UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA – PIMES
DOUTORADO EM ECONOMIA

MARIA RENATA BEZERRA MELO

# DESIGUALDADE REGIONAL DE BEM-ESTAR NO BRASIL:
## uma análise a partir do consumo das famílias

Recife
2019

MARIA RENATA BEZERRA MELO

# DESIGUALDADE REGIONAL DE BEM-ESTAR NO BRASIL:
## uma análise a partir do consumo das famílias

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutora em Economia. Área de Concentração: Métodos Quantitativos.

Orientadora: Prof.ª Dr.ª Tatiane Almeida de Menezes

Recife

2019

Catalogação na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

M528d Melo, Maria Renata Bezerra
Desigualdade regional de bem-estar no Brasil: uma análise a partir do consumo das famílias / Maria Renata Bezerra Melo. - 2019.
64 folhas: il. 30 cm.

Orientadora: Prof.ª Dr.ª Tatiane Almeida de Menezes.
Tese (Doutorado em Economia) – Universidade Federal de Pernambuco, CCSA, 2019.
Inclui referências e apêndices.

1. Desigualdade. 2. Consumo. 3. Regressão quantílica incondicional. I. Menezes, Tatiane Almeida (Orientadora). II. Título.

330 CDD (22. ed.) UFPE (CSA 2019 – 049)

MARIA RENATA BEZERRA MELO

**DESIGUALDADE REGIONAL DE BEM-ESTAR NO BRASIL:**
**uma análise a partir do consumo das famílias**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutora em Economia. Área de Concentração: Métodos Quantitativos.

Aprovada em: 25/01/2019.

**BANCA EXAMINADORA**

_____
Prof.ª Dr.ª Tatiane Almeida de Menezes (Orientadora)
Universidade Federal de Pernambuco

_____
Prof. Dr. Raul da Mota Silveira Neto (Examinador Interno)
Universidade Federal de Pernambuco

_____
Prof. Dr. Carlos Roberto Azzoni (Examinador Externo)
Universidade de São Paulo

_____
Prof.ª Dr.ª Gisléia Benini Duarte (Examinadora Externa)
Universidade Federal Rural de Pernambuco

_____
Prof. Dr. Robson Douglas Tigre (Examinador Externo)

## AGRADECIMENTOS

Para que essa conquista fosse alcançada, foram necessárias muitas renúncias e dedicação. Felizmente não estive sozinha nesta caminhada e gostaria de agradecer a todos os que colaboraram de alguma forma para que esse momento se concretizasse.

Primeiramente a Deus, pelos dons que me deu e por tudo o que proporciona em minha vida. Nem sei se sou tão merecedora dessas dádivas, mas aceito de bom grado seus desígnios.

À minha família, pelos exemplos e apoio constantes. Em particular aos meus pais, que, cada qual a seu modo, sempre se orgulharam de mim e confiaram em meu potencial. Obrigada por estarem a meu lado e acreditarem tanto em mim!

Aos professores do PIMES/UFPE, em especial a Prof.ª Dr.ª Tatiane Almeida de Menezes, pela dedicação concedida durante a orientação, sempre disponível e disposta a ajudar. Agradeço também aos outros membros da Banca de Defesa pela disponibilidade: Prof. Dr. Raul Silveira Neto, Prof.ª Dr.ª Gisleia Benini, Prof. Dr. Carlos Azzoni e Dr. Robson Tigre, os quais proporcionaram discussões e sugestões que servirão para o meu crescimento pessoal e o engrandecimento da pesquisa.

À secretaria composta por Jaqueline e Maria Luíza, pelo auxílio e presteza com que me atenderam durante a realização do curso.

À CAPES, pelo auxílio financeiro, o qual possibilitou que eu me dedicasse exclusivamente à pesquisa e à escrita desta tese.

À turma de doutorado, pelos momentos divididos juntos e que tornaram mais leve meu trabalho. Agradeço de forma especial à Julia Araújo, a qual me ajudou com a programação estatística mesmo quando ainda não me conhecia pessoalmente.

A Alessandra e Ulisses, pela amizade, companheirismo e por todo o incentivo propiciados. A Rafael Fernandes, por toda a assistência a mim concedida, não apenas com correções e todo o suporte no dia da defesa, mas por me encorajar nos momentos em que eu mais precisava. Você foi meu sustentáculo nos momentos finais dessa etapa.

Enfim, a todas as pessoas que, direta ou indiretamente, contribuíram para a execução desta tese e conclusão do Doutorado, meu muitíssimo obrigada!

## RESUMO

Esta tese é constituída por dois artigos científicos que procuraram analisar a magnitude, evolução e fatores associados à desigualdade de bem-estar entre as famílias das regiões Nordeste e Sudeste do Brasil sob a ótica do consumo. Para isso, utilizou-se metodologias inovadoras nessa área de estudo. A fonte de informação para ambos foi a Pesquisa de Orçamentos Familiares (POFs) dos biênios 2002-2003 e 2008-2009 do IBGE. O primeiro artigo aplicou métodos baseados na distribuição relativa proposta por Handcock e Morris (1999) para comparar o bem-estar das famílias das regiões Nordeste e Sudeste a partir do exame das distribuições do consumo *per capita*. Tais métodos combinam ferramentas gráficas para uma análise exploratória dos dados, índices de entropia relativa e índices de polarização relativa. Os resultados encontrados apontaram que as famílias da região Sudeste têm um nível de consumo superior ao das famílias da região Nordeste ao longo de toda a distribuição de consumo. A comparação dos índices de entropia entre os biênios revela que houve uma redução na divergência entre as distribuições de consumo das regiões, enquanto o índice de polarização relativa na mediana aponta que quase não há diferença na forma entre as distribuições, principalmente no biênio 2008-2009. Portanto, podemos concluir que essa redução na divergência entre as distribuições foi devido à uma aproximação das medianas, revelando uma convergência no padrão de vida entre as regiões. No segundo artigo, investigou-se as fontes de desigualdade na distribuição do consumo familiar *per capita* no âmbito das regiões Nordeste e Sudeste do Brasil. Para isto, empregou-se o método *Recentered Influence Function* (RIF) *Regression* proposto por Firpo, Fortin e Lemieux (2009). Os resultados apontam que as famílias da região Sudeste exibem um nível de consumo maior do que o das famílias da região Nordeste ao longo de toda a distribuição e que essas lacunas se devem majoritariamente a diferenças nas características produtivas e demográficas da família, notadamente na escolaridade dos chefes. Observou-se também uma redução da disparidade regional de consumo, com reduções maiores nos quantis superiores. Houve, contudo, um aumento da distância em termos de residência em áreas urbanas e de dotação de capital humano entre os mais pobres das duas regiões. Embora tenhamos verificado uma redução do gap, a disparidade continua ainda muito significativa. Assim, reforça-se a importância de políticas públicas, voltadas principalmente para os mais pobres da região Nordeste, que visem aumentar o acesso à educação e ao mercado de trabalho, promovendo autonomia econômica dessa população.



**Palavras-chave:** Desigualdade. Consumo. Distribuição relativa. Regressão quantílica incondicional.

## ABSTRACT

This thesis consists of two scientific articles that sought to analyze the magnitude, evolution and factors associated with the inequality of well-being among the families of the Northeast and Southeast regions of Brazil from the point of view of consumption. For this, innovative methodologies were used in this area of study. The source of information for both was the Household Budget Survey (POFs) of the biennia 2002-2003 and 2008-2009 of IBGE. The first article applied methods based on the relative distribution proposed by Handcock and Morris (1999) to compare the welfare of households in the Northeast and Southeast regions by examining distributions of per capita consumption. Such methods combine graphical tools for exploratory data analysis, relative entropy indices, and relative polarization indices. The results found that the families of the Southeast region have a level of consumption superior to the families of the Northeast region throughout the distribution of consumption. The comparison of entropy indices between the bienniums reveals that there was a reduction in the divergence between the consumption distributions of the regions, while the relative polarization index at the median indicates that there is almost no difference in the form between the distributions, especially in the 2008-2009 biennium. Therefore, we can conclude that this reduction in the divergence between the distributions was due to an approximation of the medians, revealing a convergence in the standard of living between the regions. In the second article, we investigated the sources of inequality in the distribution of family consumption per capita in the Northeast and Southeast regions of Brazil. For this, we used the method Recentered Influence Function (RIF) Regression proposed by Firpo, Fortin and Lemieux (2009). The results indicate that the families of the Southeast region show a higher level of consumption than the families of the Northeast region throughout the distribution, and that these gaps are mainly due to differences in the productive and demographic characteristics of the family, especially in the schooling of the heads. There was also a reduction in the regional disparity of consumption, with larger reductions in the upper quantiles. There was, however, an increase in distance in terms of residence in urban areas and the provision of human capital among the poorest of the two regions. Although we have seen a narrowing of the gap, the disparity still remains very significant. Thus, the importance of public policies, aimed mainly at the poorest in the Northeast region, is reinforced, aiming to increase access to education and the labor market, promoting the economic autonomy of this population.



**Keywords:** Inequality. Consumption. Relative Distribution. Unconditional quantile regression.

## LISTA DE FIGURAS

Figura 1 – Funções Densidade de Probabilidade Kernel do logaritmo do consumo, Nordeste e Sudeste: (a) 2002-2003; e (b) 2008-2009 ..... 24

Figura 2 – Funções Densidade de Probabilidade Kernel do logaritmo do consumo: (a) Nordeste; e (b) Sudeste ..... 26

Figura 3 – Distribuição relativa do logaritmo do consumo total entre as regiões Nordeste e Sudeste ..... 27

Figura 4 – Distribuição relativa do logaritmo da despesa em habitação, saúde, transporte e alimentação, respectivamente, entre as regiões Nordeste e Sudeste. *Continua*..... 28

Figura 4 – Distribuição relativa do logaritmo da despesa em habitação, saúde, transporte e alimentação, respectivamente, entre as regiões Nordeste e Sudeste. *Conclusão*. . 29

Figura 5 – Decomposição da distribuição relativa do log-consumo total entre as regiões Sudeste e Nordeste: (a) e (d) densidade relativa global; (b) e (e) efeito localização; (c) e (f) efeito forma ..... 31

Figura 6 – Distribuição do consumo familiar *per capita* das regiões Nordeste e Sudeste ..... 46

Figura 7 – Distribuição do consumo *per capita* das regiões Nordeste e Sudeste segundo algumas características do chefe e do domicílio, 2008-2009. *Continua*..... 49

Figura 7 – Distribuição do consumo *per capita* das regiões Nordeste e Sudeste segundo algumas características do chefe e do domicílio, 2008-2009. *Conclusão*. ..... 50

Figura 8 – Decomposição quantílica incondicional para o diferencial do consumo familiar *per capita* entre as regiões Nordeste e Sudeste ..... 55

## LISTA DE TABELAS

Tabela 1 – Média das despesas com consumo das famílias (R$) e participação de cada subgrupo de despesa no consumo total (%).....23

Tabela 2 – Índices de polarização. Nordeste *versus* Sudeste .....32

Tabela 3 – Análise descritiva da variável consumo familiar *per capita* .....45

Tabela 4 – Média e teste de igualdade de médias das variáveis explicativas.....47

Tabela 5 – Estimação por Mínimos Quadrados Ordinários .....51

Tabela 6 - Decomposição Oaxaca-Blinder para o consumo médio.....53

Tabela 7 – Decomposição quantílica incondicional para o consumo familiar *per capita*.....56

**SUMÁRIO**


**1 INTRODUÇÃO** ..... 9

**2 DESIGUALDADE REGIONAL DE BEM-ESTAR: UMA NOVA ABORDAGEM EMPÍRICA PARA O BRASIL** ..... 11

**2.1 Introdução** ..... 11

**2.2 Por que usar o consumo para medir os padrões de vida?** ..... 12

**2.3 Metodologia** ..... 14

***2.3.1 Base de Dados*** ..... *14*

***2.3.2 Distribuição Relativa*** ..... *15*

***2.3.3 Decomposição da distribuição relativa em efeito localização e efeito forma*** ..... *19*

***2.3.4 Índice de Entropia*** ..... *20*

***2.3.5 Índice de Polarização*** ..... *20*

**2.4 Resultados** ..... 22

***2.4.1 Evidências iniciais*** ..... *22*

***2.4.2 Distribuição relativa do consumo entre as regiões Nordeste e Sudeste*** ..... *24*

***2.4.3 Decomposição da distribuição relativa em efeito localização e efeito forma e índices de entropia e polarização*** ..... *29*

**2.5 Considerações finais** ..... 33

**3 GAP REGIONAL NOS GASTOS EM CONSUMO: APLICAÇÃO DO MODELO RIF** ..... 35

**3.1 Introdução** ..... 35

**3.2 Desigualdade no consumo: evidências empíricas** ..... 37

**3.3 Metodologia** ..... 39

***3.3.1 Base de dados*** ..... *39*

***3.3.2 Método RIF*** ..... *41*

**3.4 Resultados** ..... 44

***3.4.1 Análise descritiva*** ..... *44*

***3.4.2 Estimação por Mínimos Quadrados Ordinários*** ..... *51*

***3.4.3 Decomposição de Oaxaca–Blinder*** ..... *52*

***3.4.4 Regressões quantílicas incondicionais*** ..... *54*

**3.5 Considerações Finais** ..... 58

**REFERÊNCIAS** ..... 60

**APÊNDICE A – REGRESSÃO QUANTÍLICA INCONDICIONAL** ..... 64

# 1 INTRODUÇÃO

Desde o final do século XVIII que a análise do orçamento familiar é utilizada para documentar e divulgar a pobreza. Embora existam componentes não-econômicos dos padrões de vida, como saúde, acesso à educação e liberdade política, o consumo é a melhor medida do componente econômico do padrão de vida (DEATON; GROSH, 1998). Embora o consumo seja uma medida de padrões de vida mais útil e precisa do que a renda, não há muitos trabalhos acadêmicos com foco na desigualdade de consumo.

Esta tese é constituída por dois artigos científicos que procuram analisar a distribuição do consumo das famílias das regiões Nordeste e Sudeste do Brasil, buscando depreender a magnitude, evolução e fatores associados à desigualdade de bem-estar entre as regiões. Para isso, utiliza metodologias inovadoras nessa área de estudo.

O primeiro artigo teve como finalidade comparar o bem-estar das famílias das regiões Nordeste e Sudeste a partir do exame das distribuições do consumo *per capita*. Para isso, foram utilizados métodos baseados na distribuição relativa proposta por Handcock e Morris (1999). Tais métodos consistem em uma abordagem estatística não paramétrica que combina ferramentas gráficas para uma análise exploratória dos dados, índices de entropia relativa usados para mensurar a dissimilaridade entre distribuições e índices de polarização relativa que identificam se há uma maior densidade nas caudas das distribuições. A fonte de informação é a Pesquisa de Orçamentos Familiares (POF) dos biênios 2002-2003 e 2008-2009 do IBGE. Os resultados encontrados apontam que as famílias da região Sudeste tem um nível de consumo superior ao das famílias da região Nordeste ao longo de toda a distribuição de consumo. A comparação dos índices de entropia entre os biênios revela que houve uma redução na divergência entre as distribuições de consumo das regiões, enquanto o índice de polarização relativa na mediana aponta que quase não há diferença na forma entre as distribuições, principalmente no biênio 2008-2009. Portanto, podemos concluir que essa redução na divergência entre as distribuições foi devido à uma aproximação das medianas, revelando uma convergência no padrão de vida entre as regiões.

No segundo artigo, investigam-se as fontes de desigualdade na distribuição do consumo familiar *per capita* no âmbito das regiões Nordeste e Sudeste do Brasil, utilizando os dados da Pesquisa de Orçamento Familiar (POF) do IBGE nos biênios 2002-2003 e 2008-2009. Para a análise dos fatores associados às desigualdades regionais no consumo, emprega-se o

método *Recentered Influence Function* (RIF) *Regression* proposto por Firpo, Fortin e Lemieux (2009). Trata-se de uma abordagem de regressão quantílica incondicional, a qual explora como as disparidades na distribuição de características domiciliares (efeito composição) e nos retornos a essas características (efeito estrutura) contribuem para a desigualdade de consumo entre as localidades em diferentes pontos da distribuição. Dentre as características, encontram-se o sexo, idade, escolaridade, cor e setor de emprego do chefe, localização do domicílio e composição demográfica do agregado familiar. Os resultados apontam que as famílias da região Sudeste exibem um nível de consumo maior do que o das famílias da região Nordeste ao longo de toda a distribuição e que essas lacunas se devem tanto a diferenças nas características (ligadas a demanda e a oferta de trabalho, assim como especificidades demográficas da família) quanto nos retornos a essas características. Entretanto, as diferenças nas características das famílias, notadamente na escolaridade dos chefes, são as principais fontes de desigualdade entre as duas regiões. Observou-se também uma redução da disparidade regional de consumo, com reduções maiores nos quantis superiores. Houve, contudo, um aumento da distância em termos de residência em áreas urbanas e de dotação de capital humano entre os mais pobres das duas regiões. Embora tenhamos verificado uma redução do gap no consumo entre as regiões, a disparidade continua ainda muito significativa. Assim, reforça-se a importância de políticas públicas, voltadas principalmente para os mais pobres da região Nordeste, que visem aumentar o acesso à educação e ao mercado de trabalho, promovendo autonomia econômica dessa população.

# 2 DESIGUALDADE REGIONAL DE BEM-ESTAR: UMA NOVA ABORDAGEM EMPÍRICA PARA O BRASIL

## 2.1 Introdução

Entre 2002 e 2009, período analisado neste artigo, o Brasil foi marcado por um acentuado crescimento econômico, com o aumento das contratações nos postos de salário base – parte considerável sendo representada por ocupações formais – e uma política de valorização do salário mínimo, a qual proporcionou ganhos reais aos trabalhadores com aumentos acima da inflação (OLIVEIRA *et al*., 2017).

Também se praticou uma política de valorização do crédito e incentivos fiscais para a produção e aquisição de bens duráveis, como eletroeletrônicos e automóveis. Além disso, houve uma expansão das políticas de transferências de renda, incluindo as pensões e aposentadorias, o Benefício de Prestação Continuada (BPC) e o Programa Bolsa Família. Todos esses aspectos contribuíram para o chamado "boom do consumo" no País (POCHMANN, 2012).

Outro aspecto que chama a atenção no período é a mobilidade social no País. Segundo Neri (2011), entre 2001 e 2009, houve uma transição das classes sociais mais pobres para a classe média, a chamada classe C. Entre 2003 e 2009, as classes "AB" e "C" aumentaram sua população em 6,6 milhões e 29 milhões, respectivamente. Por outro lado, houve redução do número de pessoas pertencentes as classes mais pobres "D", 2,5 milhões, e "E", 20,5 milhões.

O impacto desse crescimento econômico na redução da pobreza e da desigualdade já foi analisado em diversos trabalhos, especialmente sob a ótica da renda. Foi observada uma redução da desigualdade de renda, com um declínio do índice de Gini de 0,58 em 2003 para 0,55 em 2009 (NERI, 2011). A pobreza também apresentou uma trajetória de queda durante a primeira década dos anos 2000, não sendo interrompida pela crise financeira de 2008 (IPEA, 2012).

No entanto, diferenças nos padrões de vida medidas a partir do consumo são muito mais exatas do que a partir da renda. Porém, não existem muitos trabalhos acadêmicos com foco na desigualdade de consumo. Sendo assim, o objetivo deste artigo foi examinar a desigualdade de bem-estar das famílias das regiões Nordeste e Sudeste através da comparação das distribuições do consumo *per capita*, buscando compreender a sua evolução no período. A

fonte de informação é a Pesquisa de Orçamentos Familiares (POF) dos biênios 2002-2003 e 2008-2009, disponibilizada pelo IBGE, e a variável de análise é a soma das despesas monetárias e não-monetárias das famílias como *proxy* do consumo.

