

Diferenciais de Salários por Raça e Gênero no Brasil: Aplicação dos Procedimentos de Oaxaca e Heckman em Pesquisas Amostrais Complexas*

Alexandre Pinto de Carvalho ♦

Marcelo Côrtes Néri ■

Denise Britz do Nascimento Silva †

Palavras-chave: diferencial de salários, raça, gênero, procedimento de Heckman, decomposição de Oaxaca, pesquisas amostrais complexas.

RESUMO

Este artigo decompõe o diferencial de salários por cor e sexo dos trabalhadores brasileiros utilizando os microdados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2003. A metodologia consiste em estimar a equação de salários (Mincer, 1974) com a correção do viés de seleção das informações (Heckman, 1979). Adicionalmente, a decomposição do diferencial da média do logaritmo do salário/hora é obtida pelo procedimento de Oaxaca(1973) considerando dois efeitos: características produtivas e discriminação. A análise empírica tem como foco o uso adequado de procedimentos de modelagem estatística em pesquisas amostrais complexas, conforme os trabalhos de Skinner, Holt e Smith (1989) e Pessoa e Silva (1998).

Os resultados indicam a importância da incorporação do plano amostral, e da correção do viés de seleção da informação de salários, para a melhoria da qualidade das estimativas dos coeficientes das equações de salário e das medidas de discriminação. Verifica-se que a estimativa do coeficiente de discriminação entre homens e mulheres de cor branca obtida a partir da equação de salários com correção é 0,30 enquanto sem correção o valor encontrado é de 0,37. A análise da equação de participação revela que quanto maior a escolaridade, mais propensos estão os indivíduos a trabalhar. A discriminação é responsável por 97% do diferencial de salários entre homens de cor branca e as mulheres de cor preta ou parda.

* Trabalho apresentado no XV Encontro Nacional de Estudos Populacionais, ABEP, realizado em Caxambú-MG – Brasil, de 18 a 22 de Setembro de 2006.

♦ Mestre pela Escola Nacional de Ciências Estatísticas/IBGE.

■ Centro de Políticas Sociais/FGV e EPGE/FGV.

† Diretoria de Pesquisas e Escola Nacional de Ciências Estatísticas/IBGE.

Diferenciais de Salários por Raça e Gênero no Brasil: Aplicação dos Procedimentos de Oaxaca e Heckman em Pesquisas Amostrais Complexas*

Alexandre Pinto de Carvalho [♦]

Marcelo Côrtes Néri [▪]

Denise Britz do Nascimento Silva [‡]

1 Introdução

Nos últimos anos, a estimação da equação de salários¹ tem sido amplamente utilizada em estudos relativos à discriminação no mercado de trabalho brasileiro segundo cor e sexo. Os resultados revelam os efeitos dos determinantes do salário segundo características individuais e do posto de trabalho. O propósito destes trabalhos é avaliar a situação em que indivíduos com atributos produtivos semelhantes, exceto pela cor ou sexo, têm salários tão diferenciados. Este conhecimento auxilia na formulação de políticas públicas em nosso mercado de trabalho. No entanto, alguns dos estudos encontrados na literatura utilizam dados provenientes de pesquisas amostrais complexas, como a Pesquisa Nacional por Amostra de Domicílios (PNAD) realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), mas não incorporam o plano amostral em suas análises. Rodrigues (2003) exemplifica como o fato de não incorporar os pesos e o plano amostral no ajustamento de modelos de regressão gera estimativas incorretas tanto para os coeficientes como para as suas respectivas variâncias.

O objetivo principal desta dissertação é decompor o diferencial da média do logaritmo do salário/hora com ênfase na cor e sexo dos trabalhadores brasileiros (homens de cor branca, mulheres de cor branca, homens de cor preta ou parda e mulheres de cor preta ou parda). A metodologia, fundamentada no modelo de capital humano, consiste em estimar a equação de salários (Mincer, 1974) com a correção do viés de seleção das informações dos salários através do procedimento de Heckman (1979). O procedimento da decomposição do diferencial de salários, que tem como base o trabalho de Oaxaca (1973), considera dois efeitos: características produtivas individuais e da discriminação. A análise empírica tem como foco o uso adequado de procedimentos de modelagem estatística em pesquisas por amostragem complexa, conforme os trabalhos de Skinner, Holt e Smith (1989) e Pessoa e Silva (1998).

O artigo encontra-se organizado em mais cinco seções, além desta introdução, das considerações finais e da bibliografia. A seção 2 apresenta uma análise não-controlada do diferencial da média dos salários entre os homens de cor branca, homens de cor preta ou parda, mulheres de cor branca e mulheres de cor preta ou parda. Avalia-se a inserção e participação no mercado de trabalho e, adicionalmente, caracterizam-se os grupos de cor e sexo segundo estatísticas descritivas da escolaridade, idade, experiência de trabalho e local da residência. A seção 3 introduz a importância do uso adequado dos dados na inferência estatística com base em pesquisas amostrais complexas como a PNAD 2003.

* Trabalho apresentado no XV Encontro Nacional de Estudos Populacionais, ABEP, realizado em Caxambú-MG – Brasil, de 18 a 22 de Setembro de 2006.

[♦] Mestre pela Escola Nacional de Ciências Estatísticas/IBGE

[▪] Centro de Políticas Sociais/FGV e EPGE/FGV

[‡] Diretoria de Pesquisas e Escola Nacional de Ciências Estatísticas/IBGE

¹ Modelagem estatística através da Análise de Regressão.

Constitui, basicamente, um resumo dos trabalhos de Skinner, Holt e Smith (1989), Pessoa e Silva (1998), Leite (2001) e Rodrigues (2003). Nas seções 4 e 5, calcula-se a decomposição do diferencial da média do logaritmo do salário/hora através da modelagem estatística. O procedimento consiste em estimar a equação de salários para cada grupo de trabalhadores com o uso da análise de regressão (equação de salários ou equação minceriana), através do método da Máxima Pseudo-Verossimilhança (Pessoa e Silva, 1998), incorporando o plano amostral. Em seguida, com base no trabalho de Oaxaca (1973), o diferencial é decomposto em dois efeitos: das características individuais e da discriminação.

A seção 6 destaca que para determinadas pessoas seria vantajoso trabalhar se o salário recebido (ou salário potencial) fosse maior que o custo de oportunidade de não trabalhar (ou salário reserva). Nesta seção aplica-se o procedimento de Heckman (1979) para corrigir o viés de seletividade da informação dos salários para os indivíduos que, apesar de estarem devidamente representados na amostra da PNAD 2003, não trabalharam na data de referência da pesquisa supostamente devido a uma avaliação do salário potencial e do custo de oportunidade envolvido nesta escolha. A seção 7, por sua vez, apresenta as considerações finais.

2 Análise Não Controlada do Diferencial de Salários

A análise não controlada do diferencial de salários é obtida através de estatísticas descritivas do rendimento do trabalho principal (salário) e indicadores do mercado de trabalho dos grupos de cor e sexo (cor/sexo). Estes resultados são obtidos sem o uso de modelagem estatística. São considerados quatro grupos no mercado de trabalho: mulheres de cor branca, mulheres de cor preta ou parda, homens de cor branca e homens de cor preta ou parda. A fonte de dados utilizada é a Pesquisa Nacional por Amostra de Domicílios (PNAD) realizada pelo IBGE em 2003 (maiores detalhes na seção 3).

Em 2003, a população economicamente ativa² (PEA) no Brasil era em torno de 142 milhões de brasileiros³ dos quais 79 milhões tinham trabalho, 8 milhões procuravam trabalho e 55 milhões estavam fora da força de trabalho. A população ocupada era composta por aproximadamente 42 milhões de homens e mulheres que declararam sua cor branca e 37 milhões de homens e mulheres de cor preta ou parda.

Na tabela 2.1, verifica-se que, embora a população economicamente ativa de mulheres seja superior a dos homens, a sua taxa de participação é inferior. A inserção no mercado de trabalho, representada pela taxa de desocupação, revela que existe um efeito entre as variáveis cor e sexo. Destaca-se a situação das mulheres de cor preta ou parda que apresentam a menor taxa de participação (47,91%) e a maior taxa de desocupação (14,47%).

