

Universidade de São Paulo

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

**Uso de uma medida de divergência simétrica no
estudo da desigualdade de renda.**

MATEUS FERRACINI

Ribeirão Preto

2018

Universidade de São Paulo
Faculdade de Economia, Administração e Contabilidade de
Ribeirão Preto
Departamento de Economia
Programa de Pós-graduação em Economia - Área: Economia
Aplicada

MATEUS FERRACINI

Uso de uma medida de divergência simétrica no estudo da
desigualdade de renda.

Orientador: Walter Belluzzo Junior

Ribeirão Preto

2018

Prof. Dr. Vahan Agopyan

Reitor da Universidade de São Paulo

Prof. Dr. Dante Pinheiro Martinelli

Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão

Preto

Prof. Dr. Renato Leite Marcondes

Chefe do Departamento de Economia

Prof. Dr. Sérgio Naruhiko Sakurai

Coordenador do Programa de Pós-Graduação em Economia - Área: Economia

Aplicada

MATEUS FERRACINI

Uso de uma medida de divergência simétrica no estudo da
desigualdade de renda.

Dissertação de Mestrado submetida ao Programa de Pós-graduação em Economia - Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para a obtenção do título de Mestre em Ciências. A original encontra-se disponível na FEA-RP/USP.

Orientador: Walter Belluzzo Junior

Ribeirão Preto

2018

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ferracini, Mateus
Uso de uma medida de divergência simétrica no estudo da desigualdade de renda./
Universidade de São Paulo
Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto
Programa de Pós-graduação em Economia - Área: Economia Aplicada; Orientador: Walter Belluzzo Junior
Ribeirão Preto, 2018- 55 p. : il.

Dissertação (Mestrado) – Universidade de São Paulo, 2018.
Versão Corrigida. A original encontra-se disponível na FEA-RP/USP
1. Desigualdade de renda. 2. Divergência J. I. Orientador: prof. Dr. Walter Belluzzo Junior. II. Universidade De São Paulo - Campus Ribeirão Preto. III. Faculdade de Economia, Administração e Contabilidade. IV. Uso de uma medida de divergência simétrica no estudo da desigualdade de renda no estado de São Paulo

Nome: Mateus Ferracini

Título: Uso de uma medida de divergência simétrica no estudo da desigualdade de renda.

Dissertação de Mestrado submetida ao Programa de Pós-graduação em Economia - Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para a obtenção do título de Mestre em Ciências. A original encontra-se disponível na FEA-RP/USP.

Aprovada em:

Banca Examinadora

**Prof. Dr. Walter Belluzzo
Junior (Orientador)**
FEA-RP/USP

Prof. Dr. Alexandre Chibebe Nicolella
FEA-RP/USP

**Prof. Dr. Naércio Aquino Menezes
Filho**
FEA-SP/USP

Prof. Dr. Sergio Pinheiro Firpo
INSPER

Agradecimentos

Agradeço primeiramente à minha família, especialmente minha mãe Liliana, por todo o apoio ao longo do período de mestrado e aos amigos que fiz nesse período, que tornaram as (muitas) horas de estudo mais agradáveis. Agradeço também a minha namorada Leticia por ter me apoiado nos momentos pessoais mais difíceis, sem sua ajuda não estaria onde me encontro hoje.

Por fim, agradeço a todos os professores pelas aulas que tive ao longo do mestrado e que são os responsáveis pelo conhecimento que agreguei ao longo desses dois anos de mestrado. Em especial gostaria de agradecer ao Professor Walter Belluzzo pela sua orientação, e ao professor Márcio Laurini por sempre estar disponível para nos ajudar.

Resumo

FERRACINI, M. **Uso de uma medida de divergência simétrica no estudo da desigualdade de renda.** 2017. Dissertação (Mestrado) - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2017.

Essa dissertação tem como objetivo apresentar uma medida de desigualdade distinta das usualmente utilizadas na literatura de desigualdade. Tal medida, aqui denominada índice J, além de possuir a característica de ser simétrica também possibilita conduzir análises distintas daquelas feitas com os índices usualmente utilizados.

Para avaliar a aplicabilidade do índice J foram utilizados dados provenientes da PNAD referente aos anos de 2007, 2011 e 2015 . Apenas dados referentes ao estado de São Paulo foram incluídos. Além de avaliar a evolução da desigualdade no período, decomposição intra e entre grupos foram conduzidas. Testes de hipótese para a desigualdade entre grupos, uma das possibilidades apresentada pelo índice J, foram conduzidos. Também foi avaliado como a presença de erros de medidas não condicionais à renda influenciaria o resultado.

Os resultados apresentados pelo índice J apontam para uma diminuição da desigualdade de renda no período analisado, sendo que a variável educação se apresentou como a característica com maior capacidade de explicar a desigualdade total a partir da desigualdade entre grupos, dentre as variáveis analisadas. A simulação de Monte Carlo conduzida para o teste de hipótese para desigualdade entre grupos também apontou à variável educação como a mais provável de gerar desigualdade. A presença de erros de medida não condicionais à renda não influenciou no resultado final do índice J, porém a simulação de tais erros contribuiu para amenizar o problema de discretização dos dados provenientes da PNAD.

O índice J apresentou-se como uma alternativa viável aos índices usualmente utilizados na literatura de desigualdade, possibilitando algumas análises distintas e que podem auxiliar no estudo da desigualdade de renda.

Palavras-chaves: Desigualdade de renda, Desigualdade entre grupos, Índice J

Abstract

FERRACINI, M. **A symmetric divergence measure applied to the study of income inequality.** 2017. Dissertation (Master Degree) - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2017.

This dissertation aims to present a measure of inequality distinct from those usually used in the inequality literature. Such a measure, here called the J index, besides having the characteristic of being symmetrical also makes it possible to conduct analyzes different from those made with the indexes usually used.

To evaluate the applicability of the J index, data from the PNAD for the years 2007, 2011 and 2015 were used in the analysis. Only data referring to the state of São Paulo were included. In addition to assessing the evolution of inequality in the period, within and between group decomposition were conducted. Hypothesis tests for the inequality between groups, one of the possibilities presented by the index J, were conducted. It was also evaluated how the presence of errors of measures not conditional to the income would influence the result.

The results presented by the index J point to a decrease in income inequality in the analyzed period, with the education variable being the characteristic with the greatest capacity to explain the total inequality from the inequality between groups, among the variables analyzed. The Monte Carlo simulation conducted for the hypothesis test for inequality between groups also pointed to the education variable as the most likely to generate inequality. The presence of measurement errors did not influence the final result of the J index, but the simulation of such errors contributed to soften the problem of discretization of PNAD data.

The index J was presented as a viable alternative to the indexes usually used in the literature of inequality, allowing some different analyzes and that can help in the study of income inequality.

Key-words: Income inequality, Between group inequality, J index

Lista de ilustrações

Figura 1 – <i>Evaluation functions</i> para os índices $Theil_t, Theil_t, I_J$	26
Figura 2 – Densidades estimadas - ano a ano.	36
Figura 3 – Densidades estimadas sobrepostas.	37
Figura 4 – Box-Plot da renda mensal ajustada.	37
Figura 5 – Densidades estimadas - Raça/Cor	42
Figura 6 – Densidades estimadas - Sexo	43
Figura 7 – Densidades estimadas - Educação	44
Figura 8 – Densidades estimadas - Trabalho	45
Figura 9 – Resultado da simulação de Monte Carlo	47
Figura 10 – Densidades estimadas com erro de medida	51

Lista de tabelas

Tabela 1 – Observações referentes a cada ano após tratamento dos dados	33
Tabela 2 – Dados descritivos - Raça/Cor	35
Tabela 3 – Dados descritivos - Sexo	35
Tabela 4 – Estatística descritiva - Educação	35
Tabela 5 – Dados descritivos - Tipo de trabalho	35
Tabela 6 – Renda ajustada (Valores em R\$ de 2015)	36
Tabela 7 – Evolução dos índices de desigualdade de renda	38
Tabela 8 – Resultados decomposição aditiva	40
Tabela 9 – Contribuição de cada característica na decomposição entre grupos - Cor/raça	41
Tabela 10 – Contribuição de cada característica na decomposição entre grupos - Sexo	43
Tabela 11 – Contribuição de cada característica na decomposição entre grupos - Educação	44
Tabela 12 – Contribuição de cada característica na decomposição entre grupos - Trabalho	45
Tabela 13 – Resultado da simulação de Monte Carlo para o teste de desigualdade entre grupos	48
Tabela 14 – Evolução dos índices de desigualdade de renda - Com erro de medida (a = -0,2, b = 0,2)	50
Tabela 15 – Evolução dos índices de desigualdade de renda - Com erro de medida (a = 0, b = 0.2)	50

Sumário

	Sumário	11
1	Introdução	13
2	Revisão da Literatura	14
2.1	Propriedades das medidas de desigualdade	15
2.2	Medidas baseadas no conceito de bem-estar	17
2.3	O índice de Gini	18
2.4	Medidas baseadas na teoria da informação	19
3	Metodologia e Dados	21
3.1	Metodologia	22
3.1.1	Índice J	22
3.1.2	Máxima desigualdade entre grupos	26
3.1.3	Estimação das densidades	28
3.1.4	Simulação dos erros de medida	32
3.2	Dados	32
4	Resultados	34
4.1	Análise exploratória e estimação das densidades.	34
4.2	Resultados do índice J.	38
4.3	Teste para desigualdade entre grupos.	46
4.4	Simulação dos erros de medida.	49
5	Considerações Finais	52
	REFERÊNCIAS	54

1 Introdução

A importância do tema desigualdade na teoria econômica é evidente. Diversos estudos já relataram a relação entre desigualdade e seu impacto em diversos setores da sociedade, como por exemplo criminalidade, saúde e educação.

O termo desigualdade é extremamente amplo e está associado diversas vezes a aspectos estritamente quantitativos e unidimensionais como por exemplo renda, porém pode estar associado a uma medida mais complexa e muitas vezes multidimensional, como ilustrado por Maasoumi e Nickelsburg (1988). Um exemplo de medida multidimensional é apresentado em Justino et al. (2004), onde os autores analisam conjuntamente a desigualdade de renda e de outros atributos como saúde e educação no Brasil.

Os métodos estatísticos usados para mensurar desigualdade são parte central da literatura, e ainda que existam diversos métodos bem estabelecidos, não existe um consenso sobre qual seria o ideal. Dentre as diferentes abordagens existentes temos aquelas que possuem respaldo na teoria da informação (TI).

Todos os índices presentes na literatura de desigualdade são, de alguma forma, funções da distribuição de probabilidade (pmf) associada ao fenômeno que está sendo analisado, e conseqüentemente dos momentos dessa distribuição. A TI e o conceito de entropia possibilita uma abordagem diferente na forma de mensurar a desigualdade, já que duas entropias são iguais se e somente se são decorrentes da mesma pmf.

Nesse trabalho nos baseamos na metodologia adotada por Rohde (2016). Abordamos os aspectos teóricos relacionados a medida de divergência J que a difere das medidas usualmente utilizadas na literatura de desigualdade que também possuem a TI como arcabouço teórico. Com o intuito de avaliar a aplicabilidade do índice J fazemos uso do mesmo para avaliar a evolução da desigualdade de renda no estado de São Paulo no período entre 2007 e 2015. Decomposição intra e entre grupos são realizadas com o intuito de melhor compreender a estrutura da desigualdade de renda e métodos não paramétricos usados para estimar a função densidade de probabilidade (pdf) da variável de interesse.

A opção pelo uso do índice de divergência J , em detrimento de métricas mais usuais,

justifica-se pelo fato desse índice apresentar algumas propriedades teóricas que possibilitam realizar análises e interpretações alternativas potencialmente interessantes. Um desses pontos é a possibilidade de se testar a significância do valor apresentado pela decomposição entre grupos, e também a contribuição de cada característica de um determinado grupo no valor da desigualdade entre grupos.

Devido à possibilidade de erro de medida na variável estudada, foram conduzidas simulações para avaliar o impacto que eventuais erros de medida teriam sob os resultados. Os dados usados são referentes à Pesquisa Nacional por Amostra de Domicílios (PNAD) e abrangem os anos de 2007, 2011 e 2015.

Os resultados apontam para uma diminuição da desigualdade de renda no período de 2007 a 2015, tal diminuição foi acompanhada por uma menor desigualdade entre grupos para os grupos analisados, essa diminuição se encontra em linha com o resultado obtido pelos índices Theil T e Theil L, mais usuais na literatura de desigualdade. A análise de Monte Carlo para o teste de desigualdade entre grupos apontou que a característica educação, além de ser a maior responsável dentre as analisada pela desigualdade entre grupos, também é a que possui maior probabilidade de rejeitar a hipótese nula de não desigualdade entre grupos. A introdução de erro de medida não afetou a desigualdade observada, todavia a introdução de tais erros contribuiu para amenizar o problema de discretização dos dados provenientes da PNAD.

2 Revisão da Literatura

Esse capítulo tem como objetivo apresentar as principais medidas de desigualdades existentes na literatura referente ao assunto. Um enfoque especial será dado às medidas que possuem como arcabouço teórico a TI, apresentadas pioneiramente por Theil (1967) e constituem um marco no desenvolvimento das análises de medidas de desigualdades, como apontado por Cowell (2003).

Primeiramente serão apresentadas as propriedades desejáveis para uma medida de desigualdade arbitrária. Em seguida, apresentaremos as medidas de desigualdade que

possuem como arcabouço teórico uma extensão do critério de bem-estar, como apresentado em Atkinson (1970). O índice de Gini que, além amplamente utilizado na literatura de desigualdade, é provavelmente o índice mais difundido à audiência não especializada é apresentado posteriormente. Por fim, serão apresentados os índices derivados dos conceitos da TI e suas principais características.

2.1 Propriedades das medidas de desigualdade

A presente seção tem como objetivo definir as propriedades desejáveis para um determinado índice de desigualdade e também estabelecer uma estrutura sobre a qual é possível desenvolver uma análise de tais índices de desigualdade. Para isso tomaremos como base a notação e terminologia abordada em Cowell (2000)

Seja $x \in \mathfrak{X}$, onde \mathfrak{X} é um intervalo em \mathbb{R} , a renda de um indivíduo. Seja também \mathfrak{F} o espaço de todas as funções de distribuição univariada válidas no contexto analisado. Sendo assim, temos que a distribuição de renda é um elemento $F \in \mathfrak{F}$, onde F possui suporte em \mathfrak{X} . Nesse caso $F(x)$ possui a interpretação usual de uma função de distribuição, ou seja, é a porcentagem da população com renda menor ou igual a x . Definimos então uma medida de desigualdade como uma função $I : \mathfrak{F} \mapsto \mathbb{R}$

Dado a estrutura estabelecida anteriormente, Cowell (2000) define um conjunto de características padrões para uma medida de desigualdade.

