrícolas têm aumento de renda relativo aos demais tipos de famílias rurais do estado.

A média dos anos de estudo também é maior para os integrantes das famílias não-agrícolas sendo estimada em 6,89 anos ficando acima da média para

---

5. Os Estados que formam um único extrato são: Acre, Amazonas, Amapá, Alagoas, Distrito Federal, Espírito Santo, Goiás, Maranhão, Mato Grosso, Mato Grosso do Sul, Paraná, Piauí, Rio Grande do Norte, Roraima, Santa Catarina e Tocantins.

6. As nove Regiões Metropolitanas são: Região Metropolitana de Belém, Região Metropolitana de Belo Horizonte, Região Metropolitana de Curitiba, Região Metropolitana de Fortaleza, Região Metropolitana de Porto Alegre, Região Metropolitana de Recife, Região Metropolitana de Salvador e Região Metropolitana de São Paulo.

o meio rural do estado como um todo, 4,58 anos, e aproximadamente 53% acima da média para os integrantes das famílias exclusivamente agrícolas que foi de 3,99 anos de estudo. Para a taxa de analfabetismo, essa atinge 7,9% da população rural de Ceará entre 10 e 65 anos de idade. Para as famílias agrícolas a proporção de analfabetos aumenta para 13,5%; e para as famílias não-agrícolas a taxa é significativamente inferior que as duas anteriores sendo calculada em aproximadamente 2,6% da população pertencente à essa classe de família.

Em relação a taxa de desemprego também se observa melhores condições, levando em consideração comparação relativa, das famílias com rendas Não-Agrícolas com aproximadamente 11,4% da população entre 10 e 65 anos desempregada. Para as famílias rurais do estado como um todo, a proporção chega a 13,9% e para as famílias ligadas exclusivamente à agricultura a proporção é de 14,5%.

**TABELA 1 – Alguns indicadores das condições de vida para as famílias rurais do Estado do Ceará como um todo; famílias exclusivamente agrícolas e famílias não-agrícolas para o ano de 2013**

Indicador	Estimativa	Erro Padrão Linearizado	Intervalo de Confiança (95%)	
	Média			
Renda <i>per capita</i>	342,024	14,043	321,259	393,810
Renda <i>per capita</i> ( $R_{D=0}$ )	259,249	16,244	229,963	325,675
Renda <i>per capita</i> ( $R_{D=1}$ )	480,458	25,924	371,966	509,569
Anos de estudo médio	4,587	0,191	4,481	5,267
Anos de estudo médio ( $R_{D=0}$ )	2,997	0,185	3,586	4,345
Anos de estudo médio ( $R_{D=1}$ )	6,898	0,232	5,590	6,541
	<b>Proporção</b>			
Analfabetismo	0,079	0,013	0,055	0,110
Analfabetismo ( $R_{D=0}$ )	0,135	0,022	0,080	0,170
Analfabetismo ( $R_{D=1}$ )	0,018	0,012	0,001	0,052
Desemprego	0,139	0,015	0,091	0,154
Desemprego ( $R_{D=0}$ )	0,145	0,022	0,093	0,155
Desemprego ( $R_{D=1}$ )	0,114	0,018	0,083	0,177

Fonte: dados da PNAD 2013/Elaboração própria.

Na Tabela 2 são apresentados os primeiros resultados dos indicadores de pobreza FGT antes da aplicação da metodologia *Propensity Score Matching*. De

acordo com os resultados, a proporção de pobres (P0) é de aproximadamente 59,6% no meio rural da região Nordeste como um todo. Para o meio rural de Ceará, o indicador é de 63,7%. Considerando as famílias com renda exclusivamente agrícolas, a proporção de pobres sobe para 70% da população e as Não-Agrícolas apresentam pobreza de 55,8%. Essas tendências se repetem nos demais indicadores.

No hiato da pobreza, o cálculo para as famílias o Nordeste rural como um todo fica em 0,311 e acima do meio rural de Ceará com 0,344. Considerando as famílias exclusivamente agrícolas e as Não-Agrícolas, a diferença no indicador chega a 19,4 pontos percentuais.

A severidade da pobreza, medida de desigualdade de renda entre os mais pobres da população, é 0,202 no Nordeste rural; 0,229 no meio rural de Ceará como um todo; este sobe para 0,308 quando se consideram apenas as famílias com renda exclusivamente agrícolas e 0,131 nas famílias Não-Agrícolas. Esses primeiros resultados mostram, de maneira geral, que as rendas não-agrícolas têm efeito negativo sobre os indicadores de pobreza. Porém, essas diferenças podem acumular os efeitos de outras variáveis sobre as rendas das famílias como grau de escolaridade número de membros, entre outros, como já foi discutido anteriormente. Esses efeitos acumulados serão retirados por meio da aplicação do PSM.

**TABELA 2 – Indicadores de pobreza FGT (Foster-Greer-Thorbecke) para o meio rural nordestino; famílias rurais do Ceará como um todo; famílias exclusivamente agrícolas e famílias não-agrícolas cearenses para o ano de 2013**

Indicadores FGT	Nordeste	Ceará	$(R_{D=0})$	$(R_{D=1})$	$\Delta_{*R_{D=0}-R_{D=1}}$
P0	0,596	0,637	0,700	0,558	0,143
P1	0,311	0,344	0,431	0,237	0,194
P2	0,202	0,229	0,308	0,131	0,177

Fonte: dados da PNAD 2013/Elaboração própria.

Na Tabela 3, são mostrados os resultados do modelo *logit* que foi usado para cálculo do *Propensity Score*. De acordo com os resultados, a maior parte das variáveis apresenta efeito estatisticamente significativo ao nível de confiança de 10% sobre as chances de a familiar ser ou não ser do tipo Não-Agrícola. A variável *dummy* para tipo de família (se é formada por empregados ou não) e

razão de dependência da família não se mostraram estatisticamente significativas, apresentando assim diferença dos resultados estimados por Lima (2008).

A variável PEA e PEA ao quadrado mostram que a medida que a idade média dos indivíduos da família aumenta, a probabilidade de a familiar ser não agrícola aumenta, mas com ganhos marginais cada vez menores. Sendo assim, a idade mostra o efeito de U invertido contendo um ponto de máximo na qual a idade apresentaria o maior efeito sobre a probabilidade de ser ou não do tipo Não-Agrícola. As demais variáveis apresentam sinais esperados e seus coeficientes são estatisticamente significativos a 10%.

O aumento da média dos anos de estudo e do número de componentes têm efeitos positivos sobre a familiar ser Não-Agrícola. Possuir rendas do não trabalho mostra efeito negativo devido a provável dependência da renda da família de programas sociais e aposentadoria, dependência essa que pode ser resultado de algum desequilíbrio no processo produtivo familiar. As famílias do tipo conta própria apresentam o maior efeito em módulo significando que a maior parte dos trabalhadores que têm renda Não-Agrícola são autônomos.

Por último, o Local de moradia apresenta efeito negativo, como já era de se esperar, devido à distância do acesso a serviços e estruturas produtivas que ficam cada vez mais difíceis de serem realizadas à medida que acontece um afastamento dos centros urbanos.

**TABELA 3 – Resultados do modelo Logit para pareamento dos indivíduos com características observadas (score das variáveis independentes) semelhantes**

Variável	Coefficiente	Erro Padrão Robusto	Estatística $z$	$prob > z$
Idade da PEA restrita	0,025697	0,0237904	1,08	0,005
Idade da PEA restrita <sup>2</sup>	-0,002244	0,0003179	-7,06	0,018
Média dos anos de estudo	0,634878	0,0124086	51,16	0,000
Nº de componentes da família	0,374755	0,0316066	11,86	0,100
Possui renda do não trabalho	-0,2384792	0,0971718	-2,45	0,014
Família do tipo contra própria	3,635898	0,1126182	32,29	0,000
Família composta por empregados	4,5889645	0,1092791	41,99	0,201
Local de moradia	-0,5053067	0,079454	-6,36	0,002
Razão de dependência	0,0166865	0,0449165	0,37	0,710
Constante	-4,971037	0,4580113	-10,85	0,000

Fonte: dados da PNAD 2013/Elaboração própria.

Após a realização do cálculo do *Propensity Score* por meio do modelo

*logit*, foi obtido o Efeito Médio do Tratamento (ATT) da renda per capita observando a diferença entre a renda tratada e não tratada com as famílias com scores semelhantes. Na Tabela 4, o ATT é apresentado com duas metodologias, *Kernel* e *Nearest Neighbor*. Pelo método de *Kernel*, considerando 2.443 observações tratadas e 3.095 no grupo de controle, já que por esse método é considerado uma média ponderada de todas as observações no grupo de controle, o ATT calculado foi de R\$ 190,28 e estaticamente significativo com estatística *t* acima de 10.

O ATT pelo método de *Nearest Neighbor* (ou vizinho mais próximo) apresentou diferença de R\$ 195,41 sobre as 2.443 observações no grupo de tratamento e 911 no grupo de controle usadas no pareamento. Esse resultado também foi estatisticamente significativo com estatística *t* de 7,025. De acordo com os resultados em ambos os métodos, a diferença entre as rendas Não-Agrícolas e exclusivamente agrícolas é positiva. Basicamente, tomando como exemplo um caso hipotético de uma família exclusivamente agrícola, em um primeiro momento, que passa a ter todos os membros empregados em atividades Não-Agrícolas posteriormente (passando assim a ser Não-Agrícola) tem seu rendimento per capita aumentando em 61,24% em média.

**TABELA 4 – Efeito médio do tratamento (ATT) no grupo tratado considerando os métodos de Kernel e Nearest Neighbor (Vizinho mais próximo)**

Método	Obs. Tratamento	Obs. Controle	ATT	Erro Padrão	Estatística <i>t</i>
Método de Kernel	2.443	3.095	190,287	18,552	10,256
Método de Nearest Neighbor	2.443	911	195,413	27,816	7,025

Fonte: dados da PNAD 2013/Elaboração própria.

Depois de considerar o impacto do ATT sobre as rendas exclusivamente agrícolas, foram recalculados os indicadores de pobreza FGT com as novas rendas agrícolas considerando o tratamento. Esses resultados mostram o impacto das rendas não agrícolas sobre os indicadores de pobreza para as famílias rurais de Ceará tomando como base as exclusivamente agrícolas que apresentaram maiores níveis de pobreza de acordo com a Tabela 2. Os resultados pelos dois métodos são apresentados na Tabela 5.

**TABELA 5 – Cálculo do impacto do ATT sobre o grupo de controle considerando os indicadores de pobreza FGT para o estado do Ceará no ano de 2013**

Indicadores FGT	$(R_{D=0})$	$(R_{KD=0})_{\Delta_{R_{D=0}-R_{KD=0}}}$ (Kernel)	$(R_{NN D=0})_{\Delta_{R_{D=0}-R_{NN D=0}}}$ (NN)		
P0	0,700	0,448	0,252	0,433	0,267
P1	0,431	0,093	0,339	0,086	0,346
P2	0,308	0,026	0,282	0,023	0,285

Fonte: dados da PNAD 2013/Elaboração própria.

Pelo método de *Kernel*, o efeito sobre os indicadores FGT são apresentados na terceira coluna da Tabela 5. A proporção de pobres cai 0,252 (ou 25,2 pontos percentuais). A medida de déficit de renda, o hiato da pobreza (P1) tem queda de 0,339 passando de 0,431 para 0,093. A medida de desigualdade de renda entre os mais pobres da população, severidade da pobreza tem queda de 0,282, passando de 0,308 para 0,026, o que representa quase a eliminação do problema considerando meio salário mínimo para o ano de 2013.

Os resultados pelo método de *Nearest Neighbor* têm uma magnitude ainda maior, relativamente comparando-o ao método de *Kernel* (como já era esperado já que o ATT por esse método foi maior do que pelo método de *Kernel*). A proporção de famílias pobres no meio rural de Ceará cai 0,267 passando de 0,700 para 0,433 da proporção de famílias com renda per capita inferior a meio salário mínimo. O hiato da pobreza apresenta uma diferença de 0,346, ficando em 0,086 após o tratamento. Ou seja, as famílias pobres com renda considerada após o tratamento sobem significativamente para próximo da linha da pobreza. A severidade da pobreza passa de 0,308 para 0,023 apresentando variação negativa de 0,229. Esses resultados somados aos obtidos pelo método de *Kernel* mostram efeitos significativos das rendas Não-Agrícolas sobre os indicadores de pobreza rural.

## 5 CONCLUSÕES

Os resultados do modelo de probabilidade usado para o cálculo do *Propensity Score Matching*, nesse caso o modelo *logit*, apresentou resultados já esperados no que diz respeito à probabilidade de auferir ou não rendas Não-Agrí-



colas. A idade mostrou efeito de U invertido sobre a probabilidade, mostrando assim que existe uma idade ótima na qual o componente da família tem maiores chances de ser Não-Agrícola. Na média dos anos de estudo, o efeito foi positivo, ou seja, cada ano a mais de estudo sobre a média de educação da família tem efeito de aumentar a probabilidade de essa ser do tipo Não-Agrícola.

Observou-se, também, que o impacto das rendas Não-Agrícolas, tomando como base os dois métodos de pareamento utilizados, foram positivos considerando a diminuição dos indicadores de pobreza FGT. O efeito médio do tratamento (ATT) se mostrou alto quando comparado com as rendas per capita familiares exclusivamente agrícolas. Pelo método de *Kernel* o impacto da introdução das rendas não ligadas à agricultura foi de R\$ 190,28 representando um aumento de aproximadamente 65% nos rendimentos. Pelo método de *Nearest Neighbor* o impacto foi ainda maior sendo calculado em R\$ 195,41 o que representa uma elevação nos rendimentos per capita familiares de 66%.

Analisando os impactos sobre os indicadores de pobreza, observou-se que a proporção de pobres caiu 0,252; o hiato da pobreza caiu 0,339; e a severidade da pobreza diminuiu em 0,282 levando em consideração a ponderação de todos os indivíduos do grupo de controle por intervalos nos *scores* de proporção, que é a base da metodologia de *Kernel*. Pelo método do vizinho mais próximo o impacto sobre a proporção de pobres foi de 0,267; sobre o déficit de renda, esse tem uma melhora de 0,346; e a severidade da pobreza caiu 0,288. Observou-se que os impactos medidos pelo segundo método foram maiores, isso devido ao ATT que também foi maior nesse segundo método que compara observações individuais com scores mais próximos entre os grupos de tratamento e controle.

Contribuições posteriores devem buscar identificar que tipo de fonte de renda não agrícola têm maior impacto nos indicadores de pobreza e ao menor custo para respaldar a formulação de políticas públicas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BARROS, A. R. **Desigualdades Regionais no Brasil**. Rio de Janeiro: editora Campus, 2011. 368 p.
- BARROS, R. P. de, CARVALHO, M.; FRANCO, S. **Pobreza multidimensional no Brasil**. Texto para Discussão n.º 1227. Rio de Janeiro: IPEA, 2006. 108 p.