Tabela 2.1 – Taxas de participação e desocupação no Brasil segundo cor e sexo

Indicadores	Homens		Mulheres	
	Branca	Preta ou Parda	Branca	Preta ou Parda
Taxa de Participação(%)	68,87	68,56	49,41	47,91
Taxa de Desocupação(%)	7,47	9,14	11,46	14,47

Fonte: Elaboração própria a partir dos microdados da PNAD/IBGE 2003

² Pessoas com 10 anos ou mais de idade.

³ Exceto as pessoas que declararam a cor ignorada ou indígena em 2003.

As estatísticas descritivas para escolaridade, experiência de trabalho, idade, local da residência (área urbana ou rural), salário efetivamente recebido no mês, jornada de trabalho semanal são utilizadas para caracterizar os grupos de trabalhadores segundo cor e sexo. Em seguida, define-se a média da razão entre o salário e a jornada de trabalho mensal para o cálculo do diferencial não-controlado entre os trabalhadores.

Os resultados da tabela 2.2 revelam que existe uma disparidade entre os salários dos trabalhadores brasileiros segundo cor e sexo. Como exemplo, embora as mulheres de cor branca tenham um nível de escolaridade superior ao dos homens de cor branca, o seu salário é inferior, R\$657 versus R\$1.010. O mesmo resultado é encontrado entre homens e mulheres de cor preta ou parda. Além disso, os salários das pessoas de cor branca são superiores ao de pessoas de cor preta ou parda em nosso mercado de trabalho. Nota-se que a jornada de trabalho dos homens e das mulheres é diferente e, por isso, faz-se necessário calcular o diferencial através de uma medida padronizada denominada salário/hora. O salário/hora é a razão entre o salário recebido no mês e jornada de trabalho mensal (i.e., jornada de trabalho semanal multiplicada 4.2 semanas).

Tabela 2.2 – Estatísticas descritivas das características produtivas dos trabalhadores

Indicadores	Homens		Mulheres	
	Branca	Preta ou Parda	Branca	Preta ou Parda
Salário (R\$)	1.010,02	478,38	656,76	348,58
Jornada de trabalho semanal (horas)	45,40	44,55	38,10	37,31
Escolaridade (anos de estudo)	7,97	5,65	9,29	7,17
Experiência (anos)	23,54	24,25	20,71	22,23
Idade (anos)	37,51	35,90	36,00	35,40

Fonte: Elaboração própria a partir dos microdados da PNAD/IBGE 2003.

O diferencial não-controlado é apresentado na tabela 2.3 através da comparação entre homens de cor preta/parda, mulheres de cor branca e mulheres de cor preta/parda com o grupo dos homens de cor branca – definido como grupo base. O propósito desta análise é avaliar se a diferença entre o salário/hora é estatisticamente diferente de zero – o teste é realizado pela estatística t. Os homens de cor branca ganham R\$ 3,26 a mais que os homens de cor preta. O menor diferencial é registrado entre homens e mulheres de cor branca (R\$1,34). Vale ressaltar que todas as diferenças são estatisticamente diferentes de zero.

Os homens de cor branca ganham, em média, o dobro dos homens e mulheres de cor preta/parda e 28% a mais que as mulheres de cor branca.

Tabela 2.3 – Diferença da média do salário/hora segundo cor e sexo

Diferença entre o salário/hora	Estimativa	Estatística t
Homens de cor branca - Homens de cor preta	3,26	26,50
Homens de cor branca - Mulheres de cor branca	1,34	8,52
Homens de cor branca - Mulheres de cor preta ou parda	3,54	28,70

Fonte: Elaboração própria a partir dos microdados da PNAD/IBGE 2003.

Os resultados da análise não-controlada do diferencial de salários revelam as disparidades segundo cor e sexo. Em geral, as pessoas de cor preta/parda ganham menos que o grupo base. No entanto, tais comparações não são suficientes para permitir uma avaliação do quanto deste diferencial é explicado pela discriminação ou pelas características individuais (escolaridade, experiência de trabalho, local de residência e idade). As próximas seções têm por objetivo testar a existência da discriminação através da modelagem estatística, além de decompor este diferencial na parcela referente à discriminação ou às características individuais.

3 Análise de Dados Amostrais Complexos

O plano amostral da PNAD fornece resultados para a totalidade do país, para cada uma das regiões geográficas e para cada unidade da federação. A amostra tem um desenho complexo, em múltiplos estágios, é estratificado e conglomerado e apresenta probabilidades distintas de seleção das diversas unidades amostrais (detalhes em Silva, Pessoa e Lila, 2002 e Leite, 2001).

É muito comum que pesquisadores utilizem os microdados da pesquisa ignorando este tipo de informação. Ao calcular as estimativas de interesse, assumem que os dados são provenientes de uma amostra aleatória simples (AAS), onde as observações são independentes e identicamente distribuídas (IID) ou utilizam os dados como se fossem informações coletadas a partir de um censo. As estimativas pontuais dos parâmetros são influenciadas pela ocorrência de pesos amostrais distintos, enquanto as estimativas de variância (precisão) dos estimadores dos parâmetros do modelo são influenciadas, também, pelos efeitos de estratificação e conglomeração (LEITE, 2001).

No caso da modelagem estatística, a utilização de métodos adequados para realização de inferência em dados amostrais complexos permite a correta estimação dos parâmetros do modelo, do grau de precisão das estimativas e dos respectivos intervalos de confiança. Ressalta-se, então, a importância da incorporação do plano nos procedimentos de inferência com base em dados amostrais complexos como os da PNAD.

Os procedimentos estatísticos utilizados neste trabalho estão em consonância com os métodos apresentados em Skinner, Holt e Smith (1989) e Pessoa e Silva (1998) e foram executados utilizando-se o pacote estatístico STATA (2003). Detalhes sobre os procedimentos de ajuste (método de estimação e testes de qualidade do modelo) dos modelos apresentados neste artigo podem ser encontrados em Carvalho (2006).

4 Análise Controlada do Diferencial de Salários e Decomposição Segundo o Procedimento de Oaxaca

O capital humano é o conjunto das habilidades do indivíduo ligadas à capacidade produtiva, e incorporadas no conhecimento e qualificação para determinadas atividades Becker (1993). Até 1950, as análises sofisticadas do investimento em educação e treinamento por Adam Smith, Alfred Marshall e Milton Friedman não incluíam em suas discussões a produtividade. Então, Theodore W. Schultz, entre outros, iniciaram uma exploração pioneira nas implicações do investimento do capital humano nas questões econômicas.

Esta seção tem por finalidade testar a existência de discriminação dos trabalhadores segundo cor e sexo através da análise controlada do diferencial de salários. Os resultados da análise controlada são obtidos pelo uso de modelagem estatística considerando o plano amostral da PNAD 2003. Adicionalmente, decompõe-se o diferencial de salários em efeitos provenientes das características individuais (escolaridade, experiência ou local de residência) e da discriminação.

A metodologia consiste em estimar a equação de salários (Mincer, 1974) e decompor o diferencial de salários através da metodologia apresentada por Oaxaca (1973). Para isso, incorpora-se o esquema amostral complexo da PNAD 2003 conforme introduzido na seção 3. Desta forma, a sinergia entre as técnicas econométricas e as ferramentas estatísticas, fundamentadas na teoria econômica, resulta em medidas concretas que propiciam um maior conhecimento do nosso mercado de trabalho. A principal ferramenta utilizada é a Análise de Regressão⁴ com modelos estimados pelo Método de Máxima Pseudo-Verssomialhança como em Carvalho (2006). A análise de regressão é utilizada no estudo da relação de uma variável resposta com um conjunto de variáveis denominadas explicativas. O seu propósito é estimar ou prever o valor médio populacional da variável dependente em termos dos valores conhecidos⁵ das explicativas.

A *teoria do capital humano* aplicada à análise de regressão fornece um arcabouço para avaliar como decisões individuais influenciam nos retornos dos rendimentos. Os trabalhos de Becker (1962), Mincer (1974) e Ramos (1996) descrevem a ligação entre o ciclo da vida de um indivíduo e os investimentos em capital humano. Ben-Porath (1973) formaliza a equação de produção do capital humano baseado nos trabalhos de Becker e Mincer.