- Definição 1 - Princípio da transferência: $I(G) > I(F)$ se a distribuição G pode ser obtida a partir de F por uma transformação *mean-preserving spread*.
- Definição 2 - \mathfrak{F} -Independence: Seja $\tau : \mathbb{R} \mapsto \mathbb{R}$ uma função contínua estritamente monotônica e defina $\mathfrak{X}^{(\tau)} := \{\tau(x) : x \in \mathfrak{X}\}$. Definimos \mathfrak{F} -Independence se $\forall \tau \in \mathfrak{F} : I(F^{(\tau)}) = I(F)$.
- Definição 3 - Separabilidade aditiva: existem funções $\phi : \mathfrak{X} \mapsto \mathbb{R}$ e $\psi : \mathbb{R}^2 \mapsto \mathbb{R}$ tal que $I(F) = \psi\left(\mu(F), \int \phi(x)dF(x)\right)$, com $\mu(F) := \int x dF(x)$

As definições apresentadas anteriormente abrangem um conjunto de medidas maior do que o de interesse na análise de desigualdade. Um exemplo, é que nem todas as transformações abarcadas pela definição 2 fazem sentido no contexto da análise de desigualdade. Sendo assim, estamos interessados em um subconjunto daquele contemplado pelas definições apresentadas.

Tal subconjunto origina um grupo de axiomas desejáveis às medidas de desigualdade. Como explicitado por Rohde (2016) e Cowell (2000), temos que:

- axioma 1 - Independência de escala: $I(x) = I(\lambda x), \forall \lambda > 0$
- axioma 2 - Anonimato: $I(\mathbf{x}) = I(\mathbf{y})$, onde $\mathbf{y} \in \mathfrak{X}$ é uma permutação de \mathbf{x} .
- axioma 3 - Princípio da população: Seja \mathbf{z} um vetor $(n \times g)$ -dimensional onde $g \in \mathbb{N}_{++}$, com $z_i \in X$ ocorrendo com frequência g , então $I(\mathbf{x}) = I(\mathbf{z})$
- axioma 4 - Princípio de Pigou-Dalton: Seja $x_i < x_j$ e $x_i - \delta \geq \inf(x)$, onde $\delta > 0$. Então $I(x_1, \dots, x_i, \dots, x_j, \dots, x_n) < I(x_1, \dots, x_i - \delta, \dots, x_j + \delta, \dots, x_n)$.
- axioma 5 - *Diminishing transfers*: O efeito marginal de uma transferência é decrescente.
- axioma 6 - Aditivamente decomponível: O índice pode ser escrito como $I(\mathbf{x}) = I^{\mathbf{W}} + I^{\mathbf{B}}$, onde $I^{\mathbf{W}}$ representa a desigualdade total intragrupos e $I^{\mathbf{B}}$ a desigualdade total entre grupos.

Dentre os axiomas apresentados anteriormente, a decomposição aditiva é de especial interesse na literatura já que nos possibilita compreender melhor o resultado apresentado pelo índice proposto em uma determinada análise. É interessante ressaltar que na literatura de desigualdade a decomposição aditiva pode ser definida perante hipóteses um pouco diferentes, como em Bourguignon (1979) e Shorrocks (1980).

2.2 Medidas baseadas no conceito de bem-estar

Como descrito por Cowell e Flachaire (2015) os índices baseados no conceito de bem-estar englobam um conjunto de ferramentas usadas na análise distribucional. Dentre eles podemos citar as medidas de desigualdade e também os índices de pobreza.

A abordagem baseada no conceito de bem-estar se deu a partir do trabalho de Atkinson (1970), no qual o autor resgata a ideia subjacente ao artigo de Dalton (1920). De acordo com Atkinson (1970), praticamente todo o trabalho empírico sobre análise de desigualdade fazia uso de um índice, como por exemplo o índice de Gini, sem que existisse nenhuma explicação acerca da escolha de tal índice em detrimento de outro qualquer. A ideia de que por trás de qualquer medida existe o conceito de bem-estar, já havia sido introduzida por Dalton (1920) e até então sido negligenciada, no contexto da análise de desigualdade.

Perante tal contexto, Atkinson (1970) propõe uma medida de desigualdade que leva em consideração uma função de bem-estar. Essa proposta foi influenciada pelo trabalho de Rothschild e Stiglitz (1969) onde os autores abordam o tema de aversão ao risco. Atkinson (1970) considera uma função de bem-estar social dada por

$$W = \int_0^{\bar{x}} U(x)f(x)dx. \quad (1)$$

Note que a equação (1) é simétrica e aditivamente separável.

Atkinson (1970) impõe uma restrição na forma funcional da função de utilidade em (1) para que o índice seja invariante a mudanças proporcionais. Sendo assim, temos que $U(x)$ é dado por

$$\begin{aligned} U(x) &= A + B \frac{x^{1-\epsilon}}{1-\epsilon}, & \epsilon \neq 1 \\ U(x) &= \exp(x), & \epsilon = 1 \end{aligned} \quad (2)$$

onde $\epsilon \geq 0$, e $A, B \in \mathbb{R}$.

O parâmetro ϵ representa a aversão à desigualdade, e vale notar que a equação (2) apresenta as mesmas características de uma função de aversão relativa ao risco constante (CRRA). Para uma função de distribuição discreta e dado as equações (1) e (2) definimos o índice de Atkinson como

$$A_I(x) = 1 - \left[\sum_i \left(\frac{x_i}{\mu} \right)^{1-\epsilon} f(x_i) \right]^{1/(1-\epsilon)}. \quad (3)$$

Note que um maior valor de ϵ , maior aversão à desigualdade, implica em mais peso para transferências de renda para a cauda inferior da distribuição e menos peso para a cauda superior, explicitando assim o fato do parâmetro refletir aversão à desigualdade.

2.3 O índice de Gini

Como apontado por Litchfield (1999) o índice de Gini não satisfaz o axioma 6, apresentado na seção 2.1, quando os vetores de renda entre os grupos se sobrepõe, como é usual. Outras formas de decompor o índice de Gini foram propostas, como apresentado por Pyatt, Chen e Fei (1980). Todavia, tais métodos não fazem com que o índice de Gini atenda a todos axiomas citados anteriormente e pode ser de difícil interpretação.

Mesmo perante tais limitações, o índice de Gini é amplamente utilizado. Definimos o índice de Gini como Gini (1912)

$$I_{Gini}(X) := \frac{1}{2n^2\bar{x}} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|. \quad (4)$$

O índice de Gini está limitado entre 0 e 1, onde 1 indica o maior nível de desigualdade e zero ausência de desigualdade.

Um ponto interessante do índice de Gini é sua relação com a curva de Lorenz. Essa curva é uma ferramenta gráfica utilizada na visualização e comparação da desigualdade de renda, ela nos mostra a relação entre a proporção acumulada da renda e a proporção

acumulada da população. Duclos (2002) define a curva de Lorenz como

$$L(p) = \frac{1}{\mu} \int_0^p Q(q) dq \quad (5)$$

onde $Q(q) = F^{-1}(p)$, ou seja, é o quantil p da distribuição. A relação entre o índice de Gini e a curva de Lorenz é dada por

$$I_{Gini}(\cdot) = 2 \int_0^1 (p - L(p)) dp \quad (6)$$

2.4 Medidas baseadas na teoria da informação

Essa seção tem como objetivo descrever a relação da TI com alguns dos principais índices de desigualdade. Para tal tomamos como base a notação e os conceitos presentes em Cowell (2011).

O uso de conceitos advindos da TI se dá em diversas áreas da economia, como apresentado por Ullah (1996). Os índices de desigualdade que possuem como arcabouço teórico a TI se deu a partir do trabalho de Theil (1967), onde o autor desenvolve a ideia de que podemos interpretar a desigualdade como um resultado proveniente da informação existente na distribuição de renda.

A TI tem como objetivo mensurar a informação proveniente de um determinado evento dentre um conjunto de eventos possíveis. Suponha um conjunto de eventos e_1, e_2, \dots, e_n os quais possuem probabilidade p_1, p_2, \dots, p_n , com $p_i \in [0, 1]$, $i = 1, 2, \dots, n$. O objetivo é assinalar um valor para uma função $h(p_i)$, onde $h'(p_i) < 0$ e portanto para dois eventos quaisquer e_1 e e_2 se $p_1 > p_2$ então $h(p_2) > h(p_1)$. Isso significa que eventos menos prováveis contêm mais informação, sendo assim temos que $h(0) = \inf$ e $h(1) = 0$, como apontado por Cover e Thomas (2012).

Outra propriedade desejável da função $h(\cdot)$ está relacionada à propriedade de

eventos independentes. Sejam e_1 e e_2 eventos independentes, então temos que

$$h(p_1 p_2) = h(p_1) + h(p_2). \quad (7)$$

Dado $h'(p_i) < 0$, $h(1) = 0$, $h(0) = \inf$, e a equação (7) temos que $h = -\log(p)$, como apontado por Cowell (2003).

Um dos principais conceitos advindos da TI é o de entropia. Entropia pode ser entendida como uma medida de incerteza sobre uma variável aleatória. Definimos entropia, $H(x)$, como

$$H(x) = - \sum_{i=1}^n p_i \log p(i). \quad (8)$$

Vale notar, que pela equação (8), a entropia é função da pmf de X, logo não depende do valor do evento x e sim da probabilidade associada a esse evento.

Outro conceito importante é o de entropia relativa, ou divergência de Kullback–Leibler, que pode ser interpretada como uma medida de distância entre duas distribuições. A entropia relativa, $D(p||q)$, é interpretada como uma medida de ineficiência quando assumimos que a distribuição é q enquanto na realidade a verdadeira distribuição é p , como apontado por Cover e Thomas (2012). Definimos entropia relativa entre duas pmf $p(x)$ e $q(x)$ como

$$D(p||q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}. \quad (9)$$

A entropia relativa apresenta a propriedade de ser sempre não negativa, e possui valor zero se e somente se $p = q$. Vale notar que a entropia relativa não é uma medida de distância matemática já que não satisfaz os critérios de desigualdade triangular e simetria.

Os conceitos de entropia e entropia relativa foram aplicados por Theil (1967) no contexto de distribuição de renda. O conceito de probabilidade relacionada a um evento tem como contraparte a parcela da renda, a distribuição usada como comparação toma a

forma de perfeita igualdade, ou seja, um vetor de médias. Sendo assim definimos o índice de Theil como

$$Theil_T(\mathbf{x}) = \sum_{i=1}^N \frac{x_i}{\bar{x}} \ln \left(\frac{x_i}{\bar{x}} \right) \quad (10)$$

onde x_i representa a renda de cada indivíduo. Note que a equação (10) é caso em que $D(\mathbf{x}||\bar{x})$, como a entropia relativa não é simétrica temos também o caso onde $D(\bar{x}||\mathbf{x})$. Esse último caso da origem ao chamado segundo índice de Theil. Definimos o segundo índice de Theil como

$$Theil_L(\mathbf{x}) = - \sum_{i=1}^N \ln \left(\frac{x_i}{\bar{x}} \right). \quad (11)$$

O índice de Theil é um caso especial de uma classe mais ampla de índices denominada índices de entropia generalizado (GE). Podemos definir o índice GE como

$$GE_\alpha(\mathbf{x}) = \frac{1}{\alpha(\alpha - 1)} \left[\frac{1}{n} \sum_{i=1}^n \left[\frac{x_i}{\bar{x}} \right]^\alpha - 1 \right] \quad (12)$$

com $\alpha \in \mathbb{R}$. Note que para $\alpha \rightarrow 1$ e $\alpha \rightarrow 0$ a equação (12) é análoga as equações (10) e (11) respectivamente. O parâmetro α na equação (12) captura a sensibilidade a mudanças distribucionais. Valores maiores de α impactam mais a cauda superior da distribuição, enquanto valores menores tem maior impacto na cauda inferior.

Como apontado por Cowell (2003) o índice GE não é interessante apenas por generalizar a ideia proposta por Theil, mas também por satisfazer todos axiomas apresentados nas seção 2.1.

3 Metodologia e Dados

Este capítulo tem como objetivo descrever com detalhes a metodologia usada nesse trabalho. O trabalho se propõe a utilizar o índice de divergência J para analisar a evolução da desigualdade de renda no Brasil. A medida de divergência J foi apresentada

primeiramente por Jeffreys (1946), na qual o autor aponta a característica do índice ser uma medida de discrepância entre duas distribuições de probabilidade.

As distribuições de probabilidade foram estimadas através de métodos não paramétricos. Esse métodos também serão descritos com detalhes nessa seção. Como a variável de interesse desse trabalho normalmente apresenta erro de medida, simulações foram conduzidas com o intuito de avaliar o comportamento do índice proposto frente a tais erros de medida.

Os dados usados nesse trabalho são provenientes Pesquisa Nacional por Amostra de Domicílios (PNAD). Os procedimentos referentes ao tratamento dos dados usados também são abordados com detalhe nesta seção

3.1 Metodologia

3.1.1 Índice J

Devido ao índice apresentar como característica a concepção de distância entre distribuições, é importante definir o conceito de distância. Definimos distância, matematicamente, como uma função que define a distância entre cada par de elementos de um determinado conjunto. Seja X um conjunto e $d : X \times X \rightarrow \mathbb{R}_+$ onde $\forall x, y, z \in X$ as seguintes propriedades são satisfeitas

1. $d(x, y) \geq 0$
2. $d(x, y) = 0 \iff x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, z) \leq d(x, y) + d(y, z)$

Funções que satisfazem todas propriedades apresentadas anteriormente são chamadas de métricas, as que satisfazem as propriedades 1 a 3 são chamadas semi-métricas e as que atendem apenas as propriedades 1 e 2 são conhecidas como pré-métricas.