Para atingir o objetivo, foram utilizados métodos baseados na distribuição relativa proposta por Handcock e Morris (1999). Trata-se de uma abordagem estatística não paramétrica que combina ferramentas gráficas para uma análise exploratória dos dados, índices de entropia para mensurar a dissimilaridade entre distribuições e índices de polarização relativa que identificam se houve uma maior densidade nas caudas das distribuições. Desta forma, é possível fazer uma análise dos dados livre de pressupostos paramétricos recorrentes sobre a forma matemática da função de probabilidades da variável resposta.

Além desta introdução, este artigo possui quatro seções adicionais. A segunda seção trata da comparação entre as medidas de desigualdade de consumo e de renda, argumentando os motivos que fazem do consumo uma medida de bem-estar superior à renda. A terceira seção discute a base de dados e as metodologias empregadas. A quarta seção analisa os resultados obtidos. Por fim, a última seção é destinada à conclusão do trabalho.

## 2.2 Por que usar o consumo para medir os padrões de vida?

Existem, basicamente, duas formas de medir disparidades de bem-estar entre as famílias. A primeira é fazendo uso de indicadores com métrica em unidades monetárias (*money metric indicators*). Os exemplos mais comuns de indicadores desse tipo são as desigualdades de renda e das despesas. A segunda não utiliza uma métrica monetária para a avaliação do indicador, baseando-se na quantidade e qualidade de serviços disponíveis à população. Indicadores de saúde, habitação e educação, tais como expectativa de vida, densidade habitacional e atraso escolar, são exemplos desse segundo tipo de medição (FIRPO, 2000).

Nos últimos anos, um grande número de trabalhos acadêmicos têm focado, dentro desse primeiro grupo de indicadores, na desigualdade de renda (MELO; MONTE, 2015; SOARES, 2006). A desigualdade no consumo (gastos das pessoas) recebeu muito menos atenção. Isso se deve, em parte, devido à indisponibilidade de dados: informações sobre os recursos domésticos, como salários e renda, estão mais frequentemente disponíveis nos

conjuntos de dados, com amostras tipicamente maiores do que em conjuntos de dados contendo informações sobre o consumo (ATTANASIO; HURST; PISTAFERRI, 2015).

Estudar a desigualdade de renda é analisar as diferenças nos padrões de vida potenciais, pois o rendimento, especialmente após os impostos, mede a capacidade que uma família ou indivíduo tem para alcançar um determinado padrão de vida, considerando que se gasta toda a renda. No entanto, as pessoas geralmente não gastam toda a sua renda. Uma parte pode ser destinada à poupança, para o pagamento de dívidas, doação para caridade ou para ajudar os filhos em vários estágios da vida. Além disso, é possível emprestar (tomar emprestado), reduzindo (expandindo) o padrão de vida para aquém (além) do que a renda permite. Todas essas finalidades do dinheiro reduzem a capacidade do destinatário do rendimento de alcançar um potencial padrão de vida representado por sua renda. Já o consumo reflete melhor o padrão de vida atual das pessoas, pois vai representar exatamente o que elas tiveram acesso (SARLO, 2016).

A renda familiar *per capita*, principal variável utilizada nas análises de bem-estar, apresenta uma flutuação cíclica que pode dificultar uma avaliação mais exata do grau de carência familiar, principalmente quando é coletada por intervalos curtos de tempo, pois apresenta apenas um quadro estático da restrição de recursos das famílias. Nos casos em que um ou mais membros da família não é capaz de obter renda por curtos períodos, podem levar a medidas de desigualdade de renda (bem-estar) sobreestimadas. Sob a hipótese de suavização do consumo ao longo do horizonte de vida, se o consumo familiar *per capita* for utilizado como indicador de bem-estar ao invés da renda familiar *per capita*, evita-se os efeitos gerados por choques sobre a parcela transitória da renda e, assim, obtém-se uma medida de desigualdade bastante robusta a variações momentâneas do cenário econômico, isto é, medidas de desigualdade do consumo revelam quão permanentemente desigual são as distribuições dos recursos disponíveis às famílias (FIRPO, 2000).

Outra questão importante sobre a renda é que geralmente os entrevistados são mais relutantes em compartilhar informações sobre seus rendimentos e ativos do que sobre seu consumo, ficando mais propensos a dar respostas intencionalmente imprecisas a perguntas sobre seus rendimentos em comparação com respostas a perguntas sobre o seu consumo. Em países onde o rendimento é tributável, os entrevistadores têm de persuadir os entrevistados de que a informação que eles fornecem não será transmitida às autoridades fiscais. Além disso, os agregados familiares ricos podem se recusar a conceder entrevistas à equipe da pesquisa ou o

entrevistado, que pode ser um membro da família ou um empregado, pode ter mais conhecimento sobre o consumo da casa do que sobre suas fontes e níveis de renda (DEATON; GROSH, 1998).

## 2.3 Metodologia

Nesta seção, apresentam-se os procedimentos adotados para comparar as distribuições de consumo das regiões Nordeste e Sudeste, nos biênios 2002-2003 e 2008-2009. Os métodos escolhidos são baseados na distribuição relativa proposta por Handcock e Morris (1999), os quais consistem em: i) ferramentas gráficas para uma análise exploratória dos dados; ii) índices de entropia para mensurar a dissimilaridade entre as distribuições em dois pontos no tempo; e iii) índices de polarização relativos que identificam uma maior densidade nas caudas da distribuição. Antes de exibir essas técnicas, porém, apresenta-se a base de dados empregada neste estudo.

### *2.3.1 Base de Dados*

Os resultados foram obtidos utilizando os microdados das Pesquisas de Orçamentos Familiares (POFs) do Instituto Brasileiro de Geografia e Estatística (IBGE) dos biênios 2002-2003 e 2008-2009. As POFs visam, principalmente, mensurar as estruturas de consumo, dos gastos, dos rendimentos e parte da variação patrimonial das famílias. Dessa forma, é possível traçar um perfil das condições de vida da população brasileira a partir da análise de seus orçamentos domésticos (IBGE, 2010).

As POFs dos anos 1980 e 1990 foram realizadas nas regiões metropolitanas de Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre, no Município de Goiânia e no Distrito Federal. Já as duas últimas POFs (2002-2003 e 2008-2009) possuem abrangência geográfica nacional, incluindo as áreas urbana e rural, motivo pelo qual optou-se por trabalhar com elas.

O termo família é utilizado neste trabalho para representar o conceito de unidade de consumo, que é a unidade básica da pesquisa. No domicílio é identificada a unidade de consumo

que compreende um único morador ou conjunto de moradores que compartilham da mesma fonte de alimentação ou compartilham as despesas com moradia. Esta definição segue as recomendações e práticas internacionais referentes a pesquisas similares.

As pesquisas foram realizadas em 12 meses para possibilitar a estimação de orçamentos familiares que atentem às mudanças que ocorrem ao longo do ano. Quatro períodos de referência foram definidos: sete dias, 30 dias, 90 dias e 12 meses. Os períodos de referência para cada informante foram estabelecidos como o tempo que antecede a data de realização da coleta no domicílio, portanto, não correspondem às mesmas datas para cada domicílio selecionado. O período de referência de sete dias é uma exceção, pois é contabilizado no decorrer da entrevista. Com um período de coleta de 12 meses e períodos de referência de até 12 meses, temos informações que se distribuem em um período total de 24 meses. Durante esse tempo, podem ocorrer mudanças absolutas e relativas nos preços. Dessa forma, os valores levantados na pesquisa devem ser valorados a preços de uma determinada data. A data de referência fixada para a compilação, análise e apresentação dos resultados da POF 2002-2003 foi de 15 de janeiro de 2003 e da POF 2008-2009 foi 15 de janeiro de 2009.

No referente à metodologia para a construção da variável consumo, foram utilizadas apenas as despesas com aquisições de bens e serviços realizadas pelas unidades de consumo para atender diretamente às necessidades e desejos pessoais de seus componentes no período da pesquisa. As despesas de consumo estão organizadas nas seguintes categorias: alimentação, habitação, vestuário, transporte, higiene e cuidados pessoais, assistência à saúde, educação, recreação e cultura, fumo, serviços pessoais e outras despesas diversas não classificadas anteriormente.

Para que fosse possível a comparação entre os anos, os valores das despesas da POF 2002-2003 foram inflacionados para serem expressos em reais de janeiro de 2009. Para isso, os dados de 2002-2003 foram multiplicados por 1,391437, mesmo procedimento adotado por Hoffmann (2010). A correção dos valores monetários foi feita utilizando o Índice Nacional de Preços ao Consumidor (INPC) do IBGE.

### *2.3.2 Distribuição Relativa*

Os métodos baseados na distribuição relativa de Handcock e Morris (1999) são um instrumental importante de análise, pois constituem uma abordagem não paramétrica que possibilita a análise dos dados sem nenhuma suposição *a priori* sobre a distribuição de probabilidades da variável de interesse. Além disso, o foco passa a ser a distribuição como um todo e não apenas uma medida central, como a média. Nesse sentido, fornece um arcabouço consistente para a análise da desigualdade no consumo.

A distribuição relativa equivale à transformação dos dados de duas distribuições (referência e comparação) em uma única distribuição, a qual contém toda a informação necessária para realizar comparações entre elas (GUIMARÃES; OLIVEIRA, 2008).

Para a aplicação empírica deste trabalho, considere $Y_0$ uma variável aleatória que representa o consumo *per capita* de uma população de referência (famílias do Nordeste). Assume-se que $Y_0$ é não negativo com função de distribuição acumulada (FDA) $F_0(y)$ e função densidade de probabilidade (FDP) $f_0(y)$. Considere ainda a mesma medida em uma população de comparação $Y$ (famílias do Sudeste), com FDA $F(y)$ e FDP $f(y)$.

A distribuição relativa (R) de $Y$ para $Y_0$ pode ser definida como uma variável aleatória gerada pela transformação ou reescalonamento da variável do grupo de comparação ($Y$) através da função de distribuição acumulada do grupo de referência ($F_0$):

$$R = F_0(Y) \tag{2.1}$$

A variável R tem a propriedade de ser contínua no intervalo $[0, 1]$ e os seus valores são representados por *r*. O dado relativo *r* mede a posição relativa de $Y$ na distribuição de $Y_0$, ou seja, como as famílias do Sudeste estariam alocadas na distribuição do consumo *per capita* das famílias do Nordeste.

Por ser uma variável aleatória, *R* possui uma função de distribuição acumulada e uma função densidade de probabilidade. A FDA relativa, $G(r)$, é dada por:

$$G(r) = F\left(F_0^{-1}(r)\right) = F\left(Q_0(r)\right) \qquad 0 \leq r \leq 1 \tag{2.2}$$

onde $Q_0(r)$ é a função quantílica do consumo *per capita*, originada da inversa da FDA $F_0^{-1}(r)$. A função quantílica, $Q(r)$, define o valor do consumo *per capita* no quantil $r$ da distribuição. Por exemplo, no caso da mediana com $r = 0{,}5$, o valor de $Q(0{,}5)$ define o consumo *per capita* que divide a população, com 50% consumindo abaixo desse valor e 50% acima.

A FDP relativa de $R$ ou densidade relativa é

$$g(r) = \frac{f\left(F_0^{-1}(r)\right)}{f_0\left(F_0^{-1}(r)\right)} = \frac{f\left(Q_0(r)\right)}{f_0\left(Q_0(r)\right)} \qquad 0 \leq r \leq 1 \qquad (2.3)$$

onde $r$ representa o quantil ou a proporção dos valores.

A densidade relativa pode ser expressa em termos da escala da medida original do consumo *per capita*. Suponha que o r-ésimo quantil de $R$ seja denotado pelo valor $y_r$. Logo, a densidade relativa pode ser expressa como:

$$g(r) = \frac{f(y_r)}{f_0(y_r)} \qquad y_r = Q_0(r) \geq 0 \qquad (2.4)$$

A interpretação da densidade relativa $g(r)$ é direta: a razão entre a densidade da população comparação (famílias do Sudeste) e a densidade da população referência (famílias do Nordeste) em um dado nível de atributo $y_r$ (consumo *per capita* referente ao quantil $r$). A taxa gerada pela distribuição relativa possibilita comparar diretamente duas distribuições.

Há três formas de interpretar os resultados para cada quantil da distribuição segundo Rodrigues, Rios-Neto e Pinto (2010): i) quando a densidade relativa for maior que 1, $g(r) > 1$, pode-se dizer que há uma sobrerrepresentação das observações do grupo de comparação em relação ao grupo de referência; ii) quando a densidade relativa for menor que 1, $g(r) < 1$, há uma sub-representação das observações do grupo de comparação em relação ao grupo de referência; e iii) quando a densidade relativa for igual a 1, $g(r) = 1$, a densidade de pessoas nos grupos de referência e comparação é a mesma para o quantil em questão, indicando uma equivalência distribucional.

No caso de duas distribuições serem idênticas, a distribuição relativa resultante seria a uniforme. Teríamos uma situação em que o consumo ao longo das distribuições para ambas as regiões seriam iguais. Como exemplo, 20% das famílias do Sudeste localizadas na cauda inferior da distribuição seriam iguais aos 20% das famílias do Nordeste que possuem os menores consumos.

Em geral, a distribuição relativa é invariante com a escala das distribuições ou a uma transformação monótona. Portanto, não faz diferença a comparação ser realizada entre o consumo *per capita* ou o logaritmo do consumo *per capita*, pois a distribuição relativa resultante é a mesma.

Para a estimação das densidades de probabilidade do consumo *per capita* das regiões Nordeste e Sudeste, utilizou-se uma metodologia não-paramétrica de suavização dos dados denominada kernel, na qual não é preciso uma suposição sobre a forma matemática da função de distribuição do consumo.

Para a estimação de densidades kernel é necessária, além da função kernel, a largura do intervalo ao qual a densidade será suavizada. A escolha da função kernel não é crucial para a performance do método, pois não influencia significativamente no formato da curva de probabilidade estimada. No entanto, a largura do intervalo tem implicações importantes, uma vez que envolve uma escolha entre viés e variância: um aumento da largura do intervalo origina uma menor variância (um aumento na precisão das estimativas), entretanto, corre-se o risco de obter uma densidade viesada no caso de uma suavização excessiva dos dados (DINARDO; TOBIAS, 2001).

Neste trabalho, seguimos a proposta de Silverman (1986), o qual recomenda a utilização da função kernel Epanechnikov, $k(x)$, com a largura ótima do intervalo definida por $h^*$ a seguir:

$$\widehat{f_n}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n}k\left(\frac{x - X_i}{h}\right) \tag{2.5}$$

$$k(x) = \frac{3}{4}(1 - x^2)I(x) \tag{2.6}$$

$$h^* = 1{,}3643\ \delta\ n^{-0{,}2}\ \sigma \tag{2.7}$$

onde $x$ é o valor da função a ser suavizado, $I(x)$ é a função indicadora, $\delta$ é uma constante para a kernel epanechnikov (1,7188), $\sigma$ é o desvio-padrão amostral e $n$ é o tamanho da amostra.

A distribuição relativa simplifica a comparação entre duas curvas de densidade de probabilidade tendo em vista que sintetiza as diferenças entre elas através de uma única curva formada pelas taxas de densidade relativas. Os gráficos que a acompanham facilitam ainda mais a interpretação dos resultados.

Outra vantagem da distribuição relativa é que as diferenças entre duas distribuições podem ser decompostas em variações na mediana (efeito localização) e na estrutura (efeito forma), permitindo análises sobre o efeito ao longo do tempo das mudanças sobre o consumo *per capita*. A seguir, será exposta a metodologia para a decomposição da distribuição relativa.

### *2.3.3 Decomposição da distribuição relativa em efeito localização e efeito forma*

Duas distribuições podem diferir em relação às suas localizações e suas formas. Desejamos saber como a distribuição do consumo *per capita* das famílias das regiões Sudeste e Nordeste se diferenciam. Há quatro casos possíveis: i) a localização é diferente, mas a forma é a mesma; ii) a forma é diferente, mas a localização é igual; iii) localização e forma são diferentes e iv) localização e forma são iguais.

Para efetuar a decomposição da distribuição relativa em efeito localização e efeito forma de acordo com Handcock e Morris (1999), considere uma variável hipotética $Y_h$ cuja localização de sua distribuição seja igual à localização da distribuição de comparação, mas a forma permaneça a mesma da distribuição de referência. Neste trabalho, a variável $Y_h$ foi ajustada para ter a mesma mediana do grupo de comparação, isto é, $Y_h = Y_0 + \rho$, onde $\rho = Q\left(\frac{1}{2}\right) - Q_0\left(\frac{1}{2}\right)$, ou seja, $\rho$ é a diferença entre a mediana de $Y$ e a mediana de $Y_0$.

Agora é possível construir as duas distribuições relativas que representam as diferenças na localização e na forma com as variáveis de três grupos: grupo de referência $Y_0$; grupo de comparação $Y$; e grupo hipotético $Y_h$. Tem-se, portanto, além da distribuição relativa do grupo de comparação em relação ao grupo de referência, $R_0 = F_0(Y)$ (equivale à equação 2.1):

i) a distribuição relativa do grupo hipotético em relação ao grupo de referência, $R_0^h = F_0(Y_h) = F_0(Y_0 + \rho)$, que isola o efeito de diferenças na localização; e

ii) a distribuição relativa do grupo de comparação em relação ao grupo hipotético, $R_h = F_h(Y)$, que isola outros efeitos, como diferenças na forma.

Em termos de taxa de densidade relativa, temos:

$$\frac{f(y_r)}{f_0(y_r)} = \frac{f_h(y_r)}{f_0(y_r)} \cdot \frac{f(y_r)}{f_h(y_r)} \tag{2.8}$$

O lado esquerdo da equação representa a densidade relativa global, $g_0(r)$, a primeira razão do lado direito da equação representa a taxa de densidade para diferenças na localização, $g_0^h(r)$, e a segunda razão do lado direito da equação representa a taxa de densidade para diferenças na forma da distribuição de consumo, $g_h(r)$:

$$g_0(r) = g_0^h(r) \cdot g_h(r) \tag{2.9}$$

Os resultados provenientes dessa técnica são apresentados através de gráficos e quantificados por meio de medidas-síntese, como o índice de entropia e o índice de polarização, descritos nas seções seguintes.

### *2.3.4 Índice de Entropia*

É possível mensurar a dissimilaridade entre duas distribuições utilizando um índice de entropia. Este artigo investiga a divergência das distribuições de consumo das famílias do Nordeste e do Sudeste nos biênios 2002-2003 e 2008-2009.

Handcock e Morris (1999) propõem a utilização da medida de Kullback-Leibler para medir a dissimilaridade entre duas distribuições. Esta medida tem a vantagem de possuir uma interpretação simples em termos da distribuição relativa, além de ser decomponível em localização e estrutura. O índice de entropia é dado por:

$$D(F; F_0) = \int_{-\infty}^{\infty} log\left(\frac{f(y_r)}{f_0(y_r)}\right) dF(y) = \int_0^1 log\big(g(r)\big) g(r) dr \tag{2.10}$$

O índice pode ser decomposto em três componentes:

$$D(F; F_0) = D(F_h; F_0) + D(F; F_h) \tag{2.11}$$

onde $D(F; F_0)$ é a divergência total entre as distribuições do grupo de referência e de comparação, $D(F_h; F_0)$ representa a dissimilaridade entre as distribuições ocasionadas por diferenças na localização, e $D(F; F_h)$ expressa a distinção entre as distribuições ocasionadas por diferenças na forma.