Diversos estudos procuram obter informações sobre os determinantes do salário dos trabalhadores a partir de suas características individuais (escolaridade, experiência, cor/raça ou local de moradia) e informações sobre o mercado de trabalho (setor de atividade ou ocupação). O entendimento do processo pelo qual as pessoas desenvolvem suas habilidades na escola e no trabalho são fundamentais para investigar não apenas o porquê das diferenças de salário, mas para fomentar idéias no desenvolvimento econômico e social do país. A literatura existente sobre o tema revela que outros fatores também influenciam o salário como, por exemplo, habilidade inata ou discriminação.

Coelho e Corseuil (2002) apresentam um resumo dos estudos sobre Diferencial de Salários nos últimos trinta anos no Brasil. Os autores encontraram diferentes abordagens no uso da equação de salário. Em alguns artigos, o foco da análise é a avaliação das características dos indivíduos na determinação do salário, enquanto outros têm como objetivo mensurar diferenciais entre grupos de trabalhadores de acordo com suas características sociodemográficas.

⁴ O termo Regressão, por sua vez, amplamente utilizado em estudos do mercado de trabalho foi introduzido Galton (1886) – Gujarati (2000).

⁵ Conhecido ou fixo no sentido de não estocástico.

Rodrigues (2003) revela a estrutura salarial do Brasil a partir da Pesquisa de Padrão de Vida (PPV), realizada em 1996 pelo IBGE, levando em consideração os efeitos do esquema amostral da pesquisa. A contribuição do trabalho é a avaliação do diferencial de salário não-controlado (razão entre os salários do grupo) e o diferencial de salário controlado quando estimado considerando-se, ou não, o esquema amostral da pesquisa.

A equação de salários (Mincer, 1974), estimada pela Análise de Regressão, é um importante instrumento para decompor o diferencial de salários entre os grupos de cor. A próxima seção discute a forma funcional da variável resposta e as variáveis explicativas utilizadas. O autor integrou a teoria do investimento em capital humano dentro de um contexto empírico, compatível com a teoria econômica. Desde então, o método passou a ser amplamente utilizado em estudos do mercado de trabalho e a equação foi denominada como “função salário do capital humano”, ou popularmente conhecida como “equação minceriana”. No caso, o logaritmo da razão entre o salário e a jornada de trabalho (hora) é a variável de interesse (também denominada por variável resposta) para o estudo do diferencial de salários.

Seja Z_i o salário/hora do i -ésimo indivíduo, $Y_i = \ln(Z_i)$ o logaritmo de Z_i e \underline{X}_i o vetor de variáveis explicativas do i -ésimo indivíduo. Considerando um modelo com apenas uma variável explicativa tem-se $Y_i = \beta_0 + \beta_k X_{ki} + \varepsilon_i$, $\varepsilon_i \sim N(0,1)$. A forma funcional desta expressão é denominada de modelo *semilog*, pois a variável resposta aparece na escala logarítmica. Assim, o coeficiente de inclinação da variável explicativa X_{ki} mede a variação proporcional constante em Y para uma dada variação absoluta no valor da variável explicativa, tal que $\beta_k = \frac{\text{Variação relativa na variável resposta}}{\text{Variação absoluta no regressor}}$. Se multiplicarmos a

variação relativa em Y por 100, a expressão fornecerá a variação percentual, ou taxa de crescimento/decrécimo em Y para uma variação absoluta em X , o regressor.

Oaxaca (1973) destaca a experiência, a escolaridade, a posição na ocupação, o setor de atividade, os grupos de ocupação, a migração e o estado civil como controles importantes da fonte da discriminação segundo cor e sexo. O controle definido pela ocupação elimina alguns dos efeitos das barreiras ocupacionais como fonte de discriminação. Como resultado, estaríamos subestimando os efeitos da discriminação quando incluimos tais fatores na análise. Com isso, ele apresentou um outro conjunto de equações de salários que não considera a ocupação e o setor de atividade dos trabalhadores denominado por “equação de salários de características individuais” que é adotado neste trabalho. A partir das evidências apresentadas por Oaxaca, e considerando o princípio da parcimônia, as variáveis explicativas aqui utilizadas são: educação, experiência e local de moradia (área urbana ou rural). A escolaridade dos trabalhadores é definida pelos anos de estudo completos. Tal informação encontra-se definida nos microdados da PNAD 2003. A experiência do trabalhador é obtida pela expressão: $\text{experiência} = \text{idade em anos completos} - \text{anos de estudo completos} - 6$.

Além do uso de modelagem, é possível descrever um indicador para avaliar o grau de discriminação entre os grupos de cor e sexo. O coeficiente de discriminação apresentado por Becker (1962) é definido como a porcentagem do diferencial salarial entre dois tipos de mercados perfeitamente substitutos. Para os casos nos quais os dois fatores não são necessariamente substitutos perfeitos, Becker definiu o coeficiente de discriminação D como uma simples diferença entre os salários e a razão dos salários na ausência de discriminação. Atribui-se como grupo base (L) os homens de cor branca e define-se também $L-1$ grupos definidos pelos homens de cor preta/parda ($l=1$), mulheres de cor branca ($l=2$) e mulheres de cor preta/parda ($l=3$).

Assim, $D = \frac{\bar{Z}_L / \bar{Z}_l - (\bar{Z}_L / \bar{Z}_l)^0}{(\bar{Z}_L / \bar{Z}_l)^0}$, $l = 1, 2, 3 \text{ e } L$, sendo:

\bar{Z}_l a média do salário/hora no l -ésimo grupo, (\bar{Z}_L / \bar{Z}_l) a razão entre as médias do salário/hora do grupo base (L) e o l -ésimo grupo, $(\bar{Z}_L / \bar{Z}_l)^0$ a razão entre as médias do salário/hora do grupo base (L) e o l -ésimo grupo na ausência da discriminação;

Oaxaca (1973) apresenta uma generalização da medida de Becker admitindo substitutos perfeitos como caso especial (sendo mais flexível para trabalhos empíricos). Para isso, aplica-se o logaritmo na expressão do coeficiente de discriminação, tal que:

$$\ln(D) = \ln \left[\frac{\bar{Z}_L / \bar{Z}_l - (\bar{Z}_L / \bar{Z}_l)^0}{(\bar{Z}_L / \bar{Z}_l)^0} \right] = \ln(D + 1) = \ln(\bar{Z}_L / \bar{Z}_l) - \ln(\bar{Z}_L / \bar{Z}_l)^0 .$$

Como não conhecemos $(\bar{Z}_L / \bar{Z}_l)^0$, para estimar D faz-se necessário obter valores para $(\bar{Z}_L / \bar{Z}_l)^0$. A estimativa é realizada considerando ou não a ausência de discriminação. Se não existisse discriminação, a estrutura salarial de um determinado grupo padrão poderia ser aplicada aos demais grupos, ou a estrutura destes grupos poderia ser aplicada ao grupo base.

A estimação pelo Método de Máxima Pseudo-Verossimilhança da equação de salários para determinados grupos de trabalhadores, fornece uma estimativa da estrutura salarial que pode ser aplicada a outros grupos. Este procedimento consiste em estimar separadamente para cada grupo de cor/sexo a equação de Mincer. A definição do modelo é dada pela expressão:

$$\ln(Z_{ii}) = \underline{X}'_i \underline{\beta}_l + \varepsilon_{ii}, \quad \varepsilon_{ii} \sim N(0,1) \quad (4.1)$$

sendo $\ln(Z_{ii})$ o logaritmo do salário/hora para o i -ésimo indivíduo pertencente ao l -ésimo grupo de cor e sexo; \underline{X}'_i a matriz das variáveis explicativas (escolaridade, experiência) do i -ésimo indivíduo do l -ésimo grupo de cor e sexo; $\underline{\beta}_l$ s parâmetros a serem estimados para o l -ésimo grupo de cor e sexo; ε_{ii} o erro aleatório.