Sejam $p(x)$ e $q(x)$ duas pmf, definimos então o índice J como

$$I_J(p||q) = \sum_{x \in X} (q(x) - p(x)) \ln \left(\frac{q(x)}{p(x)} \right) \quad (13)$$

Note que a equação (13) apresenta uma relação estreita com os índices apresentados nas equações (10) e (11). Seja $q = \bar{x}$ e $p = \mathbf{x}$, então temos que

$$I_J(\mathbf{x}||\bar{x}) = Theil_T(\mathbf{x}) + Theil_L(\mathbf{x}). \quad (14)$$

Reescrevendo a equação (13) considerando $p = \mathbf{x}$ e $q = \bar{x}$ temos que

$$I_J(\mathbf{x}||\bar{x}) = \frac{1}{\theta} \sum_{i=1}^n (x_i - \bar{x}) \ln \left(\frac{x_i}{\bar{x}} \right) \quad (15)$$

onde $\theta = \sum_{i=1}^n x_i$.

Uma característica importante do índice J é o fato dele ser uma semi-métrica, diferente dos índices apresentados nas equações (10) e (11) que, por satisfazerem apenas as propriedades 1 e 2 apresentadas anteriormente, são considerados pré-métricas. Logo o índice J possui a característica de ser simétrico, ou seja, $I_J(p||q) = I_J(q||p)$.

O índice J atende aos axiomas citados na seção 2, e portanto atende aos critérios apontados por Bourguignon (1979) para decomposição aditiva. Como exposto por Rohde (2016) o índice J possui decomposição intra e entre grupos.

Seja k o número de grupos onde cada grupo possui n^j indivíduos, com $j = 1, \dots, k$. Defina $p^j = n^j/n$ e $q^j = p^j \bar{x}^j / \bar{x}$, onde \bar{x}^j é a renda média do grupo j . Defina também as seguintes equações

$$L^j(\bar{x}^j||x^j) = \sum_{i=1}^{n^j} -\frac{1}{n^j} \ln \left(\frac{x_i^j}{\bar{x}^j} \right)$$

$$L^B(\bar{x}||\bar{x}^j) = \sum_{j=1}^k -\frac{n^j}{n} \ln \left(\frac{\bar{x}^j}{\bar{x}} \right)$$

$$T^j(x^j|\bar{x}^j) = \sum_{i=1}^{n_j} \frac{1}{n^j} \left(\frac{x_i^j}{\bar{x}^j} \right) \ln \left(\frac{x_i^j}{\bar{x}^j} \right)$$

$$T^B(\bar{x}^j|\bar{x}) = \sum_{j=1}^k -\frac{n^j \bar{x}^j}{n \bar{x}} \ln \left(\frac{\bar{x}^j}{\bar{x}} \right)$$

Após extensas manipulações algébricas a equação (13) pode ser reescrita como

$$I_J(\mathbf{x}|\bar{x}) = \underbrace{\sum_{j=1}^k \{q^j T^j(x^j|\bar{x}^j) + p^j L^j(\bar{x}^j|x^j)\}}_{I_J^W} + \overbrace{T^B(\bar{x}^j|\bar{x}) + L^B(\bar{x}|\bar{x}^j)}^{I_J^B} \quad (16)$$

onde o primeiro termo do lado direito da equação corresponde ao componente intra grupos (I_J^W), e os dois últimos termos correspondem ao componente entre grupos (I_J^B)¹.

Uma das vantagens apresentadas pelo índice de divergência J é a possibilidade de testar a significância do componente I_J^B . Como apresentado por Rohde (2016) temos que

$$I_J^B(\bar{x}^j|\bar{x}) \approx \sum_{j=1}^k \frac{(q^j - p^j)^2}{p^j} = \frac{1}{n} \chi_k^2 \quad (17)$$

onde k representa o número de grupos, q^j representa a frequência observada e p^j a frequência esperada. Onde a aproximação apresentada na equação (18) é obtida através de uma expansão de Taylor de 1ª ordem². Sendo assim, Rohde (2016) propõe que podemos formular um teste *ad hoc* para a hipótese nula de não existência de desigualdade entre grupos utilizando I_J^B como estatística teste. Conforme demonstrado por Rohde (2016) rejeitamos a hipótese nula ao nível de 5% quando

$$I_J^B(\bar{x}^j|\bar{x}) \geq \frac{1}{2n} \{1.645 + \sqrt{2(k-1) - 1}\}^2 \quad (18)$$

Um ponto importante a ser observado é que o resultado do teste proposto é diretamente afetado por dois fatores, o tamanho da amostra (n) e o número de grupos (k).

¹ A demonstração da decomposição aditiva para o índice J pode ser encontrada em Rohde (2016).

² Borrell e Talih (2011) e Rohde (2016) apresentam com mais detalhes a relação do componente I_J^B com o teste chi-quadrado de Pearson.

Sendo que tais fatores atuam em direções opostas, ou seja, um maior tamanho amostral contribui para a rejeição da hipótese nula enquanto um maior número de grupos contribui para a não rejeição da hipótese nula, dado o valor de I_J^B .

O índice J também possui uma propriedade matemática que pode nos auxiliar na interpretação de alguns resultados. Além dos axiomas apresentados na seção 2 o índice J também satisfaz mais uma propriedade

- *Axioma 7*: O índice J é uma função aditiva de n divergências no $\mathbb{R}_{++}^1 \times \mathbb{R}_{++}^1$.

Tal propriedade permite avaliarmos a contribuição de cada indivíduo para a desigualdade total, ou mesmo a contribuição de cada grupo para o valor total da desigualdade entre grupos.

Como apresentado por Rohde (2016), o axioma 7 permite reescrever a equação (15) como $I(x) = (1/n) \sum_{i=1}^n f(x; \bar{x})$, onde a *evaluation function* $f(x; \bar{x})$ no caso do índice J gera as divergências pontuais. A figura 1 mostra o comportamento das diferentes funções $f(x; \bar{x})$ para os índices de Theil e para o índice J.

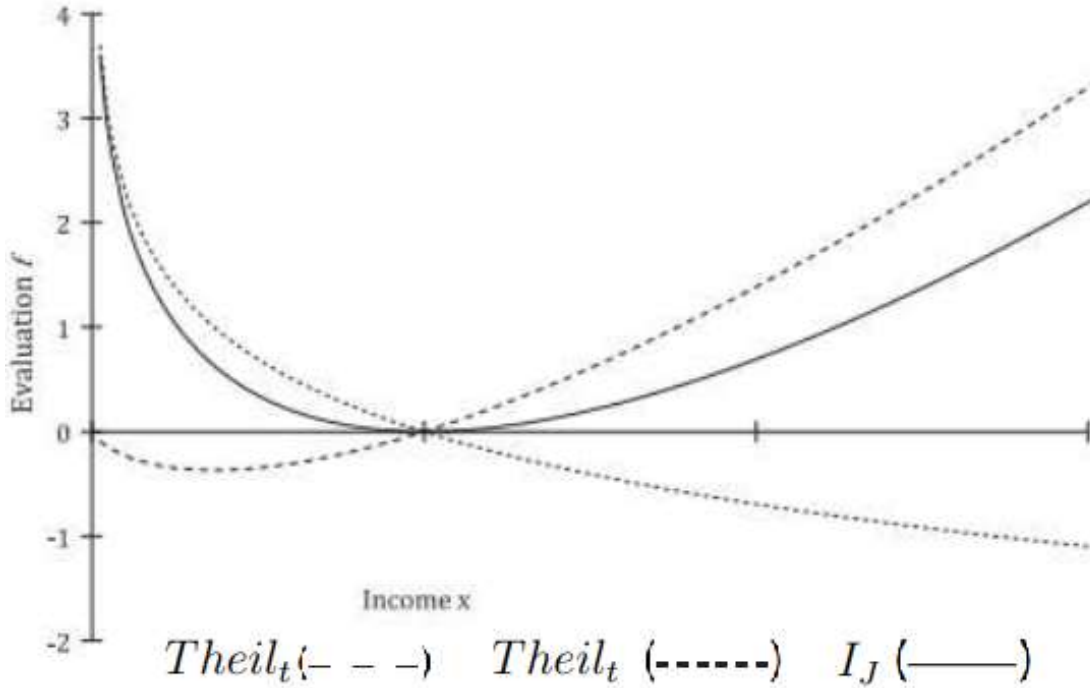
Como a figura 1 mostra, as *evaluation functions* dos índices de Theil assumem valores negativos em alguns pontos, esse comportamento impossibilita que esses índices gerem divergências pontuais pois infringe a propriedade de que uma medida de distância é sempre não negativa.

Assim como os índices de Theil, o índice J também é um caso específico de uma classe de índices. Como apontado por Cowell (2011) no caso da medidas de desigualdade as propriedades $h(0) = \inf$ e $h(1) = 0$, apresentadas na seção 2.4, fazem pouco sentido e podem ser descartadas. Sendo assim a função $h(\cdot)$ pode ser da forma $h(p) = (1 - p^\gamma)/\gamma$. Considerando essa função temos uma medida de entropia generalizada simétrica definida como Ullah (1996)

$$I_\gamma(\mathbf{x}||\bar{x}) = \frac{1}{n\gamma} \sum_{i=1}^n \left\{ \left(\frac{x_i}{\bar{x}} - 1 \right) \left[\left(\frac{x_i}{\bar{x}} \right)^\gamma - 1 \right] \right\}, \quad \gamma \neq 0 \quad (19)$$

Sendo que quando $\gamma \rightarrow 0$ temos o índice J apresentado na equação (15).

Figura 1 – *Evaluation functions* para os índices $Theil_t, Theil_t, I_J$.



Fonte: Rohde (2016)

3.1.2 Máxima desigualdade entre grupos

A possibilidade de decomposição dos índices de entropia generalizada em desigualdade entre grupos e intra grupos possibilita um melhor entendimento do objeto de estudo. Uma das métricas usuais é quanto da desigualdade total é explicada pela desigualdade entre grupos.

Sendo assim, usando a notação de Elbers et al. (2008), seja I o índice de desigualdade e $I_B(\Pi)$ o parâmetro de desigualdade entre grupos para uma determinada partição Π . Definimos então a parte da desigualdade explicada pelas diferenças entre grupos como

$$R_B(\Pi) = \frac{I_B(\Pi)}{I} \quad (20)$$

Como apontado por Elbers et al. (2008) um ponto importante a ser notado é que o valor de $I_B(\Pi)$ é diretamente relacionado a granularidade da partição Π . Sendo assim, para

duas características arbitrárias quaisquer x e y sempre temos que $R_B(\Pi_{x \cap y}) \geq R_B(\Pi_x)$ e $R_B(\Pi_{x \cap y}) \geq R_B(\Pi_y)$.

Dado o fato que o valor de $R_B(\Pi)$ é diretamente atrelado a granularidade da partição Π existe muita discussão quanto à interpretabilidade e relevância de tal métrica. Ou seja, dizer que uma dada característica x é responsável por apenas 10% da desigualdade pode não ser um argumento significativo.

Dentro de tal conjuntura, Elbers et al. (2008) propõe uma medida para comparar $I_B(\Pi)$ contra um cenário contrafactual construído como base nos mesmos dados, mesmo número de grupos e seus respectivos tamanhos, porém onde a renda é distribuída de forma a maximizar a $I_B(\Pi)$. Como apresentado em Elbers et al. (2008) podemos definir tal medida como

$$\hat{R}_B(\Pi) = \frac{I_B(\Pi)}{\max\{I_B|\Pi(j(n), J)\}} \quad (21)$$

O denominador da equação (21) consiste na máxima desigualdade entre grupos possível dado J sub-grupos de tamanho $j(n)$, como descrito em Elbers et al. (2008). Um ponto importante a ser ressaltado é que a métrica proposta na equação (21) não é necessariamente crescente para partições de maior granularidade. Isso porque o termo $\max\{I_B|\Pi(j(n), J)\}$ pode variar mais relativamente ao termo $I_B(\Pi)$, para uma determinada permutação de Π quando olhamos partições de maior granularidade.

Elbers et al. (2008) sugere impor algumas restrições à métrica proposta na equação (21). Os autores sugerem que o número de grupos e suas proporções sejam o mesmo que em $I_B(\Pi)$ e que os grupos sejam organizados de acordo com sua renda média observada. Isso faz com que o cálculo de $\hat{R}_B(\Pi)$ seja direto, já que as restrições impostas limitam o número de permutações a 1, ao mesmo tempo que cria um cenário contrafactual mais realista.

3.1.3 Estimação das densidades

Os índices de desigualdade por serem uma medida resumo acabam não mostrando fatos importantes sobre a estrutura da desigualdade. Sendo assim é importante fazer uso de outras técnicas para compreender melhor a estrutura da desigualdade.

A análise da pdf de uma variável aleatória é um recurso poderoso na descrição das propriedades da variável de interesse. Entre os métodos existentes para a estimação dessas densidades estão os métodos não paramétricos.

Os métodos não paramétricos possuem a vantagem de não impor nenhuma estrutura no que diz respeito a forma funcional da pdf, porém normalmente necessitam de uma amostra de tamanho maior, quando comparado aos métodos paramétricos, para que se obtenha resultados robustos.

Dentre os diversos métodos não paramétricos temos o baseado em *kernels*. Seja X uma variável aleatória com densidade f_X , e suponha que tenhamos uma amostra independente e identicamente distribuída (iid) de tamanho n , X_1, \dots, X_n , da variável aleatória em questão.

Nosso objetivo é estimar f_X a partir da amostra observada. Como apontado por Silverman (1986), dado a definição de pdf, uma vez que a variável aleatória X possui densidade f_X , temos que

$$f_X(x) = \lim_{h \rightarrow 0} \frac{1}{2h} P_x(x - h < X < x + h).$$

Definindo uma função de ponderação $w(x)$, onde

$$w(x) = \begin{cases} \frac{1}{2} & \text{se } |x| < 1 \\ 0 & \text{c.c} \end{cases} \quad (22)$$

é possível escrever o seguinte estimador para f_X

$$\hat{f}_X(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} w\left(\frac{x - X_i}{h}\right) \quad (23)$$

onde h é o parâmetro de *bandwith*. O estimador apresentado na equação (23) é conhecido como estimador *naive*.

É possível generalizar o estimador proposto, já que o mesmo não possui algumas propriedades técnicas desejáveis. Tal generalização é realizada através da substituição da função de ponderação por uma função *kernel*. Definimos a função *kernel*, $K(x)$, como uma função que satisfaz a seguinte condição

$$\int_{x \in \text{Supp}(X)} K(x) dx = 1. \quad (24)$$

Normalmente a função *kernel* é uma pdf simétrica, porém não é necessário que seja. Dessa forma definimos o estimador *kernel* de f_X com *kernel* $K(\cdot)$ como

$$\hat{f}_X(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right). \quad (25)$$

onde h representa o parâmetro de *bandwith*.