### *2.3.5 Índice de Polarização*

A polarização se caracteriza por um aumento das famílias que consomem em níveis baixos juntamente com um aumento das famílias que consomem em níveis elevados, isto é, há

um aumento da densidade nas caudas superior e inferior da distribuição do consumo, gerando uma densidade relativa com formato de U.

Handcock e Morris (1999) criaram um índice que captura desvios na mediana e nas caudas inferior e superior entre duas distribuições e se baseia no conceito de desvio mediano absoluto ($y_i = |x_i - Q(0{,}5)|$). Assim, define-se o índice de polarização relativa na mediana (PRM) da seguinte forma:

$$PRM(F; F_0) = 4E\left(\left|R_h - \frac{1}{2}\right|\right) - 1 \tag{2.12}$$

O índice PRM mede o desvio mediano absoluto em torno da mediana da distribuição relativa proveniente de diferenças apenas na forma, reescalonado para produzir um índice que varie de -1 a 1. Como a densidade de $R_h$ é $g_h(r)$, podemos reescrever a equação como:

$$PRM(F; F_0) = 4\int_0^1 \left|r - \frac{1}{2}\right| g_h(r)\, dr - 1 \tag{2.13}$$

A densidade relativa $g_h(r)$ é ponderada pela distância entre o dado relativo e o centro, $\left|r - \frac{1}{2}\right|$, dando um peso maior para as massas nas caudas (inferior e superior) do que ao centro da distribuição. Quando a PRM assume o valor zero indica que não há diferença na estrutura das duas distribuições. Valores positivos apontam para uma maior polarização ou a uma maior densidade nas caudas da distribuição do grupo de comparação, enquanto valores negativos assinalam uma polarização negativa ou uma maior concentração da distribuição do grupo de comparação em torno da mediana.

O índice PRM pode ser decomposto em duas partes, sendo possível identificar contribuições para a polarização decorrentes de diferenças na distribuição abaixo e acima da mediana, denominados como índice de polarização relativa na cauda inferior (equação 2.14) e índice de polarização relativa na cauda superior (equação 2.15), respectivamente:

$$PRI(F; F_0) = 8\int_0^{\frac{1}{2}} \left|r - \frac{1}{2}\right| g_h(r) dr - 1 \tag{2.14}$$

$$PRS(F_1; F_0) = 8\int_{\frac{1}{2}}^1 \left|r - \frac{1}{2}\right| g_h^1(r) dr - 1 \tag{2.15}$$

A análise dos valores positivos e negativos do PRI ou PRS é similar ao PRM. No caso de assumir o valor zero, no entanto, o componente da distribuição relativa é uniforme abaixo ou acima da mediana.

## 2.4 Resultados

Nesta seção são apresentados os resultados referentes à distribuição relativa do consumo entre as regiões Nordeste e Sudeste do País nos biênios 2002-2003 e 2008-2009. Antes, porém, são expostas, na subseção 2.4.1, algumas análises iniciais sobre as categorias de consumo *per capita* utilizadas. Em seguida, além das distribuições relativas, são apresentadas as decomposições das distribuições relativas em efeito localização e efeito forma e os Índices de Entropia e Polarização.

### *2.4.1 Evidências iniciais*

Como mencionado na seção de metodologia, as despesas de consumo estão organizadas nas seguintes categorias: alimentação, habitação, vestuário, transporte, higiene e cuidados pessoais, assistência à saúde, educação, recreação e cultura, fumo, serviços pessoais e outras despesas diversas não classificadas anteriormente. A Tabela 1 descreve a média da despesa com consumo relacionada a cada uma das categorias e a respectiva participação na despesa total, nos biênios 2002-2003 e 2008-2009, nas regiões Nordeste e Sudeste e no Brasil como um todo.

A partir da Tabela 1, observa-se que a média do consumo *per capita* da região Sudeste é sempre superior à média da região Nordeste, assim como é maior do que o consumo médio do País. Isto se verifica para todos os subgrupos de despesa, notadamente para a média de gastos com educação no biênio 2002-2003, a qual é quase três vezes maior do que a constatada na região Nordeste. A região Sudeste também se destaca no consumo das categorias fumo, cultura, saúde, habitação e transporte.

Tabela 1 – Média das despesas com consumo das famílias (R$) e participação de cada subgrupo de despesa no consumo total (%)

| Subgrupos de despesa | Brasil | | Nordeste | | Sudeste | |
|---|---|---|---|---|---|---|
| | Média | % | Média | % | Média | % |
| | | | 2002-2003 | | | |
| Alimentação | 1549,38 | 15,52 | 1205,14 | 21,15 | 1760,12 | 13,83 |
| Habitação | 3215,88 | 32,21 | 1736,12 | 30,46 | 4111,83 | 32,32 |
| Vestuário | 443,81 | 4,45 | 314,40 | 5,52 | 503,38 | 3,96 |
| Transporte | 1538,39 | 15,41 | 781,18 | 13,71 | 1904,42 | 14,97 |
| Higiene | 172,82 | 1,73 | 140,65 | 2,47 | 192,05 | 1,51 |
| Saúde | 1584,68 | 15,87 | 858,87 | 15,07 | 2207,29 | 17,35 |
| Educação | 304,29 | 3,05 | 145,04 | 2,54 | 423,58 | 3,33 |
| Cultura | 201,95 | 2,02 | 101,07 | 1,77 | 267,87 | 2,11 |
| Fumo | 58,23 | 0,58 | 28,79 | 0,51 | 76,91 | 0,60 |
| Serviços Pessoais | 87,35 | 0,87 | 46,21 | 0,81 | 118,28 | 0,93 |
| Despesas diversas | 254,59 | 2,55 | 119,21 | 2,09 | 339,04 | 2,66 |
| Outras despesas | 571,37 | 5,72 | 222,46 | 3,90 | 818,60 | 6,43 |
| Despesa total | 9982,75 | 100,00 | 5699,13 | 100,00 | 12723,37 | 100,00 |
| | | | 2008-2009 | | | |
| Alimentação | 1827,07 | 18,08 | 1378,89 | 22,66 | 2086,13 | 16,74 |
| Habitação | 3639,90 | 36,03 | 2025,33 | 33,29 | 4641,98 | 37,24 |
| Vestuário | 488,27 | 4,83 | 357,43 | 5,88 | 528,43 | 4,24 |
| Transporte | 1796,62 | 17,78 | 1041,74 | 17,12 | 2131,97 | 17,10 |
| Higiene | 213,17 | 2,11 | 174,33 | 2,87 | 228,34 | 1,83 |
| Saúde | 733,48 | 7,26 | 392,00 | 6,44 | 986,32 | 7,91 |
| Educação | 251,37 | 2,49 | 146,61 | 2,41 | 334,67 | 2,69 |
| Cultura | 189,35 | 1,87 | 104,65 | 1,72 | 241,87 | 1,94 |
| Fumo | 50,57 | 0,50 | 26,11 | 0,43 | 66,26 | 0,53 |
| Serviços Pessoais | 105,57 | 1,04 | 63,50 | 1,04 | 138,68 | 1,11 |
| Despesas diversas | 286,29 | 2,83 | 142,79 | 2,35 | 353,84 | 2,84 |
| Outras despesas | 521,13 | 5,16 | 230,51 | 3,79 | 725,64 | 5,82 |
| Despesa total | 10102,79 | 100,00 | 6083,89 | 100,00 | 12464,12 | 100,00 |

Fonte: elaboração própria a partir dos dados da POF.

Fazendo uma análise temporal, nota-se que a região Sudeste apresentou um decréscimo de seus gastos médios totais enquanto a região Nordeste exibiu um crescimento no período, seguindo a tendência brasileira. No período considerado, ambas as regiões apresentaram redução nos gastos com saúde (mais de 54%), fumo e educação. Segundo Oliveira *et al*. (2017), houve um crescimento do consumo *per capita* no Brasil do 1º ao 95º percentil no período. Para cerca de 90% da população o consumo cresceu acima da média e a partir do 85º percentil as taxas de crescimento se reduzem, ficando abaixo da média após o 90º percentil. De

acordo com os autores, esse padrão de crescimento traz consequências tanto para o bem-estar social quanto para a desigualdade.

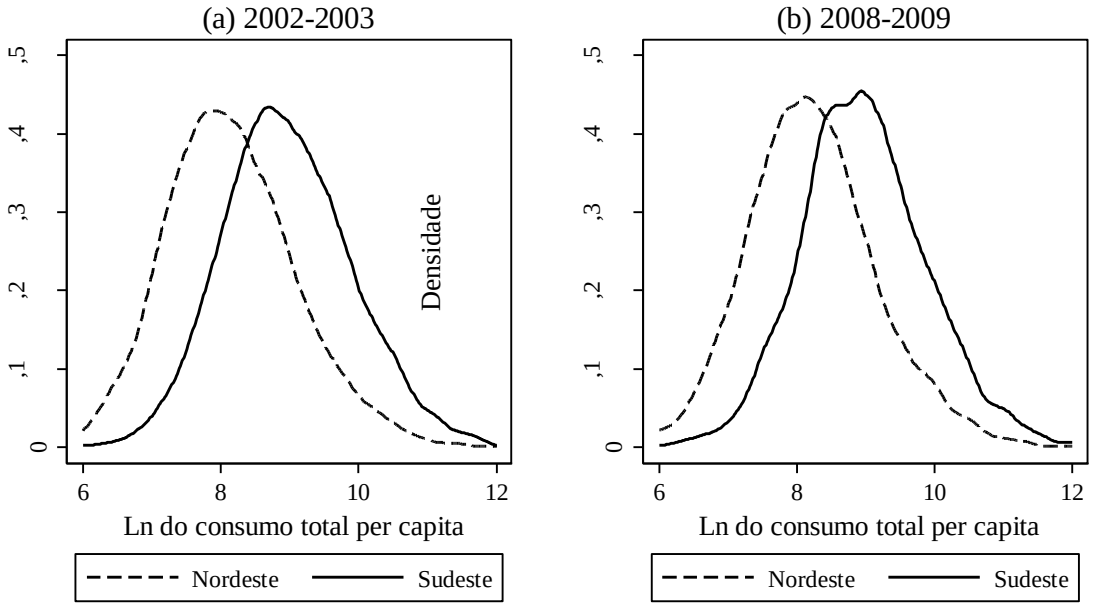
Outro aspecto analisado na Tabela 1 é a participação de cada subgrupo de despesa no consumo total. É possível observar que as categorias que apresentam um maior peso no orçamento das famílias são: habitação, alimentação, saúde e transporte. Além da despesa total, esses grupos de despesa serão estudados com maior detalhe neste artigo.

A seguir, são apresentados os resultados referentes às distribuições relativas.

### *2.4.2 Distribuição relativa do consumo entre as regiões Nordeste e Sudeste*

Esta seção tem início com a exibição das estimativas das funções densidades de probabilidade (FDPs) kernel do logaritmo do consumo total *per capita* para as regiões Nordeste e Sudeste, nos biênios 2002-2003 e 2008-2009 (Figura 1). A linha vertical e a horizontal nas FDPs identificam os componentes da distribuição relativa.

Figura 1 – Funções Densidade de Probabilidade Kernel do logaritmo do consumo, Nordeste e Sudeste: (a) 2002-2003; e (b) 2008-2009



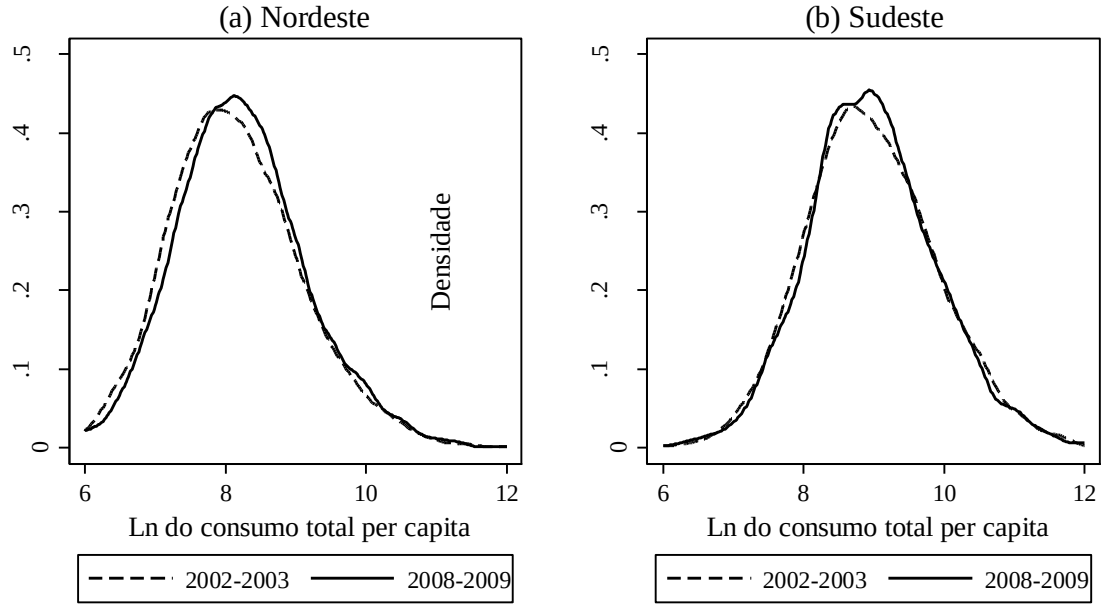

Fonte: elaboração própria a partir dos dados da POF.

Observa-se, nos gráficos (a) e (b) da Figura 1, que a mediana do logaritmo do consumo da região Sudeste, dada pelo centro da densidade de probabilidade, é maior que a da região Nordeste nos dois biênios. Para o logaritmo do consumo menor do que 8,4, observa-se que a densidade das famílias da região Nordeste é maior do que as da região Sudeste, isso significa que há uma maior incidência desses gastos mais baixos na região Nordeste. A posição favorável da região Sudeste em relação à região Nordeste fica mais evidente quando observamos que a densidade da região Sudeste é maior para o logaritmo do consumo acima de 8,4, isto é, a concentração dos gastos mais elevados se encontra na região Sudeste. No entanto, embora seja difícil de perceber, houve uma leve aproximação das medianas no segundo biênio.

A Figura 2 apresenta as densidades de probabilidade do logaritmo do consumo total *per capita* nos dois biênios, separadamente para a região Nordeste e Sudeste, para ser possível observar uma evolução no tempo. A aproximação das medianas no segundo biênio, percebida na Figura 1, deveu-se, sobretudo, ao aumento da mediana na região Nordeste. Isso pode ser um reflexo tanto do "boom do consumo" como da transição das classes sociais que houve nesse período, com uma redução do número de pessoas pertencentes as classes mais pobres e um aumento da classe média (NERI, 2011).

Em comparação com a sobreposição direta de FDPs como nas Figuras 1 e 2, a densidade relativa simplifica a comparação entre duas FDPs de várias maneiras, pois não requer que o visualizador construa as diferenças entre as duas curvas em cada ponto da escala. A densidade relativa codifica essa comparação diretamente em termos de uma proporção, que é a razão entre a densidade do grupo de comparação e a densidade do grupo de referência (CHAMBERS *et al*., 1983; CLEVELAND; MCGILL, 1984).

Figura 2 – Funções Densidade de Probabilidade Kernel do logaritmo do consumo: (a) Nordeste; e (b) Sudeste



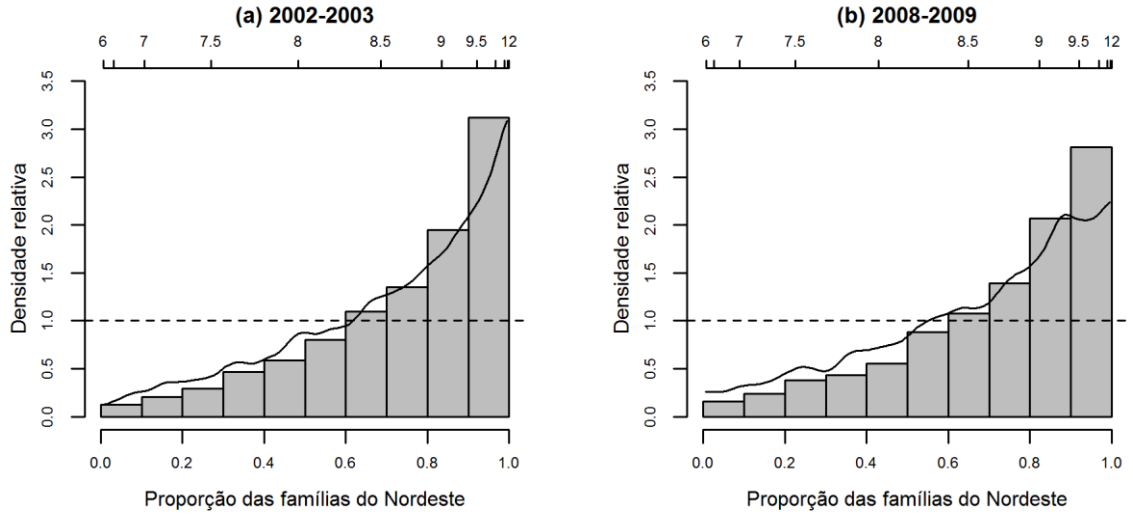

Fonte: elaboração própria a partir dos dados da POF.

Conforme o gráfico (a) da Figura 3, para valores nos seis décimos inferiores de $Y_0$ ($r < 0{,}6$), a densidade relativa é menor que 1, indicando uma frequência menor de observações da distribuição de comparação $Y$ (Sudeste) em relação ao grupo de referência $Y_0$ (Nordeste), ou seja, há menos famílias na região Sudeste nos níveis mais baixos de gastos em comparação à região Nordeste. A densidade relativa no segundo decil, por exemplo, é de, aproximadamente, 0,25 ($g(0{,}2) = 0{,}25$). Isso implica que o percentual de famílias da região Sudeste nesse nível de consumo é de cerca de 75% menor do que o das famílias da região Nordeste (encontra-se um resultado semelhante em 2008-2009). Nos decis restantes o valor é maior que 1, indicando uma frequência maior de observações de $Y$ em relação a $Y_0$, isto é, há uma prevalência de famílias da região Sudeste nos níveis mais altos da distribuição de consumo da região Nordeste. Pode-se ver que a densidade relativa no último décimo é de cerca de 3,1 (isto é, $g\,(1{,}0) = 3{,}1$). Isso indica que o percentual de famílias da região Sudeste nesse nível de consumo é de cerca de 210% maior do que o das famílias da região Nordeste. Em 2008-2009, o percentual se reduz para 190% ($g\,(1{,}0) = 2{,}9$).

É possível concluir, através dos gráficos (a) e (b) da Figura 3 que, entre os biênios 2002-2003 e 2008-2009, não houve uma mudança significativa na densidade relativa, salvo uma redução no último décimo da distribuição do consumo.

Hoffmann (2010) mostra, a partir dos dados da POF, que houve uma queda na proporção da renda apropriada pelo décimo mais rico nesse período, diminuindo de 47,1% para 44,4%, confirmando um fenômeno já constatado com os dados da PNAD. Assim sendo, podemos inferir que esse acontecimento afetou mais as famílias do Sudeste, refletindo-se numa redução da sua participação nos maiores níveis de consumo comparativamente às famílias do Nordeste.

Figura 3 – Distribuição relativa do logaritmo do consumo total entre as regiões Nordeste e Sudeste




Fonte: elaboração própria a partir dos dados da POF.

A Figura 4 exibe as densidades relativas do logaritmo da despesa em habitação, saúde, transporte e alimentação, respectivamente. É possível notar que os três primeiros grupos de despesa seguem o mesmo padrão da Figura 3. Assim sendo, quando a densidade relativa é menor que 1 (nos seis primeiros decis da distribuição) há uma sub-representação das observações do grupo de comparação (Sudeste) em relação ao grupo de referência (Nordeste) e uma sobrerrepresentação quando a densidade relativa é maior que 1, evidenciando uma superioridade da região Sudeste em relação aos gastos nessas áreas.