- BARROS, R. P. de; HENRIQUES, R.; MENDONÇA, R. A estabilidade inaceitável: desigualdade e pobreza no Brasil. HENRIQUES, R. (Org.). In: **Desigualdade e Pobreza no Brasil**. Rio de Janeiro: IPEA, 2000. 104 p.
- BARROS, R. P.; MENDONÇA, R. S. P. **Os determinantes da desigualdade no Brasil**. Texto para Discussão n.º 377. Rio de Janeiro: IPEA, 1995. 52 p.
- BECKER, S. O.; ICHINO, A. Estimation of average treatment effects based on propensity scores. **The Stata Journal**, 2, Number 4, pp. 358–377, 2002.
- CARNEIRO, M. J. Pluriatividade: uma resposta à crise da exploração familiar?. In: ENCONTRO NACIONAL DO PIPSA (PROJETO DE INTERCÂMBIO DE PESQUISA SOCIAL EM AGRICULTURA), 17, 1995, Porto Alegre, RS, **Anais do XVII PISA**. Porto Alegre, RS. 1995. p. 50-58.
- CODES, A. L. M. de. **A trajetória do pensamento científico sobre a pobreza: em direção a uma visão complexa**. Texto para Discussão n.º 1332. Brasília: IPEA, 2008. 30 p.
- DEL GROSSI, M.E. **Evolução das ocupações não-agrícolas no meio rural brasileiro, 1981- 95**. 221f. Tese (Doutorado em Ciências Econômicas) - Instituto de Economia, Universidade Estadual de Campinas, São Paulo, 1999.
- DUARTE, G. B.; SAMPAIO, B.; SAMPAIO, Y. Programa Bolsa Família: impacto das transferências sobre os gastos com alimentos em famílias rurais. **RESR**, Piracicaba, SP, vol. 47, n.º 04, p. 903-918, out/dez 2009.
- FOSTER, J.; GREER, J.; THORBECKE, E. A Class of Decomposable Poverty Measures. **Econometrica**, v. 52, n.3, May 1984, p. 761-766.
- FULLER, A. M. “From part-time farming to pluriativity: a decade of change in rural Europe”. **Journal of Rural Studies**, Londres: v. 6, n. 4, 1990. p. 361-373
- GASSON, R. Use of sociology in agricultural economics. **Journal of Agricultural Economics**, Ashford, v.22, 1974. p.29-38.
- GRAZIANO DA SILVA, J. O novo rural brasileiro. **Nova economia**, Belo horizonte, maio de 1997, p. 43-81.
- GREENE, W. H. **Econometric Analysis**. 7th. New York: Pearson, 2008. 1178 p.
- GTDN. **Uma Política De Desenvolvimento Para O Nordeste**. 2ª Edição. Recife: SUDENE, 1967. 91 p.

HOFFMANN, R. **Distribuição de renda: medidas de desigualdade e pobreza.** São Paulo: EDUSP, 1998. 204 p.

HOFFMANN, R.; NEY, M. G. Atividades não-agrícolas e desigualdade no meio rural brasileiro. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE ECONOMIA, ADMINISTRAÇÃO E SOCIOLOGIA RURAL (SOBER). 45, 2007, Londrina, PR. **Anais do XLV Congresso da Sober.** Londrina: UEL, 2007. p. 40-58.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Micro dados Da Pesquisa Por Amostra De Domicílios PNAD de 2013.**

KAGEYAMA, A. **Pluriatividade e Ruralidade: aspectos metodológicos.** Economia Aplicada. São Paulo: FIPE/FEA-USP, V.2, N.3, jul-set 1998, p.515-551.

LIMA, J. R. F. de. **Efeitos da pluriatividade e rendas não-agrícola sobre a pobreza e desigualdade rural na região Nordeste.** 157f. Tese (Doutorado em Economia Aplicada) - Departamento de Economia Rural, Universidade Federal de Viçosa, Minas Gerais, 2008.

LIMA, J. R. F. de.; SOUZA, S. F. de; SILVA, A. G. da. **A evolução da pobreza nas famílias rurais da região Nordeste: 2003-2009.** Teoria e Evidência Econômica, Passo Fundo, v. 17, n. 36, jan./jun. 2011. p. 80-97.

MARIANO, J. L.; NEDER H. D. **Desigualdade de renda e pobreza entre Famílias no meio rural do nordeste.** Economia e desenvolvimento, Recife-Pe, v. 5, n. 2, 2006. p. 221-242.

MARIANO, J. L.; NEDER, H. D. Renda e Pobreza entre Famílias no meio Rural do Nordeste. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE ECONOMIA E SOCIOLOGIA RURAL, 42. Cuiabá, MT. **Anais do XLII Congresso da Sober.** Cuiabá, Dinâmicas Setoriais e Desenvolvimento Regional. 2004. p 28-45.

MATTEI, L. **Pluriatividade e Desenvolvimento Rural no Estado de Santa Catarina.** 211f. Tese (Doutorado em Economia Aplicada) - Instituto de Economia, Universidade Estadual de Campinas, Campinas, 1999.

NASCIMENTO, C. A. do. **Pluriatividade, pobreza rural e políticas públicas: Uma análise comparada entre Brasil e União Européia.** Fortaleza: Banco do Nordeste. 2008. (BNB Teses e Dissertações, 11). 282 p.

NASCIMENTO, C. A., CARDOZO, S. A. Redes urbanas regionais e a

pluriatividade das famílias rurais no Nordeste e no Sul do Brasil, 1992-1999 e 2001-2005. **Revista Econômica do Nordeste**, Fortaleza, v.38, n. 34, out-dez, 2007. p. 637-658.

OFICINA INTERNACIONAL DEL TRABAJO (OIT). Por que la agricultura sigue siendo importante. **Empleo en el Mundo 2004-2005: empleo, productividad y reducción de la pobreza**. Ginebra: OIT, 2005. p. 135-193.

ROCHA, S. **A pobreza no nordeste: a década de 1990 vista de perto**. Revista Econômica do Nordeste, Fortaleza, v. 34, n. 1, jan./mar. 2003. p. 7-41.

ROCHA, S. **Pobreza no Brasil: afinal, do que se trata?** 3ª edição. Rio de Janeiro: editora FGV, 2006. 190 p.

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.

SCHNEIDER, S. **A pluriatividade no meio rural brasileiro: características e perspectivas para investigação**. GRAMMONT, 1ª ed. Quito/Equador: Ed. Flacso – Serie FORO, v. 1, 2009. p. 132-161.

SCHNEIDER, S.; FIALHO, M. A. V. **Pobreza rural, desequilíbrios regionais e desenvolvimento agrário no Rio Grande do Sul**. Teoria e Evidência Econômica, Passo Fundo - RS, v. 8, n. 15, 2000. p. 117-149.

SEN, A. **Desenvolvimento como liberdade**. São Paulo, Companhia das Letras. 2010. 461 p

VEIGA, J. E. da. **Pobreza Rural, Distribuição da Riqueza e Crescimento: a experiência brasileira**. In: Distribuição de Riqueza e Crescimento Econômico, Ministério do Desenvolvimento Agrário, Edson Teófilo (org.) et al. NEAD - Núcleo de Estudos Agrários e Desenvolvimento Rural, Conselho Nacional de Desenvolvimento Rural Sustentável, Brasília - DF: 2000, p. 173-200.

WORD BANK. **Pro-poor growth in the 1990s: lessons and insights from 14 countries**. Washington: DC, 2005. 104 p.

# MODELOS ECONOMÉTRICOS PARA PREVISÃO DA INFLAÇÃO NA REGIÃO METROPOLITANA DE FORTALEZA

Elano Ferreira Arruda\*

Antônio Clécio de Brito\*\*

## Resumo

A inflação é um importante indicador de bem-estar para a sociedade. Os processos inflacionários representam corrosão do poder de compra e aumento da variabilidade dos preços relativos, além de prejudicar o horizonte de previsão dos agentes econômicos. Apesar disso, poucos estudos se ocupam em selecionar mecanismos preditivos eficientes para esse indicador numa perspectiva estadual. Portanto, para avaliar e selecionar um mecanismo para prever a inflação da região metropolitana de Fortaleza, o presente estudo compara previsões para a inflação mensal dessa região, no ano de 2014, geradas a partir de técnicas Naives, de modelos ARMA, VAR e da curva de Phillips num contexto estadual. Os resultados indicam que a curva de Phillips se mostrou o melhor mecanismo para prever esse indicador.

**Palavras-Chave:** Previsão de Inflação, Curva de Phillips, Séries Temporais.

## Abstract

Inflation is an important indicator of well-being for society. Inflationary processes represent the corrosion of purchasing power and increased variability of relative prices, in addition to harming the forecast horizon of economic agents. Nevertheless, there have been only a few studies discussing the selection of efficient predictive mechanisms for this indicator in a state perspective. Therefore, in order to evaluate and select a better mechanism for inflation in the metropolitan area of Fortaleza, this study compares forecasts for monthly

\* Doutor em Economia. Professor do Departamento de Economia Aplicada e Pesquisador do Programa de Pós-Graduação em Economia da Universidade Federal do Ceará. Endereço: Av. da Universidade, 2700. 2º Andar. Benfica – Fortaleza/Ceará. CEP: 60020181. Telefone: (85) 3366.7751. E-mail: elano@ufc.br.

\*\* Bacharel em Economia. Faculdade de Economia, Administração, Atuária, Contabilidade e Secretariado – FEAACS, Universidade Federal do Ceará. Endereço: Av. da Universidade, 2700. 2º Andar. Benfica – Fortaleza/Ceará. CEP: 60020181. Telefone: (85) 3366.7751. E-mail: cleciobritoufc@gmail.com.

inflation in this region, in 2014, generated from Naives techniques, and ARMA and VAR models, as well as the Phillips curve in a state context. The results indicate that the Phillips curve show the best econometric mechanisms to predict this indicator..

**Key words:** Inflation Forecast, Phillips Curve, Time Series.

**JEL:**E31, E32, E37.

## 1 INTRODUÇÃO

A Inflação é uma das principais variáveis econômicas no contexto de formulação de política monetária, além de ser um importante indicador das condições de oferta da economia e de bem-estar para a sociedade. Processos inflacionários representam corrosão do poder de compra dos indivíduos, aumento da variabilidade dos preços relativos, além de prejudicar o horizonte de previsão dos agentes econômicos.

Na tentativa de modelar e prever o comportamento dessa variável, destaca-se a curva de Phillips, que preconiza uma relação inversa entre inflação e o desemprego ou, ainda, uma relação direta entre inflação e a atividade econômica. Empiricamente, essa relação tem sido utilizada para descrever o lado da oferta da economia em modelos macroeconômicos (CALVO, 1983; TAYLOR, 1980), para medir o impacto dos ciclos econômicos na inflação e o seu grau de persistência ou inércia (MENDONÇA, SACHSIDA e MEDRANO, 2012; SACHSIDA, 2013; ARRUDA, OLIVEIRA e CASTELAR, 2015) e para realizar previsões (STOCK E WATSON, 1999; ARRUDA, FERREIRA e CASTELAR, 2011).

No que se refere a utilização de modelos da curva de Phillips para previsões de inflação, Stock e Watson (1996 e 1999) realizam dois estudos pioneiros onde empregam vários modelos empíricos que envolvem taxa de juros, moeda e preços das mercadorias para prever a inflação americana, demonstrando que a curva de Phillips é um mecanismo robusto para gerar previsões para essa variável.

Para o Brasil, Arruda, Ferreira e Castelar (2011) realizam um minucioso estudo na tentativa de selecionar um melhor mecanismo preditivo para a inflação, utilizando vários modelos lineares e não lineares de séries temporais e da curva de Phillips. Os autores encontram que os modelos de séries temporais ARMA, VAR apresentam bom desempenho na previsão de inflação, no entan-

to, a curva de Phillips se mostrou o melhor mecanismo para gerar prognósticos futuros sobre essa variável, melhorando, inclusive, os resultados da previsão por janelas rolantes do relatório FOCUS do Banco Central<sup>1</sup>.

Portanto, apesar da importância da inflação como indicador das condições de oferta, de bem-estar e de corrosão do poder de compra dos agentes, existem poucos estudos que se ocupam a selecionar mecanismos preditivos eficientes para esse indicador, numa perspectiva estadual. Nesse sentido, esse trabalho se propõe a contribuir nessa linha, em um estudo de caso para a inflação da região metropolitana de Fortaleza.

Assim, o presente estudo tem como objetivo avaliar e selecionar um mecanismo robusto para gerar previsões de inflação para a região metropolitana de Fortaleza, a partir da utilização de técnicas *Naives*, de modelos de séries temporais como ARMA e VAR e de modelos da curva de Phillips, num contexto estadual. Vale destacar que, como afirma Patterson (2001), se o objetivo da análise econométrica é realizar previsões, os modelos não devem ser selecionados simplesmente por seu bom ajustamento e sim por sua eficiência preditiva; isto é, o melhor modelo é aquele que gera previsões com menor erro de previsão. Nesse estudo, os mecanismos preditivos serão selecionados pelo critério do erro quadrático médio de previsão (EQM) e, adicionalmente, utilizar-se-á o teste de Diebold e Mariano (1995) para examinar se a diferença no desempenho preditivo dos melhores modelos é estatisticamente significativa.

Além dessa introdução, esse trabalho possui mais quatro seções. Na seção dois, destacam-se os aspectos teóricos referentes à curva de Phillips e sua utilização junto com modelos de séries temporais para a realização de previsões para a inflação. Na seção seguinte, apresentam-se os arcações das técnicas e modelos utilizados no trabalho, além de uma análise e descrição do banco de dados. A seção quatro reserva-se a apresentação e análise dos resultados e, por fim, são tecidas as considerações finais do estudo.

## 2 ASPECTOS TEÓRICOS

### 2.1 A Curva de Phillips

A curva de Phillips inicialmente foi formulada como um *trade-off* entre inflação de salários e desemprego, ou simplesmente expressava uma curva de oferta agregada positivamente inclinada. Segundo Phillips (1958), se o desem-

1. As previsões são atualizadas a cada semana de modo a possibilitar os agentes econômicos à inclusão de novas informações relevantes para a construção de suas previsões; ou seja, é um método onde os agentes possuem um conjunto de informação que é atualizado semanalmente, ao contrário dos métodos econométricos, onde se estima um modelo para uma amostra e realizam-se previsões a partir daí.

prego fosse elevado haveria um excesso de oferta de mão de obra na economia, pressionando os salários para baixo. Assim, na medida em que os processos inflacionários fossem maiores, os salários reais seriam menores e, consequentemente, as firmas seriam motivadas a contratar mais mão de obra. Em outras palavras, existiria um *trade-off* entre inflação e desemprego; pelo qual, quanto maior o desemprego, menor seria a inflação e vice-versa.

Todavia, o cenário de estagflação dos anos 80 nos Estados Unidos levou à seguidas críticas quanto ao ajustamento das políticas de controle de demanda para lidar com a inflação e desemprego, mudando a percepção de que a curva de Phillips fosse, de fato, uma regra consistente e infalível em que mais inflação sempre estivesse associada a menos desemprego. Autores como Phelps (1967) e Friedman (1968) já haviam demonstrado, antes desses eventos, uma insatisfação com uma modelagem em que variáveis nominais fossem afetadas por variáveis reais e propuseram a inclusão de um componente que captasse a taxa esperada de inflação, afim de melhorar a especificação da curva de Phillips. Sob essas novas condições, a curva de Phillips negativamente inclinada passou a ser considerada apenas como uma relação de curto prazo, uma vez que, em longo prazo, essa relação seria vertical.

Atualmente, a curva de Phillips tem sido especificada em função de um indicador de ciclos econômicos (como o hiato do produto), ou de custo marginal das empresas; de uma variável para expectativas (*forward-looking* – expectativas racionais; ou *backward-looking* – expectativas adaptativas, ou ambas); além de um indicador de repasse cambial para a inflação.

Portanto, no presente estudo, que tem como objetivo selecionar um mecanismo para previsão da inflação,  $\pi_t$ , da região metropolitana de Fortaleza, utilizar-se-á a especificação da curva de Phillips presente em Arruda, Ferreira e Castelar (2011); ou seja, uma curva de Phillips com uma medida de ciclos econômicos (dada pelo hiato do produto defasado,  $h_{t-p}$ ), uma variável de expectativas adaptativas, ou inflação inercial (inflação defasada,  $\pi_{t-p}$ ) e, por fim, uma variável de repasse cambial (dada pela soma da oscilação cambial doméstica com a inflação ao produtor americana defasada ( $\Delta e_{t-p} + \pi_{t-p}^*$ ), nos moldes de Correa e Minella (2010)). O modelo tem a especificação a seguir:

$$\pi_t = \beta_1 \pi_{t-p} + \beta_2 h_{t-p} + \beta_3 (\Delta e_{t-p} + \pi_{t-p}^*) + \varepsilon_t \quad (1)$$

## 2.2. Previsão de inflação

A elaboração de previsões sobre a trajetória futura de agregados econômi-



cos; em especial, sobre taxa de inflação, constitui uma das ocupações principais dos formuladores de política. Essa análise possui papel central na tomada de decisões dos agentes econômicos, dado que os processos inflacionários corroem o poder de compra das famílias, aumentam a dispersão dos preços relativos, diminuindo o grau de integração dos mercados, e, por fim, afeta gravemente o horizonte de planejamento dos agentes econômicos.