Como apontado por Li e Racine (2007) e Silverman (1986) a escolha da função *kernel* possui pouco impacto no resultado final da análise. Dessa forma, o presente trabalho usou o *kernel* gaussiano, ou seja,

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-(1/2)x^2}. \quad (26)$$

Tal *kernel* foi escolhido por apresentar boas propriedades teóricas.

É importante avaliarmos a performance do estimador proposto, uma das métricas possíveis para realizar tal tarefa é o erro quadrático médio (MSE). Definimos o MSE como

$$MSE(\hat{f}) = E[\hat{f}_X(x) - f_X(x)]^2. \quad (27)$$

Outra medida proposta por Rosenblatt et al. (1956), mais usual no contexto abordado, é o erro quadrático integrado médio (MISE). Definimos o MISE como

$$MISE(\hat{f}_X) = E \int [\hat{f}_X(x) - f_X(x)]^2 dx \quad (28)$$

Como apontado por Silverman (1986) é possível reescrever a equação (28) em função do viés e da variância do estimador. Note que

$$\begin{aligned} MISE(\hat{f}_X) &= \int MSE(\hat{f}) \\ &= \int \{E\hat{f}(x) - f(x)\}^2 dx + \int VAR\hat{f}(x) dx \end{aligned} \quad (29)$$

onde o primeiro termo é a integral do viés ao quadrado e o segundo a integral da variância. Vale também notar que $E\hat{f}(x)$ e $VAR\hat{f}(x)$ são representados respectivamente pelas seguintes equações

$$E\hat{f}(x) = \int \frac{1}{h} K\left(\frac{x-y}{h}\right) f(y) dy \quad (30)$$

$$VAR\hat{f}(x) = \frac{1}{n} \int \frac{1}{h^2} K\left(\frac{x-y}{h}\right)^2 f(y) dy - \left\{ \frac{1}{h} \int K\left(\frac{x-y}{h}\right) f(y) dy \right\}^2 \quad (31)$$

logo é possível obter a expressão exata do MISE substituindo as equações (30) e (31) na equação (29).

Ainda que o resultado da estimativa sofra pouco em decorrência do *kernel* escolhido, o mesmo é bastante sensível à escolha do parâmetro de *bandwith* o que torna a discussão acerca desse tema crucial. Não existe consenso a respeito da melhor metodologia para a escolha do parâmetro h , sendo que muitas vezes o critério é subjetivo e varia de acordo com o propósito da análise que esta sendo realizada e preferência de quem conduz tal análise.

Dentre as diversas metodologias existentes para a escolha do parâmetro de *bandwith* estão aquelas baseadas nos dados, um desses métodos é o de *least-squares cross-validation*. Como apontado por Li e Racine (2007) o *least-squares cross-validation* é um método totalmente automático de seleção do parâmetro de *bandwith* que consiste no princípio de selecionar um valor de h que minimize um determinado critério de eficiência do estimador, normalmente MSE ou MISE, dado uma determinada amostra. De forma geral, o método de *least-squares cross-validation* determina o valor ótimo de h , $\forall x \in Supp f_X(x)$.

Ainda que o método de *least-squares cross-validation* seja totalmente orientado aos dados, alguns problemas podem ocorrer. Silverman (1986) aponta o fato que muitas vezes os dados utilizados são discretizados ou arredondados, e que tal fato pode levar o método a performar mal fazendo com que, muitas vezes, o procedimento apresente um valor ótimo degenerado para o parâmetro h , ou seja, $h = 0$. Tal problema pode ser contornado impondo uma restrição sobre o intervalo em que o parâmetro h é minimizado.

Scott e Terrell (1987) propõe um método denominado *biased cross-validation* (BCV) para a estimação do parâmetro h . O método é baseado na otimização da função de erro quadrático integrado médio assintótico (AMISE) em relação ao parâmetro h . O AMISE é definido como

$$AMISE(h) = \frac{R(K)}{nh} + (p!)^{-2} h^{2p} \mu_p^2 R(f^{(p)}) \quad (32)$$

onde a função $R(\cdot)$ é definida como

$$R(\psi) = \|\psi\|_2^2 = \int_{-\infty}^{\infty} \psi(x)^2 dx. \quad (33)$$

Note que podemos enxergar a função (33) como uma medida da rugosidade de ψ . Dado os critérios apresentados definimos BCV como

$$BCV(h) = \frac{R(K)}{nh} + \frac{\mu_2^2}{2n^2h} \sum_{i < j} \phi(c_{ij}) \quad (34)$$

onde μ_k é o k -ésimo momento do *kernel* $K(\cdot)$.

Perante algumas condições o estimador proposto na equação (34) é assintoticamente normal, com esperança e variância dadas respectivamente por

$$E[BCV(h)] = AMISE(h) + O(n^{-1}),$$

$$VAR[BCV(h)] = \frac{\mu_2^4 R(\phi) R(f)}{8n^2h} + O(hn^{-2})$$

Devido ao fato dos dados usados nesse trabalho apresentarem o problema de discretização descrito anteriormente, foi necessário impor um intervalo como restrição para a otimização da função (34) em relação ao parâmetro h . Dessa forma, temos que

$$h^* \in \left[\frac{hmax}{10}, hmax \right] \quad (35)$$

onde $hmax$ é dado por

$$hmax = n^{\frac{-1}{5}} 1.444 \sqrt{VAR(x)}$$

3.1.4 Simulação dos erros de medida

A existência de erros de medida na análise de variáveis como renda é frequente e um caso já conhecido da teoria econômica. Tal fato pode afetar diretamente as medidas de desigualdade, que podem ser viesadas por esse tipo de erro.

Para avaliar a importância de erros de medida foram conduzidas simulações para identificar como possíveis erros de medida afetariam os resultados encontrados. Foi considerado apenas erro de medida do tipo clássico na variável renda, logo seja x_i a renda observada do indivíduo i e λ_i o erro de medida do mesmo, então a verdadeira renda (VR) do indivíduo i é dada por

$$VR_i = x_i + \lambda_i \quad (36)$$

com o parâmetro λ_i dado por

$$\lambda_i = x_i \times u_i \quad \text{onde } u_i \sim U(a, b). \quad (37)$$

3.2 Dados

Esta seção tem como objetivo descrever com detalhes os dados que foram utilizados nesse projeto e os procedimentos aplicados aos mesmos para a realização da análise. Esse

Tabela 1 – Observações referentes a cada ano após tratamento dos dados

	Número inicial de observações	excluídos (%)	Número final de observações
2007	44774	45.98	20585
2011	38183	46.26	17662
2015	39962	45.47	18169

projeto utilizou dados provenientes da (PNAD) referentes aos anos de 2007, 2011 e 2015.

Todos dados usados na análise são referentes aos dados de indivíduos, não foram incluídos dados referente aos domicílios. Dados referentes a renda mensal, sexo, idade, raça/cor, estado de residência, educação e se trabalha no âmbito público ou privado foram extraídas da PNAD para todos anos usados na análise. As variáveis usadas correspondem respectivamente aos códigos V4720, V0302, V8005, V0404, UF, V4745 e V4715 do questionário de pessoas da PNAD.

Apenas pessoas residentes no estado de São Paulo foram incluída na análise. Observações incompletas foram removidas para que todas análises pudessem ser realizadas. Apenas pessoas com idade 15 e 65 anos foram incluídas na análise.

Valores menores que 1 para a variável de renda foram desconsiderados por, provavelmente, serem erros do questionário. Valores de renda acima de 250.000 foram considerados *outliers* e não foram incluídos. Foi criada a variável renda ajustada, que consiste na variável renda ajustada pelo Índice Nacional de Preços ao Consumidor (INPC) para valores de 2015.

Os resultados referentes a característica tipo de trabalho (V4715) foram reagrupados. Observações com resposta “Militar” e “Funcionário público estatutário” foram agrupadas em um grupo “Público”, já as demais características foram agrupadas em outro grupo (“Privado”).

A tabela 1 mostra o número de observações restantes, para cada ano, após termos efetuado os procedimentos de tratamento dos dados descritos anteriormente. É possível observar que o valor não apresenta muita variabilidade entre os anos, logo o resultado da análise provavelmente não é afetado pelo tratamento dos dados.

4 Resultados

Esse capítulo tem como objetivo apresentar os resultados encontrados pelo índice proposto na seção 3.1.1. Na seção 4.1 é realizado uma breve análise descritiva dos dados provenientes da PNAD, com o intuito de compreender melhor a estrutura dos dados analisados. Os resultados dos métodos não paramétricos descritos na 3.1.3 para a análise da pdf da variável renda ajustada também são apresentados nessa subseção.

A seção 4.2 apresenta os resultados do índice J aplicado à análise de desigualdade de renda para o estado de São Paulo. Os resultados das decomposições intra e entre grupos também são apresentados. Com o intuito de avaliar a significância do resultado obtido através da decomposição aditiva para a desigualdade entre grupos, o teste proposto por Rohde (2016), e que foi apresentado na seção 3.1.1, foi aplicado aos resultados obtidos. Devido ao fato do resultado do teste proposto ser influenciado pelo tamanho amostral, foi conduzido uma simulação de Monte Carlo para avaliar a sensibilidade do resultado frente ao tamanho amostral, os resultados do teste da desigualdade entre grupos é apresentado na seção 4.3.

Por fim a, seção 4.4 mostra os impactos que possíveis erros de medida poderiam causar na análise proposta. Mais uma vez os métodos não paramétricos apresentados anteriormente são usados para avaliar o impacto de possíveis erros de medida na estrutura da pdf de interesse.

4.1 Análise exploratória e estimação das densidades.

Devido ao fato de diversas observações terem sido excluídas para realizar as análises propostas, e também ao fato de que desejamos investigar a evolução do índice J de desigualdade nos período de 2007, 2011 e 2015 é interessante vermos a composição da base de dados utilizada com mais detalhes.

As estatísticas descritivas foram realizadas de acordo com as decomposições aditiva do índice J realizadas. Sendo assim, os resultados apresentados nessa seção estão relacionados às variáveis raça/cor, sexo, educação, tipo de trabalho, e também a variável

renda ajustada. As tabelas 2 a 5 mostram respectivamente para cada parâmetro citado anteriormente a participação de cada característica na base analisada (%) e também a renda média ajustada para cada característica ($\bar{\mu}_i$).

Tabela 2 – Dados descritivos - Raça/Cor

	2007		2011		2015	
	%	$\bar{\mu}_i$	%	$\bar{\mu}_i$	%	$\bar{\mu}_i$
Amarela	1.02	3703.59	1.08	3817.28	1.25	4681.64
Branca	66.94	2279.35	62.45	2454.24	59.90	2694.84
indigena	0.24	1571.41	0.22	1876.77	0.32	1871.10
Parda	24.98	1281.92	28.69	1584.79	30.66	1670.47
Preta	6.81	1314.83	7.56	1633.87	7.86	1799.87

Tabela 3 – Dados descritivos - Sexo

	2007		2011		2015	
	%	$\bar{\mu}_i$	%	$\bar{\mu}_i$	%	$\bar{\mu}_i$
Feminino	43.81	1577.91	43.98	1756.02	45.15	1940.00
Maculino	56.19	2288.62	56.02	2470.48	54.85	2655.87

Tabela 4 – Estatística descritiva - Educação

	2007		2011		2015	
	%	$\bar{\mu}_i$	%	$\bar{\mu}_i$	%	$\bar{\mu}_i$
fundamental_completo	11.58	1408.67	11.39	1511.14	10.01	1558.93
fundamental_incompleto	25.93	1225.09	18.49	1348.61	17.13	1449.00
medio_completo	32.73	1699.28	36.38	1748.27	37.11	1770.46
medio_incompleto	6.55	1134.88	6.30	1279.04	6.20	1338.33
sem_instrucao	3.42	975.17	4.23	1260.89	2.55	1486.97
superior_completo	13.74	4992.03	17.27	4802.61	20.61	4850.61
superior_incompleto	6.05	2426.47	5.93	2278.07	6.40	2360.57

Tabela 5 – Dados descritivos - Tipo de trabalho

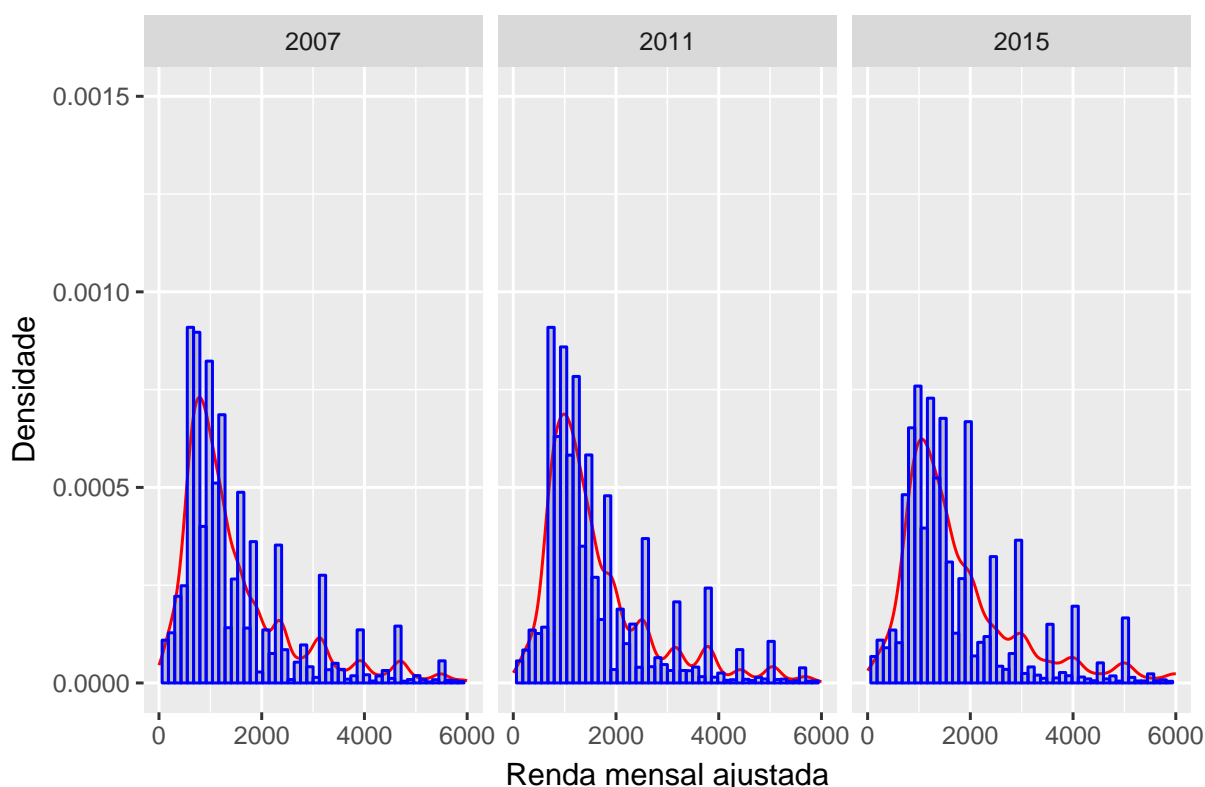
	2007		2011		2015	
	%	$\bar{\mu}_i$	%	$\bar{\mu}_i$	%	$\bar{\mu}_i$
Privado	93.94	1892.00	94.25	2086.36	93.43	2248.07
Público	6.06	3298.45	5.75	3301.28	6.57	3536.42

Como mostrado nas tabelas a composição dos dados não sofre muita alteração no decorrer dos anos, porém vemos que a renda ajustada média dos grupos teve uma variação considerável. Sendo assim é interessante termos uma ideia da evolução da renda no período analisado, a tabela 6 mostra a evolução e algumas estatísticas da renda ajustada para valores de 2015.