A partir da equação 4.1 é possível inferir para os diferentes grupos de cor e sexo quais os retornos para cada característica produtiva, através dos coeficientes estimados ($\underline{\beta}_l$). Assim,

seja $G = \frac{\bar{Z}_L - \bar{Z}_l}{\bar{Z}_l}$, tal que $\ln(G + 1) = \ln(\bar{Z}_L) - \ln(\bar{Z}_l) = \bar{X}'_L \hat{\beta}_L - \bar{X}'_l \hat{\beta}_l$, sendo \bar{Z}_L e \bar{Z}_l as médias dos logaritmos dos salários/hora para o grupo-padrão e para o l -ésimo grupo de cor e sexo. Obtém-se, então:

- O efeito devido às diferenças das características produtivas: $\ln \left(\frac{\bar{Z}_L}{\bar{Z}_l} \right)^0 = \Delta \bar{X}' \hat{\beta}_l$ (4.2)

- O efeito devido às diferenças da discriminação $\ln(G + 1) = -\bar{X}'_L \Delta \hat{\beta}$. (4.3)

5 Estimação das Equações de Salários para Grupos de Cor e Sexo

São apresentados, primeiramente, os resultados da estimação da equação de salários ajustada para a população da PNAD 2003. A distinção entre os modelos 1 e 2 (tabela 5.1) está na utilização das variáveis cor e sexo, além das demais características como escolaridade, experiência e local de moradia.

Esta análise compara indivíduos com atributos semelhantes, por exemplo, escolaridade, experiência e local de moradia, exceto a cor e sexo. Os resultados revelam que mulheres de cor branca e homens e mulheres de cor preta e parda ganham menos do que o grupo base (homens brancos). Para ambos os modelos os sinais dos coeficientes da escolaridade e experiência estão de acordo com os resultados da Teoria do Capital Humano. A interpretação dos coeficientes de escolaridade (linear e quadrático) revela que o salário aumenta conforme a escolaridade a taxas crescentes. No entanto, os sinais dos coeficientes da experiência indicam que o salário aumenta conforme a experiência de trabalho a taxas decrescentes.

Tabela 5.1 –Equação de Salários – Estimativas dos Coeficientes

Características produtivas	Modelo (1)		Modelo (2)	
	Estimativa	Estatística t	Estimativa	Estatística t
Escolaridade	0.05335	15.07	0.04619	13.67
Escolaridade ²	0.00622	30.21	0.00640	32.46
Experiência	0.04951	85.31	0.04905	86.65
Experiência ²	-0.00055	-49.45	-0.00056	-51.65
Área Urbana	0.18990	27.37	0.20074	30.07
Sexo (1=homens; 0=mulheres)	-		0.29914	56.03
Cor (1=cor branca; 0=caso contrário)	-		0.26297	43.59
Constante	-0.79963	-45.32	-1.06887	-60.37
R ²	0.40		0.44	
Prob>F	0.00		0.00	

Para a realização do procedimento de Oaxaca, e obtenção dos efeitos descritos na seção 4, é necessário ajustar as equações de salários para os diferentes grupos de cor se sexo. Utilizou-se o programa estatístico STATA 8.0, procedimento *svyregress*, para o ajuste do modelo de regressão com duas finalidades: ajustar o modelo desconsiderando o desenho amostral (tabela 5.2) e utilizando a forma adequada de estimação, com a especificação dos pesos, estratos e unidades primárias de amostragem (tabela 5.3).

A comparação entre as tabelas 5.2 e 5.3 revela que o fato de não incorporar o desenho amostral no ajuste dos modelos não altera os valores dos coeficientes. O mesmo não ocorre, entretanto, com as estatísticas de teste (t). Como exemplo, enquanto a estatística de teste para a variável escolaridade na tabela 5.2 é 13.18 na tabela 5.3 é 10.41 e isso significa que o desvio padrão no primeiro caso é menor quando estimamos os coeficientes considerando que os dados fossem obtidos através de uma amostra aleatória simples. A interpretação dos coeficientes de escolaridade e experiência de trabalho é realizada em conjunto com os termos quadráticos. O efeito marginal dos anos de estudo das mulheres de cor branca (denominado por Ef marg MB) sobre a média do ln(salário/hora) das mulheres é obtido pela expressão (tabela 5.3): $Ef\ marg\ MB = -0.0191 * anos\ de\ estudo + 0.0098 * (anos\ de\ estudo^2)$.

Adicionalmente, os gráficos 5.1 e 5.2 apresentam os efeitos marginais sobre o ln(salário/hora) para a escolaridade e experiência, respectivamente. Vale ressaltar que, como o efeito marginal é a derivado da equação de salários ele é uma função linear.

Tabela 5.2 - Equações de salários estimadas com o peso e sem a especificação do plano amostral. Estimativas dos parâmetros e valores das estatísticas de teste (negrito)

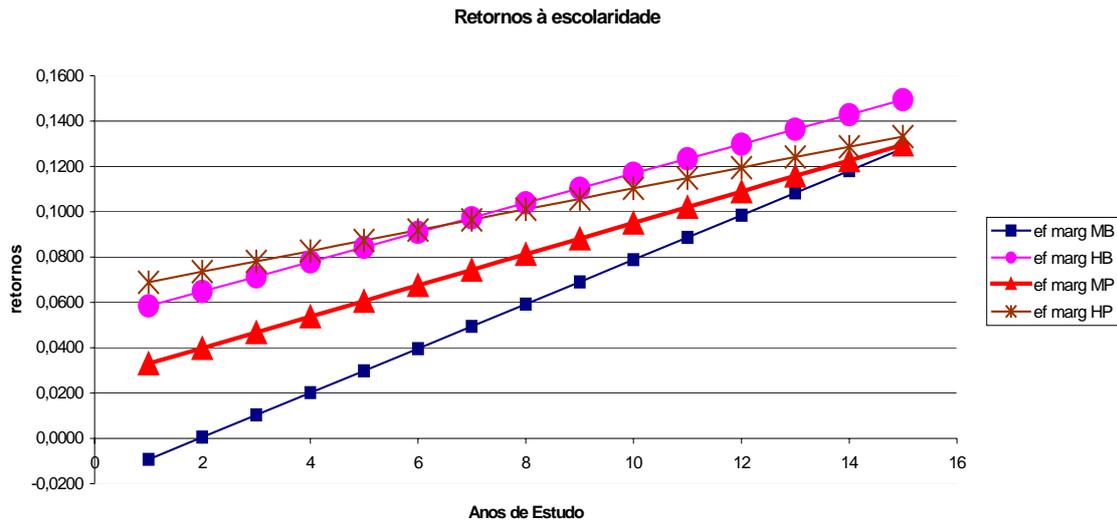
Variáveis	Cor branca		Cor Preta ou Parda	
	Homens	Mulheres	Homens	Mulheres
Anos de estudo	0,0519 13,18	-0,0191 -3,65	0,0643 19,26	0,0261 5,38
Anos de estudo ²	0,0065 28,10	0,0098 34,96	0,0046 18,74	0,0069 22,87
Experiência	0,0548 60,78	0,0382 36,00	0,0524 58,18	0,0460 37,25
Experiência ²	-0,0006 -35,28	-0,0004 -18,84	-0,0006 -40,71	-0,0005 -22,48
Área Urbana	0,2719 21,02	0,1972 9,77	0,3408 32,72	0,3494 18,31
Constante	-0,8393 -39,59	-0,5315 -16,63	-1,0332 -60,84	-1,1365 -40,20
R ²	0,44	0,42	0,34	0,32
Prob>F	0,0001	0,0001	0,0001	0,0001
Amostra	44.311	31.115	46.257	28.129
População	22.080.841	15.297.084	19.311.338	11.455.643

Tabela 5.3 - Equações de salários estimadas com o peso e com a especificação do plano amostral. Estimativas dos parâmetros e valores das estatísticas de teste (negrito)

