

DISCRIMINAÇÃO SALARIAL POR RAÇA E GÊNERO NO MERCADO DE TRABALHO DAS REGIÕES NORDESTE E SUDESTE

Jacqueline Nogueira
Mestre em Economia Caen/UFC

Emerson Marinho
Professor do Caen/UFC

RESUMO

Este artigo analisa a discriminação salarial por raça e gênero dentro das densidades das distribuições de salário dos setores de atividade, comparando as regiões Nordeste e Sudeste. Para este objetivo, utilizaram-se os dados da PNAD 2002 e uma metodologia semiparamétrica - estimador de núcleo ponderado e outra paramétrica - regressão quantílica. O primeiro método mostra com base em simulações contrafactuais que existe discriminação contra mulheres e negros no mercado de trabalho em ambas as regiões e que esta é maior contra mulheres, pois enquanto a desigualdade salarial entre raças é causada principalmente por diferenças na educação, o diferencial de salários entre homens e mulheres se deve a uma possível discriminação, visto que elas são mais escolarizadas que eles. Além disso, verificou-se numa representação visualmente clara que a discriminação contra a raça negra é maior no Sudeste para todos os setores de atividade, enquanto que não se conseguiu identificar em qual região a discriminação contra mulheres é maior, visto que ela depende do setor considerado. Já os resultados da regressão quantílica mostraram que a discriminação salarial cresce com os salários. Concluem-se a partir deste estudo que para a redução das desigualdades de raça e gênero devem-se formular políticas públicas que incentivem a educação da população negra e combatam a discriminação no trabalho, principalmente, entre os cargos de maiores salários.

Palavras-chave: Discriminação, simulação contrafactual e densidade de distribuições de salários.

CLASSIFICAÇÃO JEL: J15, J16, J71.

ÁREA12: Economia do trabalho.

WAGE DISCRIMINATION BY RACE AND GENDER IN THE LABOR MARKET OF THE REGIONS
NORTHEAST AND SOUTHEAST: AN APPLICATION OF COUNTERFACTUAL SIMULATIONS AND
QUANTILE REGRESSION.

ABSTRACT

This paper analyses the discrimination of wages by race and gender inside of the density of wage distributions according to sectors of occupation, comparing the Northeast and Southeast regions. For this aim, we use data from PNAD 2002 and a semi parametric method -weighted kernel estimator and another parametric - quantile regression. The first method showed grounded in counterfactual simulations that have discrimination against women and black workers in labor market in both regions, and this discrimination is higher for women. The kernel method provided a visually clear representation that discrimination against black workers is greater in the Southeast for all sectors, while it can not identify where discrimination against women is greater because it depend on the sector. The quantile regression showed that discrimination increases for higher wages.

Key-words: Discrimination, couterfactual simulation and density of wage distributions.

DISCRIMINAÇÃO SALARIAL POR RAÇA E GÊNERO NO MERCADO DE TRABALHO DAS REGIÕES NORDESTE E SUDESTE.¹

Jacqueline Nogueira (Mestre em Economia Caen/UFC/ jqnc@click21.com.br).
Emerson Marinho (Professor do Caen/UFC/emarinho@ufc.br)

1. INTRODUÇÃO

Há presença de discriminação no mercado de trabalho quando indivíduos igualmente produtivos, empregados num mesmo setor de atividade e região são remunerados de forma distinta baseada no grupo, classe ou categoria a que pertencem. Apesar dos estudos encontrados na literatura tratarem principalmente da discriminação baseada na raça e no gênero dos trabalhadores, indivíduos pertencentes a certos grupos, tais como deficientes físicos, homossexuais, portadores de HIV, entre outros também são atingidos pela discriminação no ambiente de trabalho.

Acredita-se que a discriminação no mercado de trabalho seja uma das principais fontes da desigualdade de rendimentos entre raça e gênero no Brasil e, portanto, da pobreza desses indivíduos. Henriques (2001) calcula, com base na PNAD de 1999, que embora negros e pardos representassem 45% da população brasileira, correspondiam a 64% da população pobre e 69% da população indigente. Soares (2000) estima que devido à discriminação homens negros ganham em torno de 5% a 20% menos que homens brancos. Já as chances das mulheres saírem da condição de pobreza também são reduzidas pela discriminação no mercado de trabalho, na medida que mulheres com as mesmas capacidades produtivas que os homens recebem salários menores. Leme e Wajnman (2000) estimam que no Brasil os homens ganham em média 40% mais que as mulheres, o que afeta a oferta de trabalho feminino comprometendo o papel das mulheres como provedoras de renda.

Alguns estudos mostram ainda que a discriminação racial nas regiões Nordeste e Sudeste têm perfis diferentes, enquanto a de gênero não varia muito entre as regiões [ver Campante et al (2002) e Cavalieri e Fernandes (1998)]. Portanto, dada a heterogeneidade existente entre as regiões brasileiras, o conhecimento da discriminação no plano regional poderia combater suas causas com mais eficácia.

Isto posto, este artigo pretende analisar a discriminação por raça e gênero com base em toda a distribuição de salários e ao mesmo tempo abordar seus aspectos regionais e setoriais. A discriminação será identificada visualmente através de exercícios contrafactuais, tais como qual seria a distribuição de salário dos trabalhadores negros caso eles tivessem a mesma escolaridade dos trabalhadores brancos? E qual seria a distribuição de salários das mulheres caso elas tivesse a escolaridade dos homens? Por último, será estimada uma regressão quantílica, a fim de verificar se existe alguma tendência na discriminação à medida que os indivíduos dos grupos discriminados (negros e mulheres) avançam para intervalos mais elevados da distribuição de salários.

Além desta introdução, o trabalho está dividido em quatro partes. Na seção seguinte, é feita uma análise exploratória dos dados, na terceira seção, tem-se a discussão da metodologia adotada. Na quarta e na última seção, tem-se a discussão dos principais resultados e as principais conclusões.

¹ Este artigo é parte integrante da monografia vencedora do prêmio Ipea-Caixa na categoria mercado de trabalho.

2. DESCRIÇÃO DA BASE DE DADOS

Utilizou-se como base de dados a PNAD (Pesquisa Nacional de Amostra por domicílio) para o ano de 2002, ano mais recente na época da realização deste trabalho. Os dados coletados foram: rendimento do trabalho principal², horas trabalhadas, raça, anos de estudo concluídos, gênero, estados da federação e atividades do trabalho principal. Todos os dados foram ponderados pelos pesos da PNAD, a fim de gerar estimativas mais próximas dos valores populacionais.

Os dados referem-se apenas aos indivíduos com rendimento positivo e que trabalham entre 40 e 44 horas semanais (no máximo 176 horas mensais), que é o limite legal de horas trabalhadas na semana. Com o objetivo de isolar os efeitos da segmentação setorial e regional, os dados foram divididos em subamostras de raça³ e gênero, segundo os setores de atividade de cada região. Cada região possui três grandes setores de atividade: agricultura (agrícola), indústria (indústria, indústria de transformação e construção) e serviços (comércio e reparação, alojamento e alimentação, transporte armazenagem e comunicação, educação, saúde e serviços sociais, serviços domésticos, outros serviços coletivos, sociais e pessoais). Portanto, tem-se uma amostra total com 21.674 observações e 24 subamostras, o número de observações de cada uma destas subamostras encontra-se na Tabela A1 do Anexo.

Dentre as características individuais capazes de influenciar a produtividade, a experiência e, principalmente, a educação são aquelas que possui maior peso. No entanto, dada a dificuldade de informações diretas sobre a experiência de um trabalhador utilizar-se-á somente a educação como representante da produtividade. Os anos de estudo concluídos utilizados como *proxy* para a produtividade foram agrupados em 5 classes: sem instrução, um a quatro anos, cinco a oito, nove a onze e doze ou mais anos de estudo.

As regiões Nordeste e Sudeste foram escolhidas para traçar um perfil regional da discriminação porque elas concentraram a maioria da população negra e branca ocupada. De acordo com os dados da PNAD 2002, elas possuem 76,15% da população negra e 66,21% da população branca. Um outro fato que chama a atenção é que a população do Nordeste é composta por uma maioria negra (os negros correspondem a 69,17% da população do Nordeste), enquanto que no Sudeste a maior parte da população é branca (os brancos são 67,15% da população do Sudeste). Por outro lado, a estimação da discriminação por gênero não seria alterada pela escolha destas regiões, pois a proporção de mulheres ocupadas é praticamente a mesma entre as regiões.

TABELA 1: Distribuição dos Trabalhadores, por Gênero e Raça, segundo as Regiões Brasileiras - 2002.

Região	Masculino		Feminino		Negra		Branca	
	Número	%	Número	%	Número	%	Numero	%
Norte	745.476	64,44	411.365	35,56	817.678	70,68	339.163	29,32
Nordeste	3.354.502	66,80	1.666.973	33,20	3.473.273	69,17	1.548.202	30,83
Sudeste	6.827.718	60,29	4.496.549	39,71	3.719.694	32,85	7.604.573	67,15
Sul	2.499.755	59,43	1.706.485	40,57	619.685	14,73	3.586.555	85,37
Centro-Oeste	941.366	60,32	619.111	39,67	814.999	52,23	745.478	47,77
Total	14.368.817	61,75	8.900.483	38,25	9.445.329	40,59	13.823.971	59,41

Fonte: Elaboração própria a partir dos dados da PNAD 2002.

