

AVALIAÇÃO DOS RETORNOS À ESCOLARIDADE PARA TRABALHADORES DO SEXO MASCULINO NO BRASIL *

Vanessa da Fonseca Pereira^a
João Eustáquio de Lima^b
João Ricardo Ferreira de Lima^c
Marcelo José Braga^d
Talles Girardi de Mendonça^e

RESUMO: Este artigo avalia os retornos a anos adicionais de estudos para trabalhadores do sexo masculino no Brasil. Com base nos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD), de 2009, e aplicando-se o procedimento de Heckman, mensuram-se as diferenças nos retornos salariais à escolaridade nos meios urbano e rural. Como aspecto de destaque metodológico e contribuição para o conjunto de estudos sobre retornos à escolaridade, tem-se a incorporação das informações do plano amostral nas estimativas. Os resultados mostram diferenciação salarial em função de idade, raça e região geográfica. Em relação à escolaridade, os retornos salariais no meio urbano são maiores que no rural. Adicionalmente, a consideração do plano amostral gera resultados expressivamente diferentes daqueles considerando a PNAD como amostra simples. Torna-se, portanto, necessário incorporar esses aspectos às novas pesquisas baseadas em fontes de dados como a PNAD.

PALAVRAS-CHAVE: Retornos à educação; qualificação; amostra complexa.

CLASSIFICAÇÃO JEL: J24; J31.

* Artigo recebido em 05/07/2011 e aprovado em 22/02/2013.

^a Estudante do Doutorado em Economia Aplicada pela Universidade Federal de Viçosa (UFV) e analista A da Empresa Brasileira de Pesquisa Agropecuária (Embrapa). Contato: vanessa.pereira@embrapa.br.

^b Doutor em Economia Rural pela Michigan State University e professor titular do Departamento de Economia Rural da UFV. Contato: jelima@ufv.br.

^c Doutor em Economia pela UFV e pesquisador A da área de Economia Aplicada da Embrapa Semiárido. Contato: joao.ricardo@cpatsa.embrapa.br.

^d Doutor em Economia Aplicada pela UFV e professor adjunto do Departamento de Economia Rural da UFV. Contato: mjbraga@ufv.br.

^e Doutor em Economia Aplicada pela UFV e professor adjunto do Departamento de Ciências Econômicas Universidade Federal de São João Del Rei (UFSJ). Contato: tallesgm@ufsj.edu.br.

EVALUATING THE RETURNS FROM SCHOOLING FOR BRAZILIAN MALE WORKERS

ABSTRACT: This study aims at evaluating the returns to additional years of education for male workers in Brazil. It applies Heckman's procedure to measure the differences between wage returns to schooling in urban and rural areas, based on data from the National Household Sample Survey (PNAD), for the year 2009. A significant methodological contribution to the literature on returns to schooling has been made by incorporating data sample into the estimation. Results show wage differentiation due to age, race and geographic region; as to education, wage returns in urban areas are bigger than in rural areas. Additionally, the use of a complex sample has produced results that are considerably different from those based on simpler sample. Therefore, it is deemed necessary to incorporate those aspects into new researches based on data from PNAD.

KEYWORDS: Returns to education; qualification; complex sample.

1. INTRODUÇÃO

Os estudos pioneiros de Mincer (1958) e Becker (1962) despertaram interesse constante e crescente aos pesquisadores nas estimativas dos retornos aos investimentos em educação e treinamento. Com base nesses trabalhos, o vasto conjunto literário da Teoria do Capital Humano postula que a educação, à medida que melhora a habilidade dos trabalhadores, dotando-lhes de maior nível de conhecimento, torna-os mais eficientes no desempenho de suas funções. O aumento da eficiência aumenta a sua produtividade, o que leva o empregador a aumentar a sua remuneração. Por conseguinte, há uma tendência de que os diferenciais salariais reflitam os padrões educacionais dos trabalhadores.

Rosenzweig (1995) enfatizou dois canais pelos quais a escolaridade é capaz de elevar a produtividade. O primeiro diz respeito ao acesso às fontes de informação, como jornais e manuais de instrução. O segundo refere-se às melhorias na habilidade de aprender e decifrar novas informações, advindas de fontes externas ou de sua própria experiência.