Tabela 6 – Renda ajustada (Valores em R\$ de 2015)

	2007	2011	2015
Mínimo	16.00	1.00	10.00
1º quartil	784.00	947.00	1000.00
Mediana	1176.00	1301.00	1500.00
Média	1977.00	2156.00	2333.00
3º quartil	2038.00	2274.00	2500.00
Máximo	78395.00	190755.00	100000.00

Figura 2 – Densidades estimadas - ano a ano.



Para compreender melhor a estrutura da distribuição de renda no período foram estimadas as respectivas pdf de cada ano, para tal o método descrito na seção 3.1.3 foi aplicado. As figuras 2 e 3 mostram o ajuste da densidade estimada para cada ano, já a figura 4 traz o *box-plot* da renda ajustada para cada ano analisado.

Vale notar que o ajuste das densidades estimadas aos histogramas é extremamente sensível ao número de intervalos escolhidos na criação dos histogramas. Outro fator importante de ser ressaltado é que a variável analisada, ainda que seja uma variável contínua, se apresenta discretizada na base que foi utilizada. Isso torna o ajuste da densidade estimada ao histograma uma tarefa ainda mais complexa.

Figura 3 – Densidades estimadas sobrepostas.

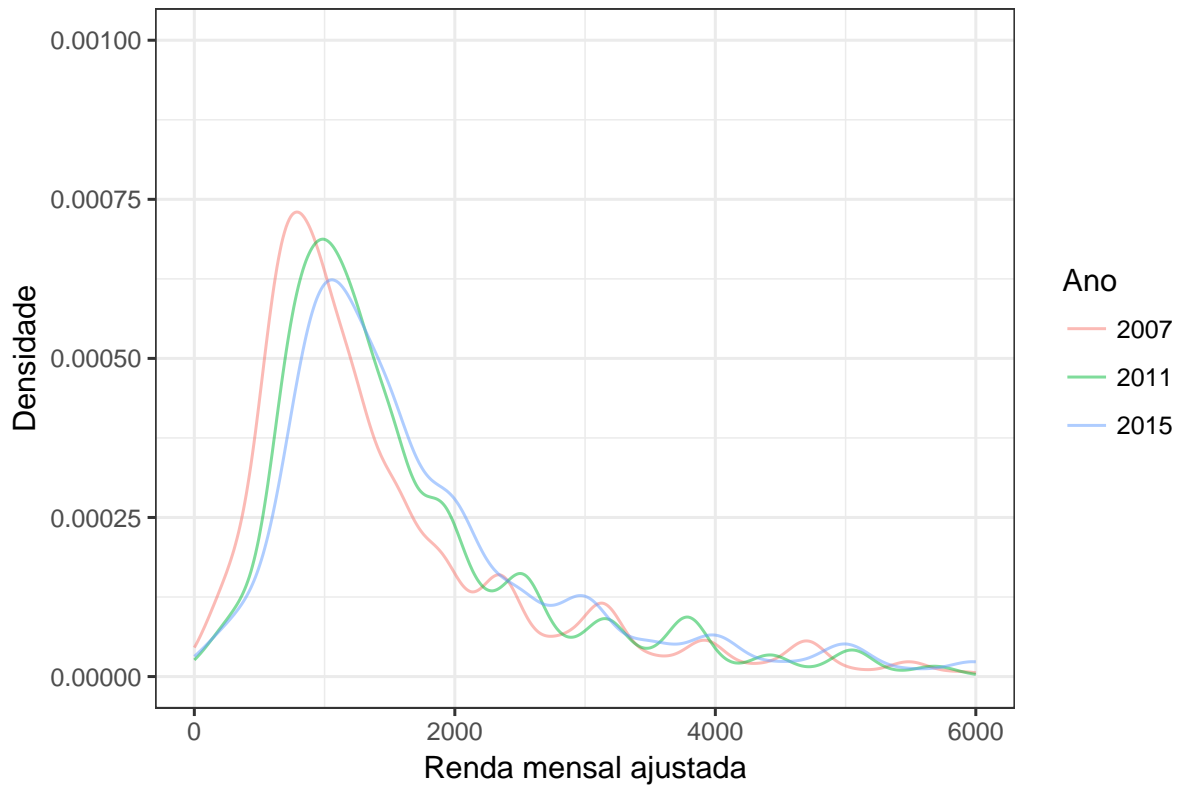
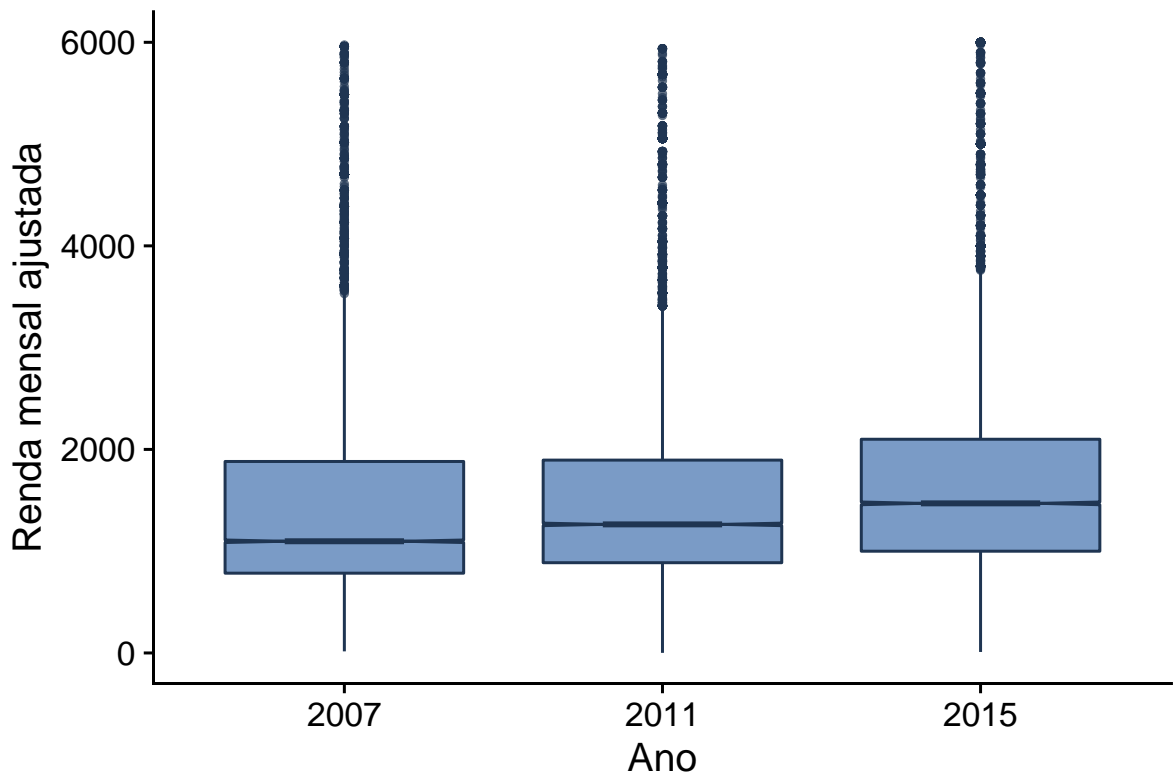


Figura 4 – Box-Plot da renda mensal ajustada.



Devido ao fato da variável renda se apresentar discretizada na base de dados utilizada foi necessário restringir o intervalo de otimização como apresentado na seção 3.1.3. A restrição imposta foi a mesma que a apresentada na equação (35), e a otimização para todas as curvas se deu no extremo inferior do intervalo de otimização. Tal fato provavelmente é decorrente dos dados utilizados serem altamente discretizados e também em decorrência do tamanho amostral. Sendo assim devemos olhar as densidades estimadas com determinada cautela, já que a estrutura dos dados utilizados dificulta a aplicação da metodologia proposta, mesmo com as adequações referentes ao intervalo de otimização.

Ainda que a estrutura dos dados imponha algumas dificuldades técnicas, as densidades estimadas possuem um formato condizente com o que é apresentado na literatura de desigualdade de renda, nos permitindo realizar assim algumas análises. O primeiro ponto a se observar é a mudança na estrutura da densidade estimada. As figuras 2 e 3 nos mostram que as densidades estimadas possuem um formato muito similar entre os anos analisados, sugerindo assim que não houve uma mudança na estrutura da desigualdade de renda. Dentro desse contexto é importante notar que o INPC teve um acúmulo de 74,47% entre 2007 e 2015 (inclusive os anos 2007 e 2015), sendo que entre os anos de 2007 e 2011 o índice acumulado é de 31,66%.

Outro fato interessante é que ambas as figuras anteriores apresentam diversos máximos locais. Tal fato pode ser decorrente principalmente de dois fatores: o alto grau de discretização dos dados ou a existência de *clusters* específicos.

4.2 Resultados do índice J.

Para analisar a evolução da desigualdade de renda no período estudado foi calculado o índice J, esse foi comparado a ambos índices de Theil. A tabela 7 apresenta os resultados para os períodos analisados.

Tabela 7 – Evolução dos índices de desigualdade de renda

	2007	2011	2015
Índice J	0.90	0.82	0.79
Theil L	0.42	0.36	0.36
Theil T	0.49	0.46	0.43

Os resultados apresentados na tabela 7 apontam para uma queda na desigualdade no período avaliado. Um ponto interessante a ser ressaltado é a queda na desigualdade no período entre os anos de 2007 e 2011 mesmo frente a uma queda na renda média, como apresentado na tabela 6.

Para avaliar com maior profundidade os resultados obtidos, decomposição entre grupos e intra grupos foram realizadas. As decomposições foram conduzidas para as variáveis educação, sexo, cor/raça e tipo de trabalho. Além das partições oriundas dessas variáveis, três novas partições (P1, P2, P3) foram derivadas a partir da combinação dessas variáveis. O intuito dessas novas partições é agrupar indivíduos com características semelhantes dentro de um mesmo *cluster*. A composição das partições criadas é dada por

- P1 \rightarrow Educação \times Cor
- P2 \rightarrow Educação \times Cor \times Sexo
- P3 \rightarrow Educação \times Cor \times Sexo \times tipo de trabalho

sendo que o critério utilizado para a criação das partições foi o valor da estatística $R_B(\Pi)$ para variáveis educação, sexo, cor/raça e tipo de trabalho. Sendo assim, optou-se por agrupar primeiramente características que obtiveram um maior valor para a estatística $R_B(\Pi)$ calculada.

A tabela 8 apresenta os resultados da decomposição aditiva para cada uma das partições citadas previamente, e também apresenta os valores das estatística $R_B(\%)$ e \hat{R}_B apresentadas na seção 3.1.2.

A tabela 8 mostra que de forma geral ocorreu uma diminuição tanto da desigualdade entre grupos quanto das estatísticas $R_B(\%)$ e \hat{R}_B . Outro fato interessante se refere a educação ser o mais impactante no que diz respeito a explicar a desigualdade existente, com valores de $R_B(\%)$ entre 30, 64% e 32, 14% para os períodos analisados. Outro ponto importante é a pouca variação existente entre os valor de $R_B(\%)$ e \hat{R}_B para a partição gerada pela variável educação, isso nos mostra que o valor $\max\{I_B|\Pi(j(n), J)\}$ se aproxima muito do índice J de desigualdade calculado.

Tabela 8 – Resultados decomposição aditiva

	2007				2011				2015			
	I_W	I_B	$R_B(\%)$	$\hat{R}_B(\%)$	I_W	I_B	$R_B(\%)$	$\hat{R}_B(\%)$	I_W	I_B	$R_B(\%)$	$\hat{R}_B(\%)$
cor	0.83	0.07	7.50	8.80	0.78	0.04	5.41	6.29	0.73	0.06	7.09	8.21
sexo	0.87	0.03	3.65	6.73	0.79	0.03	3.39	6.71	0.77	0.02	3.02	5.94
educacao	0.61	0.29	32.14	34.93	0.57	0.25	30.64	34.21	0.54	0.25	31.70	34.69
trabalho	0.88	0.02	2.50	4.71	0.81	0.01	1.71	3.11	0.78	0.02	1.94	3.43
P1	0.60	0.31	33.93	34.98	0.56	0.26	31.88	32.84	0.53	0.26	33.13	34.30
P2	0.54	0.37	40.45	41.28	0.51	0.31	38.20	38.96	0.49	0.31	38.66	39.60
P3	0.53	0.37	40.91	41.71	0.50	0.32	38.66	39.40	0.48	0.31	39.16	40.07

Outro ponto interessante a ser ressaltado é que mesmo na partição de maior granularidade (P3) o valor da desigualdade entre grupos corresponde a no máximo 40,91% da desigualdade total, ou seja, o maior valor de $R_B(\%)$ foi de 40,91% sendo que nesse caso o valor correspondente de \hat{R}_B foi de 41,71%. Sendo que a maior variação dos valores de $R_B(\%)$ e \hat{R}_B ocorreu da partição P1 para a P2, isso aponta para o fato de as características educação e cor/raça estão altamente correlacionadas o que faz com que a partição gerada por P1 seja similar a partição gerada apenas pela característica educação implicando em um valor de $R_B(\%)$ e \hat{R}_B muito similar. Já a diferença do valor de $R_B(\%)$ e \hat{R}_B existente entre P1 e P2 indica o contrário, levando nos a crer que a relação entre sexo e educação é menor o que implica em uma partição do espaço relativamente mais granular.