A literatura sobre previsão de inflação apresentou grandes avanços nos últimos anos, sugerindo novas técnicas que podem ser implementadas ao seu processo preditivo. Os principais mecanismos utilizados são técnicas *Naives*, modelagens de séries temporais e modelos da curva de Phillips.

Ascari e Marrocu (2003) realizam um exercício para previsão da taxa de inflação americana utilizando diferentes modelos de séries temporais e modelos baseados na curva de Phillips, com dados mensais para o período de 1950.1 – 2007.7. Esses autores concluíram que as melhores previsões realizadas foram obtidas através da estimação de modelos da curva de Phillips.

Para Suíça, Lack (2006) destaca a eficiência preditiva do vetor auto regressivo (VAR) para previsão da taxa de inflação, mostrando que a previsibilidade desse modelo para horizontes longos é mais robusta do que a de outros modelos univariados de séries temporais. Resultados na mesma linha foram obtidos por Rumler e Valderrama (2008) para dados da economia austríaca. Ou seja, os modelos vetoriais autoregressivos e os modelos da curva de Phillips apresentam forte eficiência preditiva para a inflação.

Vale destacar que, como argumenta Mishkin (2007), caso a curva de Phillips esteja mais horizontal, há indícios de forte persistência inflacionária e, portanto, maior volatilidade. Em importante estudo para a economia americana, Stock e Watson (2007) mostram que, em cenários de baixa volatilidade da inflação, os modelos econométricos de séries temporais mais sofisticados não conseguem melhorar simples previsões *Naives*. Resultados nessa mesma linha foram encontrados por Cechetti *et al.* (2007), que realizam o mesmo exercício para a dinâmica inflacionária dos países do G7<sup>2</sup> nos últimos trinta anos.

Para a economia brasileira, Alencar (2006) comparou a eficiência preditiva de curto prazo de modelos puramente estatísticos com a daqueles amparados pela teoria econômica. Os resultados indicam que modelos de séries temporais ARMA e VAR apresentam bom desempenho preditivo frente aos modelos da curva de Phillips.

---

2. Estados Unidos; Inglaterra; França; Japão; Itália; Alemanha e Canadá

Ainda para o Brasil, Arruda, Ferreira e Castelar (2011), adotam modelos lineares e não lineares de séries temporais e da curva de Phillips para fins de previsão da inflação brasileira, utilizando como *benchmark* o modelo autoregressivo de ordem 1, AR (1). Os autores encontram que uma curva de Phillips ampliada com especificação não linear é a que apresenta o melhor desempenho preditivo para a inflação do Brasil, melhorando inclusive as previsões de mercado do boletim FOCUS do banco central.

Portanto, observa-se que não há um consenso sobre qual o melhor mecanismo para a geração de previsões para a inflação. Há autores que argumentam em favor de modelos univariados de séries temporais, outros advogam em favor de modelos VAR e, ainda, há os que apresentam a curva de Phillips como um bom mecanismo gerador de previsões para a inflação.

Diante do exposto e da ausência de estudos dessa natureza com dados estaduais, o presente trabalho pretende contribuir nessa direção com a aplicação de técnicas *Naives*, de modelos de séries temporais e da curva de Phillips para gerar previsões para a inflação da região metropolitana de Fortaleza, seguindo principalmente a linha de Stock e Watson (2007) e de Arruda, Ferreira e Castelar (2011).

### 3 METODOLOGIA E ANÁLISE DOS DADOS

O presente trabalho utiliza como indicador de inflação, a variação mensal do Índice Nacional de Preços ao Consumidor Amplo (IPCA) da região metropolitana de Fortaleza, calculado e disponibilizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), entre janeiro de 1995 e dezembro de 2014.

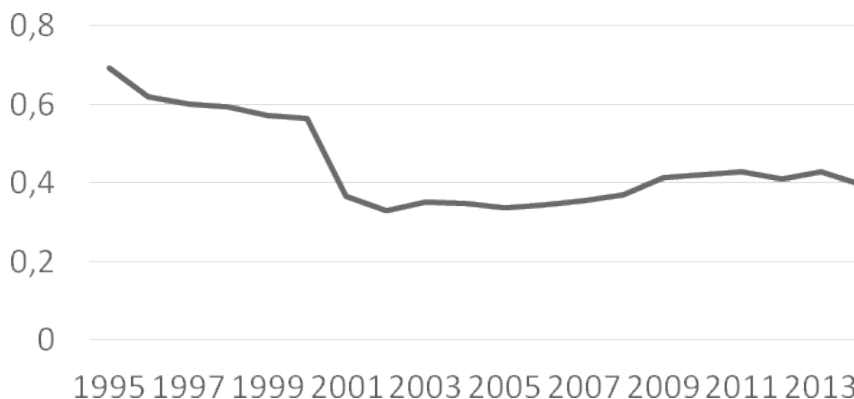
#### 3.1 Modelos Naive

Inicialmente, nos moldes sugeridos por Stock e Watson (2007), procedeu-se uma análise evolutiva da volatilidade da inflação da região metropolitana de Fortaleza, medida pelo seu desvio-padrão, entre 1995 e 2014. O gráfico 1 sintetiza o comportamento da volatilidade dessa variável.

A análise do gráfico revela uma forte redução da volatilidade da inflação da região metropolitana de Fortaleza entre 1995 e 2002, na ordem de 74,28% (com redução anual média de 9,3% a.a.), mantendo-se praticamente constante nos anos seguintes até apresentar um leve aumento a partir de 2005. Em termos

gerais, observa-se que a volatilidade da inflação de Fortaleza nesse período vem caindo, em média, 2,8% a.a.

### GRÁFICO 1: Evolução da volatilidade do IPCA da região metropolitana de Fortaleza



Fonte: Elaboração própria a partir dos resultados obtidos.

Assim, seguindo a linha de Stock e Watson (2007) e diante da observação de uma redução na volatilidade da inflação, o presente estudo fará uso de três indicadores *Naives*, quais sejam: i) a inflação do mês de previsão ocorrida no ano anterior (*Naive 1*); ii) a média das inflações do mês de previsão ocorridas nos dois anos anteriores (*Naive 2*) e iii) a média das inflações do mês de previsão ocorridas nos três anos anteriores (*Naive 3*). O objetivo dessa análise é comparar essas previsões ingênuas com as demais, de modo a observar o seu desempenho frente a outros modelos de séries temporais em cenários de volatilidade reduzida.

### 3.2 Modelos de previsão ARMA

Os trabalhos de previsão econômica dominantes durante a década de 1960 e 1970 eram baseados em modelos de equações simultâneas, realidade esta que mudaria com os choques do petróleo vividos durante a década de 1970 e com a crítica de Lucas (1972) que apontava que os parâmetros estimados em um modelo econométrico dependiam da política econômica vigente, não sendo, portanto, invariantes na presença de mudanças políticas.

A metodologia de Box e Jenkins (1978), por sua vez, abriu caminho para uma nova geração de modelos de previsão. Com o objetivo de estimar um modelo estatístico que pudesse ser interpretado como tendo sido gerado pelos

dados amostrais, essa metodologia permitiu que valores futuros de uma série fossem previstos tomando por base seus valores presentes e passados, substituindo a ênfase dada a construção de modelos de equações simultâneas, pela análise das propriedades estocásticas em uma série temporal. Em virtude da sua simplicidade e da qualidade de suas previsões, esses modelos univariados são usualmente utilizados como *benchmarks* em estudos sobre previsão (STOCK e WATSON, 1999 e 2007; ATKENSON e OHANION, 2001; ARRUDA, FERREIRA e CASTELAR, 2011).

Nos modelos ARMA ( $p, q$ ) o índice  $p$ , refere-se ao número de defasagens da própria variável dependente ( $y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p}$ ), e  $q$ , refere-se ao número de defasagens dos choques exógenos ou termos de erro ( $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-p}$ ). Segundo Hamilton (1994), um modelo ARMA ( $p, q$ ) tem a seguinte forma:

$$y_t = \mu + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2)$$

Em que,  $\varepsilon_t$  é  $iid \sim N(0, \sigma^2)$ . Se o modelo tiver apenas o componente autoregressivo, ou seja:  $\theta_1 = \theta_2 = \dots = \theta_p = 0$ , ele pode ser estimado por mínimos quadrados ordinários. Caso o modelo contenha o componente de médio móvel não é possível escrever a soma dos quadrados dos erros simplesmente como função dos valores de  $y$  observados e dos parâmetros do modelo autoregressivo e o método de estimação a ser utilizado é o da máxima verossimilhança. No presente estudo, para gerar previsões para o IPCA da região metropolitana de Fortaleza, foram utilizadas até seis defasagens nos modelos ARMA ( $p, q$ ).

### 3.3 Modelos de previsão VAR

Os modelos econômicos em geral são expressos por meio de diversas variáveis. Portanto, o uso de modelos univariados, como os dos modelos ARMA, podem ser limitados para expressar modelos econômicos. Dessa forma, Stock e Watson (2004) argumentam que, em análises econômicas que envolvam previsão de inflação, deve ser considerada a possibilidade de que outras variáveis econômicas o crescimento da oferta de crédito, a taxa de crescimento do PIB e as taxas de juros possam conter informações adicionais relevantes sobre o futuro da inflação. Nesse sentido, os modelos vetoriais autoregressivos (VAR), propostos por Sims (1980 e 1986), ganharam força a partir da década de 1980.

Como afirma Caiado (2002), modelos VAR podem ser utilizados em

alternativa ao método de modelos ARMA, uma vez que levam em consideração a existência de relações interdependentes entre as variáveis e permitem avaliar o impacto dinâmico das perturbações aleatórias sobre o sistema de variáveis, o que os tornam particularmente úteis e eficientes na previsão do comportamento futuro de séries temporais inter-relacionadas. Por essas razões, torna-se importante investigar se modelos multivariados de séries temporais tendem a gerar previsibilidade mais consistente do que modelos univariados.

Segundo Hamilton (1994)<sup>3</sup>, um vetor auto regressivo de ordem p (VAR (p)) é a forma reduzida de um modelo dinâmico de equações simultâneas e pode ser expresso como:

$$y_t = c + \sum_{i=1}^p \Phi_i y_{t-i} + \varepsilon_t \quad (3)$$

Onde,  $y_t$  representa um vetor (n x 1) com valores das n variáveis incluídas no modelo e  $\varepsilon_t$  é  $iid \sim N(0, \Sigma)$ .

Neste trabalho, serão empregadas até seis defasagens de um modelo VAR nos moldes tradicionais para análise de política monetária, qual seja, um modelo VAR com as variáveis: i) inflação da região metropolitana de Fortaleza medida pelo IPCA ( $\Pi_t$ ), variável a ser prevista; ii) a taxa de crescimento do Índice de Produção Industrial – IPI – do Estado do Ceará ( $\Delta IPI_t$ ), calculado e disponibilizado pelo IBGE, como *proxy* para o crescimento mensal do PIB do estado; iii) a taxa de juros mensal - Over/Selic ( $SELIC_t$ ), disponibilizada pelo Banco Central do Brasil (BACEN) e iv) o agregado monetário, M1, em taxa de crescimento, disponibilizado pelo BACEN ( $\Delta M1_t$ ), devidamente deflacionado pelo índice geral de preços – disponibilidade interna – IGP-DI – disponibilizado pela Fundação Getúlio Vargas.

Vale ressaltar que os testes de raiz unitária de Dickey-Fuller aumentado (ADF) e o de Kwiatkowski, Phillips, Schmidt e Shin (KPSS) foram aplicados a essas séries e os resultados mostram que as variáveis inflação e Selic são estacionárias em nível e os indicadores IPI e M1 são estacionárias em primeira diferença, conforme tabela 1.

3. Este autor pode ser consultado para mais detalhes sobre o processo de estimação e previsão dos modelos ARMA e VAR.

TABELA 1: Resultados dos Testes de Raiz Unitária modelos VAR

VARIÁVEL	ADF	KPSS
$\Pi_t$	-8,26 (-2,87)	0,23 (0,46)
$\Delta M1_t$	-3,37 (-2,87)	0,04 (0,46)
$\Delta IPI_t$	-4,65 (-2,87)	0,07 (0,46)
$SELIC_t$	-3,06 (-2,87)	0,15* (0,21)

Fonte: Elaboração própria a partir dos resultados obtidos. Valor crítico do teste a 5% de significância entre parêntesis. \*Teste aplicado com tendência e intercepto; não se pode rejeitar a hipótese nula de estacionaridade a 1% de significância.

### 3.4 Modelos da curva de Phillips

Autores como Stock e Watson (1996,1999,2007), Atkenson e Ohanian (2001), Ascari e Marrocu (2003) e Rumler e Valderrama (2008) e Arruda, Ferreira e Castelar (2011) utilizam diversas especificações da curva de Phillips para gerar previsões da taxa de Inflação.

O presente estudo fará uso de até seis defasagens de uma versão da curva de Phillips que incorpore uma medida de ciclos econômicos, uma medida de expectativas adaptativas ou inflação inercial e, ainda, uma variável para captar o repasse cambial para a inflação. O modelo empírico da curva de Phillips aqui empregado tem a forma:

$$\pi_t = \beta_1 \pi_{t-p} + \beta_2 h_{t-p} + \beta_3 (\Delta e_{t-p} + \pi_p^*) + \varepsilon_t \quad (4)$$

Em que  $\pi_t$  é a inflação da região metropolitana de Fortaleza,  $\pi_{t-p}$  é um termo de expectativas adaptativas ou inflação inercial,  $h_{t-p}$  é uma medida de ciclos econômicos (utilizou-se como *proxy* para o hiato do produto o hiato da produção industrial do Estado do Ceará, calculado a partir da aplicação do filtro de Hodrick – Prescott) e  $(\Delta e_{t-p} + \pi_p^*)$  é uma medida de repasse cambial em que  $e_t$  é o logaritmo da taxa de câmbio nominal,  $\Delta$  é o operador de primeira diferença e  $\pi_t^*$  é uma medida inflação internacional (inflação ao produtor americano – PPI americano, nos moldes de Correa e Minella (2010) e Arruda, Ferreira e Castelar (2011)). Todas essas variáveis também se mostraram estacionárias pelos testes ADF e KPSS, conforme tabela 2.

TABELA 2: Resultados dos Testes de Raiz Unitária modelos da Curva de Phillips

VARIÁVEL	ADF	KPSS
$\Pi_t$	-8,26 (-2,87)	0,23 (0,46)
$h_t$	-6,46 (-2,87)	0,01 (0,46)
$(\Delta e_{t-p} + \pi_p^*)$	-16,53 (-2,87)	0,20 (0,46)

Fonte: Elaboração própria a partir dos resultados obtidos. Valor crítico dos testes a 5% de significância entre parêntesis.

### 3.5 – Avaliando as previsões

Ao avaliar a eficiência preditiva de diferentes modelos de series temporais e da curva de Phillips, deve-se levar em consideração um critério de decisão, afim de julgar o quão preciso é o prognóstico realizado por determinado modelo. Como afirmam Pyndyck e Rubinfeld (2004), se o objetivo é prever valores futuros de uma série temporal que estejam sujeitos ao menor erro possível, um dos critérios mais adequados para se comparar previsões é chamado erro quadrado médio de previsão (EQM), dado por:

$$EQM = \frac{\sum_{t=T+h}^{T+h} (y_t^F - y_t)^2}{h} \quad (5)$$

Além de utilizar o EQM como medida de eficiência preditiva e de seleção dos melhores modelos, o presente trabalho utilizou-se o procedimento descrito por West (2006) e também empregado em Arruda, Ferreira e Castelar (2011) ilustrado pelo diagrama a seguir:

FIGURA 1: Diagrama de Previsão



Fonte: West (2006).