² Esta variável inclui salários, rendimentos dos trabalhadores por conta própria e pró-labore. No entanto, no decorrer do trabalho a denominaremos apenas por salário.

³ Considerou-se apenas as raças branca e negra (preta e parda).

Na Tabela 2, observa-se que, independente da região, o salário das mulheres é inferior ao salário dos homens a despeito do mesmo nível de escolaridade, o mesmo pode ser observado para a raça negra. Ademais, verifica-se também que existe uma disparidade de salários entre as regiões Nordeste e Sudeste, o que mais uma vez reforça a relevância de analisar a discriminação no plano regional.

TABELA 2: Relação entre Escolaridade e Salário Médio em Reais, por Gênero e Raça, segundo as Regiões Nordeste e Sudeste do Brasil – 2002.

Escolaridade	Nordeste		Sudeste		Nordeste		Sudeste	
	Masculino	Feminino	Masculino	Feminino	Branca	Negra	Branca	Negra
Sem instrução	173,18	132,09	357,96	273,96	169,96	165,44	367,66	305,20
1-4	219,59	178,36	459,49	301,28	240,57	201,50	435,50	348,84
5-8	315,71	203,99	539,53	365,55	321,17	268,60	521,69	418,39
9-11	570,04	385,30	806,03	515,57	539,18	439,47	725,41	581,42
12 ou mais	1956,89	1120,38	2423,24	1467,44	1643,85	1255,22	2022,20	1353,36

Fonte: Elaboração própria a partir dos dados da PNAD 2002.

3. METODOLOGIA

Nesta seção são apresentados os métodos para a estimação de funções densidades – estimador de núcleo e de regressão quantílica, bem como o teste de Kolmogorov-Smirnov (K-S) que verifica estatisticamente se duas amostras aleatórias são oriundas de uma mesma distribuição de probabilidade e a medida de concentração de Gini.

3.1 Estimador de núcleo

De modo semelhante ao histograma, o estimador de núcleo considera a divisão dos dados em intervalos de classes, e a cada intervalo é associado o número de observações que pertence aos respectivos intervalos. Este método difere do histograma na medida em que os intervalos são superpostos e as observações são ponderadas de acordo com sua distância em relação ao ponto médio do intervalo.

O estimador de núcleo $\hat{f}(x)$ de uma densidade univariada $f(X)$, baseada numa amostra aleatória $X_1, X_2, X_3, \dots, X_n$ de tamanho n , é definido, de acordo com Rosenblatt (1956) e Parzen (1962) como:

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

Sendo que:

$K(\cdot)$: é uma função simétrica chamada núcleo, que dá nome ao estimador, satisfazendo as seguintes propriedades: $\int_{-\infty}^{\infty} K(t)dt = 1$, $\int_{-\infty}^{\infty} tK(t)dt = 0$ e $\int_{-\infty}^{\infty} t^2 K(t)dt = k \neq 0$. Quando $K(\cdot)$

for uma função não-negativa ela será uma função densidade de probabilidade, o que implica que $\hat{f}(x)$ será também uma função densidade de probabilidade. Portanto, segundo Silverman (1986), $\hat{f}(x)$ herdará todas as propriedades de continuidade e diferenciabilidade do núcleo $K(\cdot)$.

h : é a largura dos intervalos de classes (*bandwidth*) também conhecida como parâmetro de suavização.

Assim sendo, a função densidade estimada, $\hat{f}(x)$, dependerá do formato da função núcleo e do parâmetro de suavização h . Quando o valor de h é muito pequeno detalhes espúrios da distribuição dos dados são logo observados. Por outro lado, se h é muito grande, aspectos importantes da distribuição podem ser obscurecidos como, por exemplo, a natureza bimodal da distribuição. Em geral, o critério para a determinação ótima de h é minimizar o Erro Quadrático Médio Integrado (EQMI) definido como:

$$EQMI(\hat{f}(x)) = \int E\{\hat{f}(x) - f(x)\}^2 dx = \int \{E[\hat{f}(x)] - f(x)\}^2 dx + \int \text{var } \hat{f}(x) dx \quad (2)$$

Em Silverman (1986) é mostrado que quando o valor de h é muito pequeno, o último termo de (2) (a integral da variância de $\hat{f}(x)$) torna-se muito grande embora haja uma redução da integral do viés medida por $\int \{E[\hat{f}(x)] - f(x)\}^2 dx$. Portanto, a escolha do parâmetro de suavização h implicará num *trade-off* entre viés e variância.

Em relação a função núcleo, se esta é um função simétrica, Hodje e Lehmann (1956) apud Silverman (1986) mostraram que o núcleo de *Epanechnikov*⁴ minimizará o EQMI se o parâmetro h for corretamente especificado. Em função disto, pode-se definir uma medida de eficiência⁵ para comparar qualquer outra função núcleo simétrica com o núcleo de *Epanechnikov*. O que se observa é que esta medida de eficiência pouco varia quando se consideram vários outros tipos núcleo simétricos. Neste sentido, é legítimo e justificável que se faça à escolha da função núcleo através de outras considerações como, por exemplo, graus de diferenciabilidade requeridos ou esforço computacional envolvido.

Neste trabalho, a distribuição da variável logaritmo neperiano do salário por hora mostrou-se aproximadamente com distribuição normal, portanto, adotou-se o núcleo gaussiano. Para este núcleo, Silverman (1986) mostra que o parâmetro de suavização ótimo, h_{otm} , é dado por $h_{otm} = 1,06\sigma n^{-1,5}$ sendo que σ é o desvio-padrão da distribuição gaussiana. Uma estimativa robusta de σ é dada por $A = \min(\text{desvio padrão}, \text{desvio interquartil}/1.34)$. Nestes termos, o parâmetro de suavização ótimo utilizado será:

$$h_{otm} = 0,9An^{-1,5} \quad (3)$$

O parâmetro de suavização calculado pela expressão (3) é muito adequado para distribuições unimodais e funciona razoavelmente para distribuições moderadamente bimodais. Portanto, inicialmente calculou-se h por (3) como ponto de partida para suavização e, a partir daí, aumentou-se gradativamente o valor de h até se achar que, visualmente, a densidade estivesse suavizada pois, segundo, Dinardo e Butcher (1998), geralmente é mais fácil suavizar com *o olho* do que o contrário. Em geral, as funções densidades estimadas por este método se mostram menos suavizadas quando h assume valores menores do que h_{otm} e não se alteram muito de formato quando o valor de h é maior do que h_{otm} .

Para o cálculo das densidades contrafactuais usou-se a mesma adaptação feita por Dinardo et al (1996) para o estimador de núcleo. Este método deriva da conhecida decomposição de Oaxaca (1973), que se baseia em exercícios contrafactuais. Um problema da decomposição de Oaxaca (1973) comum à de Dinardo et al (1996) é que ela ignora os efeitos

⁴ A função núcleo de *Epanechnikov* é definida como: $K_e(t) = \frac{3}{4\sqrt{5}}(1 - \frac{1}{5}t^2)$ para $-\sqrt{5} \leq t \leq \sqrt{5}$ e $K_e(t) = 0$ caso contrário.

⁵ A medida de *eficiência* para um núcleo K é definida como: $\frac{3}{5\sqrt{5}} \{\int t^2 K(t) dt\}^{-1/2} \{K(t)^2 dt\}^{-1}$.

de equilíbrio geral e dependência na ordenação de fatores explicativos. Ou seja, não se tem como captar os efeitos de uma mudança na estrutura educacional sobre o preço do capital humano, assim como o efeito de outras relações que estão sendo deixadas fora da estimação.

Desse modo, o cálculo das densidades contrafactuais é feito considerando cada observação individual como um vetor (w, p, r) onde w é o logaritmo neperiano de salário/hora, p é o atributo produtividade (que será medido somente em termos de educação) o qual assume os valores de 0 a 4, e r é a raça do indivíduo, a qual assume somente duas classificações b para branco e n para negro. A distribuição conjunta de (w, p, r) será representada por $F(w, p, r)$. A distribuição conjunta de salários (w) e atributo (p) dada uma raça (r) será representada por uma distribuição condicionada $F(w, p/r)$.