Figura 4 – Distribuição relativa do logaritmo da despesa em habitação, saúde, transporte e alimentação, respectivamente, entre as regiões Nordeste e Sudeste. *Continua*.

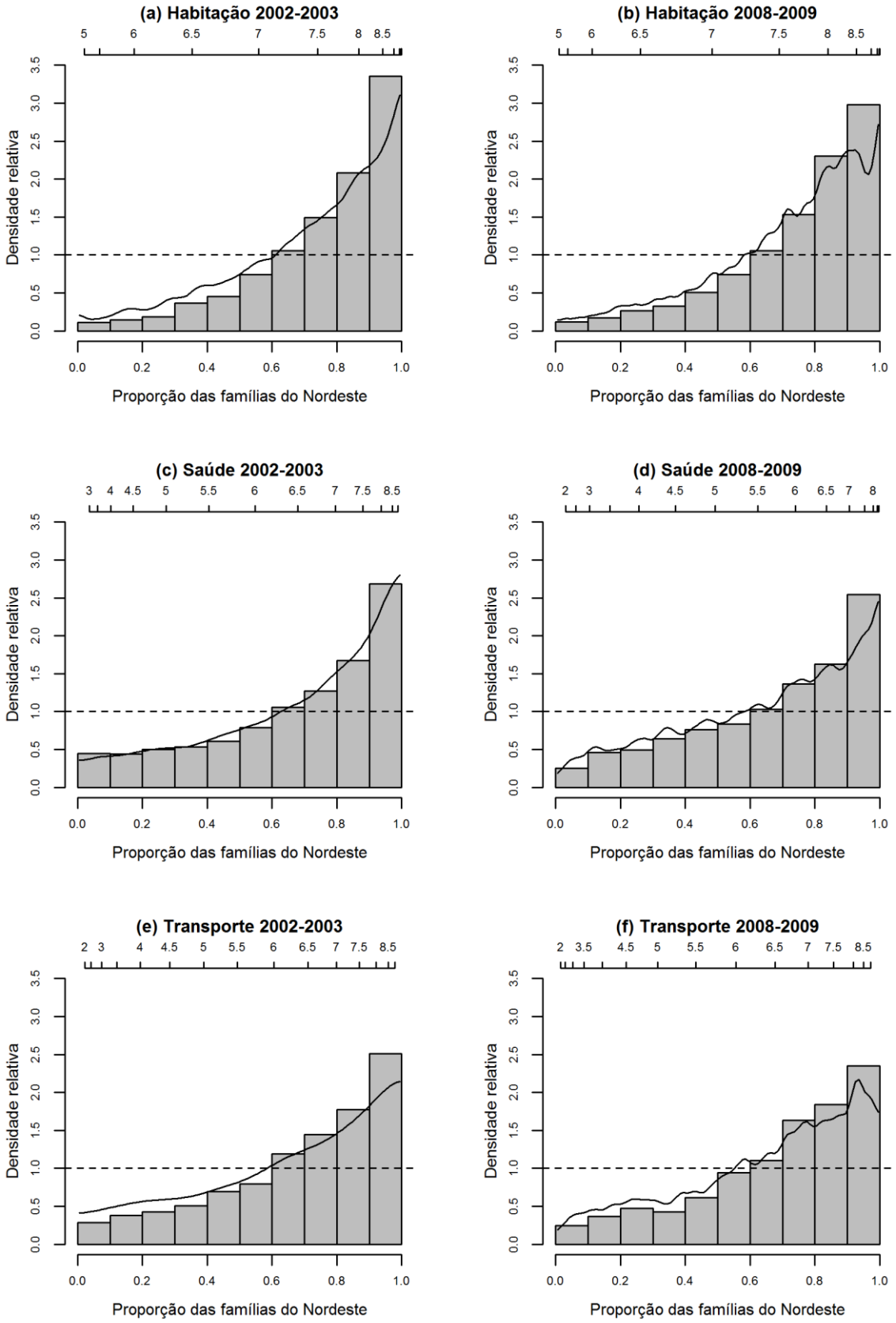
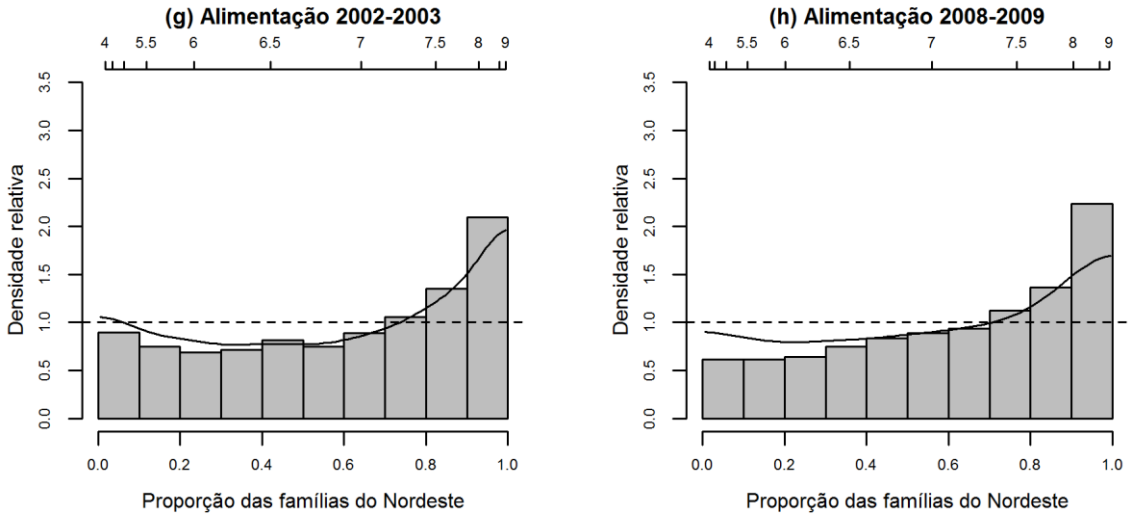


Figura 4 – Distribuição relativa do logaritmo da despesa em habitação, saúde, transporte e alimentação, respectivamente, entre as regiões Nordeste e Sudeste. *Conclusão*.




Fonte: elaboração própria a partir dos dados da POF.

Entretanto, a densidade relativa da despesa em alimentação mostra um resultado diferente. Nos sete primeiros decis, a densidade relativa está muito próxima de 1, assim, a densidade de pessoas nos grupos de referência e comparação é praticamente a mesma para os quantis em questão, indicando uma equivalência distribucional. No entanto, nos decis posteriores, notadamente no último (faixa de despesa em alimentação mais elevada), há um predomínio das famílias da região Sudeste.

Dessa forma, foi possível observar que as famílias da região Sudeste, ao longo do período analisado, estiveram numa posição mais favorável, apresentando sempre um nível de consumo mais elevado.

### *2.4.3 Decomposição da distribuição relativa em efeito localização e efeito forma e índices de entropia e polarização*

A densidade relativa no quantil $r$ de $Y_0$ é a distância entre $f_0(y_r)$ e $f(y_r)$. Tal distância é baseada no fato de que a distribuição de comparação pode estar situada à esquerda ou à direita da distribuição de referência (efeito localização). A distância, no entanto, também pode depender da curvatura da distribuição, a qual pode enfatizar ou suavizar alguns picos ao longo

dos quantis. Os picos capturam a polarização (efeito forma). Os dois efeitos são obtidos pela aplicação de uma decomposição da distância relativa (PETRARCA; RICCIUTI, 2015).

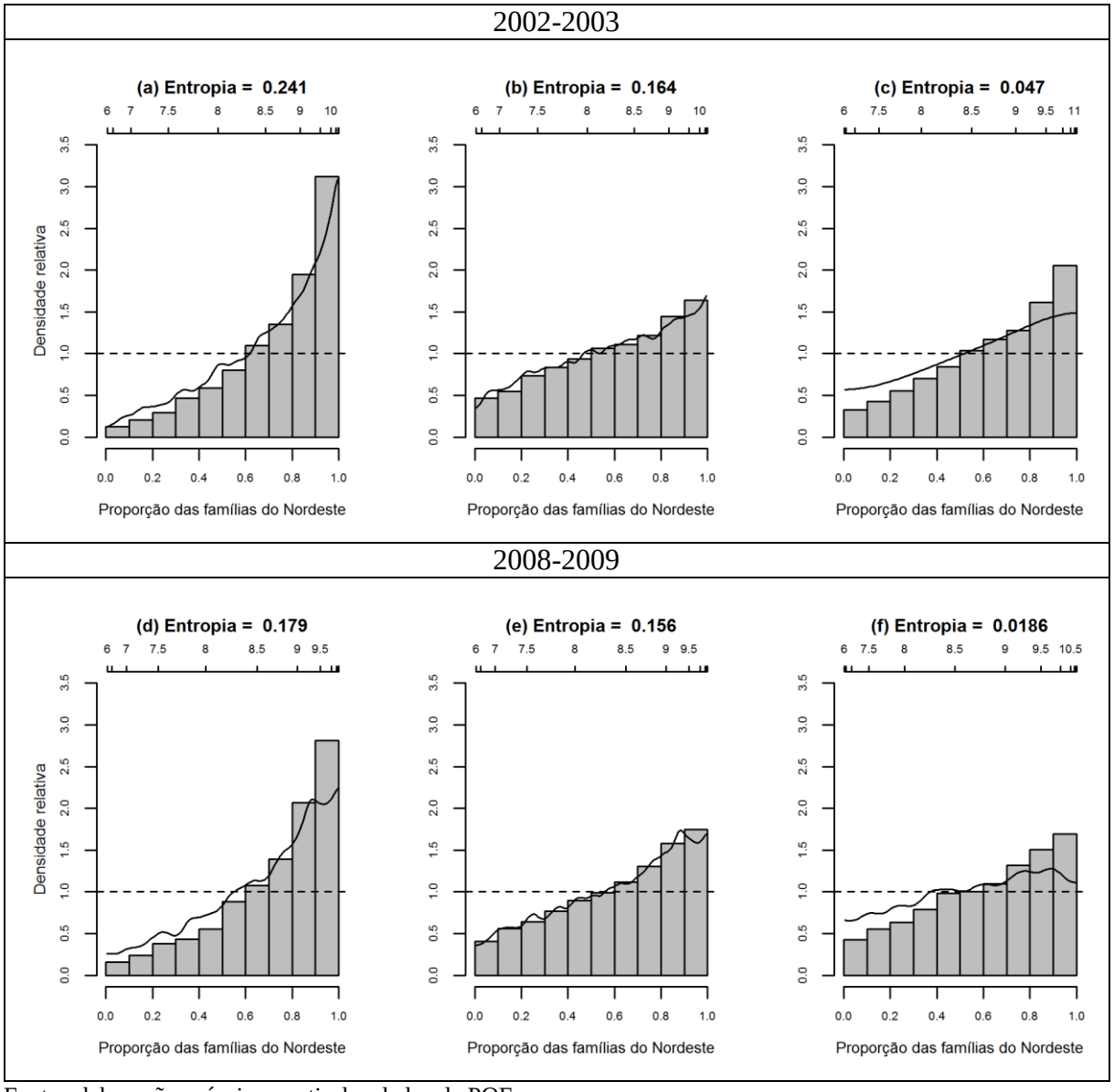
Nesta seção, são apresentados os resultados das decomposições das distribuições relativas em efeito localização e efeito forma. A Figura 5 compara a distribuição do logaritmo do consumo total *per capita* da região Sudeste em relação à distribuição da região Nordeste, tanto em 2002-2003 quanto em 2008-2009 e é exposta da seguinte maneira: os gráficos (a) e (d) mostram a densidade relativa global do grupo de comparação na escala do grupo de referência ($g_0(r)$); os gráficos (b) e (e) mostram a densidade relativa que isola o efeito de diferenças na localização ($g_0^h(r)$), exibindo como seria a densidade relativa se não houvesse diferença na forma da distribuição; e os gráficos (c) e (f) reportam a densidade relativa que expressa as diferenças na forma ($g_h(r)$). O índice de entropia relativa, indicado na parte superior dos gráficos, é a medida de dissimilaridade das distribuições apresentada na seção 2.3.4, em que os gráficos (a) e (d) apresentam a divergência total, (b) e (e) divergências devido a diferenças na localização e (c) e (f) divergências devido a diferenças na forma.

O principal efeito observado para as dissimilaridades no consumo entre as regiões é aquele decorrente de diferenças na mediana (ou localização), com a maior parte explicada da entropia total (0,164 de um total de 0,241 em 2002-2003, 0,156 de um total de 0,179 em 2008-2009). Além disso, a comparação dos índices de entropia entre os biênios revela que houve uma redução na divergência entre as distribuições de consumo das regiões (0,241 para 0,179).

Como visto na seção anterior, o gráfico (a) da Figura 5 mostra que há uma menor densidade de famílias da região Sudeste nos décimos inferiores da distribuição da região Nordeste e uma maior densidade relativa nos décimos superiores. A proporção do décimo superior chega a ser cerca de 3,1, isto é, que há um número de famílias com consumo mais abastado na região Sudeste 3,1 vezes maior do que na região Nordeste. O gráfico (b) revela que as famílias da região Sudeste tem um nível de consumo maior do que o das famílias da região Nordeste ao longo de toda a distribuição no biênio 2008-2009. Na Figura 1, gráfico (a), isso é exposto pela linha tracejada sempre no lado direito da linha cheia. Portanto, presumindo-se que a forma da distribuição permaneça inalterada, observamos mais uma vez um número de famílias na região Sudeste com um nível de consumo mais elevado que as famílias da região Nordeste. No entanto, a proporção diminui para 1,7 no último decil, sugerindo uma importância relevante do efeito de localização.

Assumindo a sobreposição das curvas na Figura 1, o gráfico (c) da Figura 5 sinaliza uma diferença na curvatura das distribuições. Há evidências de uma polarização positiva na cauda superior e uma polarização negativa na cauda inferior, dessa forma, há menos famílias com níveis de consumo muito baixos e mais famílias com consumo muito elevados na região Sudeste, dada a escala de consumo observada na região Nordeste. A análise para o biênio 2008-2009 é bem semelhante.

Figura 5 – Decomposição da distribuição relativa do log-consumo total entre as regiões Sudeste e Nordeste: (a) e (d) densidade relativa global; (b) e (e) efeito localização; (c) e (f) efeito forma




Fonte: elaboração própria a partir dos dados da POF.

A Tabela 2 apresenta as medidas numéricas resumidas dos índices de polarização relativa: na mediana (PRM), na cauda inferior (PRI) e na cauda superior (PRS). Em termos de interpretação, quando o índice PRM assume o valor zero indica que não há diferença na forma das duas distribuições, valores positivos apontam para uma maior polarização ou a uma maior densidade nas caudas das distribuição do grupo de comparação, enquanto valores negativos assinalam uma polarização negativa ou uma maior concentração da distribuição do grupo de comparação em torno da mediana. A análise dos valores positivos e negativos do PRI ou PRS é similar ao PRM. No caso de assumir o valor zero, no entanto, o componente da distribuição relativa é uniforme abaixo ou acima da mediana.

A partir do índice de polarização relativa na mediana (PRM), pode-se constatar que há uma maior desigualdade no consumo na região Sudeste em comparação à região Nordeste. Essa desigualdade tem origem na cauda superior, enquanto a cauda inferior contribui para reduzir a desigualdade. No entanto, O índice PRM, apesar de positivo, está muito próximo de zero. Isso implica que quase não há diferença na forma entre as distribuições e essa diferença é praticamente nula no biênio 2008-2009. Os valores são estatisticamente significantes.

Conclui-se, portanto, que a redução na divergência entre as distribuições se deveu prioritariamente a uma aproximação das medianas (como as funções kernel já sinalizavam), indicando uma convergência no padrão de vida entre as regiões no período considerado. Este resultado está em consonância com o de Oliveira *et al*. (2017), que, através de uma análise da Curva de Lorenz Generalizada, encontram que o bem-estar social no Brasil é maior em 2008-2009 para uma classe abrangente de funções que valorizam não só os incrementos de consumo, como também transferências progressivas.

Tabela 2 – Índices de polarização. Nordeste *versus* Sudeste

| Índices de polarização | Estimativa | Intervalo de confiança (5%) | | P-valor |
|---|---|---|---|---|
| | | 2002-2003 | | |
| Índice na mediana (PRM) | 0,035 | 0,050 | 0,065 | 0,000 |
| Índice na cauda inferior (PRI) | -0,217 | -0,189 | -0,160 | 0,000 |
| Índice na cauda superior (PRS) | 0,115 | 0,145 | 0,174 | 0,000 |
| | | 2008-2009 | | |
| Índice na mediana (PRM) | 0,005 | 0,018 | 0,031 | 0,003 |
| Índice na cauda inferior (PRI) | -0,187 | -0,162 | -0,137 | 0,000 |
| Índice na cauda superior (PRS) | 0,085 | 0,110 | 0,135 | 0,000 |

Fonte: elaboração própria a partir dos dados da POF.

## 2.5 Considerações finais

Medir o padrão de vida é essencial para mensurar a pobreza e a desigualdade e verificar suas evoluções ao longo do tempo. O consumo é a melhor medida do componente econômico do padrão de vida e, embora seja uma medida de padrões de vida mais útil e precisa do que a renda, não há muitos trabalhos acadêmicos com foco na desigualdade de consumo.

No intuito de contribuir com essa área de pesquisa, este artigo teve como objetivo comparar o bem-estar das famílias das regiões Nordeste e Sudeste a partir do exame das distribuições do consumo *per capita*. Para isso, foram utilizados métodos baseados na distribuição relativa proposta por Handcock e Morris (1999) e que consistem em ferramentas gráficas para uma análise exploratória dos dados, índices de entropia relativa e índices de polarização relativa, tendo como fonte de informação a Pesquisa de Orçamentos Familiares (POF) dos biênios 2002-2003 e 2008-2009.

Os resultados encontrados apontaram que as famílias da região Sudeste apresentam um nível de consumo mais elevado em todos os decis da distribuição, ficando sempre numa posição mais favorável ao longo do período analisado. Isto se verifica para todos os subgrupos de despesa, notadamente para a média de gastos com educação no biênio 2002-2003, a qual é quase três vezes maior do que a constatada na região Nordeste.

Observando as funções densidade de probabilidade das regiões, percebeu-se uma leve aproximação das medianas no segundo biênio, a qual deveu-se, sobretudo, ao aumento da mediana na região Nordeste. Isso pode ser um reflexo tanto do "boom do consumo" como da transição das classes sociais que houve nesse período.

Analisando os índices de entropia relativa, o principal efeito observado para as dissimilaridades no consumo entre as regiões foi aquele decorrente de diferenças na mediana (ou localização), com a maior parte explicada da entropia total. Além disso, a comparação dos índices de entropia entre os biênios revelou que houve uma redução na divergência entre as distribuições de consumo das regiões.

A partir do índice de polarização relativa na mediana (PRM), pôde-se constatar que há uma maior desigualdade no consumo na região Sudeste em comparação à região Nordeste. Essa desigualdade tem origem na cauda superior, enquanto a cauda inferior contribui para uma aproximação da distribuição à sua mediana. No entanto, O índice PRM, apesar de positivo, está

muito próximo de zero. Isso implica que quase não há diferença na forma entre as distribuições e essa diferença é praticamente nula no biênio 2008-2009.

Portanto, podemos concluir que essa redução na divergência entre as distribuições foi devido prioritariamente à uma aproximação das medianas, revelando uma convergência no padrão de vida entre as regiões no período considerado.