Por meio de mecanismos como esses, os investimentos em educação contribuem para o desenvolvimento econômico dos países. As estimativas dos retornos a esses investimentos têm maior importância em contextos onde a desigualdade de renda é acentuada, como no Brasil. Apesar de o índice de Gini da distribuição do rendimento domiciliar *per capita* ter diminuído anualmente, desde 2001, o seu valor ainda era elevado em 2009 (0,538) (IPEA, 2010). É interessante enfatizar que as diferenças nos rendimentos entre o meio rural e o urbano contribuem para manter elevada a desigualdade na distribuição de renda no Brasil.

Dessa forma, os retornos aos investimentos em capital humano no Brasil são importantes para sugerir áreas de concentração e direcionamento dos recursos públicos e privados. O conhecimento das respostas salariais à escolaridade fornece informações úteis para avaliações e alterações dos investimentos no sistema educacional. Angrist e Pischke (2009) ressaltaram que estimativas dos efeitos da escolaridade sobre os salários são úteis para prever as consequências salariais de intensificar as leis de frequência obrigatória ou de mudar os custos de frequentar um colégio, por exemplo.

Portanto, este trabalho objetivou avaliar os retornos à escolaridade para trabalhadores do sexo masculino no Brasil. Pretende-se apresentar os efeitos, em termos de retornos salariais, que são decorrentes de anos adicionais de estudos. É essencial destacar que a abordagem estatística deste estudo leva em consideração o delineamento amostral da Pesquisa Nacional por Amostra de Domicílios (PNAD). Dessa forma, têm-se estimativas pontuais e erros-padrão calculados corretamente, diferentemente de outros trabalhos que consideram a PNAD uma amostra independente e igualmente distribuída (i.i.d). A consideração da amostragem aleatória é uma das contribuições do presente trabalho.

Devido à relevância do tema, nota-se expressivo volume de trabalhos acerca dos retornos salariais no Brasil, desenvolvidos a partir dos dados da PNAD. Os principais aspectos enfocados como causas da diferenciação salarial são: idade; experiência; nível educacional; cor; gênero; região geográfica; e situação censitária¹. Uma breve exposição de trabalhos anteriormente realizados será apresentada para dar maiores subsídios à análise e ratificar a importância do presente estudo.

Langoni (1971) discorreu acerca de três estudos, dentre os quais está a sua tese de PhD², de 1970, a qual é considerada pioneira nos retornos à educação no Brasil. Nessa tese, os retornos ao investimento em capital humano faziam parte de uma análise mais ampla sobre o crescimento econômico do Brasil, do ponto de vista da eficiência alocativa. Ao utilizar dados da PNAD, de 1969, o autor encontrou um retorno aos investimentos em educação próximo a 28% para população economicamente ativa do sexo masculino.

Contudo, o crescimento dessa linha de pesquisa se deu a partir da década de 1990. Lam e Levison (1990) basearam-se nos dados da PNAD, relativos ao ano de 1985, para avaliar os padrões salariais; no que tange a escolaridade, idade e experiência; comparando homens brasileiros com os norte-americanos. Em relação às regiões geográficas brasileiras, esses autores identificaram proximidade entre as respostas salariais à educação. Já a resposta média ao ano adicional de estudo foi de 17%.

Leal e Werlang (1991) optaram por trabalhar com um pseudopainel, avaliando os retornos à educação dos homens das regiões metropolitanas brasileiras, com base nas PNAD relativas a 1976 a 1989. Para consideração da educação, os autores criaram categorias de acordo com os anos de estudo e trabalharam, então, com as variáveis *dummy* (analfabeto, primários 1 e 2, secundário e superior). As taxas de retorno estiveram em torno de 16% e variavam de acordo com o nível educacional e, também, ao longo do período de análise.

Kassouf (1994) apresentou contribuição significativa às pesquisas, ao tratar a questão do viés da seletividade amostral, que até então não era abordada. Nesse trabalho, a autora fez uma comparação do método tradicional de estimação com o procedimento de Heckman (1979), a partir da qual foi mostrada a relevância estatística da consideração do processo de seleção amostral nos estudos sobre respostas salariais. Em outro trabalho, Kassouf (1997) comparou os retornos quanto ao gênero e à situação censitária e aferiu, para os homens, retorno superior a 15%, no meio urbano, e superior a 11%, no rural.