Assim, por exemplo, a densidade de salários dos negros, $f_n(w)$, pode ser escrita como a integral da densidade de salários desses indivíduos condicionada ao seu nível de produtividade (p) e sua raça (r), $f(w/p, r)$, sobre a sua distribuição de produtividade dada sua raça $F(p/r)$, ou seja:

$$f_n(w) = \int_{p \in \Omega_p} dF(w, p/r_{w,p} = n) = \int_{p \in \Omega_p} f(w/p, r_w = n) dF(p/r_p = n) \equiv f(w, r_w = n, r_p = n) \quad (4)$$

Sendo que:

Ω_p : é o domínio de definição dos atributos produtivos;

$r_p = n$: corresponde a raça negra com nível de produtividade p ;

$r_w = n$: corresponde a raça negra com salário w ;

Supondo que a estrutura de salários dos negros, representada pela densidade condicional $f(w/p, r_{w,p} = n)$, não dependa da distribuição de produtividade (escolaridade) dos brancos, $F(p/r_p = b)$, a densidade contrafactual de salários dos negros caso eles tivessem o mesmo nível de produtividade (escolaridade) dos brancos pode ser escrita como:

$$\begin{aligned} f(w, r_w = n, r_p = b) &= \int_{p \in \Omega_p} f(w/p, r_w = n) dF(p/r_p = b) \equiv \\ &\equiv \int f(w/p, r_w = n) \psi_p(p) dF(p/r_p = n) \quad (5) \end{aligned}$$

Como $\psi_p(p) \equiv dF(p, r_p = b) / dF(p/r_p = n)$, pode-se aplicar a regra de Bayes a esta equação e reescrevê-la como $\psi_p(p) = [P(r_p = n) / P(r_p = b)] [P(r_p = b/p) / P(r_p = n/p)]$. A probabilidade não condicional $P(r_p = n)$ é igual ao número de observações da raça negra dividido pelo número de observações da raça negra e branca. A probabilidade não condicional $P(r_p = b)$ é definida similarmente. Como $\psi_p(p)$ assume um número limitado de valores pode-se estimá-la através de um simples modelo de contagem célula por célula.

Observa-se que a equação (4) é idêntica a (5), exceto pelo fator de ponderação $\psi_p(p)$. Portanto, conhecendo-se a estimativa, $\hat{\psi}_p(p)$ de $\psi_p(p)$, pode-se então estimar a densidade contrafactual pelo método do núcleo ponderado, conforme a seguinte expressão:

$$\hat{f}(w; r_w = n, r_p = b) = \frac{1}{nh} \sum_{i=1}^n \hat{\psi}(p) K\left(\frac{w - W_i}{h}\right) \quad (6)$$

A equação (6) é estimada da mesma forma que a equação (1), com a diferença de que (6) é ponderada pelo fator $\hat{\psi}_p(p)$, o que torna este estimador semiparamétrico. Portanto, a diferença visual entre a densidade de salário dos trabalhadores brancos e a densidade

contrafactual de salário dos trabalhadores negros - com a mesma produtividade (escolaridade) de trabalhadores brancos – é atribuída à discriminação contra a raça negra. Estima-se, portanto, visualmente a discriminação para cada uma das amostras da raça negra. A discriminação contra mulheres é calculada de modo similar.

3.2 Regressão Quantílica

A regressão quantílica têm sido amplamente usada para estudar determinantes de salários, efeitos de discriminação e tendências na desigualdade de salários (KOENER, 2000). Este método nada mais e que uma generalização dos quantis para o contexto de regressão linear e definida como (KOENER, 2000):

Seja Y uma variável aleatória real caracterizada por sua função de distribuição $F(y) = \text{Prob}(Y \leq y)$. O τ -ésimo quantil de Y , para $0 < \tau < 1$, é definido como:

$$Q(\tau) = \inf\{y : F(y) \geq \tau\} \quad (7)$$

Através destas definições, pode-se mostrar que os quantis podem ser formulados como uma solução de um simples problema de programação linear.

Seja uma função linear por parte, $\rho_\tau(u) = u(\tau - I(u < 0))$, sendo que $I(u < 0)$ é uma função indicadora. A minimização da esperança matemática de $\rho_\tau(Y - \xi)$ com respeito a ξ produz a solução, $\hat{\xi}(\tau)$, que é a menor solução que satisfaz $Q(\tau)$, de acordo com a expressão (8). No caso de uma amostra aleatória, $y_1, y_2, y_3, \dots, y_n$, de Y , o τ -ésimo quantil amostral é calculado através da minimização de:

$$\min_{\xi \in R} \sum_{i=1}^n \rho_\tau(y_i - \xi) \quad (8)$$

Nesses termos, os parâmetros $\beta(\tau)$'s da função quantil linear condicional aos valores observados das variáveis explicativas $x_1, x_2, x_3, \dots, x_n$, $Q_y(\tau / X = x) = x'_i \beta(\tau)$, podem ser estimados, através de programação linear, resolvendo-se o seguinte problema:

$$\begin{aligned} \hat{\beta}(\tau) &= \arg \min_{\beta \in R} \sum_{i=\{i: y_i \geq x'_i \beta\}} \tau |y_i - x'_i \beta| + \sum_{i=\{i: y_i < x'_i \beta\}} (1 - \tau) |y_i - x'_i \beta| = \\ \hat{\beta}(\tau) &= \arg \min_{\beta \in R} \sum \rho_\tau(y_i - x'_i \beta) \end{aligned} \quad (9)$$

tal que ρ_τ é uma função definida por:

$$\rho_\tau(u) = \begin{cases} \tau u, & u \geq 0 \\ (\tau - 1)u, & u < 0 \end{cases} \quad (10)$$

Assim, a função ρ_τ multiplica os resíduos por τ se eles forem não-negativos e por $(\tau - 1)$ caso contrário, para que estes sejam tratados assimetricamente.

3.3 O Teste Kolmogorov-Smirnov (K-S)

O teste de K-S verifica estatisticamente se duas amostras foram extraídas da mesma população, ou seja, caso as duas amostras sejam extraídas da mesma população é de se esperar que suas distribuições de probabilidade sejam bastantes próximas uma da outra. Desse modo, o teste K-S pode ser empregado para verificar se as verdadeiras distribuições de salários e suas contrafactuais são diferentes. No caso em que a hipótese nula é rejeitada, pode-se concluir que uma mudança na estrutura educacional altera a distribuição de salários original.

3.4 Índice de Gini

O índice de Gini é uma medida comumente usada para medir concentração de salários e seu cálculo deriva de outra medida de concentração, a curva de Lorenz. Neste sentido, calculam-se os índices de Gini para as distribuições verdadeiras de salários e para suas contrafactuais. Com isso, pode-se verificar em quanto à concentração de salários é alterada por uma mudança na estrutura educacional dos indivíduos pertencentes aos grupos supostamente discriminados.

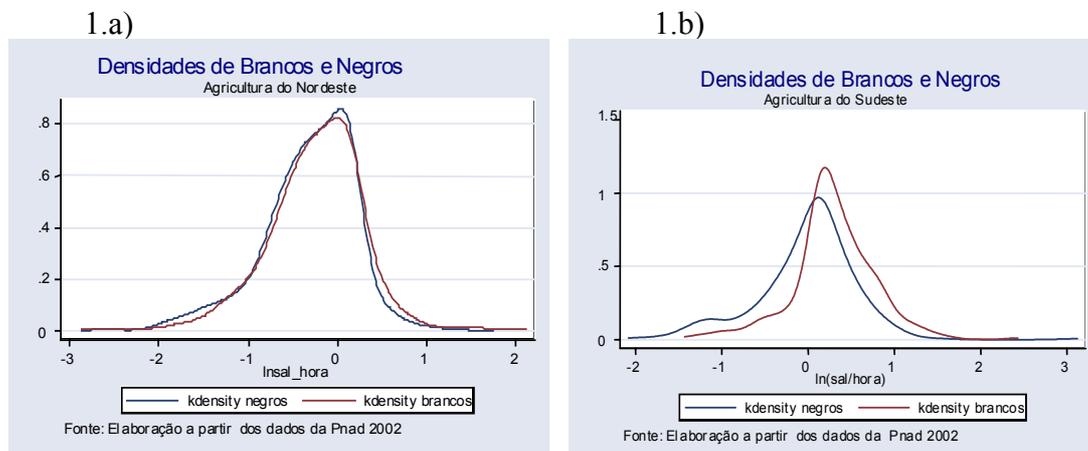
4. RESULTADOS DAS ESTIMAÇÕES DE DENSIDADES

Nesta seção estão apresentados os resultados das estimações das densidades de salários por gênero e raça, assim como os testes de Kolmogorov-Smirnov para duas amostras e os coeficientes de Gini para as densidades originais e contrafactuais dos negros e das mulheres⁶.

4.1 Resultados por Raça na Agricultura das Regiões Nordeste e Sudeste do Brasil

O salário/hora modal das densidades dos brancos e negros, Figuras 1.a e 1.b, foi igual ao salário mínimo/hora de R\$ 1,14⁷ em ambas regiões. É interessante notar que o diferencial de salários entre brancos e negros no Nordeste é praticamente imperceptível, o que provavelmente sugere um baixo diferencial de escolaridade entre raças. Já no Sudeste nota-se que a densidade de salário dos negros concentra-se mais à esquerda do que a densidade de salário dos brancos, o que mostra que existe uma diferença de salário considerável entre raças na agricultura do Sudeste.