Um ponto importante a ser ressaltado, apontado em Elbers et al. (2008), é que $R_B(\%)$ só pode ser igual a 100% em duas situações; se cada grupo é constituído de um único indivíduo ou, no caso onde existe mais indivíduos que grupos, se todos indivíduos dentro de cada um desses grupos possuírem renda idêntica.

Uma das vantagens apresentadas pelo índice J, que foi apresentada na seção 3.1.1, é fato do índice ser uma função aditiva de n divergências no $\mathbb{R}_{++}^1 \times \mathbb{R}_{++}^1$. Tal propriedade possibilita avaliarmos a contribuição de cada indivíduo para índice, e de forma semelhante nos permite avaliar a contribuição de cada grupo para o valor desigualdade entre grupos. Contudo é importante, antes de mais nada, entendermos o que significa a contribuição de cada grupo para a desigualdade entre grupos.

Como apresentado na equação (16), o resultado da desigualdade entre grupos

depende de três fatores: renda média do grupo, tamanho relativo do grupo e renda média da amostra. Sendo assim um grupo pode apresentar uma alta contribuição para a desigualdade entre grupos se possui uma renda substancialmente menor ou maior do que a média da amostra. já o tamanho relativo do grupo atua como um fator de ajuste.

Devido a tal fato é conveniente avaliar a contribuição de cada grupo em conjunto com algo que nos permita compreender como é a estrutura de cada grupo. Sendo assim, se fez a opção de utilizar as densidades estimadas referentes a cada grupo em conjunto com as estatísticas descritivas apresentadas na seção 4.1, para compreender melhor como cada um desses grupos estaria contribuindo para a desigualdade entre grupos. A tabela 9 apresenta a contribuição de cada grupo quanto à partição gerada pela característica cor/raça e a figura 5 apresenta a densidade estimada para cada grupo que compõe essa partição.

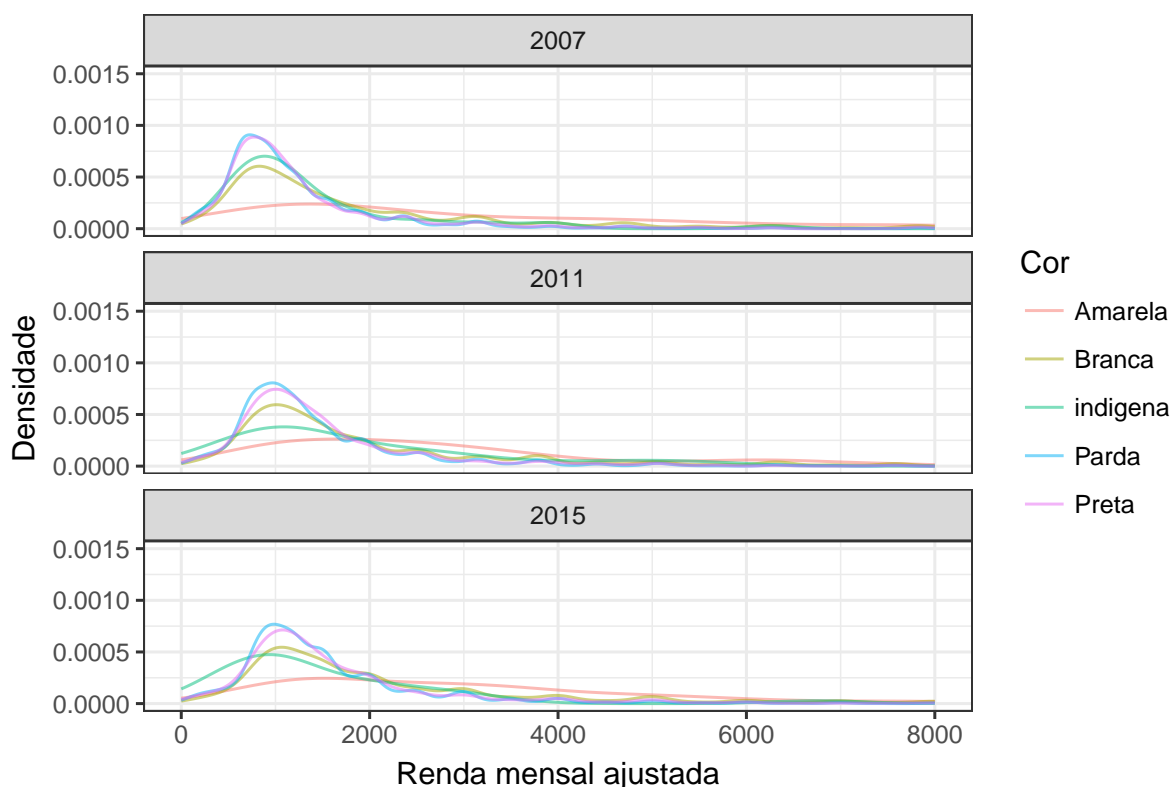
Tabela 9 – Contribuição de cada característica na decomposição entre grupos - Cor/raça

Descrição	2007	%	2011	%	2015	%
Amarela	0.0056	8.27	0.0048	10.70	0.0088	15.70
Branca	0.0145	21.50	0.0112	25.13	0.0134	23.93
indigena	0.0001	0.17	0.0000	0.09	0.0001	0.25
Parda	0.0381	56.30	0.0234	52.65	0.0291	51.82
Preta	0.0093	13.77	0.0051	11.43	0.0047	8.30
Total	0.0676	100.00	0.0445	100.00	0.0561	100.00

Um primeiro ponto importante a ser ressaltado é o fato das características Branca e Parda representarem aproximadamente 75% em todos anos analisados. Outro ponto importante a ser apontado é o aumento da participação dos grupos Parda e Preta no decorrer dos anos analisados, tal fato é apontado na tabela 2. A tabela 9 aponta que os grupos “Branca” e “Parda” são os que mais contribuem para desigualdade entre grupos, porém é importante notarmos que ambos contribuem de maneiras completamente distintas, isso devido ao fato do grupo “Parda” possuir uma renda média expressivamente menor que a média da amostra analisada enquanto o grupo “Branca” possui a maior renda média entre os grupos sendo essa renda média expressivamente superior que a média amostral.

Não obstante tais fatos, a tabela 9 nos mostra uma diminuição expressiva na contribuição dos grupos “Parda” e “Preta” entre os anos de 2007 e 2015. Tal fato atrelado

Figura 5 – Densidades estimadas - Raça/Cor



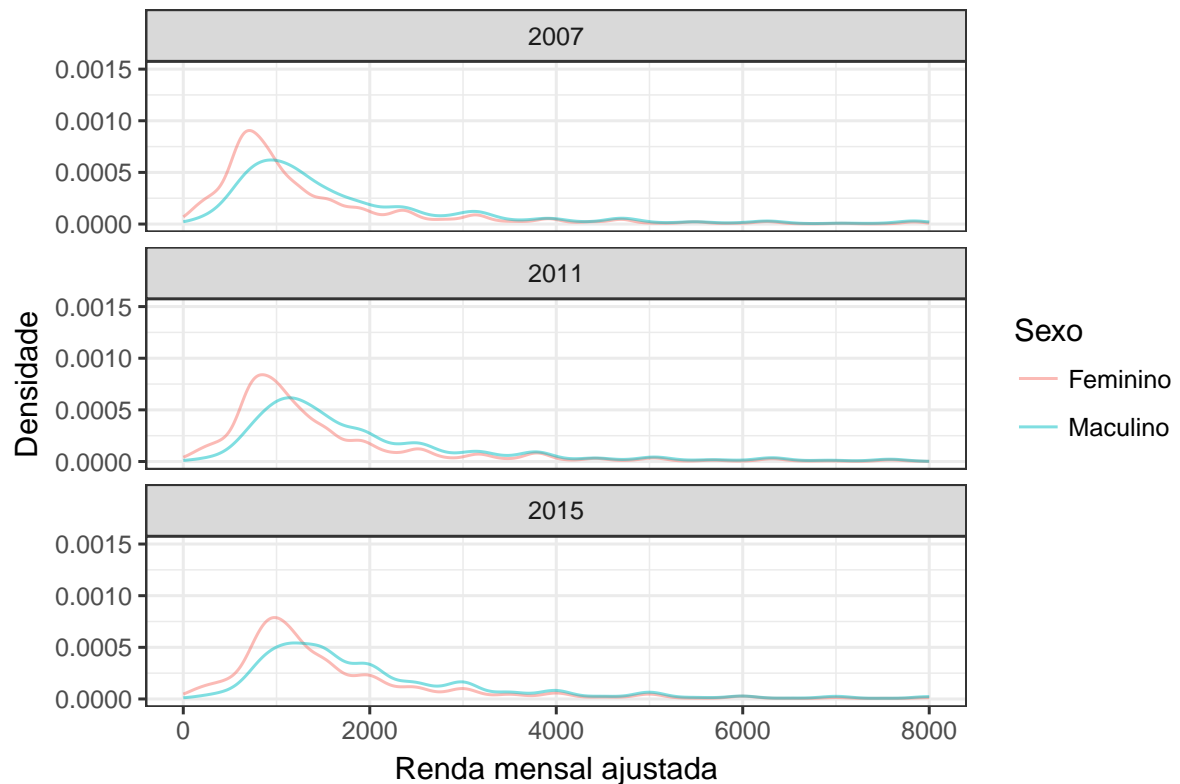
com o aumento da participação de tais grupos no decorrer dos anos analisados aponta para uma mudança na estrutura da distribuição de renda desses grupos. A figura 5 indica esse fato, é claro a mudança nas pdf dos grupos “Parda” e “Preta”. Quando comparamos os resultados para o ano de 2015 com os outros anos é de fácil percepção que as densidades de desses grupos se encontram mais próximas das densidades dos outros grupos para o ano de 2015. Outro fato importante que a figura 5 nos mostra, é a dissonância entre o grupo “Amarela” e os demais.

A tabela 10 apresenta o resultado da decomposição entre grupos para a partição gerada pela característica sexo, já a figura 6 apresenta as densidades estimadas. A figura 6 deixa claro a diferença existente, com a média do grupo sendo consistentemente maior que a do grupo feminino, ainda assim é possível verificar uma leve convergência entre as densidades. Tal fato é suportado pela ligeira queda na participação do grupo feminino na decomposição entre grupos passando de 60,67% no ano de 2007 para 58,69% no ano de 2015, isso em conjunto com uma maior participação amostral do grupo feminino em 2015 quando comparado a 2017, como apontado na tabela 3.

Tabela 10 – Contribuição de cada característica na decomposição entre grupos - Sexo

Descrição	2007	%	2011	%	2015	%
Feminino	0.0200	60.67	0.0168	60.15	0.0140	58.69
Maculino	0.0129	39.33	0.0111	39.85	0.0099	41.31
Total	0.0329	100.00	0.0279	100.00	0.0239	100.00

Figura 6 – Densidades estimadas - Sexo



Os resultados referentes a característica educação são os mais expressivos. Como apontado na tabela 8, a característica educação é responsável pelos maiores valores de R_B e \hat{R}_B , valores esses expressivamente superior quando comparado as outra características analisadas. A tabela 11 apresenta os valores da decomposição entre grupos para a característica educação, já a figura 7 apresenta as densidades estimadas.

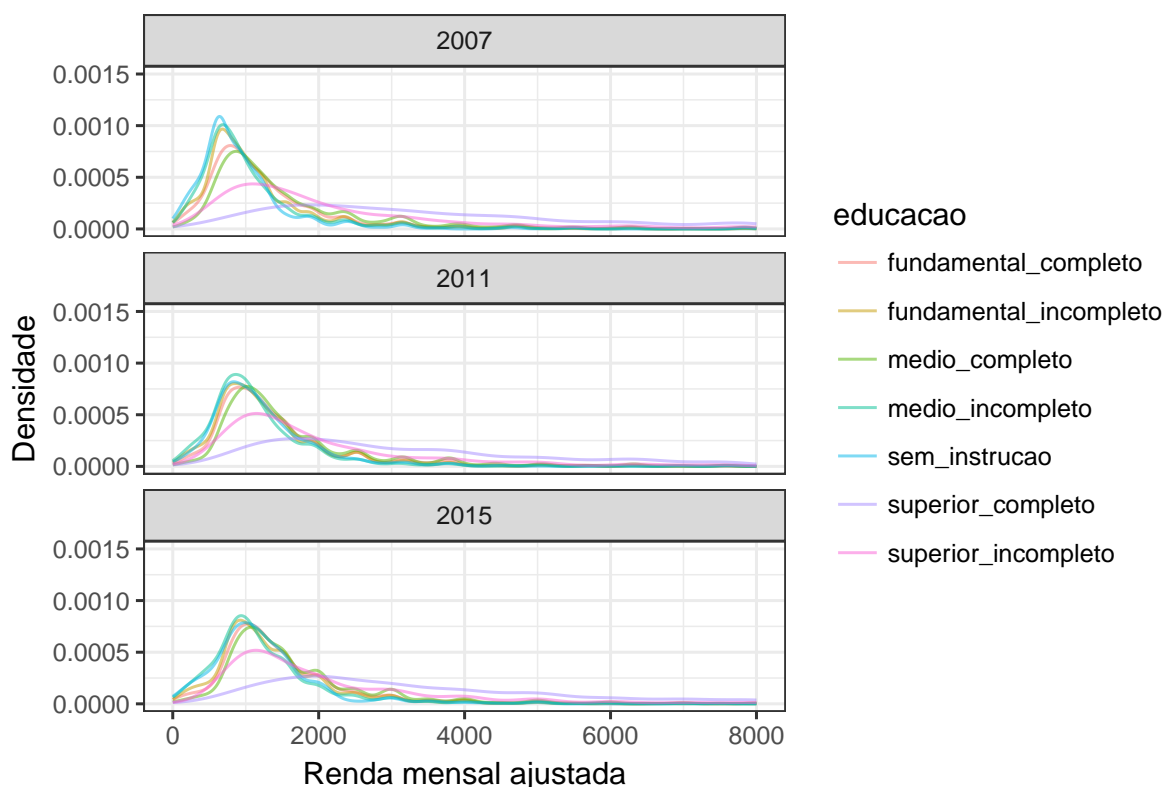
O primeiro ponto a ser ressaltado é que o grupo constituído pelas pessoas que possuem ensino superior completo contribuem com maior parte da desigualdade. A figura 7 mostra claramente como os dois grupos de maior nível educacional (superior completo e superior incompleto) possuem uma densidade bem distinta dos outros grupos. Outro fato importante é que a ligeira diminuição da contribuição relativa do grupo que possui superior completo, de 66,88% em 2007 para 64,97% em 2015, veio acompanhado de um

Tabela 11 – Contribuição de cada característica na decomposição entre grupos - Educação

Descrição	2007	%	2011	%	2015	%
Fundamental completo	0.0113	3.89	0.0121	4.81	0.0134	5.34
Fundamental incompleto	0.0472	16.28	0.0325	12.90	0.0309	12.33
Médio completo	0.0070	2.40	0.0144	5.73	0.0247	9.84
Médio incompleto	0.0155	5.34	0.0134	5.32	0.0147	5.86
Sem instrução	0.0123	4.22	0.0094	3.74	0.0042	1.66
Superior completo	0.1940	66.88	0.1698	67.42	0.1628	64.97
Superior incompleto	0.0028	0.97	0.0002	0.07	0.0000	0.00
Total	0.2900	100.00	0.2518	100.00	0.2506	100.00

participação amostral do grupo, passando de 13,74% em 2007 para 20,61% em 2015. Tal fato implica em uma diminuição da renda média do grupo como apontado na tabela 4.