Sachsida *et al.* (2004), também ao considerarem a seletividade amostral, avaliaram os retornos à escolaridade dos homens por meio de análise de seção cruzada, para o

¹ Refere-se ao fato de o domicílio situar-se na zona urbana ou rural.

² Publicada pela APEC Editora, sob o título: *As causas do crescimento econômico do Brasil*, em 1974.

ano de 1996, e com dados empilhados, para o período entre 1992 a 1999. Além de ratificarem a necessidade do processo de seleção amostral, os autores aferiram retornos de aproximadamente 16%.

Resende e Wyllie (2006) tomaram como referência a Pesquisa sobre o Padrão de Vida (PPV), também do IBGE, e compararam os padrões de diferenciais de salários entre mulheres e homens. Por meio do tratamento do viés de seletividade amostral, esses autores encontraram um retorno de 12,6% e 15,9% para mulheres e homens, respectivamente.

Destaca-se que, apesar de os procedimentos de amostragem da PNAD e da PPV possuírem desenho amostral complexo, os trabalhos supramencionados apoiaram-se em estimativas com base em amostragem aleatória simples³ ou apenas incorporando os pesos amostrais.

Nesse contexto, a partir da constatação das características específicas do plano amostral da PNAD, o trabalho de Carvalho *et al.* (2006) avaliou, entre diversos outros aspectos, a resposta dos rendimentos dos indivíduos à educação. Esses autores aplicaram o procedimento de Heckman com o enfoque nos diferenciais de salário por gênero e raça no Brasil, com base nos dados da PNAD do ano de 2003. Dentre os resultados, destaca-se a diferença entre o método que considera a amostragem simples e o que considera a amostragem complexa.

Deste modo, apesar da existência de diversos trabalhos, esta pesquisa se destaca pelo fato de considerar o efeito do plano amostral, além de adotar o critério de correção do viés de seletividade amostral. Ademais, este trabalho pode ser visto como uma medida dos retornos salariais à educação no ano de 2009, de forma a complementar os demais existentes.

Além desta introdução, o estudo possui mais três itens. O segundo apresenta o método de análise e, no terceiro, são descritos os resultados e a discussão. Por fim, apresentam-se as conclusões.

2. METODOLOGIA

Este item apresenta o método utilizado na análise. Primeiramente, o modelo de seletividade amostral é abordado, seguido pela apresentação do modelo econométrico estimado. Por fim, discorre-se brevemente acerca da fonte de dados e do tratamento neles aplicado.

³ Uma breve discussão acerca de amostragem complexa é apresentada na seção seguinte.

2.1. MODELO ANALÍTICO

De acordo com Garen (1984), os modelos de seletividade amostral partem do pressuposto de que a escolha não é uma variável exógena, mas uma ação de otimização. No presente caso, a escolha entre participar ou não do mercado de trabalho não pode ser analisada como uma decisão aleatória, visto que qualquer uma das decisões é tomada em um processo de maximização dos seus benefícios. Esse autor afirmou que tal processo garante que só sejam observadas escolhas ótimas, independentemente de qual opção tenha sido realizada. Assim, a amostra de indivíduos que participam do mercado de trabalho, foco desta pesquisa, deve ser considerada como não aleatória, o que traz à tona a necessidade de lidar com o viés⁴ de seletividade amostral.

Conseqüentemente, para a estimação dos retornos salariais à educação deve-se levar em conta, além das variáveis que explicam os rendimentos, as que se relacionam com a escolha entre participar ou não do mercado de trabalho. O procedimento de estimação é formado por duas equações, sendo uma que descreve a participação no mercado de trabalho e a equação principal, que representa as respostas salariais à educação e ao conjunto de variáveis relevantes selecionadas. Tomando-se como base a amostra inicial de todos os indivíduos, os que auferem rendimentos e os que nada recebem, observa-se que a distribuição dos retornos salariais possui um “truncamento” no ponto em que a renda passa de zero para valores positivos. Assim, para a equação de participação no mercado de trabalho, são considerados todos os indivíduos e, para a equação principal, consideram-se somente os que auferem algum rendimento.