# 3 GAP REGIONAL NOS GASTOS EM CONSUMO: APLICAÇÃO DO MODELO RIF

## 3.1 Introdução

Historicamente, o Brasil apresenta desigualdades regionais marcantes. Essas disparidades são originárias da forma como ocorreu a colonização, pela maneira como foi conduzida a industrialização e por fatores políticos. O processo de industrialização teve início no final do século XIX na região Sudeste, a qual, devido ao cultivo e à exportação do café, representava o polo dinâmico da economia. O crescimento econômico que se estabeleceu no País nos anos 1930 e de maneira mais acelerada e persistente após os anos 1950 beneficiou, de um modo mais acentuado, esta região (CACCIAMALI, 2002). O resultado foi uma elevação da concentração de renda no País nos anos 1960, desencadeando um intenso debate sobre as principais causas que motivaram o aumento da desigualdade, com uma contribuição especial nesse sentido de Celso Furtado.

Há uma continuidade na tendência concentradora nas décadas de 1970 e 1980, porém um pouco menor do que na década de 1960. No entanto, o problema da inflação e da dívida externa na década de 1980 fez o debate sobre as causas da desigualdade no Brasil perder importância, reaparecendo apenas na década de 1990, quando o tema foi novamente colocado em pauta pela literatura de crescimento econômico, com a discussão sobre convergência de renda (AZZONI, 1997).

Inferiu-se que as desigualdades regionais são explicadas por fatores relacionados a oferta de trabalho, assim como fatores relacionados a demanda por trabalho. Dentre os fatores do lado da demanda por trabalho estão as diferenças na estrutura produtiva, nas instituições e no acesso à tecnologia. Já no lado da oferta de trabalho, incluem-se as diferenças nas características produtivas dos indivíduos presentes em cada região, notadamente para o nível de escolaridade (OLIVEIRA, SILVEIRA NETO, 2017).

Neste sentido, é possível verificar, do lado da demanda por trabalho, um processo de desconcentração da atividade econômica entre as regiões do País, sobretudo industrial, desde meados da década de 1990. Além disso, nota-se uma intensificação do processo de liberalização comercial e no aumento de privatizações, resultando na incorporação de novas tecnologias na produção e, segundo Menezes Filho e Rodrigues (2003), no aumento da demanda por

trabalhadores mais qualificados. No lado da oferta de trabalho também tem havido mudanças. Tais mudanças estão ligadas à ampliação da participação das mulheres no mercado de trabalho, como mostram Melo e Monte (2015), e um aumento expressivo no nível de escolarização dos trabalhadores brasileiros, de acordo com Pauli, Nakabashi e Sampaio (2012).

Há uma série de trabalhos analisando as desigualdades regionais sob o ponto de vista da renda, buscando entender não apenas a dinâmica da desigualdade, mas as circunstâncias que determinam seu nível (MELO; MONTE, 2015; OLIVEIRA; SILVEIRA NETO, 2017; SILVEIRA NETO; MENEZES, 2008). Não obstante, o consumo é uma medida muito mais exata e precisa do que a renda para medir diferenças nos padrões de vida. A renda representa um padrão de vida potencial, pois mede, após os impostos, a capacidade que uma família ou indivíduo tem para alcançar um determinado padrão de vida. No entanto, as pessoas destinam parte de sua renda à poupança, ao pagamento de dívidas, doação para caridade, além de ser possível emprestar ou tomar emprestado, reduzindo ou expandindo o padrão de vida para aquém ou além do que a renda permite. Já o consumo representa exatamente o que elas tiveram acesso (SARLO, 2016).

Tendo em vista o exposto, este artigo tem como objetivo investigar os fatores associados à desigualdade na distribuição do consumo familiar *per capita* no âmbito das regiões Nordeste e Sudeste do Brasil no período 2002 a 2009, utilizando elementos ligados à demanda e à oferta de trabalho. Para isso, usa os dados fornecidos pelas Pesquisas de Orçamentos Familiares (POFs) do IBGE dos biênios 2002-2003 e 2008-2009.

Para atingir o objetivo, será utilizado o método de decomposição baseada na *Recentered Influence Function* (RIF) *Regression* proposta por Firpo, Fortin e Lemieux (2009). Trata-se de uma abordagem de regressão quantílica incondicional, a qual ajudará a entender como as diferenças nas distribuições de características domiciliares (efeito composição) e nos efeitos marginais dessas características (efeito estrutura) contribuem para a desigualdade de consumo entre as regiões Nordeste e Sudeste do Brasil em diferentes pontos da distribuição.

Além desta introdução, este artigo possui quatro seções adicionais. A segunda seção faz uma revisão dos principais trabalhos sobre desigualdade no consumo, focando nas metodologias empregadas e nos resultados encontrados. A terceira seção traz uma descrição da base de dados e da metodologia empregada. A quarta seção realiza uma análise preliminar dos dados sobre as características sociodemográficas das famílias e examina os resultados obtidos

através da decomposição quantílica incondicional. Por fim, a última seção é destinada às considerações finais do trabalho.

### 3.2 Desigualdade no consumo: evidências empíricas

Os estudos que medem disparidades de bem-estar entre as famílias, utilizando, dentre os indicadores, a desigualdade no consumo ou nas despesas, geralmente empregam medidas tradicionais de desigualdade, como os índices de Gini, Theil e Atkinson (DEV, 2016; FIRPO, 2000; HOFFMANN, 2010; JIRASAVETAKUL, LAKNER, 2016; KAPLANOGLOU, RAPANOS, 2018; MARTINS, 2011; SARLO, 2016).

Além de calcular indicadores tradicionais de desigualdade, o trabalho de Silveira Neto e Menezes (2010) faz uma análise de dominância de Lorenz para o Brasil metropolitano e encontra evidências de diminuição da desigualdade dos gastos e do consumo para a região metropolitana do Recife no período 1996-2003, não sendo verificada a mesma tendência para São Paulo. Já Santos (2008) realiza uma ordenação das distribuições de consumo das regiões brasileiras através não apenas da Curva de Lorenz e da Curva de Lorenz generalizada, mas também da Parada de Pen, constatando que a região Nordeste possui o menor consumo entre todas as regiões.

Attanasio, Hurst e Pistaferri (2015) se diferenciam dos demais ao empregar o desvio-padrão do logaritmo do consumo para medir a desigualdade no consumo nos Estados Unidos. Ademais, para fornecer um resultado mais completo do que aconteceu com o consumo, os autores também analisam a diferença entre o percentil 90 e a mediana (topo da distribuição), bem como a diferença entre a mediana e o percentil 10 (parte inferior da distribuição).

Oliveira *et al*. (2017) avançam nessa área adotando funções de bem-estar social para verificar o impacto da expansão do consumo sobre a dinâmica do bem-estar das famílias brasileiras no período 2002-2009. As funções de bem-estar social são representadas pelas médias Sen (associadas ao índice de Gini) e geométrica (associada ao índice Atkinson) e apontam que o principal motor do aumento do bem-estar foi o crescimento do consumo e não sua redistribuição.

Outros trabalhos, assim como este artigo, buscam entender não apenas o nível e a evolução da desigualdade no consumo, mas também os fatores associados às desigualdades. Nesse sentido, Tripathi (2016) faz um estudo de decomposição da desigualdade baseada numa análise de regressão desenvolvida por Morduch e Sicular (2002), Fields (2003) e Fiorio e Jenkins (2007) em busca dos principais determinantes da desigualdade nas despesas com consumo na Índia, nos âmbitos rural e urbano separadamente. Os resultados mostram que o tamanho do agregado familiar é responsável pela maior parcela da desigualdade. Além disso, fatores como um nível mais alto de educação, participação de trabalhadores ocupados em empregos menos produtivos (como trabalhador casual e trabalhador agrícola), um maior nível de terra possuído pelas famílias e famílias com unidades habitacionais alugadas também estão contribuindo para um maior nível de desigualdade tanto na área rural quanto na área urbana.

Hassine e Zeufack (2015) examinam a estrutura e a dinâmica da desigualdade de consumo na Tanzânia, além de verificarem o grau de desigualdade de oportunidade, no período 2001 a 2012. Para a investigação empírica das fontes das disparidades regionais e urbano-rural na Tanzânia, os autores usam a abordagem de regressão da *Recentered Influence Function* (RIF) proposta por Firpo, Fortin e Lemieux (2009). Os resultados revelam que as diferenças na distribuição das características demográficas dos agregados familiares e no acesso aos serviços básicos, seguidas por diferenças no setor de emprego do chefe são as que mais importam para a desigualdade entre as famílias urbanas e rurais. Em relação à diferença nos retornos, os fatores que mais contribuem para a lacuna rural-urbana são a posse de ativos e o setor de emprego do chefe. Já para o gap entre as regiões metropolitanas e não metropolitanas, as divergências na distribuição das características demográficas familiares e dotação de capital humano, além do acesso desigual a ativos privados e empregos produtivos limitaram a capacidade dos pobres de aproveitar as oportunidades geradas pelo crescimento econômico e de melhorar seu padrão de vida.

A abordagem da regressão recentrada na função influência (RIF), proposta por Firpo, Fortin e Lemieux (2009), fornece um procedimento baseado numa regressão simples para a realização de uma decomposição detalhada das diferentes estatísticas de distribuições, tais como quantis, variância e coeficiente de Gini. Quando aplicado aos quantis, o modelo de regressão RIF é chamado de regressão quantílica incondicional. Os outros métodos convencionais de decomposição, como o procedimento de reponderação de DiNardo, Fortin e Lemieux (1996) e a abordagem de decomposição baseada em quantis de Machado e Mata (2005), estendem a decomposição tradicional de Oaxaca-Blinder para além da média e

permitem a análise de toda a distribuição, no entanto, envolvem uma série de pressupostos e dificuldades computacionais (FORTIN; LEMIEUX; FIRPO, 2011). Assim sendo, o método RIF foi escolhido para investigar as fontes das desigualdades de bem-estar neste artigo, contribuindo com a literatura brasileira, a qual não possui trabalhos desta natureza utilizando dados de consumo.

## 3.3 Metodologia

Nesta seção, apresenta-se o método escolhido para a análise das fontes das desigualdades no bem-estar (medido pelo consumo): o *Recentered Influence Function* (RIF) *regression*, proposto por Firpo, Fortin e Lemieux (2009). Antes, porém, exibe-se a descrição da base de dados empregada neste estudo.

### *3.3.1 Base de dados*

A base de dados deste estudo é originária das Pesquisas de Orçamentos Familiares (POFs) realizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) nos biênios 2002-2003 e 2008-2009. As finalidades principais das POFs são: disponibilizar informações sobre a composição orçamentária doméstica e sobre as condições de vida da população, incluindo a percepção subjetiva da qualidade de vida, bem como gerar bases de dados e estudos sobre o perfil nutricional da população (IBGE, 2010).

As POFs foram realizadas em quatro períodos distintos: 1987-88, 1995-96, 2002-03 e 2008-09. A escolha das duas últimas pesquisas deve-se a dois fatores. Primeiro, as duas primeiras POFs foram realizadas apenas nas regiões metropolitanas brasileiras de Belém, Belo Horizonte, Curitiba, Fortaleza, Porto alegre, Recife, Rio de Janeiro, Salvador, São Paulo, a cidade de Goiânia e o Distrito Federal, enquanto nas duas últimas a POF passou a ter abrangência nacional, entrevistando 44.470 domicílios em 2002-03 e 55.970 em 2008-09 em todos os estados brasileiros. Segundo, nas POFs de 1987-88 e 1995-96 foram geradas apenas informações monetárias, já nas duas últimas existem dados referentes à renda monetária e não monetária.

As duas primeiras POFs foram concebidas prioritariamente para atender à atualização das estruturas de consumo dos índices de preços ao consumidor do IBGE. As duas últimas edições, por sua vez, buscaram relacionar a estrutura das despesas não só com os níveis de rendimentos, mas também com características socioeconômicas dos moradores como condições de habitação, escolaridade, origem dos rendimentos e perfis antropométricos. Além disso, houve um grande empenho para descrever com maior precisão os gastos com saúde e o consumo de alimentos das famílias.

O domicílio, isto é, uma moradia estruturalmente separada e independente, é a unidade amostral da pesquisa. Entretanto, a unidade básica de investigação aqui considerada é a unidade de consumo (UC), a qual compreende um único morador ou um conjunto de moradores que compartilham a mesma fonte de alimentação ou compartilham as despesas com moradia. Dessa forma, um domicílio pode compreender mais de uma unidade de consumo, embora esses casos tenham correspondido em apenas cerca de 5% dos domicílios na edição de 2002-2003 e a 2% na edição seguinte. Para efeito de divulgação de resultados, o termo família é utilizado para representar o conceito de unidade de consumo.

A POF elenca os moradores de cada domicílio e das respectivas UCs. O chefe da família é aquele responsável pela maior parte dos gastos com habitação e os demais moradores são listados de acordo com a sua relação ou parentesco com o chefe da família. Como a unidade de análise proposta neste trabalho é a família (unidade de consumo), foram consideradas como características individuais as informações do chefe, ao passo que as variáveis de despesas com consumo foram agregadas por UC. As variáveis utilizadas na regressão estão expostas no Quadro 1.

A presença de gestantes, crianças, adolescentes, jovens e idosos aumenta a vulnerabilidade das famílias porque aumenta o volume de recursos necessários para a satisfação de suas necessidades básicas. Quando o número de crianças é proporcionalmente elevado em relação ao número de adultos, a família passa a ocupar uma posição desvantajosa, pois muitos dependem da renda de poucos. Por esse motivo, a variável proporção de crianças na família foi incluída na análise.

A variável consumo compreende as despesas com alimentação, habitação, vestuário, transporte, higiene e cuidados pessoais, assistência à saúde, educação, recreação e cultura, fumo, serviços pessoais e outras despesas diversas não classificadas anteriormente. A correção

dos valores monetários foi feita utilizando o Índice Nacional de Preços ao Consumidor (INPC) para todos os valores serem expressos em reais de janeiro de 2009.

Quadro 1 – Descrição das variáveis selecionadas

| | |
|---|---|
| Dependente | |
| Consumo | Logaritmo do consumo familiar *per capita* |
| Explicativas | |
| *Características do chefe da família* | |
| Homem | Variável binária indicativa de gênero do chefe, sendo 1 para masculino e 0 para feminino |
| Cor branca | Variável binária indicativa da cor/raça do chefe, sendo 1 para brancos e 0 caso contrário |
| Idade | Anos de idade |
| Escolaridade | Anos de estudo |
| Setor de atividade | Dividida em seis categorias: Agropecuária, Indústria, Construção, Comércio, Serviços e Desocupado (indica que o chefe está desocupado). |
| *Características demográficas e gerais do domicilio* | |
| Área urbana | Variável binária indicativa de área censitária, sendo 1 para área urbana e 0 caso contrário |
| Tamanho da família | Quantidade de pessoas na unidade de consumo |
| Crianças | Proporção de crianças na unidade de consumo |

Fonte: elaboração própria.

### *3.3.2 Método RIF*

Firpo, Fortin e Lemieux (2009) propõem um método de regressão para avaliar o impacto de mudanças na distribuição das variáveis explicativas sobre os quantis da distribuição (marginal) incondicional de uma variável de resultado. O método proposto consiste em executar uma regressão de uma transformação da variável de resultado – a função de influência recentrada (RIF) do quantil incondicional – nas variáveis explicativas.

O método de regressão quantílica incondicional ajudará a entender como as diferenças nas distribuições de características domiciliares observadas e nos efeitos marginais dessas características contribuem para a desigualdade de consumo entre as regiões Nordeste e Sudeste do Brasil em diferentes pontos da distribuição.

O procedimento é realizado em duas etapas. A primeira consiste em estimar regressões quantílicas incondicionais no logaritmo do consumo familiar *per capita* para cada região, em seguida, construir uma distribuição contrafactual que prevaleceria se as famílias da região Nordeste recebessem os retornos que dizem respeito à região Sudeste. A comparação entre a distribuição contrafactual e a empírica nos permite estimar a parte do hiato de bem-estar atribuível às diferenças nas características das famílias (o efeito composição), e a parte explicada pelas diferenças nos retornos a essas características (o efeito estrutura). A segunda etapa envolve a divisão dos componentes no intuito de verificar a contribuição de cada variável característica específica na desigualdade.

O método pode ser facilmente estimado como uma regressão de Mínimos Quadrados Ordinários (MQO) da seguinte forma:

$$RIF(y, Q_\theta) = X\beta + \varepsilon \tag{3.1}$$

onde $X$ é a matriz de regressores, $\beta$ é o vetor de coeficientes da regressão, incluindo o intercepto, $\varepsilon$ é o termo de erro, $y$ é o logaritmo do consumo familiar *per capita* e $RIF(y, Q_\theta)$ é a função de influência recentrada do $\theta$-ésimo quantil de $y$, estimada computando o quantil amostral $Q_\theta$ e estimando a densidade de $y$ no ponto pelo método de kernel:

$$RIF(y, Q_\theta) = Q_\theta + \frac{(\theta - I\{y \leq Q_\theta\})}{f_Y(Q_\theta)} \tag{3.2}$$

onde $f_Y$ é a função densidade marginal de $y$ avaliada em $Q_\theta$ e $I$ é uma função indicadora se a variável resposta é menor ou igual ao quantil $Q_\theta$. $RIF$ pode ser estimada através da substituição de $Q_\theta$ pelo $\theta$-ésimo quantil da amostra e estimação de $f_Y$ pela densidade kernel.

Estima-se a Equação (3.1) para os quantis de 10 a 90 e utiliza-se as estimativas das regressões quantílicas incondicionais para se decompor a desigualdade de bem-estar em um componente atribuível a diferenças na distribuição de características e um componente devido a diferenças na distribuição dos coeficientes da seguinte forma:

$$\hat{Q}_\theta^A - \hat{Q}_\theta^B = \{\hat{Q}_\theta^A - \hat{Q}_\theta^*\} + \{\hat{Q}_\theta^* - \hat{Q}_\theta^B\} = (\bar{X}^A - \bar{X}^B)\hat{\beta}_\theta^A + \bar{X}^B(\hat{\beta}_\theta^A - \hat{\beta}_\theta^B) \tag{3.3}$$

onde $\hat{Q}_\theta$ é o $\theta$-ésimo quantil incondicional do logaritmo do consumo familiar *per capita*, $\bar{X}$ representa o vetor de médias das covariáveis e $\hat{\beta}_\theta$ é a estimativa do efeito parcial do quantil

incondicional. Os sobrescritos $A$, $B$ e $*$ designam, respectivamente, os valores da região Sudeste, Nordeste e contrafactuais. $\hat{Q}_\theta^* = X^B\hat{\beta}^A$ é o quantil contrafactual da distribuição contrafactual incondicional que representa a distribuição do bem-estar que teria prevalecido para o grupo $B$ (famílias da região Nordeste) se eles tivessem recebido o retorno às características do grupo $A$ (famílias da região Sudeste).

O primeiro termo do lado direito da equação (3.3) representa a contribuição das diferenças nas distribuições das características dos agregados familiares para a desigualdade no θ-ésimo quantil incondicional, denotado como efeito composição (ou parte explicada). O segundo termo do lado direito da equação representa a desigualdade devido a diferenças nos retornos das características do agregado familiar no θ-ésimo quantil incondicional (incluindo diferenças no intercepto), denominado efeito estrutura (ou parte não explicada dos diferenciais de consumo). Nos casos em que a filiação a um grupo está ligada a algumas características imutáveis dos trabalhadores, como raça ou sexo, o efeito da estrutura salarial também tem sido chamado de parte devido à discriminação.