FIGURA 1: Densidades de Salário-Hora por Raça na Agricultura – segundo as Regiões Nordeste e Sudeste do Brasil



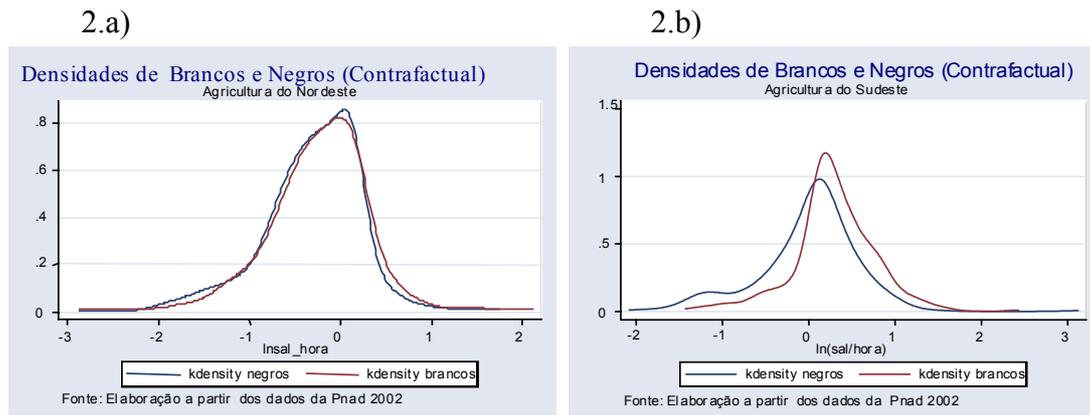
Comparando as Figuras 1.a e 2.a, e, 1.b e 2.b, constata-se que, as densidades de salários originais dos negros são muito próxima às suas densidades contrafactuais para ambas as regiões. Isso provavelmente se dá devido à pequena diferença no nível de escolaridade entre as raças. Observa-se que provavelmente não exista discriminação no Nordeste, já que as densidades de salários dos brancos e a contrafactual dos negros são praticamente as mesmas. Já no Sudeste, percebe-se uma diferença visual clara entre as densidades dos salários dos brancos e a contrafactual dos salários dos negros. Essa diferença pode ser atribuída ao efeito da discriminação racial.

⁶ Todas as estimações foram geradas pelo software Stata 8.1.

⁷ Os salários modais de todas as densidades foram calculados através do software Stata 8.1.

O teste K-S⁸ comprovou que as densidades originais dos negros se alteraram por uma mudança no nível de escolaridade, o que reforça a importância da educação na determinação dos salários. Portanto, o teste rejeitou ao nível de significância de 5% a hipótese de igualdade entre a densidade original de salário dos negros e sua contrafactual para ambas regiões. Os testes K-S para as demais distribuições também rejeitaram a hipótese de igualdade entre suas distribuições originais e contrafactuais.

FIGURA 2: Densidade de Salário-Hora para Brancos e Contrafactual para Negros na Agricultura – segundo as Regiões Nordeste e Sudeste do Brasil



No Nordeste, a concentração de salários dos negros aumentou após serem remunerados de acordo com a escolaridade dos brancos, visto que o coeficiente de Gini para a amostra original foi de 0,26⁹, enquanto para a amostra contrafactual foi igual a 0,27. Já no Sudeste, a concentração manteve-se inalterada em 0,29 após a mudança na estrutura educacional dos negros.

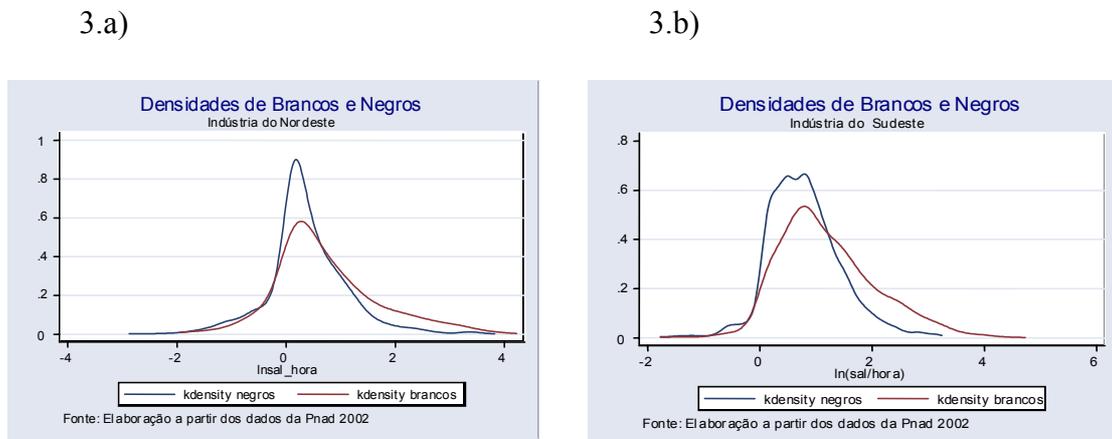
4.2 Resultados por Raça na Indústria das Regiões Nordeste e Sudeste do Brasil

As densidades de salário de brancos e negros no Nordeste, Figura 3.a, apresentaram o mesmo salário modal de R\$ 1,14, igual ao salário mínimo por hora. Já no Sudeste, Figura 3.b, a densidade de salário dos brancos apresentou um salário modal por hora de R\$ 2,27 enquanto a densidade de salário dos negros apresentou dois salários modais por hora de R\$ 1,14 e R\$ 2,27. Nas regiões Nordeste e Sudeste os salários dos negros estão mais concentrados em torno de seus salários modais por hora enquanto a cauda à direita da densidade de salários dos brancos é mais alongada, indicando que a maioria dos negros ganha em torno do salário modal por hora e que uma parcela considerável dos brancos tem seus rendimentos situados nas faixas mais elevadas de salário.

⁸ A estatística do teste K-S é definida como $D = \max |F_{n1}(X) - F_{n2}(x)|$ tal que, $F_{n1}(x)$, é a distribuição acumulada original e, $F_{n2}(x)$, a distribuição acumulada da contrafactual. O valor tabelado para amostras maiores do que 40 é dado por $1,36\sqrt{(n_1 + n_2) / n_1 n_2}$ para um nível de significância de 5%. Os valores do teste K-S para esta e as demais distribuições se encontram na Tabela A6 do Anexo.

⁹ Nas Tabelas A4 e A5 do Anexo, são apresentadas as estimativas dos índices de Gini para todas as amostras analisadas.

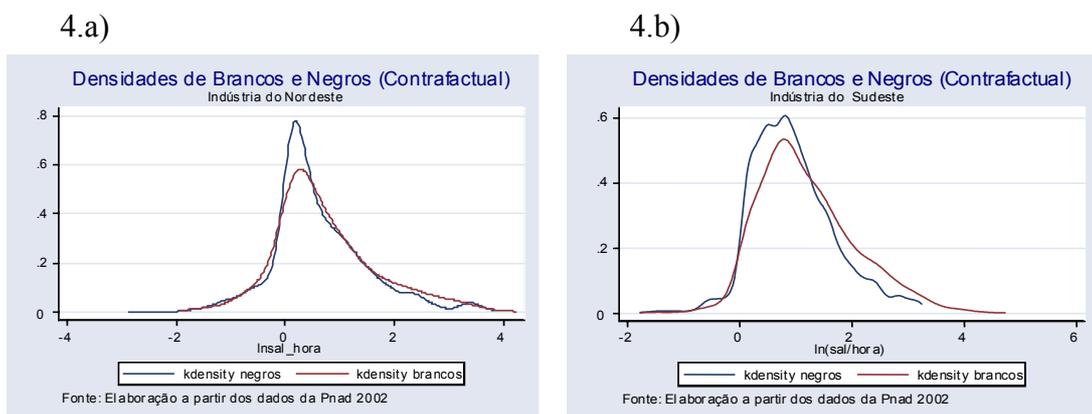
FIGURA 3: Densidades de Salário-Hora por Raça na Indústria – segundo as Regiões Nordeste e Sudeste do Brasil



Nas Figuras 4.a e 4.b constata-se que, após a aplicação da simulação contrafactual, as densidades de salário dos brancos e negros tenderam a se aproximar. Com efeito, houve um deslocamento dos salários dos negros que se concentravam em torno da moda da distribuição para a cauda à direita. No entanto, para salários maiores do que o salário modal por hora existe uma maior proporção de brancos do que de negros embora eles tenham o mesmo nível de escolaridade. Isso sugere que a discriminação contra negros acontece principalmente nos intervalos onde os salários são mais elevados.

A comparação entre a densidade dos brancos e a contrafactual dos negros nas Figuras 4.a e 4.b mostra que visualmente a discriminação contra negros no Sudeste é maior que no Nordeste.