Nesse caso, o método dos Mínimos Quadrados Ordinários (MQO) torna-se inadequado, pois considera que a amostra dos indivíduos que atuam no mercado de trabalho distribui-se aleatoriamente e estima somente a equação de rendimentos, o que gera tendenciosidade nas estimativas. Alternativamente, o viés de seleção amostral é tratado por meio do procedimento de Heckman (1979), o qual considera que a equação-objeto possui erro de especificação e, portanto, incorpora uma equação adicional para determinar a ocorrência, ou não, do evento e contornar o viés de seletividade.

⁴ É possível, ainda, que haja viés em decorrência da endogeneidade da variável que representa nível de educação, conforme mostraram Sachsida *et al.* (2004). A literatura indica que o viés de endogeneidade é mais relevante que o viés de seletividade amostral.

Com base no trabalho de Kassouf (1997), o método de Heckman é brevemente descrito a seguir. A equação de participação ou de decisão entre estar ou não, no mercado de trabalho é dada por:

$$L_i^* = \gamma'Z_i + u_i, \quad (1)$$

em que L^* é uma variável que designa a participação no mercado de trabalho e Z_i é um vetor com um conjunto de variáveis que explicam a participação no mercado de trabalho. Embora L^* seja latente, a variável L pode ser observada, de modo que se tem:

$$\begin{aligned} L_i &= 1 & \text{se } L_i^* > 0 \\ L_i &= 0 & \text{se } L_i^* \leq 0. \end{aligned}$$

Portanto, a equação (1) consiste na estimativa da probabilidade de participação do mercado de trabalho (L) em função de um conjunto de variáveis (Z). Em geral, o modelo utilizado é o Probit, de modo que L terá valor 1 quando o indivíduo participar do mercado de trabalho, e 0, caso contrário. Já em relação à variável dependente da equação principal, tem-se a definição dos salários conforme se segue:

$$W_i = \beta'R_i + v_i, \quad (2)$$

em que R_i é um vetor de variáveis explicativas do salário.

Dado que somente recebem salários os indivíduos que fazem parte da força de trabalho, a variável W será observada somente nas situações em que L^* for maior que zero. Partindo-se do pressuposto de que u_i e v_i apresentem distribuição normal, com médias zero e desvios-padrão σ_u e σ_v , com correlação igual a ρ , tem-se:

$$\begin{aligned} E(W_i | W_i \text{ observado}) &= E(W_i | L_i^* > 0) = E(W_i | u_i > -\gamma'Z_i), \\ &= \beta'R_i + E(v_i | u_i > -\gamma'Z_i) = \beta'R_i + \rho\sigma_v \lambda_i(\alpha_u), \end{aligned} \quad (3)$$

sendo:

$$\begin{aligned} \lambda_i(\alpha_u) &= \frac{\phi\left(\frac{\gamma'Z_i}{\sigma_u}\right)}{\Phi\left(\frac{\gamma'Z_i}{\sigma_u}\right)} \\ \alpha_u &= \frac{\gamma'Z}{\sigma_u}, \end{aligned}$$

em que ϕ é a função densidade de probabilidade normal padrão e Φ , a função de distribuição normal acumulada. O $\lambda_i(\alpha_u)$ é denominado Razão Inversa de Mills e representa a variável que, segundo Heckman (1979), é utilizada como explicativa na regressão principal para eliminar o viés de seletividade amostral. Logo, a partir da equação (1), a regressão pode ser representada por:

$$W|L>0 = \beta' R_i + \beta_\lambda \lambda_i(\alpha_u) + \varepsilon_i \quad (4)$$

A explicação para a inconsistência da estimação da equação (2) por MQO reside na omissão da variável $\lambda_i(\alpha_u)$, o que resulta no erro de especificação descrito por Heckman (1979). Todavia, ao trabalhar com a razão inversa de Mills como variável explicativa adicional na regressão dos salários, em conjunto com o vetor de variáveis R , a estimativa da equação para os indivíduos que fazem parte do mercado de trabalho resultará em coeficientes consistentes para os β . A razão inversa de Mills é obtida por meio das estimativas geradas para γ , a partir da equação (1).