Ou seja, a amostra total empregada nesse estudo corresponde a informações mensais entre janeiro de 1995 e dezembro de 2014 (T+1 no diagrama). Todos os modelos serão estimados utilizando dados de janeiro de 1995 a de-

zembro de 2013 (R observações no diagrama). E, por fim, serão geradas previsões para o período de janeiro a dezembro de 2014 (P resultados de previsão no diagrama).

Após a geração das previsões calculam-se os EQM de cada modelo e selecionam-se aqueles que geraram previsões mais eficazes; ou seja, com menores EQM de previsão. Após a seleção dos melhores modelos, utilizar-se-á ainda do teste de Diebold e Mariano (1995) para comparação de previsões. Esse teste consiste em uma análise das perdas associadas a cada uma das previsões a partir de um critério métrico estatístico para avaliar a igualdade de seus desempenhos preditivos. Portanto, a hipótese nula do teste é a de que os modelos possuem eficiência preditiva semelhante.

#### 4 ANÁLISE DOS RESULTADOS

Para selecionar mecanismos preditivos para a inflação da região metropolitana de Fortaleza, foram utilizados dados mensais de janeiro de 1995 a dezembro de 2013 para a estimação de modelos ARMA, VAR e da curva de Phillips. As previsões foram geradas para o ano de 2014, ou seja, o período de janeiro de 2014 a dezembro de 2014.

Os modelos foram selecionados pela eficiência preditiva, ou seja, pelo menor EQM de previsão. Nos moldes de Stock e Watson (2007) e Arruda, Ferreira e Castelar (2011) foram utilizados como *benchmarks*, a melhor previsão *Naive* e o modelo autoregressivo de ordem 1, AR (1). A tabela 3 apresenta uma síntese desses resultados e, portanto, serão selecionados como *benchmarks* o modelo AR (1) e a previsão *Naive* (3) que apresentaram EQM de previsão da ordem de 10% e 9,9%, respectivamente.

TABELA 3: Resultados dos Benchmarks

Previsão	IPCA	AR (1)	NAIVE (1)	NAIVE (2)	NAIVE (3)
2014.1	<b>0,4500</b>	<b>0,6627</b>	1,0100	0,5400	<b>0,5800</b>
2014.2	<b>0,1900</b>	<b>0,4918</b>	0,7200	0,4400	<b>0,3633</b>
2014.3	<b>0,7000</b>	<b>0,3437</b>	0,6100	0,7100	<b>0,9700</b>
2014.4	<b>1,0800</b>	<b>0,6343</b>	0,7800	0,7300	<b>0,6967</b>
2014.5	<b>0,9500</b>	<b>0,8507</b>	0,3200	0,3400	<b>0,3200</b>
2014.6	<b>0,3500</b>	<b>0,7767</b>	0,0900	-0,0900	<b>0,0167</b>
2014.7	<b>-0,1700</b>	<b>0,4349</b>	0,1900	0,3700	<b>0,3500</b>
2014.8	<b>0,0700</b>	<b>0,1386</b>	-0,1100	0,2800	<b>0,3067</b>
2014.9	<b>0,4500</b>	<b>0,2754</b>	0,4100	0,5400	<b>0,4833</b>
2014.10	<b>0,3800</b>	<b>0,4918</b>	0,4400	0,6500	<b>0,5867</b>
2014.11	<b>0,8100</b>	<b>0,4520</b>	0,9900	0,8600	<b>0,8167</b>
2014.12	<b>0,6300</b>	<b>0,6969</b>	0,7500	1,0100	<b>0,8400</b>
<b>Resultado</b>	<b>EQM</b>	<b>0,1000</b>	0,1143	0,1097	<b>0,0990</b>



Fonte: Elaboração própria a partir dos resultados obtidos

#### 4.1 Resultados dos Modelos Autoregressivos (AR)

Para iniciar a avaliação do desempenho preditivo dos modelos aqui tratados, analisou-se o desempenho dos modelos autoregressivos. A estratégia consistiu em estimar processos autoregressivos até a sexta ordem, gerar previsões e calcular o EQM. Os resultados estão sintetizados na tabela 4.

Os resultados mostram que o modelo AR (4) apresentou o melhor desempenho entre os modelos autoregressivos, com um EQM de previsão na ordem de 8,7%; ou seja, um resultado 13% melhor do que o do AR (1) e 3,3% menor do que o gerado pela previsão *Naïve* (3). Portanto, entre os modelos autoregressivos, selecionar-se-á o AR (4) para comparação com os demais modelos.

TABELA 4: Resultados dos modelos Autoregressivos (AR)

Previsão	IPCA	AR (1)	NAIVE (3)	AR (2)	AR (3)	AR (4)	AR (5)	AR (6)
2014.1	<b>0,4500</b>	<b>0,6627</b>	<b>0,5800</b>	0,6609	0,6029	<b>0,6043</b>	0,5984	0,6054
2014.2	<b>0,1900</b>	<b>0,4918</b>	<b>0,3633</b>	0,4851	0,5013	<b>0,4610</b>	0,4837	0,5012
2014.3	<b>0,7000</b>	<b>0,3437</b>	<b>0,9700</b>	0,3361	0,3499	<b>0,3807</b>	0,3801	0,3946
2014.4	<b>1,0800</b>	<b>0,6343</b>	<b>0,6967</b>	0,6593	0,6867	<b>0,6964</b>	0,6746	0,6768
2014.5	<b>0,9500</b>	<b>0,8507</b>	<b>0,3200</b>	0,8745	0,8324	<b>0,8402</b>	0,8015	0,7785
2014.6	<b>0,3500</b>	<b>0,7767</b>	<b>0,0167</b>	0,7807	0,7477	<b>0,7041</b>	0,7028	0,6827
2014.7	<b>-0,1700</b>	<b>0,4349</b>	<b>0,3500</b>	0,4163	0,4171	<b>0,4047</b>	0,4240	0,4289
2014.8	<b>0,0700</b>	<b>0,1386</b>	<b>0,3067</b>	0,1183	0,1517	<b>0,1811</b>	0,1902	0,2187
2014.9	<b>0,4500</b>	<b>0,2754</b>	<b>0,4833</b>	0,2848	0,3184	<b>0,3602</b>	0,3416	0,3483
2014.10	<b>0,3800</b>	<b>0,4918</b>	<b>0,5867</b>	0,5098	0,4752	<b>0,4971</b>	0,4626	0,4451
2014.11	<b>0,8100</b>	<b>0,4520</b>	<b>0,8167</b>	0,4530	0,4006	<b>0,3653</b>	0,3689	0,3461
2014.12	<b>0,6300</b>	<b>0,6969</b>	<b>0,8400</b>	0,7200	0,7308	<b>0,6759</b>	0,7002	0,6880
Resultado	<b>EQM</b>	<b>0,1000</b>	<b>0,0990</b>	0,0965	0,0938	<b>0,0873</b>	0,0922	0,0942

Fonte: Elaboração própria a partir dos resultados obtidos

#### 4.2 Resultados dos Modelos autoregressivos de média móvel (ARMA)

De modo semelhante, nos Modelos ARMA (p, q), foram utilizadas até 6 defasagens para prever a taxa de inflação, sendo  $p = \{1,2,3...6\}$  o número de defasagens utilizadas da variável dependente IPCA utilizada como variável explicativa e  $q = \{1,2,3...6\}$  os componentes de média móvel defasados. Os re-

sultados podem ser observados na tabela 5.

Inicialmente, observa-se que, com exceção do modelo ARMA (5,5), todos os modelos ARMA empregados geraram previsões abaixo dos *benchmarks* adotados, confirmando o bom desempenho desses modelos no processo preditivo. O modelo ARMA (6,6) apresentou o menor EQM de previsão, na ordem de 7,4%, resultado 25,5% melhor do que a previsão do AR (1), 17,2% mais eficiente do que o prognóstico da *Naive* (3) e, ainda, um erro de previsão 14,6% menor do que o gerado pelo melhor modelo autoregressivo, AR (4).

TABELA 5: Resultados dos Modelos ARMA (p,q)

Previsão	IPCA	AR (1)	NAIVE (3)	ARMA (1,1)	ARMA (2,2)	ARMA (3,3)	ARMA (4,4)	ARMA (5,5)	ARMA (6,6)
2014.1	<b>0,4500</b>	<b>0,6627</b>	<b>0,5800</b>	0,6066	0,6219	0,5686	0,5637	0,6760	<b>0,5303</b>
2014.2	<b>0,1900</b>	<b>0,4918</b>	<b>0,3633</b>	0,4719	0,4627	0,4764	0,4575	0,6247	<b>0,4970</b>
2014.3	<b>0,7000</b>	<b>0,3437</b>	<b>0,9700</b>	0,3373	0,3308	0,3422	0,3591	0,5330	<b>0,4642</b>
2014.4	<b>1,0800</b>	<b>0,6343</b>	<b>0,6967</b>	0,6963	0,7010	0,6605	0,6575	0,7250	<b>0,7126</b>
2014.5	<b>0,9500</b>	<b>0,8507</b>	<b>0,3200</b>	0,8544	0,8854	0,7918	0,8180	0,7538	<b>0,7793</b>
2014.6	<b>0,3500</b>	<b>0,7767</b>	<b>0,0167</b>	0,7334	0,7455	0,7051	0,6752	0,6101	<b>0,5882</b>
2014.7	<b>-0,1700</b>	<b>0,4349</b>	<b>0,3500</b>	0,3776	0,3741	0,3962	0,4070	0,4554	<b>0,3744</b>
2014.8	<b>0,0700</b>	<b>0,1386</b>	<b>0,3067</b>	0,1290	0,0976	0,1683	0,1414	0,3468	<b>0,1867</b>
2014.9	<b>0,4500</b>	<b>0,2754</b>	<b>0,4833</b>	0,3420	0,3302	0,3368	0,3704	0,5338	<b>0,4350</b>
2014.10	<b>0,3800</b>	<b>0,4918</b>	<b>0,5867</b>	0,5349	0,5344	0,4890	0,4759	0,5012	<b>0,4632</b>
2014.11	<b>0,8100</b>	<b>0,4520</b>	<b>0,8167</b>	0,4441	0,4417	0,4135	0,4197	0,2811	<b>0,3663</b>
2014.12	<b>0,6300</b>	<b>0,6969</b>	<b>0,8400</b>	0,7414	0,7529	0,6928	0,6930	0,5682	<b>0,5809</b>
Resultado	<b>EQM</b>	<b>0,1000</b>	<b>0,0990</b>	0,0853	0,0859	0,0889	0,0844	0,1061	<b>0,0745</b>

Fonte: Elaboração própria a partir dos resultados obtidos. Foram testadas variações no ARMA(p,q) oscilando os valores de p e q; entretanto, não houve ganhos em termos de desempenho preditivo.

### 4.3 – Resultados dos Modelos Vetoriais Autoregressivos (VAR)

Em moldes semelhantes, foram utilizadas até seis defasagens dos modelos VAR, com as variáveis inflação da região metropolitana de Fortaleza, taxa de crescimento da produção industrial (*proxy* para atividade econômica), taxa de crescimento da oferta monetária e a taxa de juros Over/Selic. Os resultados estão sintetizados na tabela 6.

Em termos gerais, vale destacar o bom desempenho dos modelos VAR para previsão, uma vez que cinco dos modelos empregados melhoraram consideravelmente as previsões oriundas do AR (1) e da *Naive* (3). O melhor desem-

penho preditivo foi observado no modelo VAR (4), que apresentou EQM de previsão da ordem de 7,9%, resultado 21% melhor do que o gerado pelo AR (1) e 20,1% mais eficiente do que o modelo *Naive* (3).

TABELA 6: **Resultados dos Modelos VAR**

Previsão	IPCA	AR (1)	NAIVE (3)	VAR(1)	VAR(2)	VAR(3)	VAR(4)	VAR(5)	VAR(6)
2014.1	<b>0,4500</b>	<b>0,6627</b>	<b>0,5800</b>	0,6703	0,6237	0,6204	<b>0,6647</b>	0,6834	0,6964
2014.2	<b>0,1900</b>	<b>0,4918</b>	<b>0,3633</b>	0,3237	0,3178	0,3185	<b>0,2938</b>	0,3234	0,3335
2014.3	<b>0,7000</b>	<b>0,3437</b>	<b>0,9700</b>	0,2854	0,3306	0,3292	<b>0,3240</b>	0,2591	0,2590
2014.4	<b>1,0800</b>	<b>0,6343</b>	<b>0,6967</b>	0,5159	0,5959	0,6898	<b>0,6337</b>	0,6514	0,6852
2014.5	<b>0,9500</b>	<b>0,8507</b>	<b>0,3200</b>	0,7734	0,8330	0,8123	<b>0,8252</b>	0,7171	0,7202
2014.6	<b>0,3500</b>	<b>0,7767</b>	<b>0,0167</b>	0,7011	0,6722	0,7044	<b>0,6299</b>	0,5294	0,5245
2014.7	<b>-0,1700</b>	<b>0,4349</b>	<b>0,3500</b>	0,3824	0,3498	0,3924	<b>0,3429</b>	0,3175	0,3720
2014.8	<b>0,0700</b>	<b>0,1386</b>	<b>0,3067</b>	0,1158	0,0617	0,0845	<b>0,1147</b>	0,1444	0,1906
2014.9	<b>0,4500</b>	<b>0,2754</b>	<b>0,4833</b>	0,2481	0,2337	0,2856	<b>0,3242</b>	0,3157	0,3775
2014.10	<b>0,3800</b>	<b>0,4918</b>	<b>0,5867</b>	0,4740	0,4738	0,4382	<b>0,4573</b>	0,5132	0,5232
2014.11	<b>0,8100</b>	<b>0,4520</b>	<b>0,8167</b>	0,3958	0,3672	0,3368	<b>0,4058</b>	0,4611	0,4278
2014.12	<b>0,6300</b>	<b>0,6969</b>	<b>0,8400</b>	0,6684	0,7065	0,7151	<b>0,7142</b>	0,7954	0,7942
<b>Resultado</b>	<b>EQM</b>	<b>0,1000</b>	<b>0,0990</b>	0,1034	0,0886	0,0882	<b>0,0791</b>	0,0804	0,0852

Fonte: Elaboração própria a partir dos resultados obtidos.

Além disso, esse desempenho foi 9,4% superior ao do melhor modelo autoregressivo, AR (4), mostrando-se mais eficiente do que a maioria dos modelos ARMA empregados, com exceção apenas do ARMA (6,6), que obteve um erro de previsão ainda 5,8% menor do que o do VAR (4). Resultados nessa direção também foram encontrados por autores como Rumler e Valderrama (2008), Arruda, Ferreira e Castelar (2011) e Lack (2006) que também destacam o bom desempenho dos modelos VAR frente aos processos ARMA para previsão da inflação.

#### 4.4 Modelos da Curva de Phillips

Por fim, foram estimados até seis *lags* dos modelos da curva de Phillips para prover prognósticos sobre a inflação da região metropolitana de Fortaleza. Uma síntese das previsões e dos respectivos EQM de previsão pode ser observada na tabela 7.