Variáveis	Cor branca		Cor Preta ou Parda	
	Homens	Mulheres	Homens	Mulheres
Anos de estudo	0,0519 10,41	-0,0191 -3,29	0,0643 15,39	0,0261 4,18
Anos de estudo ²	0,0065 22,26	0,0098 30,98	0,0046 15,61	0,0069 18,95
Experiência	0,0548 56,97	0,0382 35,64	0,0524 47,73	0,0460 35,07
Experiência ²	-0,0006 -33,58	-0,0004 -18,44	-0,0006 -32,06	-0,0005 -21,65
Área Urbana	0,2719 12,57	0,1972 7,47	0,3408 17,36	0,3494 10,43
Constante	-0,8393 -27,50	-0,5315 -14,10	-1,0332 -35,88	-1,1365 -24,73
R ²	0,44	0,42	0,34	0,32
Prob>F	0,0001	0,0001	0,0001	0,0001
Amostra	44.311	31.115	46.257	28.129
População	22.080.841	15.297.084	19.311.338	11.455.643

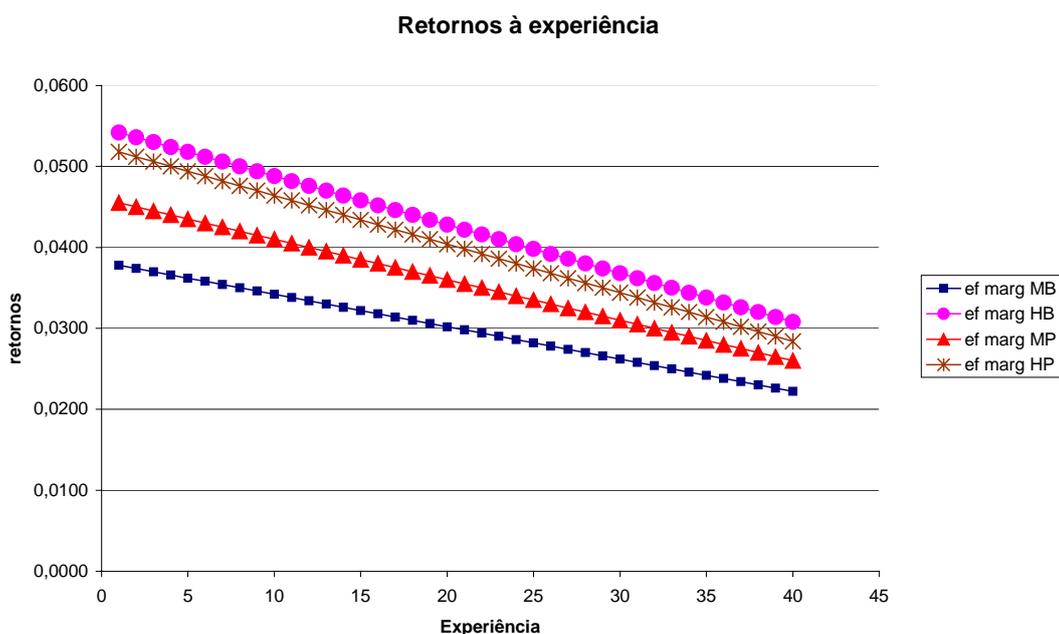
No gráfico 5.1, as mulheres de cor branca apresentam os menores retornos em todos os anos de estudo quando comparados com outros de grupos de sexo e cor. Além disso, exceto para um ano de estudo, o valor dos retornos é positivo conforme a teoria do capital humano (quanto maior a escolaridade maior o rendimento). Os homens apresentam os maiores retornos quando comparados às mulheres. Os retornos dos homens de cor branca superam o dos homens de cor preta/parda a partir dos seis anos de estudo.

Gráfico 5.1 – Retornos à escolaridade sobre o ln(salário/hora)



Nota: ef marg MB = efeito marginal das mulheres cor branca, ef marg HB = efeito marginal dos homens de cor branca, ef marg MP = efeito marginal das mulheres de cor preta/parda e ef marg HP = efeito marginal dos homens de cor preta/parda.

Gráfico 5.2 – Retornos à experiência sobre o ln(salário/hora)



5.1 Análise da Equação de Salários com efeitos de interação

Os coeficientes apresentados na tabela 5.3 foram obtidos de forma independente para grupo de cor e sexo. No entanto, para o cálculo do efeito devido às diferenças dos retornos (seção 4) é necessário estimar não só a diferença entre os coeficientes das equações por cor e sexo, mas também seu o desvio-padrão incorporando o plano amostral da pesquisa. Para tanto, ajusta-se um modelo para estimar a equação de salários com efeitos de interação entre as variáveis escolaridade, experiência, idade e local de residência e os fatores sexo e cor dos trabalhadores.

Tabela 5.4 – Ajuste do Modelo de Regressão da Equação dos Salários com a interação sexo*cor e as variáveis explicativas anos de estudo, experiência e local de residência

Variáveis	Estimativas	Estatística t
Anos de estudo (Homens de cor branca)	0,0519	10,41*
Homens de cor preta/parda	0,0124	2,13*
Mulheres de cor branca	(-) 0,0711	-10,24*
Mulheres de cor preta/parda	(-) 0,0258	-3,34*
Anos de estudo² (Homens de cor branca)	0,0065	22,26*
Homens de cor preta/parda	(-) 0,0020	-5,21*
Mulheres de cor branca	0,0033	8,7*
<i>Mulheres de cor preta/parda</i>	<i>0,0004</i>	<i>0,78</i>
Experiência (Homens de cor branca)	0,0548	56,97*
<i>Homens de cor preta/parda</i>	(-) 0,0024	<i>-1,67</i>
Mulheres de cor branca	(-) 0,0165	-11,75*
Mulheres de cor preta/parda	(-) 0,0088	-5,47*
Experiência² (Homens de cor branca)	(-)0,00059	-33,58*
Homens de cor preta/parda	(-)0,00004	-1,53
Mulheres de cor branca	0,00016	5,82*
<i>Mulheres de cor preta/parda</i>	<i>0,00005</i>	<i>1,57</i>
Área Urbana (Homens de cor branca)	0,2719	12,57*
Homens de cor preta/parda	0,0689	2,71*
Mulheres de cor branca	(-) 0,0747	-2,75*
Mulheres de cor preta/parda	0,0775	2,09*
Constante (Homens de cor branca)	(-) 0,8393	-27,5*
Homens de cor preta/parda	(-) 0,1939	-5,2*
Mulheres de cor branca	0,3078	7,28*
Mulheres de cor preta/parda	(-) 0,2972	-5,7*
R ²	0,4457	
Prob>F	0,0000	
Amostra	149.812	
População	68.144.906	

Nota: * significativo ao nível de 5% de confiança.

Os resultados da tabela 5.4 em negrito representam as estimativas dos coeficientes para o grupo base *L* (homens de cor branca). As linhas em itálico representam a diferença entre o grupo base (*L*) e o *l*-ésimo grupo (homens de cor preta/parda, mulheres de cor branca e mulheres de cor preta/parda). Como exemplo, o coeficiente para os anos de estudo dos homens de cor branca é **0,0519** (tabela 5.2 e tabela 5.3) enquanto que o para os homens de cor preta/parda é obtido por: **0,0519** + 0,0124 (tabela 5.4) = **0,0643** (tabela 5.2). A amostra considera todos os grupos, aproximadamente 149 mil trabalhadores.

5.2 Resultados da decomposição do diferencial de salários

O objetivo desta seção é detalhar os procedimentos para decomposição da média do logaritmo do salário/hora através da análise controlada (modelagem estatística). Em seguida, investigar o quanto desta diferença é explicada pelas características individuais dos trabalhadores e quanto é proveniente dos retornos dos coeficientes. Os resultados da tabela 5.5 permitem avaliar se a diferença entre a média do logaritmo do salário/hora do grupo base (homens de cor branca) e o *l*-ésimo grupo de cor e sexo é estatisticamente significativa.

Tabela 5.5 – Diferença do ln(salário/hora) entre os grupos de cor

Diferenças	Estimativa	Estatística t
Homens de cor branca - Homens de cor preta	0.59	50.72
Homens de cor branca - Mulheres de cor branca	0.15	18.83
Homens de cor branca - Mulheres de cor preta ou parda	0.69	56.74

Fonte: Elaboração própria a partir dos microdados da PNAD 2003.