FIGURA 4: Densidade de Salário-Hora para Brancos e Contrafactual para Negros na Indústria – segundo as Regiões Nordeste e Sudeste do Brasil



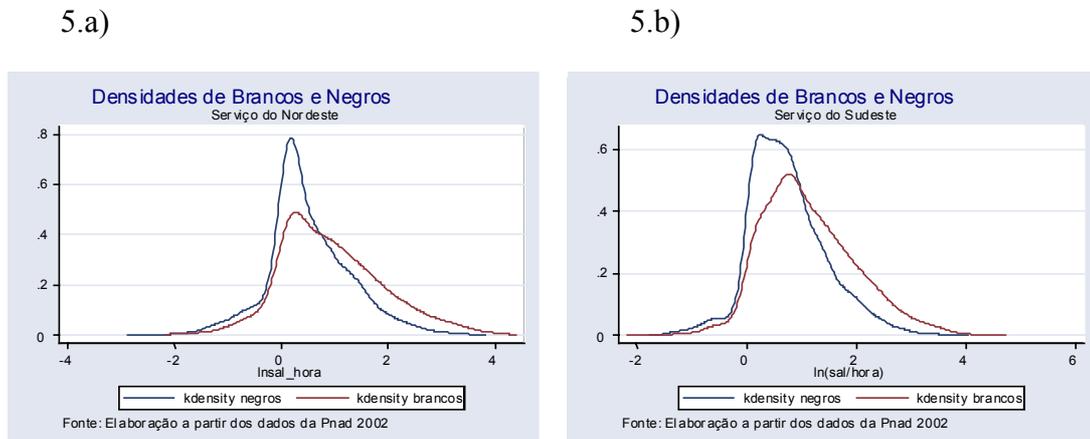
Em termos de concentração, os negros pioraram após terem seu trabalho remunerado de acordo com a escolaridade dos brancos. No Nordeste, o índice de Gini para a distribuição original de salários dos negros foi igual a 0,43, enquanto seu respectivo contrafactual foi igual a 0,54. No Sudeste, o índice de Gini original (0,38) também foi menor que o contrafactual (0,43).

4.3 Resultados por Raça no Serviço das Regiões Nordeste e Sudeste do Brasil

Em ambas regiões, conforme as Figuras 5.a e 5.b, a densidade de salário de brancos e negros teve salário modal por hora de R\$ 1,14 com exceção na região Sudeste onde o salário modal por hora dos brancos foi de R\$ 2,27. Adicionalmente, as densidades dos negros se

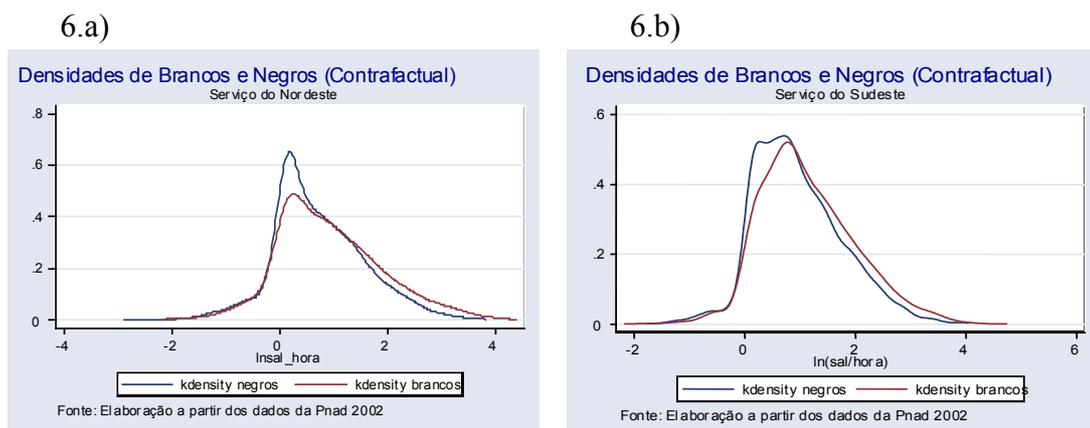
concentram mais à esquerda do que a densidade dos brancos, ou seja, os negros ganham em geral salários mais baixos que os brancos.

FIGURA 5: Densidades de Salário-Hora por Raça no Serviço – segundo as Regiões Nordeste e Sudeste do Brasil



Em ambas regiões, a simulação contrafactual acarretou uma tendência de sobreposição das densidades de salários de brancos e negros conforme mostram as Figuras 6.a e 6.b. Da mesma forma que aconteceu para as densidades do setor industrial, houve um deslocamento dos salários que se situavam em torno da moda da distribuição para a cauda à direita. No entanto, nos intervalos superiores ainda existe uma proporção maior de brancos do que de negros, o que pode indicar que a persistência da discriminação ocorre justamente para salários mais elevados. Uma comparação entre as Figura 6.a e 6.b mostra que no setor serviço o efeito visual do fator discriminação nas duas regiões é muito próximo, ou seja, a discriminação contra negros no serviço é muito próxima entre as regiões.

FIGURA 6: Densidades de Salário-Hora para Brancos e Contrafactual para Negros no Serviço – segundo as Regiões Nordeste e Sudeste do Brasil

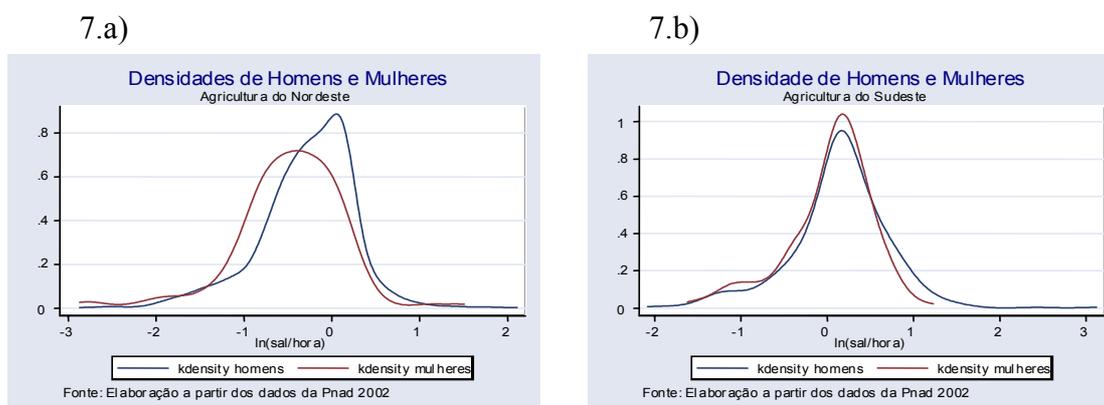


A concentração de salários dos negros se elevou após estes serem remunerados de acordo com a escolaridade dos brancos. De fato, os índices de Gini para os salários contrafactuais foram maiores do que os índices para os salários originais. No Sudeste, os índices de Gini para os salários originais e contrafactuais foram iguais 0,42 e 0,46, respectivamente. No Nordeste, os índices de Gini para os salários originais e contrafactuais foram iguais a 0,46 e 0,49, respectivamente.

4.4 Resultados por Gênero na Agricultura das Regiões Nordeste e Sudeste do Brasil

De acordo com as Figuras 7.a e 7.b, as densidades de salário estimadas por gênero na agricultura das regiões Nordeste e Sudeste mostraram que o salário modal por hora foi de R\$ 1,14 equivalente ao salário mínimo por hora, a exceção das densidades de salário das mulheres no Nordeste que foi de R\$ 0,68. Também se observa que, em ambas regiões, a densidade de salário das mulheres se concentra mais à esquerda do que a dos homens, sendo esse resultado mais claramente observado no Nordeste. Isso revela que no setor agrícola as mulheres em sua maioria ganham menos do que os homens.

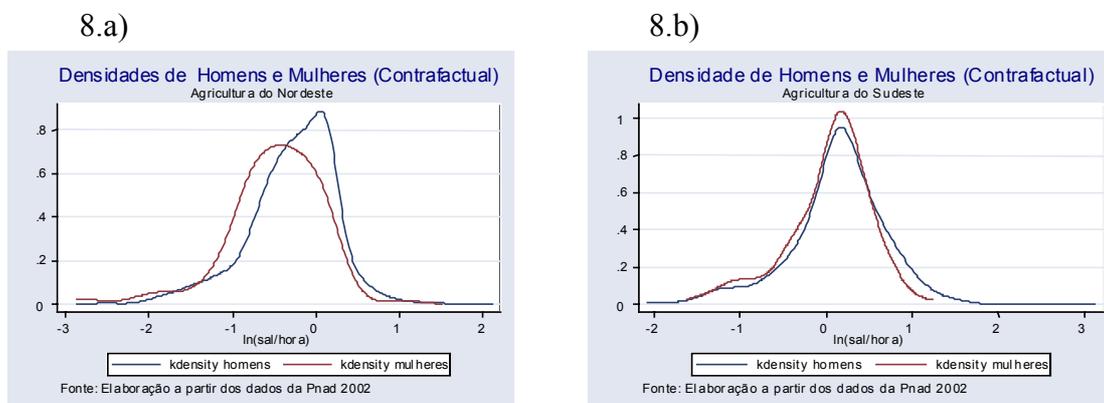
FIGURA 7: Densidades de Salário-Hora por Gênero na Agricultura – segundo as Regiões Nordeste e Sudeste do Brasil



Nas Figuras 8.a e 8.b a diferença visual entre as densidades dos homens e a contrafactual das mulheres corresponde a um possível fator de discriminação, já que na densidade contrafactual as mulheres têm a mesma escolaridade que os homens. A discriminação de salários por gênero é visualmente elevada no Nordeste, enquanto no Sudeste a discriminação é visível apenas nos intervalos mais elevados da distribuição de salário.