TABELA 7: Resultados previsões dos modelos da Curva de Phillips

Previsão	IPCA	AR (1)	NAIVE (3)	CP (1)	CP (2)	CP (3)	CP (4)	CP (5)	CP (6)
2014.1	<b>0,4500</b>	<b>0,6627</b>	<b>0,5800</b>	0,6370	0,7928	0,7130	0,6858	<b>0,6310</b>	<b>0,6601</b>
2014.2	<b>0,1900</b>	<b>0,4918</b>	<b>0,3633</b>	0,5286	0,4946	0,4822	0,4521	<b>0,4306</b>	<b>0,3968</b>
2014.3	<b>0,7000</b>	<b>0,3437</b>	<b>0,9700</b>	0,2908	0,3273	0,3530	0,4271	<b>0,4268</b>	<b>0,4461</b>
2014.4	<b>1,0800</b>	<b>0,6343</b>	<b>0,6967</b>	0,5610	0,5455	0,5595	0,5936	<b>0,6170</b>	<b>0,6088</b>
2014.5	<b>0,9500</b>	<b>0,8507</b>	<b>0,3200</b>	0,7984	0,7497	0,6884	0,7279	<b>0,6963</b>	<b>0,6846</b>
2014.6	<b>0,3500</b>	<b>0,7767</b>	<b>0,0167</b>	0,7424	0,6961	0,6654	0,5612	<b>0,6043</b>	<b>0,5843</b>
2014.7	<b>-0,1700</b>	<b>0,4349</b>	<b>0,3500</b>	0,3489	0,3423	0,3528	0,2981	<b>0,2408</b>	<b>0,2606</b>
2014.8	<b>0,0700</b>	<b>0,1386</b>	<b>0,3067</b>	0,1611	0,0456	0,0861	0,1051	<b>0,0985</b>	<b>0,0842</b>
2014.9	<b>0,4500</b>	<b>0,2754</b>	<b>0,4833</b>	0,2668	0,2793	0,3444	0,3717	<b>0,3562</b>	<b>0,3632</b>
2014.1	<b>0,3800</b>	<b>0,4918</b>	<b>0,5867</b>	0,6082	0,6097	0,5696	0,5677	<b>0,5427</b>	<b>0,5374</b>
2014.11	<b>0,8100</b>	<b>0,4520</b>	<b>0,8167</b>	0,4605	0,5554	0,4964	0,4766	<b>0,4604</b>	<b>0,4596</b>
2014.12	<b>0,6300</b>	<b>0,6969</b>	<b>0,8400</b>	0,7490	0,7776	0,7797	0,7090	<b>0,7221</b>	<b>0,7176</b>
<b>Resultado</b>	<b>EQM</b>	<b>0,1000</b>	<b>0,0990</b>	0,1052	0,1022	0,0963	0,0757	<b>0,0703</b>	<b>0,0705</b>

Fonte: Elaboração própria a partir de resultados obtidos.

Pode-se observar que, as previsões geradas pelos modelos da curva de Phillips, CP (5) e CP (6), apresentaram os menores erros de previsão entre todos os modelos empregados nesse estudo, 7,03% e 7,05%, respectivamente, corroborando que, mesmo para dados estaduais, a curva de Phillips ainda se mostra um mecanismo robusto para prover previsões eficientes para os processos inflacionários.

Em termos comparativos, observa-se que a CP (5) melhorou em cerca de 30% o desempenho do modelo AR (1), em 29% o resultado da *Naive* (3), em 19,5% a previsão do melhor modelo autoregressivo (AR (4)), em 5,6% o desempenho do ARMA (6,6) e em 11% a eficiência preditiva do modelo VAR (4). Portanto, a CP (5) apresentou os melhores prognósticos entre todos os modelos aqui empregados, todavia se faz necessário examinar se essa diferença de desempenho é estatisticamente significativa.

#### 4.5 Comparando Previsões

Para concluir a análise, procedeu-se uma avaliação comparativa das previsões aqui geradas a partir do critério estatístico de Diebold e Mariano (1995) para atestar a significância estatística da diferença entre os erros de previsão dos

melhores modelos. Portanto, comparar-se-á o EQM de previsão do modelo da curva de Phillips (CP(5)), que apresentou maior eficiência preditiva, com os melhores modelos em cada categoria, quais sejam, o AR (4), o ARMA (6,6), o VAR(4) e, ainda, a CP (6) que também apresentou melhor desempenho do que estes. Os resultados estão sintetizados na tabela abaixo.

TABELA 8: **Teste comparativo de previsões**

Previsões	CP (5)	AR(4)	CP (5)	ARMA (6,6)	CP (5)	VAR(4)	CP (5)	CP (6)
<b>EQM</b>	0,0703	0,0873	0,0703	0,0745	0,0703	0,0791	0,0703	0,0705
<b>D</b>	-0,016		-0,004		-0,008		-0,00024	
<b>S<sub>1</sub></b>	-2,132		-3,601		-2,223		-0,856	
<b>p-valor</b>	0,033		0,000		0,026		0,392	

Fonte: Elaboração própria a partir dos resultados obtidos

Os resultados do Teste confirmam que, de fato, a CP (5) gera previsões para a inflação da região metropolitana de Fortaleza com EQM estatisticamente inferior às obtidas pelos demais modelos empregados nesse estudo e, portanto, atesta que, mesmo para dados estaduais, a curva de Phillips se mostra um bom mecanismo para prever a inflação. E, ainda, ao comparar os dois melhores modelos da curva de Phillips, CP (5) e CP (6), não se pode rejeitar a hipótese nula de que esses modelos têm desempenho de previsão semelhante.

## 5 CONSIDERAÇÕES FINAIS

O presente estudo se pautou na análise e seleção de um mecanismo robusto para gerar previsões para a inflação da região metropolitana de Fortaleza entre janeiro de 2014 e dezembro de 2014 utilizando técnicas *Naive*, modelos de séries temporais (ARMA, VAR) e modelos da curva de Phillips numa perspectiva estadual.

Em termos gerais, observa-se um bom desempenho dos modelos ARMA e VAR frente aos *benchmarks* adotados. O modelo ARMA (6,6) apresentou EQM de previsão na ordem de 7,4%, resultado 25,5% melhor do que o prognóstico do AR (1), 17,2% melhor do que a previsão *Naive* (3) e, ainda, um erro de previsão 14,6% menor do que o gerado pelo melhor modelo autoregressivo, o AR (4). Com desempenho semelhante, o modelo VAR (4) aqui empregado

obteve um EQM de 7,9%, melhorando em 21% o desempenho do AR (1) e em 20,1% a previsão do modelo *Naive* (3). Além disso, seu resultado ainda foi 9,4% superior ao do melhor modelo autoregressivo, AR (4).

Entretanto, apesar do bom desempenho dos modelos de séries temporais empregados, os dois melhores resultados foram obtidos por modelos da curva de Phillips numa perspectiva estadual, quais sejam, a CP (5) e a CP (6), com erros de previsão da ordem de 7,03% e 7,05%, respectivamente. O teste de Diebold e Mariano (1995) confirmou que a curva de Phillips, CP(5), apresentou um EQM de previsão estatisticamente menor do que o dos demais modelos.

Assim, conclui-se que a curva de Phillips ainda se mostra um mecanismo eficiente para produzir previsões de inflação com EQM de previsão reduzidos frente a modelos de séries temporais, mesmo numa análise estadual. Todavia, esse estudo pode ser ampliado para uma análise comparativa entre os modelos aqui empregados e suas versões não lineares, ou, ainda, para a realização desse exercício considerando efeitos de sazonalidade nos modelos empregados.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALENCAR, B.K. *Modelos de Previsão da Inflação: uma Análise Comparativa no Curto Prazo*. Dissertação - (mestrado em economia) – Programa de pós-graduação e pesquisa em administração e economia. IBMEC, Rio de Janeiro. 2006
- ARRUDA, E. F.; FERREIRA, R. T & CASTELAR, I. Modelos Lineares e não Lineares da Curva de Phillips para a Previsão da Taxa de inflação no Brasil. *Revista Brasileira de Economia*, v. 65, n. 3, p. 237-252, 2011.
- ARRUDA, E.F; OLIVEIRA, M.T. A; CASTELAR, I. *Ciclos Econômicos, Expectativas e Inflação: uma análise a partir da curva de Phillips Novo Keynesiana*. In: Encontro CAEN-EPGE de Políticas Públicas e Crescimento Econômico, n. 7., 2015, Fortaleza.
- ASCARI, G.; MARROCU, E. *Forecasting inflation: a comparison of linear Phillips curve models and nonlinear time series models*. Working paper centre for North Economic Research, University of Cagliari and Sassari, Sardinia, 2003.
- ATKENSON, A; OHANION, L.E. Phillips curves useful for forecasting inflation? Federal Reserve Bank of Minneapolis *Quarterly Review*, p 2-11, 2001
- BOX, G.; JENKINS, G. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, 1978.

- CAIADO, J. Cointegração e causalidade entre as taxas de juro e a inflação em Portugal. *Gestin*, v.1, p 107-118, 2002.
- CALVO, G.A. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, v.12, n. 3, p. 383-98, 1983
- CECHETTI, G. S.; HOOPER, P.; KASMAN, C. B.; SCHOENHOLTZ, L. K.; WATSON, W. M. *Understanding the Evolving Inflation Process*. U.S. Monetary Policy Forum, 2007. Disponível em: [http://www.princeton.edu/~mwatson/papers/USMPF\\_Report\\_July\\_2007.pdf](http://www.princeton.edu/~mwatson/papers/USMPF_Report_July_2007.pdf). Acesso em 20 junh.2015.
- CORREA, A. S.; MINELLA, A. Nonlinear Mechanisms of the Exchange Rate Pass-Through: A Phillips Curve Model with Threshold for Brazil. *Revista Brasileira de Economia*, v. 64, n. 3, p. 231–243, 2010.
- DIEBOLD, F. X; MARIANO, R. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, v.13, p. 253-265, 1995.
- FRIEDMAN, M. The role of monetary policy. *American Economic Review*, v. 58, n. 1, p. 1-17, 1968.
- HAMILTON, J. D. *Time Series Analysis*. Princeton University Press, 1994.
- LACK, C. Forecasting Swiss inflation using VAR models. *Swiss National Bank Economic Studies*, v.2, 2006.
- LUCAS, Robert Jr., Expectations and the neutrality of money; *Journal of Economic Theory*, v. 4, n. 2, p. 103-124, Apr. 1972.
- MENDONÇA, M. J. C.; SACHSIDA, A.; MEDRANO, L. Inflação versus Desemprego: Novas Evidências para o Brasil. *Economia Aplicada*, v. 16, p. 475–500, 2012.
- MISHKIN, S. F. Inflation Dynamics. *International Finance*, v. 10 n .3, p. 317-334, 2007.
- PATTERSON, K. *An Introduction to Applied Econometrics*. New York Macmillan Press, 2001.
- PHELPS, E. S. Phillips curves, expectations of inflation and optimum unemployment over time. *Economica*, v. 34, n. 135, p 254-281, 1967.
- PHILLIPS, A. W. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957. *Economica*, v. 25, n.100, p.283-299, 1958.

PINDYCK, R.S; RUBIFIELD, D.L. *Microeconomia: Modelos e previsões* 5ª ed. Rio de Janeiro: Elsevier, 2004.

RUMLER, F; VALDERRAMA, M.T. Comparing the new Keynesian Phillips curve with time series models forecast inflation. *Working paper series from oesterreichische National Bank*, n. 148, 2008.

SACHSIDA, A. Inflação, Desemprego e Choques Cambiais: Uma Revisão da Literatura sobre a Curva de Phillips no Brasil. *Revista Brasileira de Economia*, v. 67, n. 4, p. 549-559, 2013.

SIMS, C. A. Macroeconomics and Reality. *Econometric*. v. 48, n 1, p.1-48, 1980.

\_\_\_\_\_ Are Forecasting Models Usable for Policy Analysis? *Federal Reserve Bank of Minneapolis Quarterly Review*, v. 10, nº 1, p. 2-16, 1986.

STOCK, J.H., e WATSON, M.W. Evidence on Structural Instability in Macroeconomic Time Series Relations. *Journal of Business and Economic Statistics*, v.14, p. 11-29, 1996.

\_\_\_\_\_ *Forecasting Inflation*. NBER Working Paper Series. 7023, 1999.

\_\_\_\_\_ *Econometria*. São Paulo: Pearson, 2004.

\_\_\_\_\_ Why Has U.S. Inflation Become Harder to Forecast? *Journal of Money, Credit and Banking*, v.39, n.1, p. 3-33, 2007.

TAYLOR, J. B. Aggregate Dynamics and Staggered Contracts. *Journal of Political Economy*, v. 88, p. 1–23, 1980

WEST, K. D. Forecast evaluation. Elliot, G., Granger, C. W. J., & Timmermann, A., editors, *Handbook of Economic Forecasting*, p. 99–134. Elsevier. 2006.



# PRINCIPAIS FATORES QUE IMPACTAM NA REPROVAÇÃO E EVASÃO DOS ALUNOS DOS CURSOS TECNOLÓGICOS NO INSTITUTO FEDERAL DE CIÊNCIA E TECNOLOGIA DO CEARÁ

José Nailton Martins de Sousa\*

Francisco José Tabosa\*\*

Andrei Simonassi\*\*\*

Pablo Urano de Carvalho Castelar\*\*\*\*

## Resumo

O presente trabalho tem como objetivo principal analisar os principais fatores que impactam no nível de reprovação e/ou evasão dos alunos nos cursos tecnológicos do Instituto Federal de Ciência e Tecnologia do Ceará (IFCE), Campus-Fortaleza. A metodologia empregada foi o modelo econométrico de dados em painel, com informações dos cursos tecnológicos ministrados pelo Instituto Federal do Ceará, no período de 2006 a 2013. A base de dados foi extraída do Sistema de Acompanhamento Acadêmico dos 23 campi do Instituto Federal do Ceará, que possui 22.000 alunos matriculados nos diferentes níveis de ensino. Dentre outros resultados obtidos, verificou-se que um aumento de 1% do percentual de alunos com renda familiar inferior a 2 salários mínimos ocasiona uma elevação média da Taxa de Evasão e Reprovação de aproximadamente 0,72%. A elevação do percentual de alunos originários de escolas públicas em 1% contribui para redução da Taxa de reprovação em 0,07%. De certa forma, pode-se inferir que políticas públicas que objetivem diminuir as reprovações poderiam ser direcionadas ao aumento da renda das famílias dos estudantes e ao investimento em educação pública.

**Palavras-Chave:** Evasão escolar; Determinantes da evasão escolar; Modelo de dados em painel.

**JEL:** I21; I26; C25.

---

\* Mestre em Economia pelo CAEN/UFC. Servidor do IFCE. Av. da Universidade, 2700, 2o andar – Fortaleza/CE – CEP 60000-000 – Tel(85)33667751.

\*\* Economista. Dr. Professor do MAER/UFC. Campus do Pici, s/n bloco 826 – Fortaleza/CE – CEP 60000-000 – Tel(85)33669716.

\*\*\* Economista. Dr. Professor do CAEN/UFC. Av. da Universidade, 2700, 2o andar – Fortaleza/CE – CEP 60000-000 – Tel(85)33667751.

\*\*\*\* Economista. Dr. Professor do Curso de Finanças da UFC. Rua Marechal Deodoro, 400, 4o andar – Fortaleza/CE – CEP 60020-060 – Tel(85)33667590.

## Abstract

This work aims to analyze the main factors impacting the level of failure and/or avoidance of students in technological courses of IFCE, Campus-Fortaleza. The methodology used was the econometric model of panel data with information on technological courses offered by the Federal Institute of Ceará, in the period of 2006 to 2013. The database was extracted from the Academic Support System of Camper 23 of the Federal Institute of Ceará which has 22.000 students enrolled in different educational levels. Among other results, it was found that a 1% increase in the percentage of students with family income less than two minimum wages causes elevations of Evasion and disapproval rate of approximately 0.72%. The increase in the percentage of original public school students at 1%, decreased the failure rate at 0.07%. In a way, it can be inferred that public policies that have objective reduce failures could be directed to increase the income of the families of students and investment in public education.

**Key words:** School dropout; Determinants of school dropout; Panel Data Model.

**JEL:**I21; I26; C25.

## 1 INTRODUÇÃO

A evasão escolar é uma problemática ocorrida a nível mundial que tem despertado o interesse dos pesquisadores acadêmicos e órgãos governamentais<sup>1</sup> devido à importância da educação no processo do crescimento econômico e na redução da desigualdade de renda. Conforme a teoria do capital humano, a educação é vital para o progresso econômico, aumento da competitividade do país e melhoria do bem-estar social.