Os efeitos composição e estrutura podem ser ainda mais decompostos na contribuição de cada característica do agregado familiar (ou grupo de algumas características) da seguinte forma, respectivamente:

$$\hat{Q}_\theta^A - \hat{Q}_\theta^* = \sum_k (\bar{X}_k^A - \bar{X}_k^B)\hat{\beta}_{\theta,k}^A \qquad k = 1, \ldots K \tag{3.4}$$

$$\hat{Q}_\theta^* - \hat{Q}_\theta^B = \sum_k \bar{X}_k^B \left(\hat{\beta}_{\theta,k}^A - \hat{\beta}_{\theta,k}^B\right) \qquad k = 1, \ldots K \tag{3.5}$$

onde $k$ designa cada uma das características do agregado familiar.

Algumas dificuldades de interpretação surgem quando as variáveis explicativas são categóricas (com mais de duas categorias). O efeito de uma variável categórica é geralmente modelado pela inclusão de variáveis binárias para as diferentes categorias na equação de regressão, onde uma das categorias (a categoria base) é omitida para evitar a colinearidade. Na presença de variáveis categóricas, os resultados da decomposição detalhada dependem da categoria base omitida. Isto não representa um problema crítico para a parte explicada (efeito composição), pois a soma das contribuições das variáveis binárias que compõem a variável categórica (ou seja, a contribuição total) não é afetada pela escolha da categoria base. Já para a parte não explicada da decomposição (efeito estrutura), mudar a categoria base não apenas altera os resultados para as variáveis binárias individuais, mas também altera a contribuição da

variável categórica como um todo, pois não é possível distinguir a parte que é atribuível à associação de grupo (ou à diferença nos interceptos) e a parte atribuída às diferenças nos coeficientes de inclinação da categoria base. No entanto, o efeito estrutura total independe da categoria omitida (FORTIN, LEMIEUX, FIRPO, 2011; JANN, 2008). Portanto, decidiu-se por mostrar os resultados da decomposição detalhada para a variável setor de atividade como um todo, e não por setores individuais, e não interpretá-la na parte não explicada devido às dificuldades expostas.

A seguir são apresentados os resultados da decomposição do gap nos gastos em consumo entre Nordeste-Sudeste.

## 3.4 Resultados

Nesta seção, será inicialmente apresentada uma análise descritiva da variável consumo, verificando se há diferenças nas médias do consumo familiar *per capita* entre as regiões Nordeste e Sudeste, controlando para cada variável explicativa. Logo depois será analisada a diferença na composição das famílias das regiões Nordeste e Sudeste em termos das seguintes variáveis (gênero, cor ou raça, idade, anos de estudo, setor de atividade e área censitária do chefe). Em seguida, será apresentada a estimação da equação do consumo por mínimos quadrados ordinários, a decomposição de Oaxaca-Blinder e, por fim, a decomposição quantílica para os quantis 10, 50 e 90.

### *3.4.1 Análise descritiva*

Em primeiro lugar, examinaremos a variável consumo familiar *per capita*, a qual representa a variável dependente do modelo. Conforme mostra a Tabela 3, o consumo anual médio da região Sudeste é sempre mais do que o dobro da região Nordeste. Conforme Santos (2008), a região Nordeste possui o menor consumo entre todas as regiões.

Tabela 3 – Análise descritiva da variável consumo familiar *per capita*

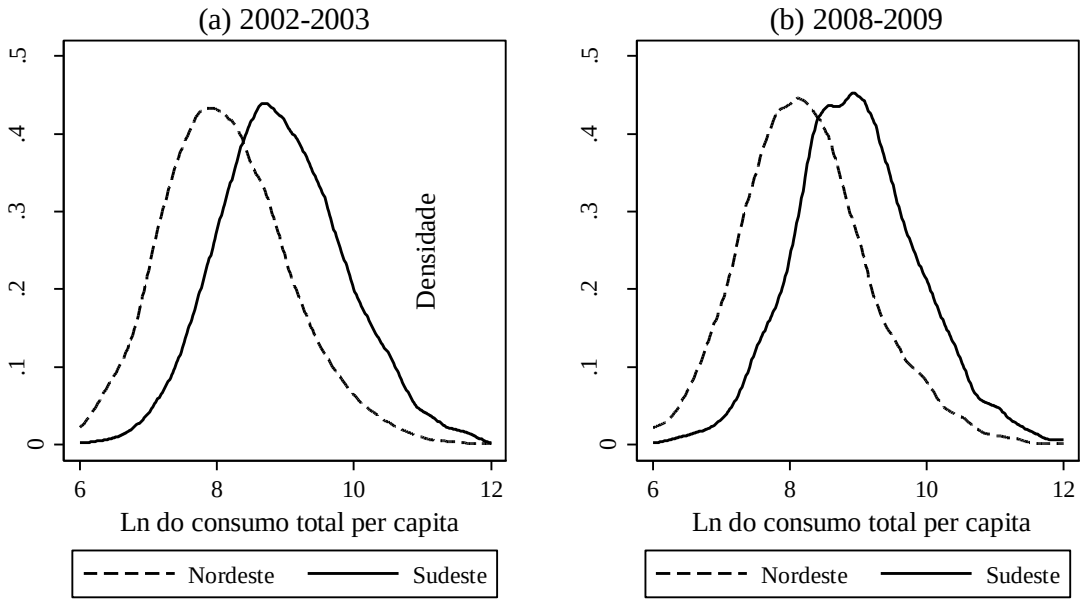
| Região | Nordeste | | Sudeste | |
|---|---|---|---|---|
| Ano | 2002-2003 | 2008-2009 | 2002-2003 | 2008-2009 |
| Média | 5.570,07 | 6.095,31 | 12.553,98 | 12.481,47 |
| Desvio-padrão | 8.555,25 | 8.876,59 | 18.230,23 | 16.523,73 |
| Mínimo | 86,59 | 90,01 | 297,43 | 84,00 |
| Máximo | 299.690,50 | 185.624,00 | 420.664,50 | 275.093,90 |

Fonte: elaboração própria a partir dos dados da POF.

Observa-se também que houve uma aproximação das médias no período analisado, com um aumento da média da região Nordeste e uma redução da média da região Sudeste, reduzindo a diferença das médias de R$ 6.983,91 para R$ 6.386,16. Isto pode estar relacionado à redução da desigualdade da despesa de consumo *per capita* no Brasil constatada por Hoffmann (2010), representada por uma queda no índice de Gini de 0,524 em 2002-2003 para 0,515 em 2008-2009.

A Figura 6 apresenta as densidades de probabilidade do logaritmo do consumo familiar *per capita* das regiões Nordeste e Sudeste nos biênios 2002-2003 e 2008-2009. Nota-se que as famílias da região Sudeste exibem um nível de consumo maior do que o das famílias da região Nordeste ao longo de toda a distribuição. Isto pode ser percebido pela linha cheia sempre no lado direito da linha tracejada.

Figura 6 – Distribuição do consumo familiar *per capita* das regiões Nordeste e Sudeste




Fonte: elaboração própria a partir dos dados da POF.

A partir da Tabela 4 avaliaremos as médias das variáveis explicativas do modelo, as quais simbolizam as características do chefe da família e as características demográficas e gerais do domicilio. Além disso, apresentam-se os resultados do teste $t$ de igualdade de médias, o qual foi realizado para verificar se as médias das regiões Nordeste e Sudeste de cada categoria são estatisticamente diferentes. Os valores das variáveis binárias representam a frequência relativa da categoria de referência. Por exemplo, a maioria dos chefes das famílias são do sexo masculino (72% na região Nordeste e 73% na região Sudeste no biênio 2002-2003, sofrendo uma redução em cada região de 4% no biênio 2008-2009), sendo que a diferença é estatisticamente significante apenas no biênio 2002-2003 a um nível de 10%.

Há uma fração substancialmente maior de chefes que se declaram com cor branca na região Sudeste (62% em 2002-2003 e 58% em 2008-2009) que na região Nordeste (29% em 2002-2003 e 27% em 2008-2009). Em ambos os grupos, a média de idade do chefe é de 46,9 no primeiro biênio, enquanto no segundo biênio já apresenta uma pequena diferença, subindo para 47,42 na região Nordeste e 48,76 na região Sudeste. Em relação aos anos de estudo, há uma diferença significativa entre as regiões, com um aumento da média de 4,41 para 5,37 na região Nordeste e de 6,60 para 7,57 na região Sudeste.

O setor de serviços é o que concentra o maior número de trabalhadores (25% na região Nordeste e 35% na região Sudeste em 2002-2003 e 26% na região Nordeste e 32% na região Sudeste em 2008-2009). No biênio 2002-2003, a agropecuária na região Nordeste concentrou a mesma porção de trabalhadores que no setor de serviços, enquanto na região Sudeste o percentual foi de apenas 9%. Não há diferença na composição dos grupos entre as regiões no setor de comércio nos dois biênios e no de construção no biênio 2002-2003. A fração de desocupados aumentou no segundo biênio, provavelmente devido à crise financeira de 2008.

Há uma forte concentração de famílias na área urbana (74% na região Nordeste contra 92% na região Sudeste). No que diz respeito ao tamanho das famílias, houve uma redução na quantidade média de pessoas (passando de 4,01 para 3,55 na região Nordeste e de 3,41 para 3,14 na região Sudeste) e também na proporção de crianças (passando de 25,53 para 22,02 na região Nordeste e de 19,42 para 16,74 na região Sudeste).

Tabela 4 – Média e teste de igualdade de médias das variáveis explicativas

| | 2002-2003 | | | 2008-2009 | | |
|---|---|---|---|---|---|---|
| | Nordeste | Sudeste | Teste de média | Nordeste | Sudeste | Teste de média |
| Homem | 0,72 | 0,73 | * | 0,68 | 0,69 | |
| Cor branca | 0,29 | 0,62 | *** | 0,27 | 0,58 | *** |
| Idade | 46,89 | 46,85 | | 47,42 | 48,76 | *** |
| Anos de estudo | 4,41 | 6,60 | *** | 5,37 | 7,57 | *** |
| Agropecuária | 0,25 | 0,09 | *** | 0,22 | 0,07 | *** |
| Serviços | 0,25 | 0,35 | *** | 0,26 | 0,32 | *** |
| Indústria | 0,08 | 0,14 | *** | 0,08 | 0,14 | *** |
| Construção | 0,07 | 0,08 | | 0,07 | 0,09 | *** |
| Comércio | 0,13 | 0,12 | | 0,13 | 0,13 | |
| Desocupado | 0,21 | 0,22 | | 0,24 | 0,26 | *** |
| Área urbana | 0,74 | 0,92 | *** | 0,74 | 0,92 | *** |
| Tamanho da família | 4,01 | 3,41 | *** | 3,55 | 3,14 | *** |
| Proporção de Crianças | 25,53 | 19,42 | *** | 22,02 | 16,74 | *** |

Fonte: elaboração própria a partir dos dados da POF.
Nota: *denota significância ao nível de 10%; ** denota significância ao nível de 5%; e ***denota significância ao nível de 1%.

Salienta-se agora o diferencial das distribuições de consumo entre as regiões condicional a algumas variáveis explicativas. O intuito é averiguar se a tendência de maiores níveis de consumo na região Sudeste ocorre também dentro das categorias e em diferentes

partes da distribuição. Os resultados servirão de base para a discussão subsequente sobre a contribuição de cada característica do agregado familiar no consumo.

A Figura 7 mostra as distribuições do consumo *per capita* das regiões Nordeste e Sudeste segundo algumas características do chefe e do domicílio para o biênio 2008-2009 (os resultados para o biênio 2002-2003 são semelhantes, por isso foram omitidos). Em todos os casos, a distribuição do consumo da região Sudeste está mais deslocada à direita do que a da região Nordeste, de modo muito semelhante à distribuição geral. Isso significa que as famílias da região Sudeste apresentam um nível de consumo maior do que o das famílias da região Nordeste ao longo de toda a distribuição, independente se as unidades de consumo estão situadas na área urbana ou rural, se são chefiadas por homens ou mulheres, brancos ou não brancos. Também independe do setor no qual o chefe trabalha e até mesmo se ele está desocupado. A decomposição permitirá analisar se essas diferenças se mantêm após considerarmos os demais fatores observados.

Realizou-se o teste de Kolmogorov-Smirnov para avaliar a igualdade das distribuições de consumo entre as regiões dentro de cada um desses grupos (gênero, cor, área censitária e setores de atividade). A conclusão foi que as distribuições das regiões Nordeste e Sudeste são diferentes em todos os casos. Esses resultados reforçam a necessidade de considerarmos essas variáveis no modelo.

Figura 7 – Distribuição do consumo *per capita* das regiões Nordeste e Sudeste segundo algumas características do chefe e do domicílio, 2008-2009. *Continua*.

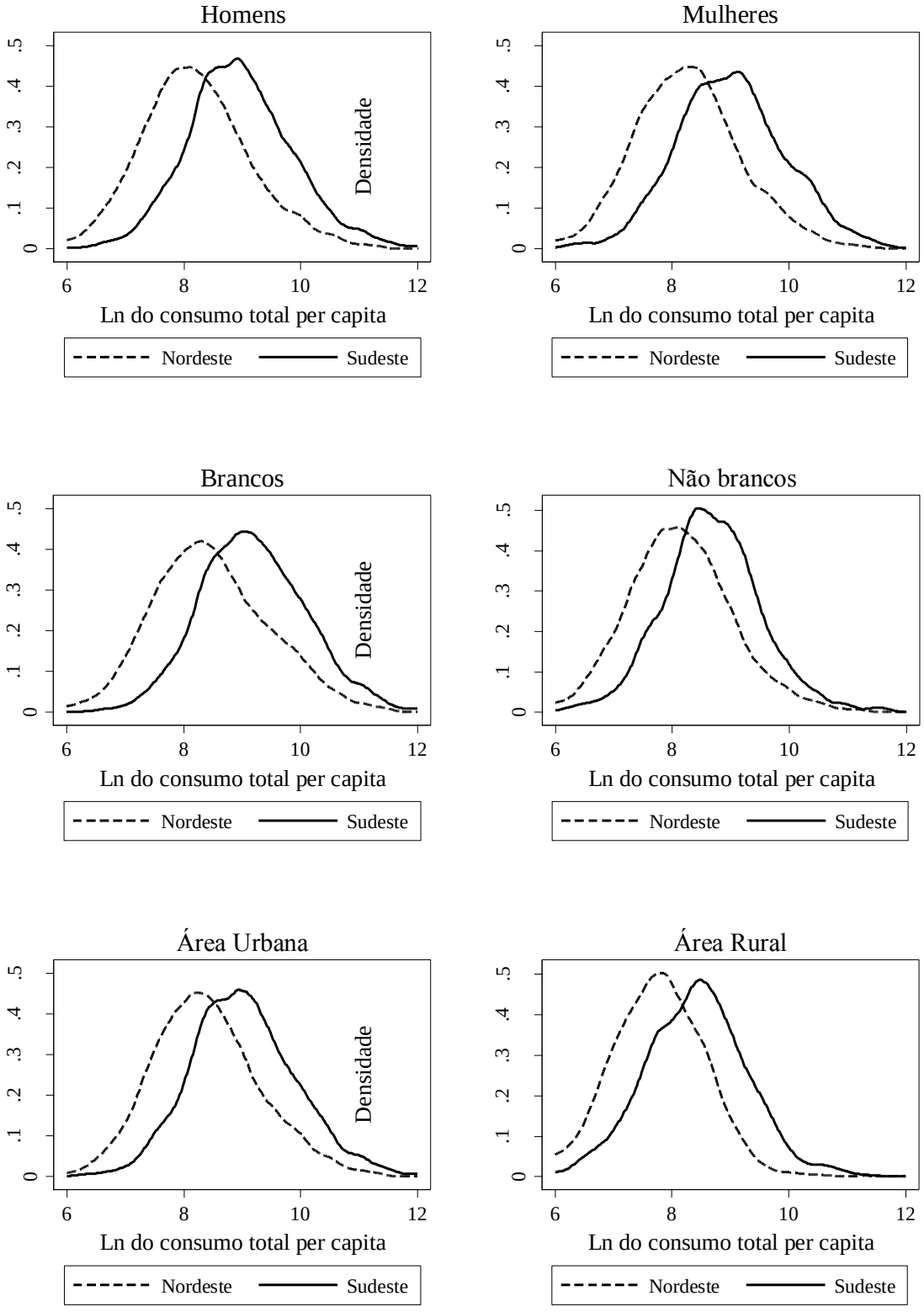
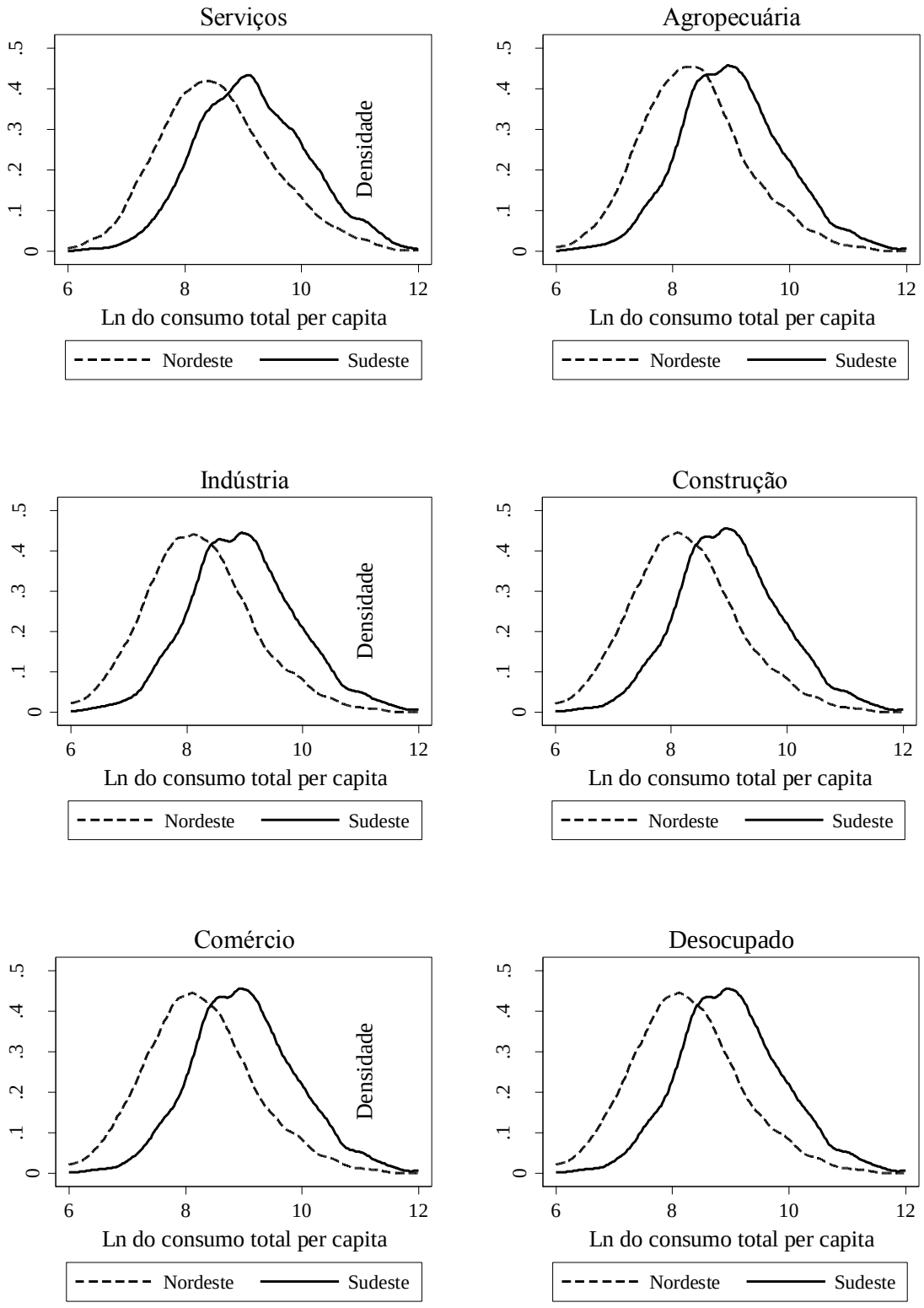


Figura 7 – Distribuição do consumo *per capita* das regiões Nordeste e Sudeste segundo algumas características do chefe e do domicílio, 2008-2009. *Conclusão*.