Observa-se que embora não se perceba uma diferença visual clara entre as densidades originais e contrafatuais das mulheres, Figuras 7.a e 8.a, e, 7.b e 8.b, o teste K-S rejeitou a igualdade entre estas duas distribuições (Ver Tabela A6 do Anexo).

FIGURA 8: Densidade de Salário-Hora para Homens e Contrafactual para Mulheres na Agricultura – segundo as Regiões Nordeste e Sudeste do Brasil



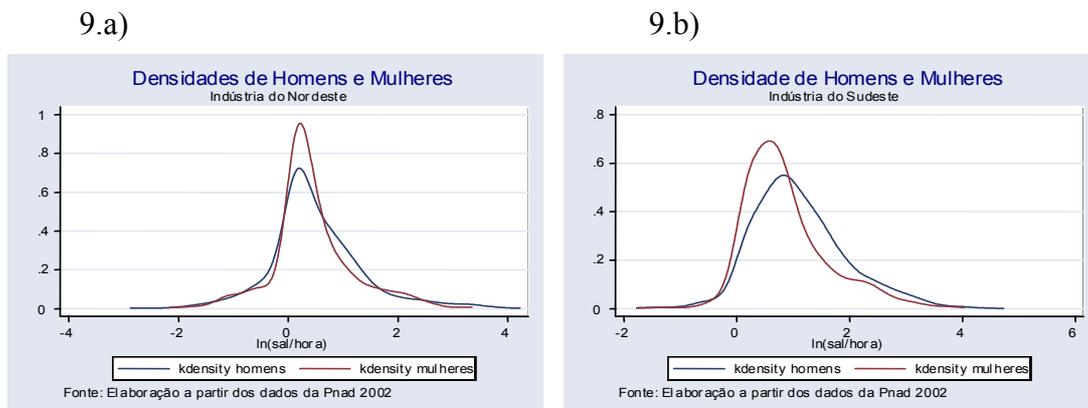
A remuneração das mulheres segundo a escolaridade dos homens não alterou a concentração de salários das mesmas na agricultura do Sudeste, pois, o índice de Gini de 0,23,

foi o mesmo para as amostras de salários originais e a contrafactual. Já no Nordeste houve uma redução da concentração, pois o índice de Gini caiu de 0,30 para 0,27.

4.5 Resultados por Gênero na Indústria das Regiões Nordeste e Sudeste do Brasil

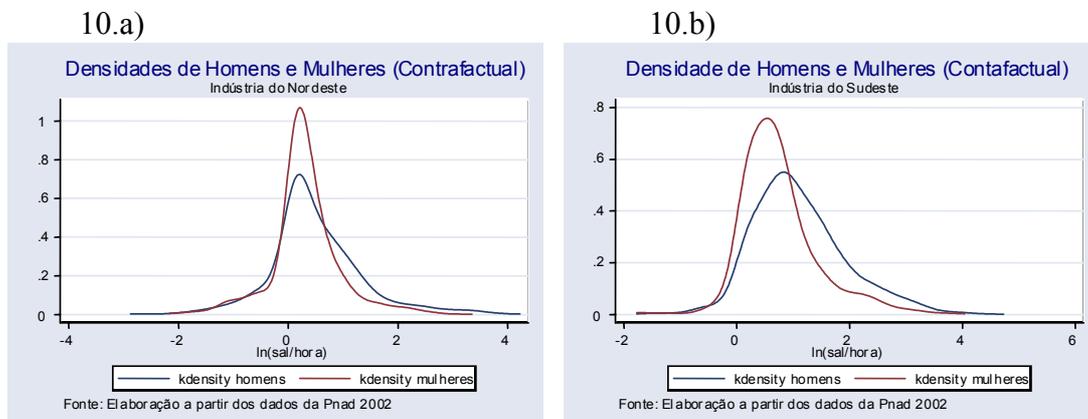
Observa-se através das Figuras 9.a e 9.b que o salário/hora modal da densidade de salário dos homens no Sudeste foi de R\$ 2,27 enquanto, nas demais, o salário/hora modal foi igual a R\$ 1,14. Assim como na agricultura, as densidades de salários das mulheres no setor industrial também se localizam mais à esquerda do que a densidade dos homens, ou seja, as mulheres em sua maioria ganham salários inferiores aos dos homens.

FIGURA 9: Densidades de Salário-Hora por Gênero na Indústria – segundo as Regiões Nordeste e Sudeste do Brasil



Comparando as Figuras 9.a e 10.a, e, 9.b e 10.b, observa-se que, houve uma transladação à esquerda da densidade das mulheres, ou seja, as mulheres na indústria recebem salários mais baixos que os dos homens a despeito de seu maior nível de escolaridade. Portanto, a simulação mostra que existe discriminação salarial contra as mulheres e que esta é visualmente mais explícita na cauda à direita da distribuição de salário. Nota-se, nas Figuras 10.a e 10.b, que visualmente a discriminação por gênero na indústria do Sudeste é maior que no Nordeste.

FIGURA 10: Densidade de Salário-Hora para Homens e Contrafactual para Mulheres na Indústria – segundo as Regiões Nordeste e Sudeste do Brasil



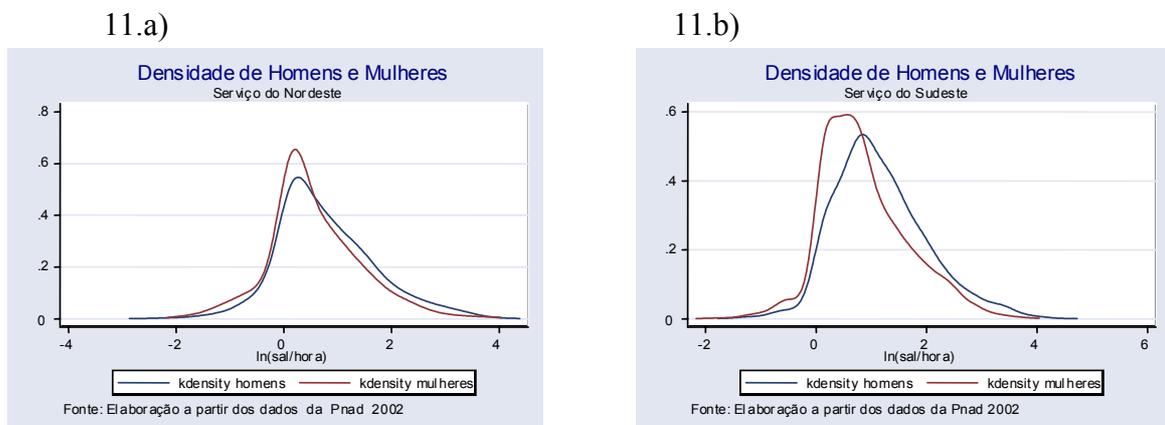
A remuneração das mulheres pela escolaridade dos homens diminuiu a concentração de salários das mulheres na indústria em ambas as regiões. No Nordeste, o índice de Gini para a amostra de salário original foi de 0,43, enquanto que para a amostra contrafactual ele foi

igual a 0,34. No Sudeste, o Gini para a amostra original foi igual a 0,47 e para a contrafactual foi de 0,43.

4.6 Resultados por Gênero no Serviço das Regiões Nordeste e Sudeste do Brasil

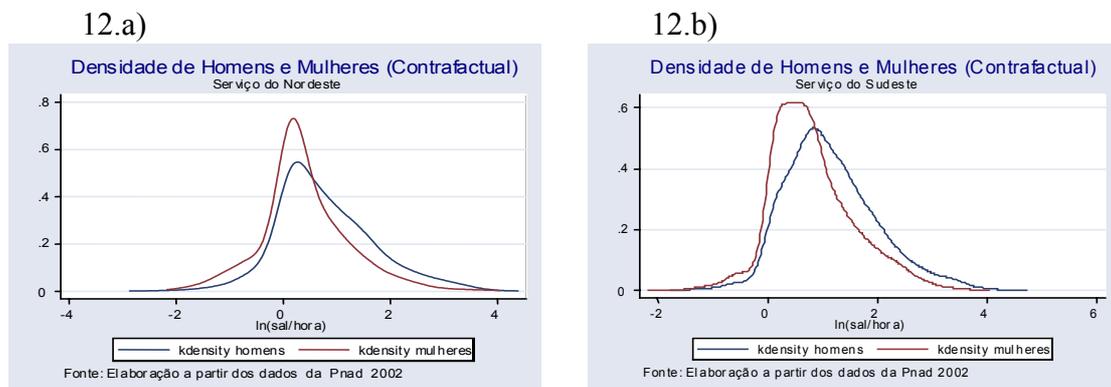
Na Figura 11.a, o salário modal para ambas as densidades foi igual ao salário mínimo por hora de R\$ 1,14. Para salários menores do que a moda observa-se em geral que existe um maior número de mulheres nestes intervalos. Este efeito é visualmente mais claro na região Sudeste, conforme mostra a Figura 11.b. O salário modal para a densidade de homens foi de R\$ 2,27, enquanto que para a densidade de mulheres foi de R\$ 1,14.