Figura 7 – Densidades estimadas - Educação

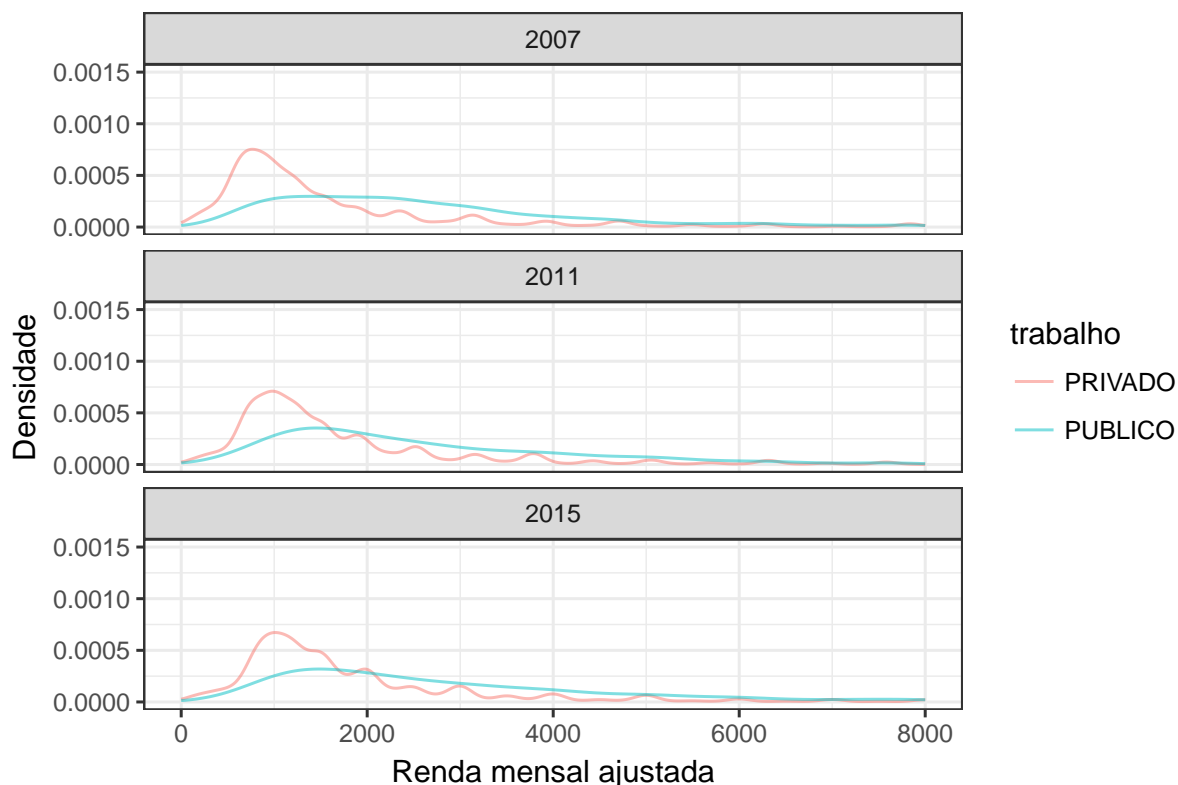


Por fim a tabela 12 apresenta os resultados para a característica trabalho. Os resultados mostram que o grupo Público é responsável por mais de 90% da desigualdade entre grupos. A figura 8 evidencia a diferença existente entre os grupos. Um ponto importante a ser ressaltado é que o grupo público representa uma parcela pequena da amostra em todos os anos, entre 5,75% e 6,57%, esse fato acaba por atenuar o valor total da desigualdade entre grupos.

Tabela 12 – Contribuição de cada característica na decomposição entre grupos - Trabalho

Descrição	2007	%	2011	%	2015	%
Privado	0.0018	7.93	0.0010	7.18	0.0013	8.15
Público	0.0207	92.07	0.0130	92.82	0.0141	91.85
Total	0.0225	100.00	0.0140	100.00	0.0154	100.00

Figura 8 – Densidades estimadas - Trabalho



Os resultados apresentados nessa seção corroboram os resultados apresentados por Ramos e Vieira (2000) e Ferreira et al. (2000). Os autores analisam dados do Brasil referentes à década de 90 e apresentam a educação como o principal fator responsável por gerar desigualdade, usando como metodologia a decomposição do índice de Theil. Os autores também apontam que fatores como raça e sexo possuem baixa contribuição para gerar desigualdade. Os resultados obtidos pela metodologia aqui proposta mostram que o fator educação como o principal responsável por gerar desigualdade no estado de São Paulo.

Devido às características apresentadas pelo índice J, é possível estimar como cada nível de educação contribui para a desigualdade entre grupos gerada pelo fator educação. Os resultados da tabela 11 mostram que o nível de instrução superior completo contribui

com a mais de 60% para a geração de desigualdade entre grupos, para a característica educação. Esse Valor passou de 66,88% em 2007 para 64,97% em 2015, isso acompanhado de uma maior parcela de pessoas com ensino superior, já que no ano de 2007 13,74% da amostra possuía ensino superior enquanto para o ano de 2015 o valor era de 20,61%. A tabela 6 mostra uma ligeira queda da renda média ajustada no período analisado, tal queda na renda média em conjunto com uma maior parcela de pessoas com nível superior e uma menor contribuição do nível de instrução superior completo para a geração de desigualdade entre grupos, para a característica educação, sugere que o aumento da oferta de mão-de-obra qualificada acarretou numa menor discrepância salarial entre os níveis educacionais contribuindo assim para a diminuição do valor apresentado pelo índice J no período.

4.3 Teste para desigualdade entre grupos.

Como apresentado na seção 3.1.1 o índice J nos possibilita testar a hipótese de existência de desigualdade entre os grupos, Rohde (2016) todavia aponta que o resultado do teste é sensível ao tamanho amostral e também ao número de grupos. Dentro desse contexto o autor aponta que o teste pode ser útil quando o número de observações é pequena ou quando o número de grupos é grande.

Devido ao fato das amostras utilizadas em nossas análises serem todas grandes, como apontado pela tabela 1, optou-se por realizar uma simulação de Monte Carlo para avaliarmos o comportamento do teste proposto por Rohde (2016). Seja n o número de observações totais de uma determinada amostra, $\alpha \in (0, 1)$ e $i \in \mathbb{N}$ a simulação de Monte Carlo empregada consiste em:

- Sorteie com reposição uma subamostra de tamanho $\lceil \alpha n \rceil$
- Calcule a estatística apresentada na equação (18)
- Determine a decisão do teste
- Repita o processo i vezes

após realizarmos os passos descritos calcula-se qual a frequência da decisão do teste, ou seja, número de vezes que a hipótese nula foi rejeitada dividido por i .

A figura 9 e a tabela 13 contém os resultados da simulação de Monte Carlo. Como esperado, o resultado do teste é extremamente sensível ao tamanho amostral, sendo que para um parâmetro α de 0,1 a hipótese nula de não existência de desigualdade entre grupos é rejeitada aproximadamente 100% das vezes.

Figura 9 – Resultado da simulação de Monte Carlo

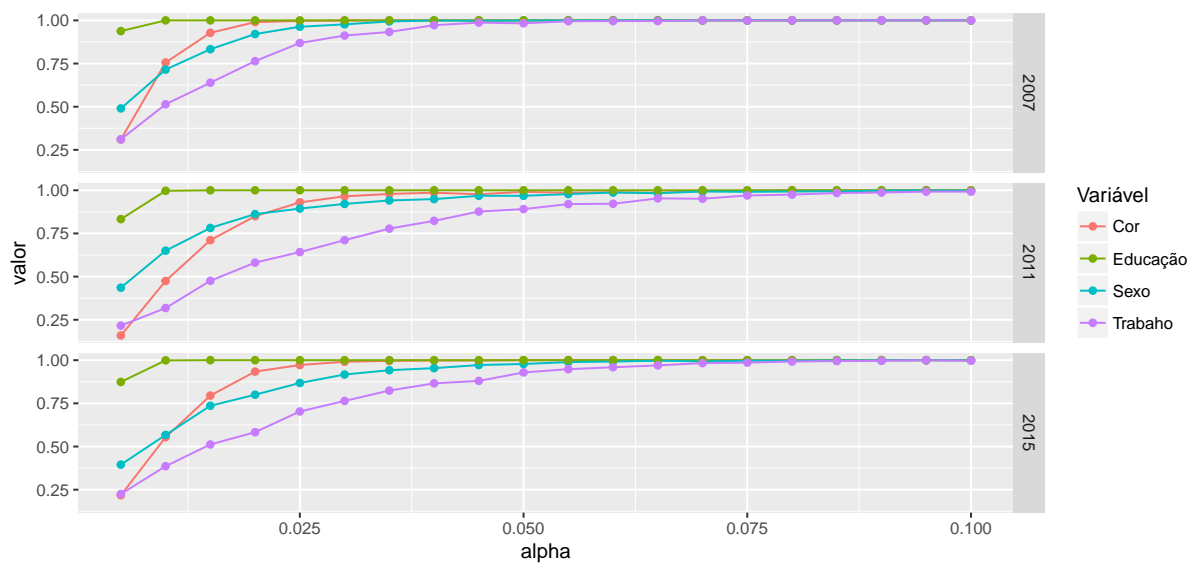


Tabela 13 – Resultado da simulação de Monte Carlo para o teste de desigualdade entre grupos

α	2007				2011				2015			
	Cor	Sexo	educacao	trabalho	Cor	Sexo	educacao	trabalho	Cor	Sexo	educacao	trabalho
0.005	0.3100	0.4900	0.9380	0.3110	0.1590	0.4360	0.8330	0.2160	0.2180	0.3950	0.8740	0.2250
0.010	0.7560	0.7150	1.0000	0.5140	0.4750	0.6500	0.9970	0.3180	0.5540	0.5670	0.9990	0.3860
0.015	0.9280	0.8330	1.0000	0.6390	0.7110	0.7820	1.0000	0.4760	0.7950	0.7360	1.0000	0.5120
0.020	0.9900	0.9210	1.0000	0.7640	0.8500	0.8620	1.0000	0.5810	0.9340	0.8000	1.0000	0.5830
0.025	0.9970	0.9630	1.0000	0.8690	0.9300	0.8940	1.0000	0.6420	0.9720	0.8680	1.0000	0.7030
0.030	0.9990	0.9770	1.0000	0.9120	0.9650	0.9210	1.0000	0.7110	0.9910	0.9170	1.0000	0.7640
0.035	1.0000	0.9940	1.0000	0.9330	0.9790	0.9410	1.0000	0.7780	0.9960	0.9420	1.0000	0.8240
0.040	1.0000	0.9990	1.0000	0.9720	0.9860	0.9490	1.0000	0.8230	0.9980	0.9540	1.0000	0.8660
0.045	1.0000	0.9980	1.0000	0.9870	0.9770	0.9680	1.0000	0.8770	0.9980	0.9720	1.0000	0.8800
0.050	1.0000	0.9990	1.0000	0.9830	0.9910	0.9680	1.0000	0.8910	1.0000	0.9780	1.0000	0.9290
0.055	1.0000	1.0000	1.0000	0.9950	0.9850	0.9780	1.0000	0.9200	1.0000	0.9890	1.0000	0.9480
0.060	1.0000	1.0000	1.0000	0.9960	0.9890	0.9870	1.0000	0.9220	1.0000	0.9920	1.0000	0.9590
0.065	1.0000	1.0000	1.0000	0.9960	0.9850	0.9830	1.0000	0.9530	1.0000	0.9970	1.0000	0.9700
0.070	1.0000	1.0000	1.0000	1.0000	0.9960	0.9930	1.0000	0.9510	1.0000	0.9930	1.0000	0.9830
0.075	1.0000	1.0000	1.0000	0.9990	0.9970	0.9910	1.0000	0.9700	1.0000	0.9970	1.0000	0.9860
0.080	1.0000	1.0000	1.0000	1.0000	1.0000	0.9930	1.0000	0.9750	1.0000	0.9970	1.0000	0.9920
0.085	1.0000	1.0000	1.0000	1.0000	0.9990	0.9940	1.0000	0.9840	1.0000	0.9990	1.0000	0.9950
0.090	1.0000	1.0000	1.0000	1.0000	0.9990	0.9980	1.0000	0.9870	1.0000	0.9990	1.0000	0.9970
0.095	1.0000	1.0000	1.0000	1.0000	1.0000	0.9990	1.0000	0.9920	1.0000	1.0000	1.0000	1.0000
0.100	1.0000	1.0000	1.0000	1.0000	0.9990	0.9980	1.0000	0.9920	1.0000	0.9990	1.0000	0.9970

Um ponto importante a ser apontado é a mudança na velocidade de convergência de 2007 em comparação aos outros anos analisados. Tal fato tem duas origens, primeiramente a amostra para os anos de 2011 e 2015 é ligeiramente menor que para o ano de 2007. O segundo ponto é a própria diminuição da desigualdade entre grupos, como apontado pela tabela 8.

É importante também ressaltar que os grupos definidos pela variável educação foram pouco afetados como apontado pela tabela 13. Isso nos leva a crer que mesmo frente a uma diminuição da desigualdade no período analisada, como apontado pela tabela 7, e um aumento no nível educacional da população, sinalizado pela maior participação de indivíduos com ensino superior completo na amostra no decorrer dos anos analisados, a educação ainda é um fator chave para entendermos a desigualdade.