Fonte: elaboração própria a partir dos dados da POF.

Antes de expor a decomposição quantílica incondicional, será apresentada a estimação do modelo por mínimos quadrados ordinários e a decomposição de Oaxaca-Blinder tradicional para compararmos os resultados obtidos.

### *3.4.2 Estimação por Mínimos Quadrados Ordinários*

Estimou-se o modelo, primeiramente, através do método de Mínimos Quadrados Ordinários (MQO) e os coeficientes estimados da equação são apresentados na Tabela 5. É possível observar que quase todas as variáveis do modelo são estatisticamente significantes, excetuando-se a variável desocupado no biênio 2008-2009 e a variável construção. Isto significa que o nível médio de consumo das famílias cujos chefes trabalham na agropecuária não difere daquelas cujos chefes trabalham no setor de construção ou que encontram-se desocupados.

Tabela 5 – Estimação por Mínimos Quadrados Ordinários

| | 2002-2003 | | 2008-2009 | |
|---|---|---|---|---|
| | Nordeste | Sudeste | Nordeste | Sudeste |
| Homem | 0,091 *** | 0,069 *** | 0,135 *** | 0,085 *** |
| Cor branca | 0,092 *** | 0,220 *** | 0,127 *** | 0,208 *** |
| Idade | 0,011 *** | 0,012 *** | 0,014 *** | 0,014 *** |
| Anos de estudo | 0,100 *** | 0,102 *** | 0,092 *** | 0,100 *** |
| Agropecuária (omitida) | | | | |
| Serviços | 0,206 *** | 0,171 *** | 0,217 *** | 0,159 *** |
| Indústria | 0,202 *** | 0,158 *** | 0,179 *** | 0,153 *** |
| Construção | 0,022 | -0,065 | 0,026 | 0,047 |
| Comércio | 0,268 *** | 0,158 *** | 0,212 *** | 0,125 *** |
| Desocupado | 0,032 * | 0,077 ** | 0,004 | -0,014 |
| Área urbana | 0,139 *** | 0,181 *** | 0,185 *** | 0,181 *** |
| Tamanho da família | -0,108 *** | -0,118 *** | -0,124 *** | -0,142 *** |
| Crianças | -0,008 *** | -0,009 *** | -0,008 *** | -0,008 *** |
| Constante | 7,490 *** | 7,807 *** | 7,335 *** | 7,685 *** |
| $R^2$ | 0,507 | 0,519 | 0,482 | 0,501 |

Fonte: elaboração própria a partir dos dados da POF.
Nota: *denota significância ao nível de 10%; ** denota significância ao nível de 5%; e ***denota significância ao nível de 1%.

Na avaliação do modelo é importante verificar o grau explicativo do conjunto de variáveis independentes sobre a variável dependente. Neste trabalho, o valor do coeficiente de determinação ($R^2$) foi sempre acima de 0,48, significando que pelo menos 48% da variação da variável dependente pode ser atribuída às variáveis explicativas.

Para avaliar a qualidade do ajustamento do modelo, foi efetuado o teste $F$ de significância global. A probabilidade associada ao teste $F$ foi inferior a 0,05 (0,000) em todos os casos, portanto, rejeitou-se a hipótese nula de insignificância global do modelo, concluindo-se que existe pelo menos uma variável independente que contribui de maneira relevante para a previsão da variável dependente.

Em relação aos coeficientes estimados, observou-se que se distinguem entre as regiões e entre os biênios. Os sinais estão de acordo com o esperado, exceto pela variável desocupado, a qual mostra que, em 2002-2003, as unidades de consumo em que os chefes estão desocupados apresentam em média um nível de consumo maior do que entre aquelas em que os chefes estão ocupados no setor agropecuário, tudo o mais mantido constante.

Ademais, existe um diferencial de consumo na média em favor dos chefes homens, brancos e que residem na área urbana relativamente aos chefes do sexo feminino, não brancos e que residem na área rural, respectivamente. Também há um efeito positivo da idade e da escolaridade sobre o consumo.

Da mesma forma, o nível de consumo é mais elevado naquelas famílias em que os chefes estão empregados no setor de serviços, industrial e de comércio em comparação com as famílias em que os chefes trabalham no setor agropecuário.

Por fim, verifica-se uma relação negativa entre o grau de consumo tanto com o tamanho da família quanto com a proporção de crianças. Em outras palavras, quanto maior o número de pessoas na família e o número de crianças, menor o percentual de consumo.

### *3.4.3 Decomposição de Oaxaca–Blinder*

Apresenta-se agora a decomposição de Oaxaca-Blinder tradicional, isto é, a decomposição do diferencial de consumo médio em um termo associado às diferenças nas

características e outro aos diferentes retornos. A Tabela 6 mostra o resultado da decomposição que utiliza os coeficientes das famílias da região Sudeste como referência.

Tabela 6 - Decomposição Oaxaca-Blinder para o consumo médio

| | 2002-2003 | | | 2008-2009 | | |
|---|---|---|---|---|---|---|
| Sudeste | 8,957 | | | 8,975 | | |
| Nordeste | 8,126 | | | 8,224 | | |
| Diferença | 0,831 | *** | | 0,751 | *** | |
| *Parte explicada* | | | % | | | % |
| Homem | 0,001 | ** | 0,22 | 0,001 | * | 0,18 |
| Cor branca | 0,072 | *** | 15,12 | 0,066 | *** | 14,47 |
| Idade | 0,000 | | -0,08 | 0,018 | *** | 4,03 |
| Anos de estudo | 0,223 | *** | 46,90 | 0,220 | *** | 48,48 |
| Setor de atividade | 0,024 | *** | 5,12 | 0,018 | *** | 3,97 |
| Área urbana | 0,033 | *** | 6,96 | 0,032 | *** | 7,02 |
| Tamanho da família | 0,071 | *** | 14,81 | 0,059 | *** | 12,95 |
| Crianças | 0,052 | *** | 10,96 | 0,040 | *** | 8,84 |
| Total | 0,476 | *** | 100,00 | 0,454 | *** | 100,00 |
| Percentual | 57,33 | | | 60,42 | | |
| *Parte não explicada* | | | % | | | % |
| Homem | -0,016 | | -4,52 | -0,034 | *** | -11,60 |
| Cor branca | 0,037 | *** | 10,56 | 0,022 | *** | 7,27 |
| Idade | 0,043 | | 12,22 | -0,001 | | -0,45 |
| Anos de estudo | 0,007 | | 2,06 | 0,045 | *** | 15,11 |
| Setor de atividade | -0,024 | | -6,65 | -0,032 | | -10,63 |
| Área urbana | 0,031 | | 8,67 | -0,003 | | -0,90 |
| Tamanho da família | -0,040 | * | -11,33 | -0,064 | *** | -21,41 |
| Crianças | -0,001 | | -0,36 | 0,015 | | 5,03 |
| Constante | 0,317 | *** | 89,34 | 0,350 | *** | 117,58 |
| Total | 0,355 | *** | 100,00 | 0,297 | *** | 100,00 |
| Percentual | 42,67 | | | 39,58 | | |

Fonte: elaboração própria a partir dos dados da POF.
Nota: *denota significância ao nível de 10%; ** denota significância ao nível de 5%; e ***denota significância ao nível de 1%.

Através da Tabela 3, vimos que as famílias da região Sudeste consomem um valor a mais do que as famílias da região Nordeste, em média, de R$ 6.983,91 no biênio 2002-2003 e de R$ 6.386,16 no biênio 2008-2009. Como é possível ver na Tabela 6, a maior parte dessas diferenças são explicadas pelas diferenças de características entre os dois grupos (57,33% e 60,42%, respectivamente). Fazendo uma análise individual, o fator que mais contribui para

essas desigualdades regionais de consumo são as diferenças nos anos de estudo dos chefes. Outras duas características que também têm influência sobre o gap de consumo, mas com importância menor, são a cor do chefe e o tamanho da família.

O restante do diferencial de consumo entre as regiões se deve a diferenças nos coeficientes dos regressores das equações. O fator individual mais importante para explicar a diferença nos coeficientes é a cor do chefe no biênio 2002-2003 e os anos de estudo no biênio 2008-2009.

### *3.4.4 Regressões quantílicas incondicionais*

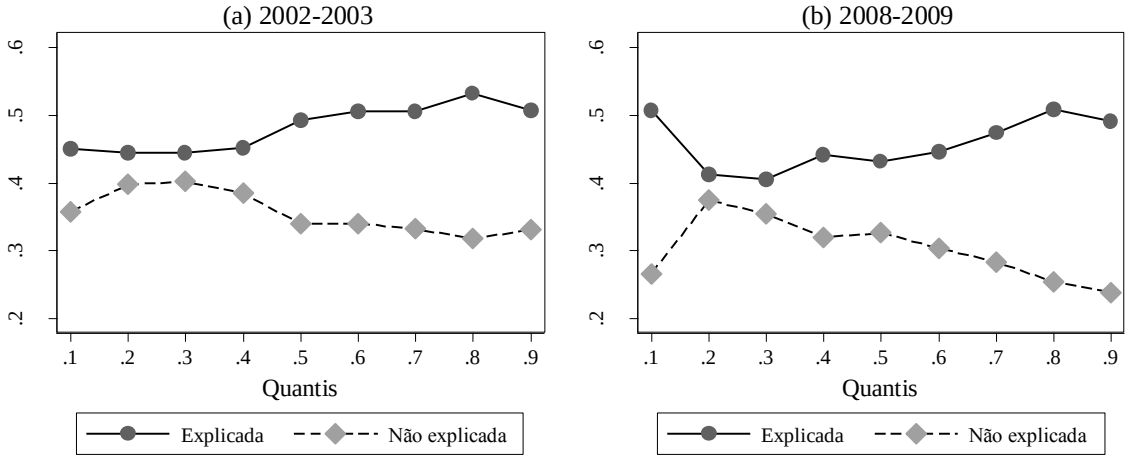
Os resultados da decomposição do gap de consumo Nordeste-Sudeste nos quantis são apresentados na Figura 8 e mostram que as diferenças nas características (parte explicada) das famílias são as principais fontes de desigualdade entre as duas regiões em todos os quantis e nos dois biênios, assim como nos resultados da decomposição Oaxaca-Blinder tradicional. Isso significa que as famílias da região Sudeste têm níveis mais elevados de consumo porque possuem características superiores às suas contrapartes da região Nordeste. Hassine e Zeufack (2015) encontram resultados similares para a Tanzânia, onde as diferenças nas dotações das famílias são as principais fontes da desigualdade de consumo rural-urbana. Isto se verifica em toda a distribuição, indicando que as famílias urbanas são melhores do que as rurais porque têm vantagens nas características.

Não obstante, existe um diferencial de consumo que é explicado apenas pelo fato de as famílias residirem em regiões diferentes (parte não explicada). A importância do efeito estrutura (diferença nos retornos) vai diminuindo à medida que avançamos para os quantis superiores, com algumas exceções (notadamente no primeiro quantil). Este resultado é semelhante ao encontrado por Oliveira e Silveira Neto (2016), os quais fazem uma análise análoga usando o logaritmo do salário-hora como variável dependente.

Observa-se uma redução da disparidade regional de consumo nos quantis 10, 50 e 90 no segundo biênio, com reduções maiores nos quantis superiores (ver Tabela 7). Isto se deveu a um crescimento proporcionalmente maior nos níveis de consumo da região Nordeste do que na região Sudeste nos quantis 10 e 50 e a um decrescimento do nível de consumo no quantil 90 na região Sudeste. Houve, inclusive, uma inversão: no biênio 2002-2003 o diferencial de

consumo era maior entre as famílias mais ricas. No entanto, no biênio 2008-2009 o gap passa a ser maior entre os mais pobres.

Figura 8 – Decomposição quantílica incondicional para o diferencial do consumo familiar *per capita* entre as regiões Nordeste e Sudeste




Fonte: elaboração própria a partir dos dados da POF.

A decomposição mostra, na Tabela 7, que esse movimento favorável de redução da desigualdade foi impulsionado por um maior efeito composição e um menor efeito estrutura nos quantis inferior e superior e o contrário aconteceu na mediana, com um menor efeito composição e um maior efeito estrutura. A seguir, discute-se a decomposição detalhada para tentar entender as principais características que explicam a desigualdade de consumo entre as regiões e quais foram determinantes para a redução verificada no período.

A análise da contribuição das características (parte explicada) específicas do agregado familiar para o hiato do consumo Nordeste-Sudeste, na Tabela 7, revela que as diferenças na escolaridade dos chefes são as que mais contribuem para a desigualdade de consumo entre as regiões e a importância aumenta à medida que caminhamos para os quantis superiores. Em termos de mudanças nos biênios, verifica-se, no quantil 10, que houve uma piora nas diferenças de pessoas residentes nas áreas urbanas e na dotação de capital humano entre as regiões, e uma aproximação na proporção de crianças nas famílias. Isto sugere que as políticas de desenvolvimento implementadas entre os mais desprovidos de recursos da região Nordeste não foram apropriadas para como combater o analfabetismo, incentivar a educação

Tabela 7 – Decomposição quantílica incondicional para o consumo familiar *per capita*

| | 2002-2003 | | | | | | 2008-2009 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quantil | 10º | 50º | | 90º | | | 10º | 50º | | 90º | | |
| Sudeste | 7,799 | 8,903 | | 10,191 | | | 7,848 | 8,932 | | 10,183 | | |
| Nordeste | 6,991 | 8,070 | | 9,353 | | | 7,076 | 8,174 | | 9,454 | | |
| Diferença | 0,808 *** | 0,832 *** | | 0,837 *** | | | 0,772 *** | 0,758 *** | | 0,729 *** | | |
| *Parte explicada* | | % | | % | | % | | % | | % | | % |
| Homem | 0,002 ** | 0,53 | 0,001 * | 0,26 | -0,001 | -0,18 | 0,002 * | 0,34 | 0,001 | 0,12 | -0,001 | -0,11 |
| Cor branca | 0,062 *** | 13,76 | 0,090 *** | 18,23 | 0,062 *** | 12,27 | 0,068 *** | 13,36 | 0,066 *** | 15,20 | 0,067 *** | 13,56 |
| Idade | 0,000 | -0,06 | 0,000 | -0,07 | -0,001 | -0,10 | 0,013 *** | 2,61 | 0,017 *** | 3,99 | 0,027 *** | 5,45 |
| Anos de estudo | 0,095 *** | 21,18 | 0,212 *** | 43,03 | 0,371 *** | 73,37 | 0,128 *** | 25,20 | 0,208 *** | 48,22 | 0,348 *** | 70,79 |
| Setor de atividade | 0,093 *** | 20,64 | 0,028 *** | 5,68 | -0,035 *** | -6,91 | 0,083 *** | 16,39 | 0,017 *** | 3,94 | -0,026 *** | -5,29 |
| Área urbana | 0,049 *** | 10,91 | 0,037 *** | 7,52 | -0,002 | -0,42 | 0,077 *** | 15,16 | 0,026 *** | 6,08 | -0,004 | -0,73 |
| Tamanho da família | 0,074 *** | 16,44 | 0,063 *** | 12,82 | 0,076 *** | 15,10 | 0,075 *** | 14,83 | 0,053 *** | 12,19 | 0,063 *** | 12,78 |
| Crianças | 0,075 *** | 16,66 | 0,061 *** | 12,42 | 0,034 *** | 6,79 | 0,062 *** | 12,17 | 0,044 *** | 10,17 | 0,017 *** | 3,51 |
| Total | 0,451 *** | 100,00 | 0,493 *** | 100,00 | 0,506 *** | 100,00 | 0,506 *** | 100,00 | 0,432 *** | 100,00 | 0,491 *** | 100,00 |
| *Parte não explicada* | | % | | % | | % | | % | | % | | % |
| Homem | 0,061 * | 17,05 | -0,008 | -2,46 | -0,130 ** | -39,18 | 0,044 | 16,45 | -0,039 * | -11,80 | -0,152 *** | -63,81 |
| Cor branca | 0,047 *** | 13,04 | 0,067 *** | 19,75 | -0,017 | -5,10 | 0,039 *** | 14,66 | 0,031 *** | 9,64 | -0,028 * | -11,86 |
| Idade | 0,014 | 3,95 | 0,002 | 0,51 | 0,051 | 15,45 | -0,049 | -18,59 | 0,006 | 1,99 | -0,059 | -24,86 |
| Anos de estudo | 0,028 | 7,86 | 0,043 ** | 12,52 | -0,104 ** | -31,41 | 0,079 *** | 29,62 | 0,081 *** | 24,93 | -0,133 ** | -55,85 |
| Setor de atividade | 0,403 *** | 112,85 | -0,046 | -13,55 | -0,254 *** | -76,73 | 0,533 *** | 200,35 | 0,015 | 4,59 | -0,330 *** | -138,61 |
| Área urbana | 0,087 | 24,34 | 0,025 | 7,27 | -0,068 * | -20,57 | 0,137 ** | 51,64 | -0,015 | -4,70 | -0,132 *** | -55,50 |
| Tamanho da família | 0,068 | 19,06 | 0,027 | 7,96 | -0,227 *** | -68,69 | -0,103 | -38,68 | -0,001 | -0,29 | -0,214 *** | -90,08 |
| Crianças | -0,119 *** | -33,20 | -0,014 | -4,02 | 0,051 | 15,36 | -0,069 ** | -25,98 | 0,024 | 7,29 | 0,086 *** | 36,21 |
| Constante | 0,235 | 65,76 | 0,282 ** | 83,00 | 0,804 *** | 242,74 | 0,270 ** | 101,51 | 0,258 ** | 79,09 | 0,859 *** | 360,86 |
| Total | 0,357 *** | 100,00 | 0,340 *** | 100,00 | 0,331 *** | 100,00 | 0,266 *** | 100,00 | 0,326 *** | 100,00 | 0,238 *** | 100,00 |
| Total | 0,357 *** | 100,00 | 0,340 *** | 100,00 | 0,331 *** | 100,00 | 0,266 *** | 100,00 | 0,326 *** | 100,00 | 0,238 *** | 100,00 |

Fonte: elaboração própria a partir dos dados da POF.

Nota: *denota significância ao nível de 10%; ** denota significância ao nível de 5%; e ***denota significância ao nível de 1%.

básica e promover a urbanização. Oliveira *et al*. (2017) observam que o maior incremento de consumo no décimo quantil de 2002 para 2009 no Brasil ocorreu nos itens de habitação, o qual pode ter sido impulsionado por uma maior urbanização na região Sudeste entre as famílias mais pobres. No quantil 50, houve um aumento da contribuição da diferença nos anos de estudo e na idade. No entanto, todas as outras características passaram a contribuir menos com a diferença, em especial a cor e a proporção de crianças. No quantil 90, por sua vez, as mudanças na idade deterioraram o gap de consumo, enquanto mudanças na proporção de crianças e na escolaridade ocorreram em favor do Nordeste.