De acordo com o relatório do Banco Mundial (2012), países de alto desempenho educacional como Japão, Coréia e Austrália gastam menos que os demais países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), enquanto os Estados Unidos, apesar de gastarem acima da média, apresentam resultados decrescentes no Programa Internacional de Avaliação de Estudantes (PISA). O Brasil investiu, em 2012, aproximadamente 5,2% do PIB com educação – superior aos 4,8% da média dos países da OCDE – sem,

1 Para uma revisão da literatura e abordagem mais aprofundada da discussão a nível mundial, sugere-se a leitura dos trabalhos Moseki (2004), e Yeide e Krobin (2009).

no entanto, reduzir os altos índices de evasão e uma das maiores taxas de repetência escolar do mundo, superando apenas alguns países africanos de baixa renda. Um desafio aos governantes, portanto, reside no fato de que a educação de qualidade não depende apenas do aumento dos gastos.

Tanto a evasão quanto o abandono escolar<sup>2</sup> geram altos custos e ineficiência econômica representando um empecilho para a melhoria dos indicadores educacionais que se refletem diretamente no crescimento econômico, mercado de trabalho e, portanto, no bem-estar de toda a sociedade. Assim, é imprescindível a compreensão dos determinantes da evasão para que se encontrem alternativas para erradicá-la. Apesar das controvérsias entre os autores, há um consenso de que tanto as instituições de ensino devem adotar processos de gestão que sejam corretivos e preventivos da evasão com o intuito de reduzir os índices de abandono e evasão.

O relatório da UNESCO (2008) destaca que as causas da evasão escolar são múltiplas e complexas, abrangendo situações específicas dos diversos países, fatores particulares do aluno e o nível das redes de ensino. Dentre os problemas das redes de ensino, destacam-se a falta de recursos e de segurança nas escolas, o excesso de alunos nas salas de aula e a falta de qualificação dos professores. Verifica-se ainda que nos países em desenvolvimento, mesmo as escolas bem equipadas são incapazes de evitar a evasão, se o aluno estiver submetido a uma situação de pobreza ou miséria.

Um avanço importante no cenário educacional brasileiro se deu em meados dos anos 1990, com o advento da Lei de Diretrizes e Bases da Educação Nacional (LDB – Lei nº 9394/96). Baseada no princípio do direito universal à educação, a referida lei garantiu um gasto mínimo com a manutenção e desenvolvimento do ensino público, divisão de responsabilidades entre a União, os Estados e os Municípios além do acesso obrigatório e gratuito ao ensino fundamental.

No Brasil, as pesquisas sobre a evasão escolar se tornaram mais frequentes a partir de 1995, quando foi constituída a Comissão Especial de Estudos sobre Evasão, através de Portaria da Secretaria de Ensino Superior do Ministério da Educação (SESU/MEC), com o objetivo de desenvolver um estudo, sobre o desempenho das Instituições Federais de Ensino Superior. Nesse estudo, a SESU divulgou indicadores globais que apontavam para uma evasão média nacional de 50% nas Instituições Federais de Ensino Superior - IFES, considerando o

2. Ressalte-se que há diferença nos dois conceitos: abandono escolar é a condição do aluno que deixa de frequentar a escola durante o andamento do ano letivo, mas, volta a se matricular no ano seguinte. No caso da evasão escolar, o aluno não retorna no período letivo posterior.

conjunto dos cursos de graduação de cada instituição. Ao mesmo tempo, apontava para os baixos índices de diplomação registrados.

De acordo com os dados da Secretaria de Educação do Ceará - SEDU o desempenho das escolas cearenses, comparado às demais do país, para o ano de 2012, revelaram que ainda são altas as taxas de evasão e repetência no ensino médio, em particular na série inicial. O Estado do Ceará, conforme explicitado por Naspolini (2001), é o estado com maior extensão territorial no semi-árido brasileiro, ambiente que favoreceu uma ocupação baseada em culturas de subsistência: a pecuária, o extrativismo, a agricultura e a indústria periférica. As condições climáticas associaram-se a um modelo social gerado por contornos políticos e econômicos baseados na exploração de mão-de-obra desqualificada. Eventualmente, diversas políticas educacionais foram postas em prática, buscando reduzir desigualdade e melhor capacitar a população da região<sup>3</sup>. Uma dessas políticas foi a instalação da escola técnica e centro federal de educação tecnológica, eventualmente transformados em Instituto Federal.

Criado oficialmente no dia 29 de dezembro de 2008, pela Lei nº 11.892, o Instituto Federal do Ceará congrega os extintos Centros Federais de Educação Tecnológica do Ceará (Cefets/CE) e as Escolas Agrotécnicas Federais dos municípios de Crato e de Iguatu. O Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE) tem hoje 25 unidades, distribuídas em todas as regiões do Estado, sendo 23 *campi* efetivamente implantados, localizados nos municípios de Acaraú, Aracati, Baturité, Camocim, Canindé, Caucaia, Cedro, Crateús, Crato, Fortaleza, Iguatu, Jaguaribe, Juazeiro do Norte, Limoeiro do Norte, Maracanaú, Morada Nova, Tabuleiro do Norte, Tauá, Tianguá, Quixadá, Sobral, Ubajara e Umirim, e dois *campi* avançados em implantação, nos municípios de Guaramiranga e Jaguaruana<sup>4</sup>.

A extensão de tal instituição convida a um estudo mais aprofundado sobre o rendimento de seus discentes, assim como a sua importância como ferramenta de disseminação da educação pela região. O presente trabalho tem como intento explorar as variáveis que impactam os índices de evasão escolar e reprovação no IFCE.

Na esfera internacional, muitos pesquisadores de vários países relataram que a evasão escolar é um problema que envolve múltiplas intervenções para erradicá-lo, tanto da parte de formuladores de políticas públicas quanto da sociedade em geral, como, por exemplo, Caetano (2005), Bratti (2007), Mahuteau

---

3. Ver Naspolini (2001), e Lima (2012).

4. Informações disponíveis em

e Mavromaras (2013), Rumberger e Lim (2008). De acordo com esses autores, as intervenções para evitar a evasão escolar podem estar relacionadas a aspectos socioeconômicos (como a pobreza, pois muitos estudantes deixam a escola para trabalhar), causas relativas ao professor (qualidade de ensino), causas relativas ao aluno (muitas vezes correlacionadas aos próprios fatores socioeconômicos, como a boa alimentação) e causas relativas às práticas pedagógicas e institucionais (formas de avaliação inadequadas e regulamentos rígidos).

No âmbito nacional, a literatura centrada na ótica dos fatores determinantes da evasão escolar, destacam-se os trabalhos de Johann (2012), Furtado e Alves (2012), Pereira, Zavala e Santos (2011), Gonçalves, Rio-Neto e César (2008), Leon e Menezes-Filho (2002), Silva, Rodrigues, Brito e Franca (2012), Monteiro e Arruda (2011), Castelar, Monteiro e Lavor (2012) e Cardoso e Verner (2006).

O presente trabalho tem como objetivo principal, como mencionado, analisar os principais fatores que impactam no nível de reprovação e/ou evasão dos alunos nos cursos tecnológicos do Instituto Federal de Ciência e Tecnologia do Ceará - IFCE, Campus-Fortaleza. A metodologia empregada será o modelo econométrico de dados em painel, com informações dos cursos tecnológicos ministrados pelo Instituto Federal do Ceará, no período de 2006 a 2013. A base de dados será extraída do Sistema de Acompanhamento Acadêmico dos 23 campi do Instituto Federal do Ceará, que tem 22.000 alunos matriculados nos diferentes níveis de ensino.

O estudo foi dividido em cinco seções. Inicialmente, na seção 2, discute-se a revisão de literatura sobre os determinantes da evasão escolar no âmbito internacional, para o Brasil, Nordeste e, por fim, Ceará. Na terceira seção, apresenta-se a base de dados e a construção das variáveis utilizadas. Na quarta seção, especifica-se a metodologia utilizada no trabalho. Na quinta seção são analisados os resultados do modelo. Por fim, na última seção, apresentam-se as considerações finais.

## 2 REFERENCIAL TEÓRICO

A literatura que aborda com o olhar da ciência econômica a questão da educação, particularmente no que cerne a evasão e reprovação escolar, é vasta. No presente trabalho, optou-se por fazer menção apenas a alguns dos trabalhos de maior relevância ao tipo de análise aqui proposta<sup>5</sup>.

---

5. Para extensas revisões de literatura sobre os temas, recomenda-se a leitura de De Witte et al. (2013), Hunt (2008), Rumberger e Lim (2008).

## 2.1 Determinantes da evasão escolar conforme a literatura internacional

Em uma revisão teórica da literatura da evasão escolar, Rumberger e Lim (2008) analisam 203 estudos no assunto e dentre os principais resultados obtidos destacam: a) o fracasso acadêmico no sentido de notas baixas no início do processo educativo é um forte aspecto de previsão de futuro abandono; b) Comportamentos, por partes dos estudantes, dentro e fora do ambiente escolar, como faltas, atos delinquentes e abuso de substâncias ilegais, também são fortes preditores de abandono, ao passo que um ambiente familiar estável e acesso a recursos sociais e financeiros influenciam de forma significativa a probabilidade do estudante completar seus estudos.

Em um estudo similar, mas especificamente para os Estados Unidos, McNeal (1997) pesquisou as causas da evasão escolar para os adolescentes americanos, e relacionou esta às possibilidades de emprego. Conforme o autor, trabalhar não implica necessariamente em abandono escolar. Controlando por diversas características dos estudantes, como desempenho acadêmico, classe social, gênero, raça, envolvimento na escola e idade, os resultados do artigo demonstram que depende do tipo de emprego que o estudante busca, onde trabalhos menos intensivos e de horas mais flexíveis estavam ligados à permanência na escola.

Em um estudo para analisar a evasão escolar em Portugal, utilizando dados do censo de 2001, Caetano (2005) relatou os fatores geográficos e econômicos que influenciam o aluno a deixar a escola. O autor ressalta que as melhores oportunidades de trabalho se encontram nas regiões mais industrializadas, e tal busca por trabalho seria, assim, uma das principais causadoras da evasão. No entanto, nas localidades rurais, a evasão está ligada às dificuldades econômicas enfrentadas pelas famílias, questões culturais e problemas no deslocamento para frequentar as aulas. Da ótica do aluno, são apontados outros fatores, como a falta de motivação, o intuito de obter a independência financeira e a dificuldade de ingressar no ensino superior.

Já o estudo realizado por Bridgeland *et al.* (2006) verifica as taxas elevadas de evasão escolar, e quais os motivos para isso, através de pesquisa com os próprios estudantes. Os resultados apontam que tais motivos são: o fato de ser reprovado em uma série; a falta de preparo anterior para compreender o material apresentado em cada ano, sugerindo quantidade maior de docentes, e que estes docentes tenham melhores qualificações, visando melhorar o currículo acadêmico; e o interesse dos alunos, não apenas a compreensão do material, mas também a importância do mesmo para o futuro. Na Espanha, Enguita,

Martínez e Gómez (2010) captaram as múltiplas dimensões necessárias para explicar repetência e o abandono da escolar que vão desde o novo papel da educação na sociedade da informação para a dinâmica interna e instrumentos do sistema de ensino, além de examinar melhor divisões sociais e a tendência generalizada de adolescentes de se sentir desligada da instituição de ensino. Os dados utilizados foram os registros de alunos que abandonaram a escola antes de obter um certificado de escolaridade obrigatória para os anos de 2007 e 2008.

Ao analisar o efeito da renda dos pais sobre o abandono escolar de jovens com 16 anos na Inglaterra e no País de Gales, com o uso de dados de um estudo britânico para o ano de 1970, Bratti (2007), através da metodologia econométrica de variáveis instrumentais, constatou que, apesar da renda familiar ter um efeito negativo sobre a evasão escolar, a magnitude desse efeito é relativamente pequena. Outros atributos que apresentam efeitos mais fortes sobre a probabilidade de o aluno deixar a escola foram a escolaridade dos pais e o interesse precoce deles pela educação dos filhos.

Utilizando dados do PISA (Programa Internacional de Avaliação de Estudantes) e da pesquisa longitudinal australiana sobre jovens de 15 a 18 anos para os anos de 2006 e 2009, Mahuteau e Mavromaras (2013) mediram o grau em que a pontuação dos estudantes australianos no PISA está associada com o afastamento antecipado do indivíduo da escola. Conforme as autoras, com o uso da metodologia multinível para estimar a probabilidade do aluno evadir, é possível saber se os resultados dessa avaliação têm potencial para fornecer um indicativo das chances do estudante abandonar a escola. Dentre outros resultados, verificaram a importância da pontuação no PISA para prever a evasão escolar e ressaltam que as desvantagens sociais e individuais desempenham papel decisivo nesta relação.

## **2.2 Determinantes da evasão escolar conforme a literatura nacional**

Ainda na década de 1990, em um estudo para o Estado de São Paulo, Mello e Souza e Silva (1994) utilizam dados da Pesquisa nacional de Amostra Domiciliar (PNAD) de 1982 para analisar as relações entre origem familiar e qualidade da educação com transição escolar, distorção idade-série, aprovação e evasão. Através das técnicas de Mínimos Quadrados Ordinários e de regressão logística, os autores apresentaram os efeitos persistentes da desigualdade social no desempenho educacional, relatando que a tendência é que estudantes mais pobres abandonem a escola com maior frequência.

Posteriormente, Costa e Meneses (1995), analisando o Estado do Ceará,

consideraram a repetência e a reprovação como um fator propulsor à evasão escolar, assim, representa um alto custo de oportunidade para o governo o aluno reprovar e evadir. A repetência é uma causa relacionada ao aluno devido aos fatores econômicos e sociais. Na maioria das vezes, a falta de cultura e alfabetização na família, desestruturação familiar, má alimentação convergem para um péssimo rendimento escolar.

Conforme Caldas (2000), ao analisar o município de Santa Maria, no Rio Grande do Sul, a evasão escolar é um problema complexo e se relaciona com outros importantes temas da pedagogia, como formas de avaliação, reprovação escolar, currículo e disciplinas escolares. Para combater a evasão escolar, portanto, é preciso atacar em duas frentes: uma de ação imediata que busca resgatar o aluno “evadido”, e outra de reestruturação interna.

Castro (2000) verificou uma relação inversa entre as políticas que reduzem as desigualdades entre diferentes regiões do Brasil e as melhoras nos índices de reprovação e evasão escolar. Argumentou que embora tenha ocorrido uma melhora nos indicadores educacionais, o grau de desigualdade regional e interestadual implicaria na continuação de índices elevados de reprovação e evasão escolar nas regiões menos favorecidas. Sugere a continuidade de políticas federais que busquem nivelar o desempenho das escolas de todo o país.

Com o uso da base de dados de 1999 do Sistema Nacional de Avaliação da Educação Básica (SAEB) e do modelo econométrico de regressão multinível, Ferrão, Beltrão e Santos (2002) pesquisaram o impacto de políticas de não-repetência no desempenho escolar dos alunos do 4º ano do ensino fundamental da região Sudeste. Dentre os principais resultados obtidos, verifica-se que o regime de progressão automática nas escolas públicas, pode contribuir para a correção da defasagem idade-série sem perda da qualidade na educação. Logo, o desempenho escolar do aluno em risco de reprovação tende a ser melhor se ele for promovido do que se ficar retido. Assim, tais políticas de não repetência evitariam a evasão escolar.

A partir dos microdados da Pesquisa Mensal do Emprego (PME) do IBGE no período de 1984-1997, Leon e Menezes-Filho (2002) estimaram um modelo discreto binário para verificar os determinantes da reprovação, do avanço e da evasão escolar condicionada à reprovação em seis regiões metropolitanas brasileiras para os alunos do 4º e 5º ano do ensino fundamental e da 3ª série do ensino médio. Concluíram que os efeitos das características familiares sobre a probabilidade de reprovar, avançar e evadir dos alunos são mais importantes nas séries avançadas e, especialmente, quando se trata do avanço escolar.