Tabela 5.6 – Diferenças entre as médias das características produtivas, entre coeficientes estimados para o grupo padrão (*L*) e o *l*-ésimo grupo de cor e sexo, e valores das estatísticas de teste (em negrito)

Características	Diferenças : (Grupo padrão) - (<i>l</i> -ésimo grupo de cor/sexo)					
	Entre as médias			Entre os coeficientes estimados		
	Mulher de cor branca	Mulher de cor preta ou parda	Homem de cor preta ou parda	Mulher de cor branca	Mulher de cor preta ou parda	Homem de cor preta ou parda
Anos de estudo	-1,32	0,80	2,32	-0,0711	-0,02578	0,0124
	-36,17	15,32	44,55	-10,2	-3,3	2,1
Anos de estudo ²	-21,83	12,77	33,41	0,0033	0,00035	-0,0020
	-36,92	15,87	44,21	8,7	0,8	-5,2
Experiência	2,84	1,31	-0,71	-0,0165	-0,00883	-0,0024
	23,89	9,92	-5,35	-11,8	-5,5	-1,7
Experiência ²	159,82	85,74	-40,09	0,0002	0,00005	-0,000039
	23,66	11,11	-4,91	5,8	1,57	-1,5
Área Urbana	-0,07	-0,03	0,06	-0,0747	0,07745	0,0689
	-21,54	-5,46	9,77	-2,8	2,1	2,7
Constante				0,3078	-0,29722	-0,1939
				7,28	-5,70	-5,20

O cálculo da diferença entre as médias das características produtivas (equação 4.2) foi obtido com a incorporação do plano amostral da PNAD. Esta análise tem por finalidade testar se a diferença entre as características dos grupos de cor e sexo são estatisticamente diferentes de zero. Verifica-se na tabela 5.6 que os homens têm, em média, uma escolaridade inferior às mulheres de cor branca. Com base nos resultados das tabelas 5.5 e 5.6 é possível a construção da decomposição do diferencial da média do $\ln(\text{salário/hora})$ de forma adequada.

A finalidade da tabela 5.7 é avaliar o quanto as características produtivas explicam o diferencial e qual parcela é proveniente da discriminação. A estimativa do coeficiente de discriminação (D), por sua vez, sintetiza os resultados da análise controlada da discriminação presente no mercado de trabalho brasileiro em 2003. A conclusão é que a discriminação no mercado de trabalho é mais evidente contra as mulheres de cor preta. Desta forma, considerando como grupo base os homens de cor branca, o maior valor registrado de D foi de 0,78 entre as mulheres de cor preta/parda e o grupo base. Em seguida, 0,37 para as mulheres de cor branca e 0,33 para os homens de cor preta e parda.

Tabela 5.7 - Efeitos da Discriminação Estimados pelas Características Pessoais

Efeitos	Mulher branca		Homem preto		Mulher preta	
	(1)a	(2)b	(1)a	(2)b	(1)a	(2)b
Diferencial de salários = (3)	0,152	100,0%	0,593	100,0%	0,688	100,0%
Anos de estudo	0,0254	16,7%	0,1489	25,1%	0,0209	3,04%
Anos de estudo ²	-0,2143	-140,9%	0,1526	25,7%	0,0878	12,76%
Experiência	0,1085	71,4%	-0,0371	-6,3%	0,0602	8,74%
Experiência ²	-0,0686	-45,1%	0,0253	4,3%	-0,0467	-6,79%
Área Urbana	-0,0138	-9,1%	0,0212	3,6%	-0,0108	-1,58%
Efeito de Discriminação						
Somatório do efeitos = (4)	-0,1629	-107,1%	0,3110	52,4%	0,1113	16,2%
Estimativa de $\ln(D+1)$ = (5)	0,3150	207,1%	0,2825	47,6%	0,577	83,8%
Estimativa de D = (6)	0,37		0,33		0,78	

Notas:

(a) é igual ao produto entre o coeficiente da variável explicativa e a diferença das características médias da tabela 5.5

(b) é igual a coluna (a) expressa como percentual da diferencial de salários.

(3) Representa o diferencial de salários entre o grupo padrão e o l-ésimo grupo de cor/sexo

(4) Somatório dos efeitos sobre o diferencial de salários

(5) É igual a (3) - (4)

(6) É igual a exponencial do item (5)

Esta seção desenvolveu os procedimentos teóricos e operacionais para o cálculo da decomposição da média do $\ln(\text{salário/hora})$ em pesquisas como a PNAD 2003. Foi identificado que a melhor maneira para estimar as equações de salários é através de um modelo com interações que já fornece as estimativas das diferenças dos coeficientes, e respectivos desvios-padrão com a incorporação do plano amostral. A próxima seção tem por finalidade avaliar e corrigir os possíveis vieses de seleção da informação do salário.

6 Decomposição do Diferencial de Salário com Correção do Viés de Seleção do Salário na Análise de Regressão

Esta seção tem por finalidade apresentar os procedimentos necessários para corrigir o viés de seleção das informações dos salários reportados pelos indivíduos. Com isso, avaliar os impactos nas equações de salários e nos resultados da decomposição de salários. Com relação à variável resposta, $\ln(\text{salário/hora})$, vale ressaltar que na PNAD, como em outras pesquisas, as informações coletadas são fornecidas pelas pessoas que tinham trabalho na época da pesquisa. Isto é, os salários observados na PNAD 2003 estão relacionados com a decisão de um indivíduo trabalhar ou não.

De fato, para determinadas pessoas seria vantajoso trabalhar se o salário recebido (ou salário potencial) fosse maior que o custo de oportunidade (ou salário reserva). Desta forma, existe um viés de seleção das informações do salário. Neste caso, as estimativas dos coeficientes da equação de salários obtidas a partir das informações dos indivíduos que trabalham na data de referência da pesquisa podem estar viesadas sob a ótica do **viés de seleção das informações**, considerando-se que o modelo utilizado na seção 5 não incorpora a informação sobre a avaliação dos indivíduos que não trabalham no que se refere ao custo de oportunidade.

Heckman (1979) apresentou o fenômeno e sua solução não apenas para pesquisas amostrais. De modo geral, o viés de seleção de informação pode ser decorrente de duas razões, como destaca Heckman: ou em virtude de seletividade das informações dos indivíduos ou devido ao desenho amostral da pesquisa. Cabe ressaltar que, ao utilizar os dados da PNAD, o efeito que deseja-se incorporar na modelagem visa corrigir o viés de seletividade da informação dos salários para os indivíduos que, apesar de estarem devidamente representados na amostra da PNAD 2003, não trabalhavam na data de referência da pesquisa supostamente devido a uma avaliação do salário potencial e do custo de oportunidade envolvido nesta escolha.

O problema de estimar a equação de salários é que não observamos os salários para toda a amostra, mas apenas para aqueles que trabalham. A metodologia apresentada é um resumo dos trabalhos de Heckman (1979) e Kassouf (1994) e sua motivação é incorporar a complexidade do desenho amostral (seção 3) na estimação dos modelos. A equação de participação avalia a probabilidade do indivíduo trabalhar segundo algumas variáveis explicativas⁶. O modelo utilizado é o *probit* com a incorporação do plano amostral. A variável dependente assume o valor 1 se o indivíduo tem rendimento (ocupado) e 0 caso contrário (descocupados ou inativos).

Os resultados da equação de salários – modelo *probit* – estão descritos na tabela 6.1. Verifica-se que conforme aumenta a experiência de trabalho maior é a probabilidade de um indivíduo participar do mercado de trabalho. O sinal negativo do termo quadrático da experiência indica que a probabilidade de participar no mercado de trabalho cresce a taxas decrescentes. O efeito do termo linear da escolaridade em todos os grupos de cor reflete que a probabilidade aumenta conforme os anos de estudo.

⁶ Ressalta-se que para identificação é necessário que algumas destas variáveis não estejam incluídas na equação de salários.