Em relação à contribuição das diferenças nos retornos (parte não explicada) sobre a desigualdade total, há distinções para cada quantil e biênio. No quantil 10, isto é, para as famílias mais pobres do biênio 2002-2003, a diferença nos retornos dos homens e dos chefes de cor branca, comparados aos retornos das mulheres e dos chefes não brancos, respectivamente, se destacam na elucidação das diferenças de consumo não explicadas por diferenças nas características produtivas e demográficas das famílias[1]. Isto significa que havia uma maior discriminação de gênero e racial na região Sudeste entre a população menos abastada no período. Já em 2008-2009, as diferenças nos retornos são maiores entre aqueles que residem na área urbana e com a mesma escolaridade. Não é mais possível constatar diferenças na discriminação de gênero entre as regiões. No quantil 50, a diferença nos retornos entre os brancos relativamente ao não brancos e com a mesma escolaridade é que são determinantes para explicar a lacuna total no biênio 2002-2003, mudando apenas a ordem de importância no segundo biênio, com a escolaridade passando a ter o papel principal. Já no quantil 90 acontece algo destoante: as diferenças nos retornos são favoráveis aos nordestinos, com exceção da proporção de crianças no biênio 2008-2009. Conforme observado por Oliveira e Silveira Neto (2016), um maior retorno à educação no quantil superior da distribuição no Nordeste favorece a redução da desigualdade regional de renda. Como existem menos trabalhadores qualificados no Nordeste comparado ao Sudeste, os movimentos de oferta e demanda elevam o prêmio salarial relativo à qualificação.

[1] Desconsidera-se a interpretação da constante e do valor referente ao setor de atividade na parte não explicada, pois dependem da categoria omitida, como explicado na metodologia.

## 3.5 Considerações Finais

Este artigo teve como objetivo investigar os fatores associados ao gap de consumo entre as regiões Nordeste e Sudeste do Brasil no período 2002 a 2009, utilizando variáveis ligadas à demanda e à oferta de trabalho. Para atingir o objetivo, foi utilizado o método proposto por Firpo, Fortin e Lemieux (2009) baseado na *Recentered Influence Function* (RIF) *Regression* e que faz uma decomposição do tipo Oaxaca-Blinder nos quantis. Para isso, fez-se uso dos dados fornecidos pelas Pesquisas de Orçamentos Familiares (POFs) do IBGE dos biênios 2002-2003 e 2008-2009.

Os resultados encontrados apontam que o consumo médio da região Sudeste é sempre mais do que o dobro da região Nordeste. No entanto, houve uma aproximação das médias no período analisado, reduzindo a diferença das médias de R$ 6.983,91 para R$ 6.386,16.

Notou-se também que as famílias da região Sudeste exibem um nível de consumo maior do que o das famílias da região Nordeste ao longo de toda a distribuição, independente se as unidades de consumo estão situadas na área urbana ou rural, se são chefiadas por homens ou mulheres, brancos ou não brancos. Também independe do setor no qual o chefe trabalha e até mesmo se ele está desocupado.

Os resultados da decomposição do gap de consumo Nordeste-Sudeste mostraram que as diferenças nas características (parte explicada) das famílias são as principais fontes de desigualdade entre as duas regiões, sugerindo que as famílias da região Sudeste têm níveis mais elevados de consumo porque possuem características superiores às suas contrapartes da região Nordeste. A característica que mais contribui para o hiato do consumo Nordeste-Sudeste é a escolaridade dos chefes e quanto mais elevado o quantil, maior a importância.

As diferenças nos retornos também contribuem de maneira relevante para a desigualdade total, mas em menor magnitude. No quantil 10, percebeu-se uma maior discriminação de gênero e raça na região Sudeste no biênio 2002-2003. Já em 2008-2009, as diferenças nos retornos são maiores entre aqueles que residem na área urbana e com a mesma escolaridade. No quantil 50, a diferença dos retornos predominou entre os brancos e com a mesma escolaridade. Por fim, no quantil 90, as diferenças nos retornos são favoráveis aos nordestinos, com exceção da proporção de crianças no biênio 2008-2009.

Observou-se uma redução da disparidade regional de consumo no segundo biênio, com reduções maiores nos quantis superiores. Isto se deveu a um crescimento proporcionalmente maior nos níveis de consumo da região Nordeste do que na região Sudeste nos quantis 10 e 50 e a um decrescimento do nível de consumo entre os mais ricos na região Sudeste.

Esse movimento favorável de redução das diferenças foi impulsionado por um menor efeito estrutura e um maior efeito composição nos quantis inferior e superior, enquanto na mediana há um menor efeito composição e um maior efeito estrutura. É notável a aproximação da proporção de crianças entre as regiões em todos os quantis. Além disso, entre os mais pobres do Nordeste, verificou-se uma melhoria em relação à região Sudeste em termos de pessoas em setores de atividade que apresentam maiores remunerações e em retornos ao tamanho da família. Entre aqueles com consumo mediano, percebeu-se uma aproximação de chefes com cor branca assim como uma redução na discriminação racial e de gênero do Nordeste em relação ao Sudeste. Outro fato importante foi a redução das diferenças nos retornos entre aqueles que moram na área urbana. Já entre os mais abastados houve uma redução nas diferenças dos anos de estudo e uma melhora nos retornos em praticamente todas as características, exceto na proporção de crianças.

Não obstante essas melhorias, nota-se um aumento da distância entre os mais pobres das duas regiões, notadamente em termos de dotação de residência em áreas urbanas e de capital humano. Isto sugere que as políticas de desenvolvimento implementadas na região Nordeste não foram apropriadas para como combater o analfabetismo, incentivar a educação básica e promover a urbanização entre os mais desprovidos de recursos.

Em suma, este artigo buscou contribuir no debate sobre as fontes das desigualdades regionais de consumo no Brasil e os resultados mostraram que essas lacunas se devem tanto a diferenças nas características (ligadas a demanda e a oferta de trabalho, assim como especificidades demográficas da família) quanto nos retornos a essas características.

Embora tenhamos verificado uma redução do gap no consumo entre as regiões, a disparidade continua ainda muito significativa. Assim, reforça-se a importância de políticas públicas, voltadas principalmente para os mais pobres da região Nordeste, que visem aumentar o acesso à educação e ao mercado de trabalho, promovendo autonomia econômica dessa população.

# REFERÊNCIAS

ATTANASIO, O.; HURST, E.; PISTAFERRI, L. The Evolution of Income, Consumption, and Leisure Inequality in the United States, 1980–2010. *In*: CARROLL, C. D.; CROSSLEY, T. F.; SABELHAUS, J. (Eds.) *Improving the Measurement of Consumer Expenditures.* University of Chicago Press, Studies in Income and Wealth, v. 74, p. 100-140. 2015.

AZZONI, C. R. Concentração regional e dispersão das rendas per capita estaduais: análise a partir de séries históricas estaduais de PIB, 1939-1995. *Estudos Econômicos*. São Paulo, v. 27, n. 3, p. 341-393. 1997.

CACCIAMALI, M. C. Distribuição de renda no Brasil: persistência do elevado grau de desigualdade. *In*: PINHO, D. B.; VASCONCELLOS, M. A. S. (Org.). *Manual de Economia*. 4 ed. São Paulo: Saraiva, v. 1, p. 406-422, 2002.

CHAMBERS, J. M.; CLEVELAND, W. S.; KLEINER, B.; TUKEY, P. A. *Graphical Methods For Data Analysis*. Wadsworth, Pacific Grove, CA. 1983.

CLEVELAND, W. S.; MCGILL, R. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79, 531-554. 1984.

DEATON, A.; GROSH, M. Chapter 17: Consumption. *In*: GROSH, M.; GLEWWE, P. (Eds.) *Designing Household Survey Questionnaires for Developing Countries*: Lessons from Ten Years of LSMS Experience. Working Papers 218, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies. 1998.

DEV, S. M. *The Problem of Inequality*. Indira Gandhi Institute of Development Research, Mumbai, 2016.

DINARDO, J.; FORTIN, N. M.; LEMIEUX, T. Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. *Econometrica*, v. 64, p. 1001-1044. 1996.

DINARDO, J.; TOBIAS, J. L. Nonparametric density and regression estimation. *Journal of Economic Perspectives*, v. 15, n. 4, p. 11-28. 2001.

FIELDS, G. S. Accounting for income inequality and its change: a new method, with application to the distribution of earnings in the United States. *In*: SOLOMON W.; POLACHEK (Eds.) *Worker Well-Being and Public Policy*. Research in Labor Economics, v. 22, p. 1-38. 2003.

FIORIO, C. V.; JENKINS, S. P. ineqrbd: Regression-based inequality decomposition, following Fields (2003). United Kingdom Stata Users' Group Meetings, 2007.

FIRPO, S. A evolução das desigualdades de renda e de consumo ao longo do ciclo da vida. *Pesquisa e Planejamento Econômico*, v. 30, n. 1, p.49-68, 2000.

FIRPO, S.; FORTIN, N. M.; LEMIEUX, T. Unconditional Quantile Regressions. *Econometrica*, v. 77, n. 3, p. 953-973. 2009.

FORTIN, N; LEMIEUX, T; FIRPO, S. Decomposition Methods in Economics. *In*: ASHENFELTER, O; CARD, D. (Eds.) *Handbook of Labor Economics*. North Holland: Elsevier, v. 4A, p. 1-102. 2011.

GUIMARÃES, R. R. M.; OLIVEIRA, A. M. H. C. *Análise da Distribuição Salarial entre o Setor Público e Privado no Brasil (1987-2005) com Aplicações para a Reforma Administrativa do Governo Federal*. 1. ed. Finanças Públicas: XII Prêmio do Tesouro Nacional 2007. Brasília: Universidade de Brasília, p. 251-308, 2008.

HANDCOCK, M. S.; MORRIS, M. *Relative Distribution Methods in the Social Sciences*. Springer-Verlag, New York, 1999.

HASSINE, N. B.; ZEUFACK, A. *Inequality of Outcomes and Inequality of Opportunity in Tanzania*. Policy Research Working Paper 7260. Washington, DC.: World Bank Group. 2015.

HOFFMANN, R. Desigualdade da renda e das despesas *per capita* no Brasil, em 2002-2003 e 2008-2009, e avaliação do grau de progressividade ou regressividade de parcelas da renda familiar. *Economia e Sociedade*, Campinas, v. 19, n. 3 (40), p. 647-661. 2010.

IBGE. *Pesquisa de Orçamentos Familiares 2008-2009*: Despesas, Rendimentos e Condições de Vida. Rio de Janeiro, 2010. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv45130.pdf. Acesso em: 27 de dezembro de 2017.

IPEA. *A década inclusiva (2001-2011)*: desigualdade, pobreza e políticas de renda, Comunicados do IPEA, n. 155, 2012.

JANN, B. The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, v. 8, n. 4, p. 453-479. 2008.

JIRASAVETAKUL, L-B. F; LAKNER, C. *The Distribution of Consumption Expenditure in Sub-Saharan Africa*: The Inequality among All Africans. Policy Research Working Paper 7557. 2016.

KAPLANOGLOU, G.; RAPANOS, V. T. Evolutions in Consumption Inequality and Poverty in Greece: The Impact of the Crisis and Austerity Policies. *Review of Income and Wealth*, Series 64, n. 1. 2018.

MACHADO, J. F.; MATA, J. Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, v. 20, p. 445-465. 2005.

MARTINS, A. S. R. *Análise da Desigualdade de Renda e de Despesa no Brasil*. 2011. 35 p. Monografia – Faculdade de Economia e Administração. Insper Instituto de Ensino e Pesquisa. São Paulo, 2011.

MELO, M. R. B.; MONTE, P. A. Decomposição da renda domiciliar: uma análise regional. *Revista Econômica do Nordeste*, v. 46, p. 137-153, 2015.

MENEZES FILHO, N.; RODRIGUES, M. Tecnologia e Demanda por Qualificação na Indústria Brasileira. *Revista Brasileira de Economia*, v. 57, n. 3. 2003.

MORDUCH, J.; SICULAR, T. Rethinking Inequality Decomposition with Evidence from Rural China, *Economic Journal*, v. 112, p. 93-106. 2002.

NERI, M. *A nova classe média*: o lado brilhante da base da pirâmide. São Paulo: Saraiva, 2011.

OLIVEIRA, L. S.; QUINTAES, V. C. C.; SANTOS, L. A.; SOUZA, D. F. *The expansion of consumption and the welfare dynamics of the Brazilian families: a decomposition analysis of poverty and inequality*. Seventh ECINEQ Meeting, New York. 2017.

OLIVEIRA, R.; SILVEIRA NETO, R. M. Estrutura produtiva ou escolaridade? Uma análise dos fatores explicativos da desigualdade de renda entre as regiões Sudeste e Nordeste por quantil para o período entre os anos de 1970 e 2010. *Revista Econômica do Nordeste*, v.47, n. 3, p. 75-91. 2016.

_____. Quarenta Anos de Disparidades Regionais no Brasil: Qual o Papel da Escolaridade e da Estrutura Produtiva? *In*: ENCONTRO NACIONAL DE ECONOMIA, 45., 2017, Natal. *Anais...* Natal: ANPEC, 2017.

PAULI, R. C.; NAKABASHI, L.; SAMPAIO, A. V. Mudança estrutural e mercado de trabalho no Brasil. *Revista de Economia Política*, v. 32, n. 3 (128), p. 459-478. 2012.

POCHMANN, M. *Nova classe media?* O trabalho na base da pirâmide social. São Paulo: Boitempo, 128p., 2012.

PETRARCA, I.; RICCIUTI, R. *Relative income distribution in six European countries*: market and disposable income. LIS Working papers 629, LIS Cross-National Data Center in Luxembourg. 2015.

RODRIGUES, C. G.; RIOS-NETO, E. L. G.; PINTO, C. C. X. Fatores associados ao declínio do desempenho escolar para a coorte de alunos da 4a série do ensino fundamental no Brasil. *In*: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 17., 2010, Caxambú. *Anais…* Minas Gerais: ABEP, 2010.

SANTOS, I. C. M. A Distribuição do Consumo: uma análise das Regiões Brasileiras através da POF 2002-2003. *In*: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 16., 2008, Caxambu. *Anais...* Belo Horizonte: ABEP, 2008.

SARLO, C. *Consumption inequality in Canada*: Is the gap growing? Fraser Institute. 2016.

SILVEIRA NETO, R. M.; MENEZES, T. A. Disparidades Regionais de Renda no Brasil: Analisando a Importância do Capital Humano. *In*: Fundação Konrad-Adeneur. (Org.). *Fórum Brasil-Europa*. Fundação Konrad-Adeneur. 2008.

_____. Nível e Evolução da Desigualdade dos Gastos Familiares no Brasil: uma Análise para as Regiões Metropolitanas no Período 1996 a 2003. *Estudos Econômicos*, São Paulo, v. 40, n. 2, p. 341-372, abril-junho 2010.

SILVERMAN, B. W**.** *Density Estimation***.** Londres: Chapman and Hall, 1986.

SOARES, S. S. D. Análise de bem-estar e decomposição por fatores da queda na desigualdade entre 1995 e 2004. *Econômica*, v. 8, n. 1, p. 83-115, 2006.

TRIPATHI, S. *Source of Inequality in consumption Expenditure in India*: A Regression Based Inequality Decomposition Analysis. Munich Personal RePEc Archive Paper n. 72117. 2016.

## APÊNDICE A – REGRESSÃO QUANTÍLICA INCONDICIONAL

| | 2002-2003 | | | | | | 2008-2009 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Quantil | 10º | | 50º | | 90º | | 10º | | 50º | | 90º | |
| | | | | | Sudeste | | | | | | | |
| Homem | 0,159 | *** | 0,085 | ** | -0,060 | | 0,178 | *** | 0,053 | ** | -0,056 | |
| Cor branca | 0,189 | *** | 0,274 | *** | 0,190 | *** | 0,214 | *** | 0,208 | *** | 0,211 | *** |
| Idade | 0,009 | *** | 0,011 | *** | 0,016 | *** | 0,010 | *** | 0,013 | *** | 0,020 | *** |
| Anos de estudo | 0,043 | *** | 0,097 | *** | 0,169 | *** | 0,058 | *** | 0,095 | *** | 0,158 | *** |
| Serviços | 0,587 | *** | 0,194 | *** | -0,231 | *** | 0,563 | *** | 0,163 | *** | -0,122 | * |
| Agropecuária (omitida) | | | | | | | | | | | | |
| Indústria | 0,661 | *** | 0,194 | *** | -0,216 | ** | 0,706 | *** | 0,132 | *** | -0,224 | *** |
| Construção | 0,348 | *** | -0,136 | * | -0,150 | * | 0,277 | *** | 0,037 | | -0,080 | |
| Comércio | 0,625 | *** | 0,162 | ** | -0,249 | ** | 0,625 | *** | 0,138 | *** | -0,301 | *** |
| Desocupado | 0,359 | *** | 0,119 | * | -0,184 | ** | 0,355 | *** | -0,003 | | -0,293 | *** |
| Área urbana | 0,268 | *** | 0,202 | *** | -0,012 | | 0,437 | *** | 0,149 | *** | -0,020 | |
| Tamanho da família | -0,123 | *** | -0,105 | *** | -0,127 | *** | -0,181 | *** | -0,127 | *** | -0,151 | *** |
| Crianças | -0,012 | *** | -0,010 | *** | -0,006 | *** | -0,012 | *** | -0,008 | *** | -0,003 | *** |
| Constante | 6,801 | *** | 7,756 | *** | 8,987 | *** | 6,568 | *** | 7,737 | *** | 8,663 | *** |
| $R^2$ | 0,187 | | 0,323 | | 0,214 | | 0,187 | | 0,323 | | 0,214 | |
| | | | | | Nordeste | | | | | | | |
| Homem | 0,075 | *** | 0,096 | *** | 0,120 | *** | 0,114 | *** | 0,109 | *** | 0,166 | *** |
| Cor branca | 0,030 | | 0,044 | ** | 0,247 | *** | 0,068 | *** | 0,090 | *** | 0,317 | *** |
| Idade | 0,008 | *** | 0,011 | *** | 0,015 | *** | 0,011 | *** | 0,013 | *** | 0,021 | *** |
| Anos de estudo | 0,037 | *** | 0,087 | *** | 0,193 | *** | 0,043 | *** | 0,080 | *** | 0,183 | *** |
| Serviços | 0,132 | ** | 0,262 | *** | 0,108 | * | 0,010 | | 0,168 | *** | 0,253 | *** |
| Agropecuária | -0,273 | *** | -0,027 | | 0,189 | *** | -0,439 | *** | -0,078 | * | 0,282 | *** |
| Indústria | 0,154 | *** | 0,265 | *** | 0,005 | | 0,031 | | 0,147 | *** | 0,167 | * |
| Construção (omitida) | | | | | | | | | | | | |
| Comércio | 0,213 | *** | 0,370 | *** | 0,092 | | 0,069 | | 0,213 | *** | 0,091 | |
| Desocupado | -0,108 | * | 0,082 | * | 0,041 | | -0,260 | *** | -0,038 | | 0,048 | |
| Área urbana | 0,150 | *** | 0,169 | *** | 0,081 | *** | 0,251 | *** | 0,170 | *** | 0,158 | *** |
| Tamanho da família | -0,140 | *** | -0,112 | *** | -0,071 | *** | -0,152 | *** | -0,127 | *** | -0,091 | *** |
| Crianças | -0,008 | *** | -0,009 | *** | -0,008 | *** | -0,009 | *** | -0,009 | *** | -0,007 | *** |
| Constante | 7,033 | *** | 7,512 | *** | 7,958 | *** | 6,912 | *** | 7,515 | *** | 7,463 | *** |
| $R^2$ | 0,168 | | 0,330 | | 0,234 | | 0,164 | | 0,316 | | 0,209 | |

Fonte: elaboração própria a partir dos dados da POF.

Nota: *denota significância ao nível de 10%; ** denota significância ao nível de 5%; e ***denota significância ao nível de 1%.