Cerqueira (2004) pesquisou os principais determinantes do desempenho escolar no Brasil, utilizando dados do Censo Escolar 1999 e da Pesquisa de Informações Básicas Municipais 1999, do IBGE. Foram estimados modelos de regressão (método dos mínimos quadrados ordinários), tendo como variável dependente uma transformação logit das seguintes taxas: taxa de distorção idade-série, taxa de repetência e taxa de abandono. Como covariáveis, foram utilizadas variáveis que procuram caracterizar, sobretudo, os aspectos relacionados à infraestrutura educacional dos municípios. Concluiu que a importância da infraestrutura escolar e da qualificação docente são elementos associados à eficácia escolar.

Gonçalves, Rio-Neto e César (2008) analisaram o progresso escolar dos alunos do ensino fundamental, em escolas de Pernambuco, Sergipe, Pará, Rondônia, Mato Grosso do Sul e Goiás, para identificar os fatores associados à evasão no período de 1999 a 2003. A metodologia utilizada foi o modelo hierárquico logístico longitudinal. Dentre as principais conclusões, destacaram a influência de variáveis relacionadas à trajetória passada (repetência antes da 4ª série e proficiência) e contemporânea (situação de trabalho) do aluno para explicar a ocorrência da evasão.

Em um elaborado estudo com dados para os municípios do Estado de São Paulo, Gremaud *et al.* (2010) investigou o papel que o desempenho escolar no ensino fundamental, em especial a qualificação do aluno em termos de competências e habilidades desenvolvidas ao final deste nível de ensino, exerce na decisão de ingresso e permanência no ensino médio. Inicialmente, os autores analisaram a relação entre a proficiência e o não ingresso no ensino médio calculando, através do modelo Logit, a probabilidade de evasão escolar, probabilidade de aprovação no ensino fundamental e de matrícula no ensino médio do ano seguinte e a probabilidade de reprovação no ensino fundamental e matrícula no ano seguinte. Uma das conclusões relevantes é que existe uma relação positiva e significativa entre as notas acima da média e a permanência na escola. Quanto menores as notas do aluno, aumenta sua probabilidade de abandonar os estudos.

Com o objetivo de quantificar e traçar o perfil da evasão nos cursos de graduação do campus de Cuiabá da Universidade Federal de Mato Grosso (UFMT), Pereira, Zavala e Santos (2011) fizeram uma análise longitudinal de dois cortes de alunos ingressantes nos anos de 1995 e de 1998, até o prazo máximo de integralização curricular do referido curso, de acordo com os prazos estipulados pelo Conselho Nacional de Educação (CNE), que resulta em média

em 10 anos. Adotaram a metodologia proposta pela Associação Nacional de Dirigentes das Instituições Federais de Ensino Superior (ANDIFES), Associação Brasileira dos Reitores das Universidades Estaduais e Municipais (ABRUEM), Secretaria de Educação Superior (SESU) - Ministério da Educação e Cultura (MEC). Os dados utilizados são oriundos do sistema acadêmico da UFMT, que além de informações sobre a situação administrativa e acadêmica do aluno, possui informações sobre algumas de suas características pessoais.

Os resultados encontrados pelos referidos autores revelam que a evasão vem aumentando, chegando a 31% do corte de 1998. Os dados por área do conhecimento revelam que as áreas de Linguística, Letras e Artes, seguida de Ciências Exatas e da Terra, e Engenharias lideram o ranking com as mais elevadas taxas de evasão. Os cursos noturnos, apesar de menor evasão, apresentam maior retenção dos alunos. Os jovens com menos de 21 anos, do sexo masculino, oriundo de escola pública, cursando até o segundo ano de curso e com coeficiente de rendimento até 4,9, têm duas vezes mais chance de evadir que os demais alunos.

Em uma pesquisa sobre a evasão escolar no Instituto Federal Sul-rio-grandense com um estudo de caso no Campus Passo Fundo, Johann (2012) buscou compreender o fenômeno da evasão apresentando alguns dados estatísticos específicos de dois cursos e de turmas ingressantes em 2009. A metodologia utilizada na pesquisa consistiu numa revisão bibliográfica, análise documental e estatística sobre os alunos evadidos e as causas da evasão escolar no IF Sul Campus Passo Fundo, fornecidos pela Secretaria de Registros Acadêmicos do Campus. Concluiu que não são apenas fatores econômicos que justificam os altos índices de evasão. A grande maioria dos alunos evadidos aponta a incompatibilidade de conciliar os estudos com o trabalho, como sendo a causa de sua evasão, mesmo tendo consciência de que o estudo poderá gerar novas oportunidades sociais e de inserção no mundo do trabalho.

Furtado e Alves (2012), com o objetivo de realizar uma análise dos fatores determinantes da evasão universitária na Universidade do Vale do Rio dos Sinos (UNISINOS) localizada no Rio Grande do Sul, utilizam uma amostra representativa de alunos dessa universidade, ingressantes no período de 2005 a 2008, e a metodologia econométrica de regressão Logit para relacionar as variáveis socioeconômicas e a evasão discente. Dentre outros resultados obtidos, concluíram que algumas variáveis pesquisadas possuíam relação crescente à evasão, como, por exemplo, a variável disciplinas canceladas, em que um número maior de cancelamentos resultou em uma probabilidade mais elevada

de evasão por parte dos alunos. Outras variáveis possuíam relação inversa a esse fenômeno, como a variável disciplinas cursadas, na qual um número maior de disciplinas cursadas por semestre reduzia as chances do estudante evadir.

Moreira (2012) identificou os fatores que influenciam a evasão escolar nos cursos técnicos do Programa Nacional de Integração da Educação Profissional com a Educação Básica na Modalidade de Educação de Jovens e Adultos (PROEJA) na Rede Federal de Educação Profissional e Tecnológica de Minas Gerais (RFEPT-MG) criado no país em 2005. A metodologia adotada foi uma pesquisa bibliográfica e com uso de questionários. Os resultados da pesquisa indicaram que os alunos abandonaram a escola em razão da influência de fatores individuais, tais como a dificuldade de conciliar o horário de estudo e trabalho (66%), a necessidade de trabalhar (64,4%), o fato de a escola ser distante da casa e/ou do trabalho (63,6%), porque tinha dificuldades financeiras para realizar o curso (54,7%) e ainda porque faltava motivação para continuar os estudos (50,8%). No que diz respeito a fatores vinculados ao contexto da escola, um percentual superior a 30% se sentiu influenciado a evadir devido à falta de assistência financeira (47,6%), de flexibilidade nos horários para cursar as matérias (45,9%), ao excesso de matérias no curso (32,3%) e à existência de professores muito exigentes (30,9%).

### **2.3 Determinantes da evasão escolar no Nordeste e Ceará**

Gomes-Neto e Hanushek (1994) investigaram as causas e as consequências da evasão e da repetência escolar para o nordeste rural do Brasil no período de 1983 a 1985. Aos autores utilizam no trabalho um modelo Probit com variáveis relacionadas a características dos estudantes, das famílias, dos professores e das escolas. Identificam, assim, a evasão e a repetência escolar como uma das principais falhas do sistema educacional brasileiro, possivelmente devido ao fato dos índices que medem esses fenômenos serem maiores no Brasil do que nos demais países.

Em estudo mais recente sobre os motivos de evasão escolar de adolescentes pobres da periferia da cidade de Fortaleza, no Ceará para o ano de 2003, Cardoso e Verner (2006) utilizaram uma pesquisa com estatísticas descritivas dos dados coletados sinalizando os fatores de risco para a evasão entre esses adolescentes. Tais fatores são: parentalidade (paternidade/maternidade) precoce, pobreza extrema com privação alimentar na infância e o aumento do custo da frequência à escola. Dentre os principais resultados obtidos, concluíram que

o abandono da escola não leva necessariamente ao trabalho e sim à inatividade. As referidas autoras sugeriram incluir adolescentes de 15 a 18 anos na população alvo de programas de transferência de renda com condicionalidades.

Já Monteiro e Arruda (2011) averiguaram a influência da violência urbana sobre as evasões ocorridas nas escolas da rede estadual da Região Metropolitana de Fortaleza em 2003, com base em uma pesquisa envolvendo 54 escolas, juntamente com informações do Censo Escolar e da Secretaria de Segurança Pública. Os autores utilizam o modelo econométrico Probit Ordenado, e os resultados apontam que a violência urbana inibe os alunos de frequentar a escola, salientando a importância do docente na permanência do aluno nas atividades escolares.

Com o objetivo de investigar os fatores que motivaram a evasão de alunos ingressantes em 2005 no curso de Licenciatura Plena em Educação Física da Universidade Federal do Piauí, Silva, Rodrigues, Brito e Franca (2012) utilizaram uma metodologia descritiva e estatística. A pesquisa empírica foi realizada no período de agosto a novembro de 2010. Verificaram que os fatores responsáveis pela evasão dos alunos foram: a falta de informações sobre o curso, o descontentamento com a profissão, a imaturidade ao escolher o curso, a pouca interação com o curso, e, a simultaneidade de dois cursos.

Com a proposta de analisar as causas do abandono escolar nos municípios cearenses, utilizando uma base de dados que compreende 521 escolas públicas de ensino médio do Estado do Ceará em três anos, 2008, 2009 e 2010, Castelar, Monteiro e Lavor (2012) adotaram a metodologia do Probit Ordenado. Os principais resultados obtidos sugerem que as elevadas taxas de abandono escolar estão associadas com maior percentual de repetência, baixo número de docentes, maior distorção idade-série, elevada taxa de criminalidade e maior PIB no município. Em relação a este último, os autores ressaltam que municípios mais ricos apresentam mais oportunidades de emprego, o que eleva o custo de oportunidade de frequentar as escolas para os alunos de baixo nível socioeconômico.

Por fim, Souza (2014) analisa o fenômeno da permanência escolar nos cursos oferecidos na Educação Profissional de Nível Técnico, com ênfase no Curso Técnico Subsequente de Redes de Computadores, nas duas primeiras turmas iniciadas no segundo semestre de 2011 e concluídas no primeiro semestre de 2013, ofertado pelo Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte – IFRN, Campus São Gonçalo do Amarante. A metodologia adotada consistiu na análise documental, aplicação de questionário

e entrevistas semiestruturadas, os quais permitiram a coleta de dados e análise qualitativa. Concluiu que fatores como, o perfil/desempenho do corpo discente, os recursos e estruturas físicas escolares, os processos e as práticas pedagógicas favorecem a permanência do aluno na referida instituição de ensino.

Com essa rica literatura, apenas parte da extensa coleção de artigos e livros dedicados ao tema, é possível ter um parâmetro para os resultados da modelagem econométrica a ser adotada, a qual é exposta a seguir.

### 3 METODOLOGIA E BASE DE DADOS

#### 3.1. Base de dados

A base de dados foi extraída do Sistema de Acompanhamento Acadêmico dos 23 *campi* do Instituto Federal do Ceará (IFCE), que possui 22.000 alunos matriculados nos diferentes níveis de ensino. Foram coletadas observações de 10 cursos tecnológicos, do período de 2006 a 2013, num total de 160 observações.

Na Tabela 1 têm-se as estatísticas descritivas das variáveis. Observa-se que a Taxa média de alunos com renda familiar inferior a 2 salários mínimos (Txrenda) no curso é de 7,95% com uma máxima de 30,80% e desvio-padrão de 0,0537 no período de 2006 a 2013.

A Taxa média de evasão e reprovação (Txevrep) obteve-se o valor médio de 3,38% com desvio-padrão de 0,15 e o valor máximo de 72,60% no período de 2006 a 2013. No mais, verifica-se por meio da variável Txoriesc que a taxa média de alunos oriundos de escola pública no curso foi de 5,78% com desvio-padrão de 0,0737 e a máxima 30,80% no período estudado.

Além disso, a variável Treprov que representa a taxa de reprovação do período estudado obteve uma média de 19,45% com desvio-padrão de 0,1275 e máxima de 56,00%. A variável Txidade que representa a taxa de alunos com idade superior a 35 anos de idade apresentou média e desvio-padrão de 26,55% e 0,133 respectivamente no período de 2006 a 2013. Note-se ainda que a Txidade que representa a taxa percentual de alunos com idade superior a 35 anos, obteve-se uma média de 26,55% com desvio-padrão de 0,12 no período.

TABELA 1: Estatística Descritiva das séries: 2006 a 2013

Variável	Média	Desvio-padrão	Min	Max
Txrenda	7,95	0,0537	0,00	30,80
Txevrep	3,38	0,1541	0,00	72,60
Txreprov	19,45	0,1275	0,00	56,00
Txidade	26,55	0,1233	0,04	69,00
Txoriesc	5,78	0,0737	0,00	57,10

Fonte: Elaboração dos autores.

### 3.2. Método de Análise

A metodologia empregada é a de efeitos fixos e aleatórios aplicados em dados em painel. Tal metodologia se torna interessante em estudos como o presente trabalho devido à possibilidade de corrigir eventuais problemas de heteroneidade não-observada<sup>6</sup>. Clarke *et al.* (2010) ressaltam a importância de tal metodologia, principalmente a utilização do modelo de efeitos aleatórios, em trabalhos que analisam a questão educacional, sendo apropriada, assim, a sua adoção.

Os modelos utilizam informações sobre 10 cursos tecnológicos ministrados pelo Instituto Federal do Ceará, no período de 2006 a 2013 com dados semestrais. Inicialmente estima-se o modelo (1) que tem por finalidade encontrar os condicionantes da Taxa de reprovação e evasão (Txevrep) que será definido da seguinte forma:

$$Txevrep_{i,t} = \beta_{0i,t} + \beta_1Trenda_{i,t} + \beta_2Txidade_{i,t} + \beta_3Txoriesc + \varepsilon_{i,t} \quad (1)$$

Na sequência, é estimado o modelo (2), que determina os condicionantes da Taxa de Reprovação (Txreprov) da seguinte maneira:

$$\mathbb{R} \quad prov_{i,t} = \beta_{0i,t} + \beta_1Trenda_{i,t} + \beta_2Txidade_{i,t} + \beta_3Txoriesc + \varepsilon_{i,t} \quad (2)$$

Tal que:

$Txevrep_{i,t}$  = representa a taxa de evasão e reprovação no curso  $i$  no período  $t$ ;

$Txreprov_{i,t}$  = é a taxa de reprovação no curso  $i$  no período  $t$ ;

$Txoriesc_{i,t}$  = representa a taxa de alunos oriundos de escola pública no curso  $i$  no período  $t$ ;

$Txrenda_{i,t}$  = corresponde a taxa de alunos com renda familiar inferior a 2

6. Sobre a adoção de efeitos fixos e aleatórios e suas vantagens, sugere-se a leitura de Baltagi (2007) e Loureiro e Costa (2009).

salários mínimos no curso  $i$  no período  $t$ ;

Txidade $_{i,t}$  = é a taxa de alunos com idade superior a 35 anos;

$i$  = cursos tecnológicos;  $t$  = períodos (2006,...,2013).

Observa-se que o termo correspondente aos efeitos individuais ( $\beta_0$ ) mantém-se constante ao longo do tempo ( $t$ ), sendo, talvez, preferível tratá-lo de forma aleatória e não determinística, a fim de se destacar a heterogeneidade individual de cada curso. Sendo assim, o vetor de “constantes” do modelo (1) passa a ser considerado como aleatório:  $\beta_{0_{i,t}} = \beta_{0_{1,t}} + \mu_i$  em que  $E(\mu_i) = 0$

De acordo com Greene (2000), a escolha de uma ou de outra especificação (fixa ou aleatória) deve ser procurada nos pressupostos comportamentais da base de dados. No entanto, de acordo com a literatura especializada, a utilização do teste de especificação de Hausman faz-se necessária. Tal teste avalia a consistência de um estimador comparado a um alternativo, permitindo a escolha apropriada do modelo mais adequado. Vale salientar que a utilização de uma especificação incorreta poderá conduzir a problemas de especificação ou consistência dos estimadores (GREENE, op cit). O teste de especificação está definido na equação (3).

$$H = (\beta_H - \beta_E)' (\sum_E - \sum_H)^{-1} (\beta_H - \beta_E) \quad (3)$$

Sendo que:  $\beta_H$  responde aos estimadores de efeitos aleatórios;  $\beta_E$  aos estimadores de efeitos fixos, e;  $(\sum_E - \sum_H)$  a matriz composta pela diferença entre as variâncias dos estimadores de efeitos fixos e aleatórios.