No entanto, o termo quadrático indica que, para os homens, a probabilidade cresce a taxas decrescentes e para as mulheres a taxas crescentes. Ser chefe do domicílio é uma variável importante para determinar a probabilidade do indivíduo participar do mercado de trabalho, independente do sexo e a da cor. Para os indivíduos que declararam ter uma relação de filho com o chefe do domicílio, o comportamento é o contrário. Isto é, o sinal negativo expressa que estes indivíduos estão menos propensos a participar do mercado de trabalho. Além das características pessoais e de responsabilidades no domicílio, foi incluída uma informação que tem impacto sobre todas as pessoas que vivem no mesmo domicílio: se há *crianças de 0 a 5 anos de idade*. Considerando esta informação, enquanto os homens estão mais propensos a trabalhar, as mulheres têm um comportamento contrário. Heckman (1979) e Kassouf (1994) justificam este comportamento pela regra de decisão entre o salário potencial e o salário reserva descrito na seção anterior.

Tabela 6.1 – Equação de Participação no Mercado de Trabalho Brasileiro em 2003

Características	Cor branca		Cor Preta ou Parda	
	Homens	Mulheres	Homens	Mulheres
Anos de estudo	0,1484 30,05	0,0752 16,49	0,1027 19,38	0,0560 11,36
Anos de estudo ²	-0,0043 -12,14	0,0017 5,51	-0,0001 -0,31	0,0044 11,26
Experiência	0,0944 63,52	0,0745 50,16	0,1087 74,94	0,0911 65,62
Experiência ²	-0,0016 -61,00	-0,0014 -48,44	-0,0017 -61,37	-0,0015 -56,98
Área Urbana	-0,2875 -13,75	0,2482 9,57	-0,1858 -8,47	0,2398 10,28
Chefe do domicílio	0,5883973 25,89	0,4216303 23,2	0,5646465 25,7	0,4643732 27,02
Filho	-0,26 -14,37	-0,12 -6,56	-0,23 -13,83	-0,05 -2,81
Criança de 0 a 5 anos	0,02 1,53	-0,18 -12,79	0,04 3,01	-0,21 -14,27
Constante	-1,1674 -43,90	-1,7858 -54,57	-1,3662 -47,77	-2,0279 -62,18
Amostra	87.601	96.528	98.593	99.112

A partir dos coeficientes da equação de participação (*modelo probit*), calcula-se a variável *lambda*, ou razão inversa de Mills (detalhes em Carvalho, 2006), que é utilizada como variável explicativa para estimação da equação de salários. Os resultados da tabela 6.2 apresentam as equações de salários com a correção de Heckman. Vale ressaltar que o coeficiente da variável *lambda* é estatisticamente significativo, indicando que sua inclusão é necessária para correção do viés de seleção da informação do salário.

Tabela 6.2 – Equação de Salários com Correção de Heckman

Características	Cor branca		Cor Preta ou Parda	
	Homens	Mulheres	Homens	Mulheres
Anos de estudo	0,0280 5,42	-0,0257 -3,98	0,0467 10,88	0,0217 3,44
Anos de estudo ²	0,0072 24,06	0,0098 30,82	0,0046 15,66	0,0067 18,10
Experiência	0,0311 24,51	0,0324 11,71	0,0255 14,54	0,0401 17,94
Experiência ²	-0,0002 -9,96	-0,0003 -6,28	-0,0002 -8,62	-0,0005 -11,94
Área Urbana	0,3169 14,42	0,1750 6,1	0,3711 18,89	0,3312 9,93
Lambda	-0,3456 -24,38	-0,1068 -2,32	-0,3569 -19,39	-0,0908 -3,10
Constante	-0,3256 -9,12	-0,3040 -2,87	-0,4560 -11,78	-0,9317 -12,92

O sinal negativo da variável *lambda* indica que os fatores latentes (não mensurados) na equação de salários aumentam a probabilidade de participação, porém diminuem os retornos do salário. Por exemplo para os homens de cor branca (gráfico 6.1), o coeficiente da escolaridade estimado pelo método tradicional era 0.052 (seção 5) e com a correção de Heckman passa a ser 0.028 (tabela 6.2). Quando realizamos a mesma análise para a variável experiência de trabalho, para todos os grupos de cor e sexo, os coeficientes também sofrem reduções. A comparação com os resultados do método tradicional (seção 5) revela que, tanto os coeficientes, quanto as variâncias, se alteram com a utilização deste procedimento. Mesmo assim, todas os coeficientes são significativos no modelo. De posse das estimativas da equação de salários corrigidas para o viés de seleção da informação, o próximo passo é decompor o diferencial entre os grupos de trabalhadores (tabela 6.3).

Gráfico 6.1 – Comparação entre os coeficientes da Equação de Salários pelo Método Tradicional e com a Correção de Heckman.

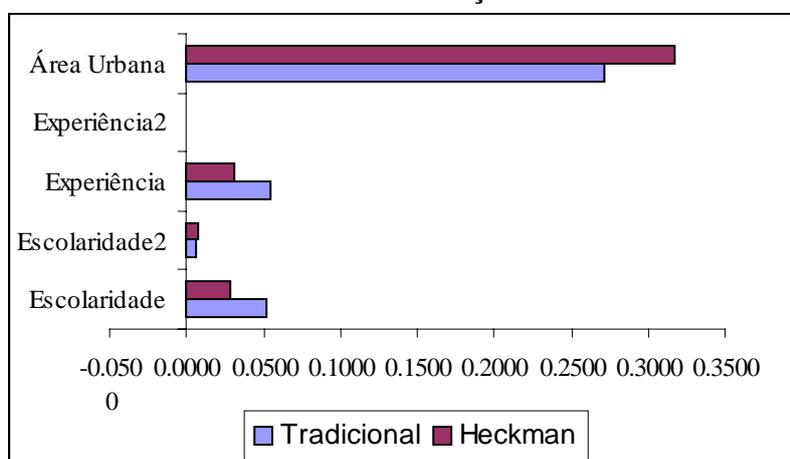


Tabela 6.3 – Diferenças entre as médias dos regressores e coeficientes estimados com a correção de Heckman

Características	(Grupo padrão) - (l-ésimo grupo de cor e sexo)					
	Entre as médias			Entre os coeficientes estimados		
	Mulher de cor branca	Mulher de cor preta ou parda	Homem de cor preta ou parda	Mulher de cor branca	Mulher de cor preta ou parda	Homem de cor preta ou parda
Anos de estudo	-1,32 -36,17	0,80 15,32	2,32 44,55	-0,0537 -4,6193	0,0187 1,9769	-0,0064 -0,5544
Anos de estudo ²	-21,83 -36,92	33,41 15,87	33,41 44,21	0,0026 4,1266	-0,0026 -4,3515	-0,0005 -0,7658
Experiência	2,84 23,89	-0,71 9,92	-0,71 -5,35	0,0013 0,3283	-0,0056 -1,8592	0,0090 2,5816
Experiência ²	159,82 23,66	-40,09 11,11	-40,09 -4,91	-0,0001 -1,5198	0,0000 -0,4806	-0,0002 -4,0455
Área Urbana	-0,07 -21,54	0,06 -5,46	0,06 9,77	-0,1420 -2,8024	0,0542 1,3020	0,0143 0,2583
Lambda	-0,36	-0,42	-0,42	0,2388 3,9667	-0,0114 -0,3489	0,2547 5,8578
Constante		-	-	0,0216	-0,60614	-0,1305

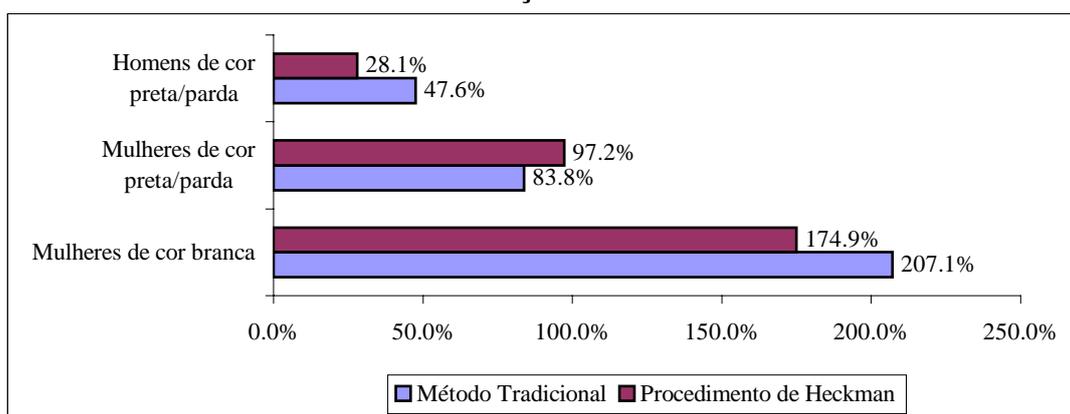
Tabela 6.4 – Efeitos da Discriminação Estimados pelas Características Pessoas – Correção de Heckman