Em contrapartida, a diminuição na velocidade de convergência dos grupos decorrente das características Cor e Sexo, analisadas em conjuntos com a diminuição no índice de desigualdade no período analisado aponta para uma possível mudança na estrutura de tais grupos.

Como citado anteriormente, Ramos e Vieira (2000) apontam o fator educação como o principal responsável por gerar desigualdade, todavia os autores não realizam nenhum teste para verificar tal hipótese. Isso devido ao fato do índice de Theil não possibilitar testar a significância do resultado obtido na decomposição entre grupos.

O uso do índice J permitiu avaliar a significância da decomposição entre grupos e, ainda que o tamanho amostral influencie diretamente o resultado, os resultados apontam a educação como o fator com maior probabilidade de gerar desigualdade entre grupos dentre os analisados.

4.4 Simulação dos erros de medida.

A presença de erros de medida é fator já conhecido na análise de renda, ou seja, é necessário levarmos em consideração. A estrutura discretizada dos dados provenientes da PNAD é um ponto que realça tal situação. Para avaliar o impacto que possíveis erros de

Tabela 14 – Evolução dos índices de desigualdade de renda - Com erro de medida ($a = -0,2$, $b = 0,2$)

	2007	2011	2015
Índice J	0.92	0.84	0.80
Theil L	0.43	0.37	0.36
Theil T	0.50	0.47	0.44

Tabela 15 – Evolução dos índices de desigualdade de renda - Com erro de medida ($a = 0$, $b = 0.2$)

	2007	2011	2015
Índice J	0.90	0.82	0.79
Theil L	0.42	0.37	0.36
Theil T	0.49	0.46	0.43

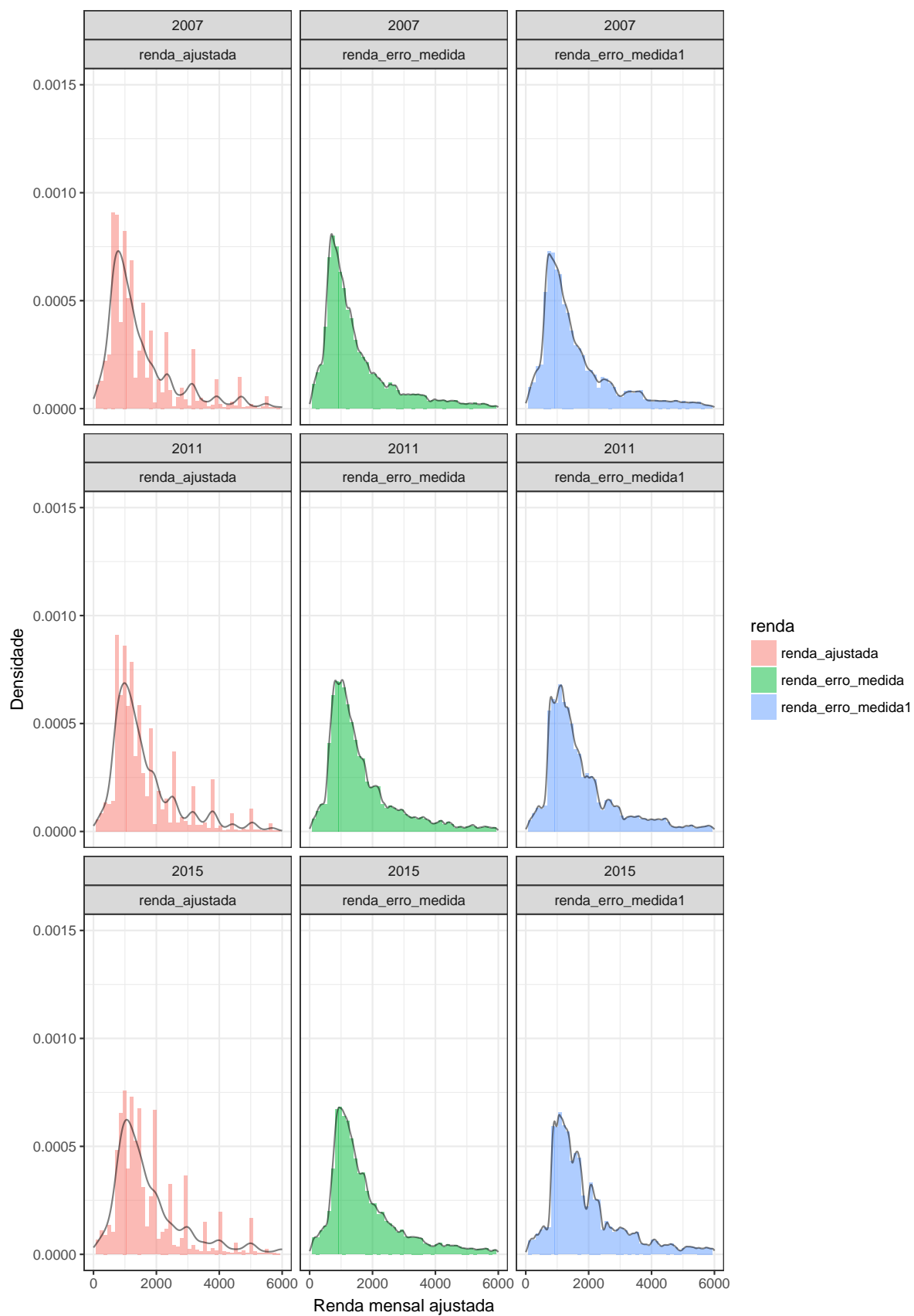
medida poderiam ter sobre o índice J tais erros foram simulados e novos resultados para os índices calculados.

Os erros simulados seguem a metodologia proposta na seção 3.1.4 e diferentes valores para os parâmetros a e b , apresentados na equação (37), são testados para avaliar a sensibilidade do índice J e também dos dois índices de Theil. Os resultados para essa análise são mostrados nas tabelas 14 e 15.

A introdução de um erro de medida afetou pouco o resultado do índice J, tal fato provavelmente esta relacionado a metodologia escolhida para a simulação do erro de medida. Um ponto importante a ser ressaltado é que a distribuição do erro de medida não é condicional a renda, o que impacta diretamente no fato dos erros de medida simulados ter tido pouco impacto no resultado final. Uma hipótese mais realista seria a de que a distribuição do erro de medida é condicional a renda, isso devido ao fato de que pessoas com uma renda maior teriam razões distintas para não informar sua verdadeira renda do que pessoas de renda mais baixa. No entanto a complexidade de se simular erros de medida condicionais, bem como a arbitrariedade envolvida na escolha da característica do erro de medida para diferentes níveis de renda acabou sendo determinante para a escolha de uma abordagem mais simplificada.

Ainda que a introdução de erro de medida não tenha afetado de forma substancial os resultados do índice J de desigualdade, ele afetou a estrutura discretizada que os

Figura 10 – Densidades estimadas com erro de medida



dados apresentavam. A figura 10 mostra como a introdução de erro medida alterou a estrutura altamente discretizada dos dados provenientes da PNAD. Isso nos leva a crer que a incorporação de erro de medida condicional a renda pode ser uma alternativa que auxilie tanto no problema de notificação incorreta quanto em situações em que os dados utilizados são altamente discretizados, como no caso da PNAD.

5 Considerações Finais

A presente dissertação teve como objetivo apresentar o índice J e sua aplicabilidade no estudo da desigualdade de renda. A aplicabilidade índice J foi analisada utilizando dados provenientes da PNAD referente aos anos de 2007, 2011 e 2015 para o estado de São Paulo. A escolha do índice J como medida de desigualdade, frente à medidas mais usuais na literatura de desigualdade, foi devido as propriedades matemáticas que tal índice possui, propriedades essa que nos permite realizar diferentes tipos de análises.

Os resultados mostram uma diminuição na desigualdade entre os anos de 2007 e 2015 passando respectivamente de 0,90 para 0,79, sendo que a queda mais acentuada ocorreu entre 2007 e 2011. Outro ponto importante de ser observado é o resultado obtido pela decomposição entre grupos. Tal resultado nos mostra que a característica educação é a que mais impacta na desigualdade entre grupos, dentre as características que foram analisadas. Dado as propriedades matemática do índice J, foi possível avaliar mais a fundo tal resultado e perceber que o grupos de maior nível educacional (superior completo) é o que mais contribui para essa desigualdade.

Uma das vantagens do uso do índice J é a possibilidade de se testar a hipótese de não existência de desigualdade entre grupos. O teste realizado é influenciado fortemente pelo tamanho amostral e também pelo número de grupos. A simulação de Monte Carlo conduzida para avaliar como o teste era afetado pelo tamanho amostral, nos mostra a característica educação com a maior probabilidade de gerar desigualdade entre grupos. Outro fato importante que foi possível observar é a diminuição na velocidade de convergência, no que tange a rejeição da hipótese nula, entre os anos de 2007 e 2015. Tal fato indica uma

possível diminuição de desigualdade entre grupos.

Os resultados encontrados corroboram com os apresentados por Ramos e Vieira (2000) e Ferreira et al. (2000), onde os autores apontam a educação como o principal fator responsável por gerar desigualdade. No estudo realizado por Araujo, Salvato e Souza (2008), que tem como escopo a região sul do Brasil, também apontou a educação como principal fator na causa da desigualdade. O uso do índice J permitiu testar tal hipótese formalmente, e também nos permitiu compreender melhor a contribuição de cada característica para a desigualdade.

A introdução de erros de medida pouco afetou os índices de desigualdade. Ainda assim, um ponto que vale ressaltar é que a simulação de tais erros de medida contribuiu para atenuar o problema dos dados provenientes da PNAD se apresentarem fortemente discretizados.

Em suma, a utilização do índice J nos permitiu avaliar não apenas a evolução da desigualdade no período analisado como também entender melhor a contribuição de determinadas características e permitir testar a significância da desigualdade entre grupos.

Referências

- ARAÚJO, J. R.; SALVATO, M. A.; SOUZA, P. Decomposição do índice de theil-t em disparidades regionais, entre gêneros, raciais e educacionais: uma análise da desigualdade de renda na região sul. *Anais do XI Encontro de Economia da Região Sul*, 2008.
- ATKINSON, A. B. On the measurement of inequality. *Journal of economic theory*, Elsevier, v. 2, n. 3, p. 244–263, 1970.
- BORRELL, L. N.; TALIH, M. A symmetrized theil index measure of health disparities: An example using dental caries in us children and adolescents. *Statistics in medicine*, Wiley Online Library, v. 30, n. 3, p. 277–290, 2011.
- BOURGUIGNON, F. Decomposable income inequality measures. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 901–920, 1979.
- COVER, T. M.; THOMAS, J. A. *Elements of information theory*. [S.l.]: John Wiley & Sons, 2012.
- COWELL, F. Theil, inequality and the structure of income distribution. 2003.
- COWELL, F. *Measuring inequality*. [S.l.]: Oxford University Press, 2011.
- COWELL, F. A. Measurement of inequality. *Handbook of income distribution*, Elsevier, v. 1, p. 87–166, 2000.
- COWELL, F. A.; FLACHAIRE, E. Statistical methods for distributional analysis. In: *Handbook of Income Distribution*. [S.l.]: Elsevier, 2015. v. 2, p. 359–465.
- DALTON, H. The measurement of the inequality of incomes. *The Economic Journal*, JSTOR, v. 30, n. 119, p. 348–361, 1920.
- DUCLOS, J.-Y. Poverty and equity: theory and estimation. *Canadá: Département d'économique and CRÉFA, Université Laval*, 2002.
- ELBERS, C. et al. Reinterpreting between-group inequality. *The Journal of Economic Inequality*, Springer, v. 6, n. 3, p. 231–245, 2008.
- FERREIRA, F. H. et al. *Os determinantes da desigualdade de renda no Brasil: luta de classes ou heterogeneidade educacional?* [S.l.]: Pontifícia Universidade Católica de Rio de Janeiro, Departamento de Economia, 2000.
- GINI, C. Variabilità e mutabilità. *Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T)*. Rome: Libreria Eredi Virgilio Veschi, 1912.
- JEFFREYS, H. An invariant form for the prior probability in estimation problems. In: THE ROYAL SOCIETY. *Proceedings of the Royal Society of London a: mathematical, physical and engineering sciences*. [S.l.], 1946. v. 186, n. 1007, p. 453–461.
- JUSTINO, P. et al. Multidimensional inequality: An empirical application to brazil. *Documento de trabalho*, v. 24, 2004.

LI, Q.; RACINE, J. S. *Nonparametric econometrics: theory and practice*. [S.l.]: Princeton University Press, 2007.

LITCHFIELD, J. A. Inequality: Methods and tools. *World Bank*, v. 4, 1999.

MAASOUMI, E.; NICKELSBURG, G. Multivariate measures of well-being and an analysis of inequality in the michigan data. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 6, n. 3, p. 327–334, 1988.

PYATT, G.; CHEN, C.-N.; FEI, J. The distribution of income by factor components. *The quarterly journal of economics*, MIT Press, v. 95, n. 3, p. 451–473, 1980.

RAMOS, L.; VIEIRA, M. L. Determinantes da desigualdade de rendimentos no brasil nos anos 90: discriminação, segmentação e heterogeneidade dos trabalhadores. *Desigualdade e pobreza no Brasil. Rio de Janeiro: IPEA*, p. 159–176, 2000.

ROHDE, N. J-divergence measurements of economic inequality. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Wiley Online Library, v. 179, n. 3, p. 847–870, 2016.

ROSENBLATT, M. et al. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 27, n. 3, p. 832–837, 1956.

ROTHSCHILD, M.; STIGLITZ, J. *Increasing Risk: A Definition and its Economic Consequences*. [S.l.], 1969.

SCOTT, D. W.; TERRELL, G. R. Biased and unbiased cross-validation in density estimation. *Journal of the american Statistical association*, Taylor & Francis, v. 82, n. 400, p. 1131–1146, 1987.

SHORROCKS, A. F. The class of additively decomposable inequality measures. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 613–625, 1980.

SILVERMAN, B. W. *Density estimation for statistics and data analysis*. [S.l.]: CRC press, 1986. v. 26.

THEIL, H. *Economics and information theory*. [S.l.], 1967.

ULLAH, A. Entropy, divergence and distance measures with econometric applications. *Journal of Statistical Planning and Inference*, North-Holland, v. 49, n. 1, p. 137–162, 1996.