A estatística do teste terá, sob a hipótese nula da ortogonalidade entre os efeitos individuais aleatórios e os regressores, ou seja, de que o estimador de efeitos aleatórios é o apropriado, uma distribuição assintótica  $\chi^2$  com  $K$  graus de liberdade. Caso a hipótese nula não possa ser aceita, o modelo de efeitos fixos é o que apresenta os melhores estimadores, enquanto que os estimadores do modelo de efeitos aleatórios continuam consistentes, mas não podem ser considerados eficientes.

#### 4 RESULTADOS E DISCUSSÃO

Na Tabela 2 encontram-se dispostos os resultados estimados por efeito aleatório, para Taxa de Evasão e Reprovação (Txevrep) e Taxa de Reprovação (Txreprov). Deve-se ressaltar que foi utilizado o teste de *Hausman*, que testa a hipótese dos efeitos de heterogeneidade individual serem ou não correlacionados com os regressores, para testar quais destes modelos, efeitos fixos ou aleató-

rios, é o mais apropriado para cada índice.

Conforme a estatística de *Hausman*, a 5% de significância, *Txevrep* e *Txreprov* apresentam, respectivamente, 0,1782 e 0,0893 – tais resultados levam a aceitar a hipótese nula de que as diferenças entre os parâmetros estimados pelos dois modelos são sistemáticas, de maneira que os efeitos-controle são estatisticamente correlacionados com os regressores, aceitando, assim, o modelo com efeito aleatórios.

Analisando inicialmente o modelo tendo como variável dependente *Txevrep* (evasão e reprovação), tem-se que o coeficiente estimado para a taxa de alunos com renda familiar inferior a 2 salários mínimos no curso foi de 0,7194, com significância de 4%, ou seja, aumento de 1% do percentual de alunos com renda familiar inferior a 2 salários mínimos ocasiona elevações da Taxa de Evasão e Reprovação de aproximadamente 0,72%. Parece intuitivo que alunos oriundos de famílias mais carentes estejam sujeitos a uma maior taxa de evasão e de reprovação, dada a fragilidade da situação financeira requerer, com frequência, uma busca por trabalho, o que influencia no rendimento acadêmico e até na incompatibilidade de horários, que não permite a progressão nos estudos.

Em relação à *Txidade*, a taxa de alunos com idade superior a 35 anos, o coeficiente estimado foi de -0,3491, e estatisticamente significativo a 5%. Assim sendo, um aumento de 1% na taxa de alunos com idade superior a 35 anos na média ocasiona uma redução de 0,34% na *Txevrep*. Uma possível explicação para tal resultado, que parece contradizer a literatura que argumenta que distorção idade-série é algo problemático em termos de evasão e reprovação, é que no caso específico do ensino técnico, o aluno mais velho está mais ciente e seguro de suas escolhas, apresentando maior maturidade e compromisso com o curso o qual decidiu se dedicar. Já com relação à taxa de alunos oriundos da escola pública (*Txoiesc*), tal variável não apresentou resultados estatisticamente significantes no modelo 1 estimado.

Na mesma Tabela 2, podem também ser observados os resultados do modelo estimado somente com os condicionantes da Taxa de reprovação nos cursos como variável dependente. Observe-se que, neste caso, a variável que corresponde à taxa de alunos com idade superior a 35 anos não apresentou significância estatística.

Note que o coeficiente estimado para a taxa de renda de alunos com renda familiar inferior a 2 salários mínimos no curso foi de 0,483, significativo a 5%. Assim, pode-se dizer que o aumento de 1% na taxa de alunos com renda



familiar inferior a 2 salários mínimos ocasiona elevações da Taxa de Reprovação de aproximadamente 0,48%. Novamente, parece razoável que alunos mais carentes em termos de renda, requeridos a complementar financeiramente para a família, além de estarem sujeitos a múltiplos outros problemas, tenham rendimento inferior.

De certa forma, pode-se inferir que políticas públicas de que tenham o objetivo de diminuir as reprovações deveriam ser direcionadas ao aumento da renda das famílias dos estudantes. Os resultados corroboram os estudos nacionais Moreira (2012), Leon e Menezes-Filho (2002), e Castro (2000).

Por fim, com relação à Taxa de alunos oriundos da escola pública ( $Tx_{oriesc}$ ), seu valor estimado foi de 0,075, com resultado significativo a 1%, ou seja, a elevação da taxa de alunos originais da escola pública em 1% contribui para aumentar a Taxa de reprovação em 0,07%.

**TABELA 2: Resultados dos Modelos de Regressão estimados com variável dependente  $Tx_{evrep}$ ,  $Tx_{reprov}$**

	<i>Tx<sub>evrep</sub></i> (modelo1)		<i>Tx<sub>reprov</sub></i> (modelo 2)	
	Coefic.	Valor-p	Coefic.	Valor-p
Txrenda	0.7194 (0.34)	0.04	0.483 (0.16)	0.00
Txidade	-0.3491 (0.11)	0.00	0.073 (0.54)	0.14
Txoriesc	-0.2426 (0.21)	0.27	0.075 (0.03)	0.00
Const.	0.3877 (0.046)	0.00	-2.304 (-0.53)	0.00
	Nº de obs: 160 Nº de grupos: 10		Nº de obs: 160 Nº de grupos: 10	
	Teste de Hausman: Prob > chi2 0.1782		Teste de Hausman: Prob > chi2 0.0893	

Fonte: Dados da pesquisa. Elaboração dos autores.

## 5 CONSIDERAÇÕES FINAIS

O presente trabalho teve como objetivo analisar os principais fatores que impactam no nível de reprovação e/ou evasão dos alunos nos cursos tecnológicos do IFCE, Campus-Fortaleza. Para isso, dispôs de modelos de efeitos aleatórios, utilizando dados semestrais em painel no período de 2006 a 2013.

Os principais resultados podem ser sumarizados da seguinte maneira: o aumento de 1% na taxa de alunos com renda familiar com renda inferior a 2 salários mínimos ocasiona elevações da Taxa de Evasão e Reprovação de aproximadamente 0,72%. Como argumentado anteriormente, alunos mais carentes em termos de renda, requeridos a complementar financeiramente para a família, além de estarem sujeitos a múltiplos outros problemas, teriam naturalmente um rendimento inferior e uma necessidade maior de evadir.

Constata-se, também, que a elevação de 1% na taxa de alunos com idade superior a 35 anos na média ocasiona uma redução de 0,34% na  $T_{xevrep}$ . Ou seja, na presença de alunos de idade superior, tem-se menor taxa de evasão/reprovação. A intuição por trás de tal resultado reside no maior senso de responsabilidade de existe no indivíduo mais maduro, mais seguro de sua escolha de curso e maior dedicação e compromisso a terminar os estudos de sua escolha.

Para os condicionantes da Taxa de reprovação, tem-se que a taxa de renda de alunos com renda familiar inferior a 2 salários mínimos no curso foi de 0,483, ou seja, o aumento em 1% na taxa de alunos com renda familiar inferior a 2 salários mínimos ocasiona elevações da Taxa de reprovação em 0,48%. De certa forma, pode-se inferir que políticas públicas de que tenham objetivo diminuir as reprovações poderiam ser direcionadas ao aumento da renda das famílias dos estudantes. Destaca-se que o estudo trata-se de um estudo de caso dos estudantes do IFCE.

Finalmente, a Taxa de alunos oriundos da escola pública ( $T_{xoiesc}$ ) apresentou resultado de 0,075, ou seja, a elevação do percentual de alunos oriundos de escola pública em 1% contribui para aumentar a Taxa de reprovação em 0,07%.

Logo, com base nas evidências empíricas apresentadas, é cabível mencionar e propor algumas considerações e/ou recomendações políticas, com o objetivo de melhorar o entendimento os condicionantes da Taxa de reprovação e evasão dos discentes por meio de acompanhamentos sistemáticos das famílias e dos estudantes o que levaria a uma formação de uma base de dados mais robusta e consequentemente proporcionaria uma investigação mais eficiente.

Uma possível extensão do trabalho seria o de especificar quais cursos apresentam as maiores taxas de evasão e reprovação, e buscar relacionar tais taxas com as variáveis aqui estudadas.

## REFERÊNCIAS BIBLIOGRÁFICAS

BALTAGI, B. H. Forecasting with Panel Data. Center for Policy Research Working Paper No. 91. Maxwell School of Citizenship and Public Affairs. Syracuse University. February 2007.

BRATTI, M. Parents' income and children's school drop-out at 16 in England and Wales: Evidence from the 1970 British Cohort Study. *Rev Econ Household*, Milan, p.15-40, 2007.

BRIDGELAND, J.M., DILULIO JR., J.J, MORRISON, K. B. The Silent Epidemic: Perspectives of High School Dropouts. *Civic Enterprises Reports in Association with Peter D. Hart Research Associates for the Bill & Melinda Gates Foundation*, March 2006.

CAETANO, L. Abandono Escolar: Repercussões Sócio-Econômicas na Região Centro: Algumas Reflexões. *Finisterra*, XL, v.79, p.163-176, 2005.

CALDAS, E.L. Combatendo a Evasão Escolar. São Paulo: Instituto Polis, Dicas Nº 172, 2000. Disponível em: <http://www.polis.org.br/publicacoes/download/arquivos/Dicas172.pdf> Acesso em: 02/02/2015

CASTELAR, P. U. C.; MONTEIRO, V. B.; LAVOR, D. C. *Um Estudo sobre as Causas de Abandono Escolar nas Escolas Públicas de Ensino Médio no Estado do Ceará*. Anais do VIII Encontro - Economia do Ceará em Debate, Fortaleza, 2012.

CASTRO, M. H. G. As desigualdades regionais no sistema educacional brasileiro, in Ricardo Henriques (org.) *Desigualdade e Pobreza no Brasil*, Rio de Janeiro, IPEA, pp. 425-458, 2000.

CERQUEIRA, C.A. Determinação de fatores ligados às taxas de distorção idade/série, taxa

De evasão escolar e taxa de repetência. In: RIOS-NETO e RIANI (Org.). *Introdução à Demografia da Educação*. Campinas: ABEP, 2004

CLARKE, P., CRAWFORD, C., STEELE, F. VIGNOLES, A. The Choice

Between Fixed and Random Effects Models: Some Considerations for Educational Research. IZA Discussion Paper No. 5287. October 2010.

COSTA, M.V.N., MENESES, Z.M. Evasão Escolar: Causas e Repercussão Social. Monografia do Curso de Especialização em Planejamento Educacional. Fortaleza: UNIFOR, 1995.

DE WITTE, K. CABUS, S., THYSSEN, G., GROOT, W., VAN DEN BRINK, H.M. A Critical Review of the Literature on High School Dropout. *Educational Research Review*. Volume 10, Pages 13-28, December 2013.

ENGUITA, M. F., MARTÍNEZ, L. M., GÓMEZ, J. R. School Failure and Dropouts in Spain. *Social Studies Collection* n. 29. Fundación La Caixa, 2010.

FERRÃO, M. E.; BELTRÃO, K. L.; SANTOS, D. P. dos. O impacto de políticas de não-repetência sobre o aprendizado dos alunos da 4ª série. *Pesquisa e Planejamento Econômico*, v. 32, n. 3, p.495-514, 2002

FURTADO, V. V. A.; ALVES, T.W. Fatores determinantes da evasão universitária: uma análise com alunos da UNISINOS. *Revista Contemporânea de Economia e Gestão*, v.10, n.2, 2012.

GREENE, W. H. *Econometric Analysis*. Macmillan Publishing Company, 2ª ed., p.777, 2010.

GREMAUD, A.P.; NICOELLA, A.C.; SCORZAFAVE, L.G.; OLIVEIRA, R.G.; SOARES, T.M.; BELLUZO JR., W. *A Relação entre o Abandono Escolar no Ensino Médio e o Desempenho no Ensino Fundamental Brasileiro. Gestão do Conhecimento, Linhas de Pesquisa 2009/2010*. Instituto Unibanco. 2010.

GOMES-NETO, J. B.; HANUSHEK, E. A. Causes and Consequences of Grade Repetition. *Economic Development And Cultural Change*, v.43, n.1, p.117-148, out. 1994.

GONÇALVES, M.E.; RIOS-NETO, E.L.G.; CÉSAR, C.C.. *Evasão no ensino fundamental brasileiro: identificação e análise dos principais determinantes*. XVI Encontro Nacional de Estudos Populacionais, Campinas, 2008.

HUNT, F. Dropping-Out from School: a Cross-Country Review of Literature. Create Pathways to Access, Research Monograph No. 16. May 2008.

JOHANN, C. C. *Evasão escolar no instituto federal sulrio-Grandense: um estudo de caso no Campus passo fundo*. Dissertação (Mestrado em Educação), Universidade de Passo Fundo, Porto Alegre, 2012.

LEON, F. L. L. de; MENEZES-FILHO, N. A. Reprovação, Avanço e Evasão Escolar no Brasil. *Pesquisa e Planejamento Econômico*, v. 32, n. 3, p.417-451, 2002.

LIMA, A.C. Ciclo de Avaliação da Educação Básica do Ceará: Principais Resultados. *Est. Aval. Educ.*, São Paulo. V. 23. N. 53. P. 38-58. Set/Dez. 2012.

LOUREIRO, A.O.F., COSTA, L. O. Uma Breve Discussão sobre os Modelos com Dados em Painel. Ipece. Nota Técnica, n. 37. Março 2009.

MAHUTEAU, S.; MAVROMARAS, K. *An Analysis of the Impact of Socioeconomic Disadvantage and School Quality on the Probability of School Dropout*. IZA, Discussion Paper n. 7566, Agosto de 2013.

MCNEAL, R. B. High School Dropouts: A closer examination of school effects. *School Science Quarterly*, v.78, n.1, p. 209-222, 1997.

MELLO E SOUZA, A., SILVA, N. V. Origem Familiar, Qualidade da Educação e Escolas Públicas e Particulares em São Paulo: relações e efeitos nas transições escolares. *Pesquisa e Planejamento Econômico*, v. 24, n.1, pp. 97-114, Abril 1994.

MONTEIRO, V. B.; ARRUDA, E. F. *O impacto da violência urbana nos indicadores de evasão escolar na Região Metropolitana de Fortaleza*. Anais do I Circuito de Debates acadêmicos, 2011.

MOREIRA, Priscila Rezende. *Evasão Escolar nos Cursos Técnicos do PROEJA na Rede Federal de Educação Profissional e Tecnológica de Minas Gerais*. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

MOSEKI, M.M. The Nature of Truancy and the Life World of Truants in Secondary School. Masters Degree Dissertation. University of South Africa. June 2004.

NASPOLINI, A. A Reforma da Educação Básica no Ceará. *Estudos Avançados*. No. 15 (42). 2001.

PEREIRA, R. S.; ZAVALA, A. A.; SANTOS, A. C. Evasão na Universidade Federal de Mato Grosso. *Revista de Estudos Sociais*, n.26, v.13, 2011.

RUMBERGER, R., LIM, S. A. Why Students Drop Out: A Review of 25 Years of Research. *California Dropout Research Project*, Policy Brief 15, University of California, 2008.

SOUZA, J. A. S. Permanência e evasão escolar: Um estudo de caso em uma instituição de ensino profissional. (Mestrado Profissional em Gestão e Avaliação da Educação Pública) - Faculdade de Educação, Universidade Federal de Juiz de Fora, Juiz de Fora, 2014.

WORLD BANK. World Development Report 2012: Gender Equality and development. World Bank 2012. Disponível em: <https://openknowledge.worldbank.org/handle/10986/4391>.

YEIDE, M., KOBRIN, M. Truancy Literature Review. Prepared for the U.S. Department of Justice. Office of Justice Programs. October 2009.