Efeitos	Mulher branca		Homem preto		Mulher preta	
	(1)a	(2)b	(1)a	(2)b	(1)a	(2)b
Diferencial de salários = (3)	0,152	100,0%	0,593	100,0%	0,688	100,0%
Anos de estudo	0,0341	22,4%	0,1083	18,3%	0,00040551	0,05891%
Anos de estudo ²	-0,2139	-140,7%	0,1551	26,1%	-0,00001751	-0,00254%
Experiência	0,0919	60,4%	-0,0180	-3,0%	-0,00022505	-0,03269%
Experiência ²	-0,0519	-34,1%	0,0095	1,6%	0,00000001	0,00000%
Área Urbana	-0,0123	-8,1%	0,0231	3,9%	0,01795235	2,60798%
Lambda	0,0382	25,1%	0,1488	25,1%	0,00103260	0,15001%
Somatório do efeitos = (4)	-0,1139	-74,9%	0,4268	71,9%	0,0191	2,8%
Efeito de Discriminação						
Estimativa de ln(D+1) = (5)	0,27	174,9%	0,17	28,1%	0,67	97,2%
Estimativa de D = (6)	0,30		0,18		0,95	

Na tabela 6.4, o efeito da discriminação explica 97% (contra apenas 3% das características pessoais) do diferencial observado entre homens de cor branca e mulheres de cor preta/parda. Entre homens e mulheres de cor branca o efeito da discriminação explica 175% do diferencial de salários.

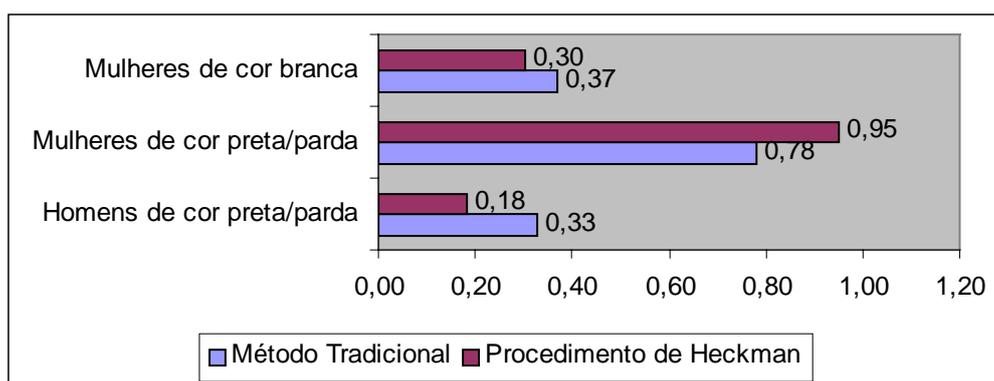
O gráfico 6.2 revela que se a decomposição fosse realizada pelo método tradicional a conclusão seria que 84% do diferencial entre homens de cor branca e mulheres de cor preta é proveniente da discriminação. No entanto, com a correção de Heckman, este indicador é de 97%, ou seja, estaríamos subestimando o efeito da discriminação. Destaca-se o efeito da discriminação entre homens e mulheres de cor branca que cai 7 pontos percentuais com a correção de Heckman.

Gráfico 6.2 - Efeito da Discriminação – Método Tradicional e de Heckman



A estimativa do coeficiente de discriminação estaria subestimada pelo método tradicional (gráfico 6.3), exceto na comparação entre os homens de cor branca e as mulheres de cor preta cujo coeficiente aumentou de 0.78 (método tradicional) para 0.95 (correção de Heckman).

Gráfico 6.3 – Estimativa do Coeficiente de Discriminação – Método Tradicional e Procedimento de Heckman



7 Considerações Finais

Em geral, as análises da equação de salários que utilizam os microdados da PNAD não incorporam as complexidades da pesquisa. Além disso, os resultados são utilizados para o estudo da decomposição do diferencial de salários. Contudo, esse procedimento pode não ser o mais adequado na avaliação da discriminação no mercado de trabalho brasileiro.

Este artigo conclui que o procedimento de Heckman corrige o viés de seleção das informações dos salários sendo adequado para obter estimativas consistentes dos parâmetros da equação. Por exemplo, verifica-se que a estimativa do coeficiente de discriminação entre homens e mulheres de cor branca obtida a partir da equação de salários com correção é 0,30 enquanto sem correção o valor encontrado é de 0,37. A análise da equação de participação revela que quanto maior a escolaridade, mais propensos estão os indivíduos a trabalhar. A discriminação é responsável por 97% do diferencial de salários entre homens de cor branca e as mulheres de cor preta ou parda. Registra-se, como uma extensão natural deste trabalho, a análise da decomposição do salário/hora para os diferentes percentis da distribuição do rendimento de trabalho. Este estudo já foi realizado por Soares (2000) como uma extensão da metodologia de Oaxaca, porém sem incorporar o plano amostral na modelagem.

8 Referências Bibliográficas

- BECKER, G.S. Investment in human capital: A theoretical analysis. *Journal of Political Economy*, vol. 70, n. 5, part 2), 9-49, Chicago, 1962
- BECKER, G. S. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. Chicago. The University of Chicago Press, 1993.
- BEN-PORATH, Y. Short-Term Fluctuations in Fertility and Economic Activity in Israel. *Demography*. vol.10, n.2.185–204. Maryland, 1973.
- CARVALHO, A. P. Decomposição do Diferencial de Salários no Brasil em 2003: Uma Aplicação dos Procedimentos de Oaxaca e Heckman em Pesquisas Amostrais Complexas, Dissertação de Mestrado, Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2006.
- COELHO, A. MORI, CORSEUIL, CARLOS H. Diferenciais salariais no Brasil: um breve panorama. Brasília, IPEA, 2002. (Textos Para Discussão, 898).
- HECKMAN, J.J. Sample Selections Bias as a Specification Error. *Econometrica*, vol. 47, nº 1, 1979.
- KASSOUF, A. L. The Wage Rate Estimation Using The Heckman Procedure. *Revista de Econometria*. Rio de Janeiro, v.14, nº1, p.89-107, 1994.
- LEITE, P.G. Análise da Situação Ocupacional de Crianças e Adolescentes nas Regiões Sudeste e Nordeste do Brasil Utilizando Informações da PNAD 1999. Dissertação de Mestrado. Escola Nacional de Ciências Estatísticas. Rio de Janeiro, 2001.
- MINCER, J. *Schooling, Experience, and Earnings*. New York 1974
- OAXACA, R. Male-Female Wage Differentials in Urban Labor Market. In *International Economic Review*, v. 14, n.3, p. 693-709. 1973.
- PESSOA, D.G.C.; SILVA, P.L.N. Análise de dados amostrais complexos. São Paulo: ABE, 1998. (13º Simpósio Nacional de Probabilidade e Estatística – Caxambu) 187p.
- RODRIGUES, S.C. Análise da Estrutura Salarial Revelada pela PPV Incorporando Peso e Plano Amostral. Rio de Janeiro, 2003. Dissertação de Mestrado, Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2006.
- SILVA, P.L.N, PESSOA D.G.C, e LILA, M.F. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. *Ciência & Saúde Coletiva*. vol.7 .no.4. Rio de Janeiro, 2002.
- SKINNER, C.J.; HOLT, D e SMITH, T.M.F. *Analysis of Complex Surveys*. Chichester: John Wiley & Sons, 1989.
- SOARES, S. S. D. O perfil da discriminação no mercado de trabalho – homens negros, mulheres brancas e mulheres negras. Texto para discussão no 769. 26p. 2000.