FIGURA 11: Densidades de Salário-Hora por Gênero no Serviço – segundo as Regiões Nordeste e Sudeste



Comparando as densidades de salário originais das mulheres e suas contrafactuals, Figuras 11.a e 12.a, e 11.b e 12.b, percebe-se que houve uma transladação à esquerda da densidade das mulheres. Se quando as mulheres são remuneradas de acordo com a escolaridade dos homens a distância entre as suas densidades aumenta, então, existem evidências de discriminação. Além disso, a distância entre as densidades de homens e as densidades contrafactuals das mulheres é maior nos intervalos superiores da distribuição de salários, o que sugere uma maior discriminação para salários mais elevados. As Figuras 12.a e 12.b não mostram claramente em qual região a discriminação contra mulheres é maior.

FIGURA 12: Densidade de Salário-Hora para Homens e Contrafactual para Mulheres no Serviço – segundo as Regiões Nordeste e Sudeste



A concentração de salários das mulheres no serviço diminuiu após a simulação contrafactual em ambas as regiões. No Nordeste, o índice de Gini para a distribuição original

de salários das mulheres é de 0,51 e seu respectivo índice contrafactual é de 0,49. No Sudeste, o índice de Gini para as amostras originais foi igual a 0,46, enquanto que para as amostras contrafactuais ele caiu para 0,45.

5. RESULTADOS DA REGRESSÃO DE SALÁRIOS POR QUARTIS

Na Tabela 3, verifica-se que com exceção da agricultura, nos demais setores de atividade econômica, a razão salário contrafactual dos negros/ salário dos brancos, assim como a razão salário contrafactual das mulheres/ salário dos homens, apresenta uma tendência decrescente à medida que se avança para os quantis mais elevados da distribuição de salários. Isso sugere evidências de que a discriminação é maior para as faixas de salário mais elevadas. Com base nestas evidências estimou-se uma equação de salário por quantis, a fim de verificar se os coeficientes estimados para as variáveis gênero e raça confirmam essa tendência de aumento na discriminação à medida que os salários aumentam.

TABELA 3: Razão Salário Contrafactual dos Negros/Salário dos Brancos e Salário Contrafactual das Mulheres/ Salário Dos Homens nos Quantis 0,25, 0,50 E 0,75, segundo os Setores de Atividade das Regiões Nordeste e Sudeste.

		0,25 Quantil		0,50 Quantil		0,75 Quantil	
		Negrocontr/ branco	Mulhercontr/ homem	Negrocontr/ branco	Mulhercontr/ homem	Negrocontr/ branco	Mulhercontr/ homem
NE	Agricultura	100%	80%	95%	86%	100%	80%
	Industria	100%	100%	91%	80%	83%	75%
	Serviço	100%	100%	78%	64%	75%	61%
SE	Agricultura	70%	94%	87%	100%	80%	89%
	Industria	93%	80%	80%	71%	70%	63%
	Serviço	83%	100%	80%	70%	80%	67%

Fonte: Valores gerados a partir do software Stata.

Desse modo, foi especificada uma equação de salário, onde a variável dependente é o logaritmo neperiano do salário/hora e as variáveis explicativas são:

- Anos de estudo concluídos, que assume os valores de 0 a 4: 0 para pessoas sem instrução, 1 para pessoas de 1 a 4 anos de estudo, 2 para pessoas de 5 a 8 anos de estudo, 3 para pessoas de 9 a 11 anos de estudo e 4 para pessoas com 12 ou mais anos de estudo;
- O quadrado dos anos de concluídos;
- Uma variável binária para sexo, que assume o valor 1 para homens;
- Uma variável binária para raça, que assume o valor 1 para a raça branca;
- Duas variáveis binárias para distinguir três setores de atividade: setor primário (tomado como base), setor secundário e setor terciário;
- Uma variável binária para região, que assume o valor 1 a região Nordeste.

Os coeficientes para as dummies de raça e gênero indicam que a discriminação aumenta para níveis de salários mais elevados, reforçando os resultados visuais encontrados nas densidades de salários. De fato no primeiro quantil, os brancos ganham 12,52 % a mais que os negros, 17,13 % no segundo quantil e 20,02 % no terceiro quantil. Já a discriminação contra mulheres se mostrou mais elevada do que a discriminação contra negros e com a mesma tendência de aumento para níveis de salários mais elevados. Os homens ganham a mais que as mulheres 26,36 %, 35,30% e 49,18% no primeiro, segundo e terceiro quantil,

respectivamente. Estes resultados corroboram com aqueles obtidos por Soares (2000) onde o termo discriminação é maior para os centésimos mais elevados da distribuição de salários.

Os sinais das outras variáveis revelam também alguns resultados interessantes. De acordo com a teoria do capital humano esperar-se-ia que os anos de estudo e anos de estudo ao quadrado tivessem sinais positivo e negativo, respectivamente, no entanto, os sinais estimados para estas variáveis apresentaram sinais contrários. Isso mostra que no Brasil o retorno à educação é crescente, e não decrescente como afirma a teoria do capital humano. [Baptista (1998), Duarte (2003)].

TABELA 4: Resultados das Estimções da Equação Quantílica de Salários.

Coeficientes	0,25 Quantil ¹⁰	0,50 Quantil	0,75 Quantil
Constante	-32,56* (0,03)	-11,67* (0,03)	8,60* (0,02)
Anos de estudo	-23,50* (0,03)	-27,99* (0,01)	-28,41* (0,02)
Anos de estudo ao quadrado	13,20* (0,01)	16,95* (0,00)	19,03* (0,00)
Sexo	26,36* (0,01)	35,30* (0,01)	49,18* (0,01)
Raça	12,52* (0,01)	17,53* (0,01)	20,02* (0,01)
Região	-28,68* (0,01)	-29,66* (0,01)	-29,35* (0,01)
Indústria	73,62* (0,02)	69,57* (0,02)	79,95* (0,02)
Serviço	73,62* (0,03)	64,94* (0,03)	75,06 (0,02)

Obs.: Os valores dos coeficientes foram ajustados pela relação $\% = \exp(\text{coef}) - 1$. Os desvios padrões encontram-se entre parênteses.

* = significativo ao nível de 1%.

Número de observações= 21674

6. CONCLUSÃO

Os resultados da estimação das densidades de salários e da regressão quantílica confirmaram que existe discriminação contra mulheres e negros no mercado de trabalho das regiões Nordeste e Sudeste do Brasil, sendo que esta é maior para mulheres do que para negros. Além disso, a discriminação apresentou uma tendência crescente para os intervalos mais altos da distribuição de salários, o que corrobora com os resultados obtidos por Soares (2000) onde o termo discriminação é maior para os centésimos mais elevados da distribuição de salários.

Em relação discriminação racial, quando na simulação contrafactual se dá aos trabalhadores negros o mesmo nível de escolaridade dos trabalhadores brancos, há uma redução na diferença visual entre suas densidades de salários. Neste sentido, embora uma parcela considerável da diferença salarial entre raças ocorra por diferenças na educação, ainda persiste uma parcela da diferença não explicada pela educação, que pode ser atribuída justamente à discriminação.

Por outro lado, quando se dá a mesma escolaridade dos trabalhadores homens para as mulheres a diferença entre suas densidades de salários aumenta, o que é uma evidência de que esta diferença não se deve à escolaridade, mas sim a outros fatores tais como a discriminação.

¹⁰ No primeiro quartil, 25% da amostra ganha até R\$ 160, enquanto 75% ganham pelo menos 160, no segundo quartil 50% dos indivíduos ganham até R\$ 250 e a outra metade ganha acima desse valor, no terceiro quartil 25% dos indivíduos ganham pelo menos R\$ 400, enquanto 75% ganham até R\$ 400.

O índice de Gini mostrou que os salários dos trabalhadores negros tornaram-se mais concentrados após serem remunerados de acordo com a escolaridade dos brancos. Já no caso das mulheres, a simulação em geral deixou suas densidades menos concentradas. Portanto, a mudança na educação dos trabalhadores negros deixou sua distribuição de salários mais rica e desigual, por sua vez, a distribuição de salários das mulheres se tornou mais pobre e menos desigual.

A comparação entre as densidades de salários contrafactual dos trabalhadores negros e a dos brancos na agricultura mostrou que a discriminação salarial contra negros praticamente inexistente no Nordeste, embora no Sudeste ela tenha se mostrado visualmente elevada. Já a discriminação por gênero na agricultura se mostrou visualmente mais elevada no Nordeste do que no Sudeste. Na indústria, a discriminação contra negros e mulheres é maior no Sudeste. Por sua vez, no setor de serviços a discriminação contra negros não pareceu ser visualmente muito diferente entre as regiões, o mesmo aconteceu para a discriminação contra mulheres.

Assim sendo, na comparação entre as regiões não se consegue no geral identificar em qual delas é mais intenso o processo de discriminação contra as mulheres. No entanto, a discriminação contra negros aparenta ser visualmente maior na região Sudeste. Como na região Nordeste existe uma maior proporção de negros estes têm seu poder de barganha aumentado, o que provavelmente leva a uma menor discriminação no trabalho.

Com base no exposto acima, pode-se afirmar que as políticas de combate à desigualdade devem levar em consideração a raça e o gênero, além das características regionais. A desigualdade salarial entre raças pode ser reduzida por políticas que incentivem a educação dos indivíduos da raça negra, enquanto que para reduzir a desigualdade de rendimentos entre gênero é imprescindível políticas que focalizem a discriminação contra mulheres no mercado de trabalho. Ademais, o estudo da discriminação por setores de atividade indica onde as políticas regionais de combate à discriminação devem ser mais efetivas.

Deve-se estar atento também para o fato de a discriminação aumentar nos intervalos de salários mais altos, o que torna mais difícil para os indivíduos discriminados ocuparem posições de chefia dentro das empresas e, portanto, se situarem num extrato social superior.

Apesar destas conclusões deve-se salientar que nem todo diferencial de salários por gênero e raça pode ser atribuído à discriminação, pois uma parcela deste diferencial pode corresponder a outros fatores que não estão sendo considerados neste estudo.

7. REFERÊNCIAS BIBLIOGRÁFICAS

BAPTISTA, Dulce. **Diferenciais de Rendimento e Discriminação por Sexo no Mercado de Trabalho Brasileiro, na Década de Noventa**. Departamento de Economia da UFMG, 1998 (Monografia de Conclusão de Curso de Graduação).

BUTCHER, Kristin F; DINARDO. John. **The Immigrant and Native-born Wage Distributions: Evidence from United State Censuses**. NBER: Working Paper Series. N° 6630, 29p, July 1998.

CAMPANTE, Filipe R; CRESPO, Anna R. V.; LEITE. Phillipe. Desigualdade Salarial entre Raças no Mercado de Trabalho Urbano Brasileiro: Aspectos Regionais. In: **Revista Brasileira de Economia**. Rio de Janeiro, Vol 58, nº 2, p 185-210, Abril-Junho 2004.

CAVALIERI, Cláudia H; FERNANDES, Reynaldo. Diferenciais de Salários por Gênero e Cor: Uma Comparação entre as Regiões Metropolitanas Brasileiras. In: **Revista de Economia Política**. São Paulo, vol.18, nº 1, Janeiro-Março 1998.

DINARDO, John; FORTIN, Nicole M; LEMIEUX, Thomas. **Labor Market Institutions and The Distribution of Wages, 1973-1992: A Semiparametric Approach**. NBER: Working Paper Series, N° 5093, 46p, April 1995.

DUARTE, Angêlo; FERREIRA, Pedro; SALVATO, Márcio. Disparidades Regionais ou Educacionais? In: **XXXI Encontro Nacional de Economia**, 2003. 33p.

HENRIQUES, Ricardo. **Desigualdade Racial no Brasil: Evolução das Condições de Vida na Década de 90**. Rio de Janeiro: IPEA, 2001, 49p (Texto para Discussão, N° 807).

KOENKER, Roger. Quantile Regression. In: Stephen Fienberg and Jay Kadane (orgs). **Internacional Encyclopedia of the Social Sciences**, p 1-11, Oct. 2000.

LEME, Maria Carolina da Silva; WAJNMAN, Simone. Tendências de Coorte nos Diferenciais de Rendimento por Sexo. In: HENRIQUES, Ricardo (Org.). **Desigualdade e Pobreza no Brasil**. Rio de Janeiro: IPEA, 2000, p. 251-270.

OAXACA, Ronald. Male-Female Wage Differentials in Urban Labor Markets. In: **International Economic Review**, vol. 14, n° 3, p. 693-709, Oct 1973.

PARZEN, Emanuel. On Estimation of a Probability Density Function and Mode. In: **The Annals of Mathematical Statistics**, vol 3, n° 3, p 1065-1076, Sep 1962.

ROSENBLATT, Murray. Remarks on Some Nonparametric Estimates of a Density Function. In: **The Annals of Mathematical Statistics**, vol 27, n°3, p 832-837, Sep 1956.

SILVERMAN, B. W. **Density Estimation for Statistics and Data Analysis**. London; New York: Chapman and Hall. 1986. 175p.

SOARES, Serguei Suarez Dillon. **O Perfil da Discriminação no Mercado de trabalho – Homens Negros. Mulheres Brancas e Mulheres Negras**. Brasília: IPEA, 2000. 25p (Texto para Discussão, N° 769).

TIEFENTHALER, J. Female Labor Force Participation and Wage Determination in Brazil, 1989. In: PSACHAROPOULOS, G.; TZANNATOS, Z. **Women's Employment and Pay in Latin America: Overview and methodology**. Washington, D.C., The World Bank, p. 89-118, 1992.

8. ANEXOS

TABELA A1: Número de Observações em cada Amostra.

	Raça		Gênero	
	Branca	Negra	Masculino	Feminino
NE Agricultura	294	1048	1213	129
NE Industria	643	1587	1717	513
NE Serviço	1907	3499	2420	2986
SE Agricultura	317	415	588	144
SE Industria	2814	1653	3396	1071
SE Serviço	4846	2651	3355	4142

Fonte: Elaboração a partir dos dados da PNAD 2002

TABELA A2: *Bandwidth* Calculados pelo Seletor Ótimo e *Bandwidth* Suavizados no Olho por Gênero, segundo Setor de Atividade das Regiões Nordeste e Sudeste do Brasil.

Setor de Atividade	Homens no Nordeste		Mulheres no Nordeste		Homens no Sudeste		Mulheres no Sudeste	
	$h_{\text{calculado}}$	$h_{\text{suavizado}}$	$h_{\text{calculado}}$	$h_{\text{suavizado}}$	$h_{\text{calculado}}$	$h_{\text{suavizado}}$	$h_{\text{calculado}}$	$h_{\text{suavizado}}$
Agricultura	0,11	0,15	0,17	0,20	0,10	0,19	1,14	0,18
Indústria	0,10	0,22	0,11	0,20	1,14	0,21	0,14	0,20
Serviço	0,16	0,27	0,12	0,26	0,14	0,20	1,14	0,18

Fonte: Elaboração dos autores

TABELA A3: *Bandwidth* Calculados pelo Seletor Ótimo e *Bandwidth* Suavizados no Olho por Raça, segundo Setor de Atividade das Regiões Nordeste e Sudeste do Brasil.

Setor de Atividade	Branco no Nordeste		Negro no Nordeste		Branco no Sudeste		Negro no Sudeste	
	$h_{\text{calculado}}$	$h_{\text{suavizado}}$	$h_{\text{calculado}}$	$h_{\text{suavizado}}$	$h_{\text{calculado}}$	$h_{\text{suavizado}}$	$h_{\text{calculado}}$	$h_{\text{suavizado}}$
Agricultura	0,15	0,18	0,12	0,15	0,08	0,14	0,12	0,18
Indústria	0,20	0,28	0,10	0,18	0,16	0,20	0,12	1,14
Serviço	0,19	0,25	0,11	0,20	0,14	0,20	0,12	0,16

Fonte: Elaboração dos autores.

TABELA A4: Índices de Gini para as Distribuições de Salário por Gênero nos Setores de Atividade das Regiões Nordeste e Sudeste do Brasil – 2002

Índices de Gini para as Distribuições de Salários	Agricultura		Indústria		Serviço	
	Nordeste	Sudeste	Nordeste	Sudeste	Nordeste	Sudeste
Homens	0,26	0,29	0,52	0,49	0,53	0,49
Mulheres	0,30	0,23	0,43	0,47	0,51	0,46
Contrafactual Das Mulheres	0,27	0,23	0,34	0,43	0,49	0,45

Fonte: Elaboração dos autores.

TABELA A5: Índices de Gini para as Distribuições de Salário por Raça nos Setores de Atividade das Regiões Nordeste e Sudeste do Brasil - 2002

Índices de Gini para as Distribuições de Salários	Agricultura		Indústria		Serviço	
	Nordeste	Sudeste	Nordeste	Sudeste	Nordeste	Sudeste
Branco	0,29	0,24	0,57	0,51	0,55	0,49
Negro	0,26	0,29	0,43	0,38	0,46	0,42
Contrafactual Dos Negros	0,27	0,29	0,54	0,43	0,49	0,46

Fonte: Elaboração dos autores.

TABELA A6: Estatística dos Testes K-S para as Distribuições de Negros e Mulheres.

Setor de Atividade	Negros				Mulheres			
	Nordeste		Sudeste		Nordeste		Sudeste	
	KS obs	N	KS obs	N	KS obs	N	KS obs	N
Agricultura	0,67*	1048	0,61*	415	0,81*	129	0,67*	144
Indústria	0,81*	1587	0,95*	1653	0,84*	513	0,96*	1071
Serviço	0,85*	3499	0,93*	2651	0,85*	2986	0,94*	4142

Fonte: Elaboração dos autores.

* = significativo ao nível de 5%.

Obs.: o tamanho das amostras original e contrafactual é o mesmo.