Entretanto, visto que os dados dessa pesquisa são oriundos de um desenho amostral complexo, o procedimento de estimação em dois estágios não é aplicável. Nesse caso, a estimação das duas equações é feita, simultaneamente, por meio do método Máxima Pseudoverossimilhança (MPV)⁵. De acordo com Nascimento Silva *et. al.* (2002), esse método é comumente empregado para ajustes de modelos paramétricos, quando se consideram o plano amostral e os pesos no processo de inferência com dados de amostras complexas. Observa-se que os ajustes aplicados à amostra complexa não excluem o tratamento do viés de seletividade amostral, o qual é feito pela estimação das equações de participação e rendimentos, porém, não mais em etapas, mas, sim, concomitantemente.

Nesse contexto, Nascimento Silva *et. al.* (2002) destacaram que o plano amostral da PNAD incorpora todos os aspectos que definem um “plano amostral complexo”, os quais são estratificação das unidades de amostragem, conglomeração⁶, probabilidades desiguais de seleção em um ou mais estágios e ajustes dos pesos amostrais para calibração com totais populacionais conhecidos. Por essas razões, esses autores enfatizaram que a realização de trabalhos que se baseiam nos dados da PNAD não pode ser fundamentada nos procedimentos usuais aplicados às amostras com observações independentes e identicamente distribuídas. Adicionalmente, afirmaram que técnicas e sistemas de análise tradicionais podem produzir resultados incorretos para as estimativas pontuais e para os respectivos desvios-padrão e níveis de significância.

⁵ Pessoa e Nascimento Silva (1998) apresentaram uma discussão mais detalhada acerca do método de MPV.

⁶ Seleção da amostra em diversos estágios, com unidades compostas de amostragem.

Segundo Guimarães (2007), as estimativas de indicadores e a geração simultânea de sua variância e do desvio-padrão devem levar em conta o delineamento da amostra. Para isso, devem ser utilizadas, juntamente aos dados da PNAD, duas variáveis que definem o desenho da amostra, quais sejam, o estrato ao qual pertence o domicílio levantado e sua unidade primária de amostragem (UPA ou PSU, em inglês *Primary Sampling Unit*). Nesse sentido, um pré-requisito para se trabalhar com os dados é que os estratos possuam ao menos duas PSU, o que atua como uma restrição. Diante da existência de estratos que contenham apenas uma unidade primária de amostragem, para obtenção de estimativas, tornam-se necessárias a implementação de procedimentos para a identificação desses estratos e a sua agregação a estratos com maior número de observações. Tais procedimentos foram aplicados aos dados utilizados neste trabalho e estão descritos no trabalho de Faria (2006).

Em relação à metodologia utilizada para estimação da variância e dos desvios-padrão, optou-se pela linearização de Taylor⁷.

Adicionalmente, para a mensuração do Efeito do Plano Amostral (EPA), foi utilizado o método MEFF (*misspecification effect*). O MEFF compara a estimativa da variância do parâmetro, obtida quando se considera o plano amostral, com outra estimativa do mesmo modelo, mas que desconsidera as características de peso, conglomerado e estratificação. Assim, quanto mais afastado estiver o valor do MEFF da unidade, mais incorreta será a estimativa feita sem a consideração do plano amostral. Valor de MEFF maior que 1 indica que a desconsideração do desenho amostral leva à subestimação da verdadeira variância, enquanto valores menores que 1 demonstram que a desconsideração do plano amostral superestima a verdadeira variância. De acordo com Leite e Nascimento Silva (2002), o procedimento de cálculo desse método é dado por:

$$MEFF = \frac{V_{verd}(\hat{\beta})}{E_{verd}(v_0)}, \quad (5)$$

em que v_0 é um estimador usual e consistente da variância do estimador sob a hipótese de observações IID; $V_{verd}(\hat{\beta})$ é a variância do estimador sob consideração do desenho amostral complexo; $E_{verd}(v_0)$ é a esperança do estimador usual sob o plano amostral complexo. O procedimento analítico desta pesquisa foi realizado com o auxílio do *software* Stata 11.1.

⁷ Nascimento Silva *et al.* (2002) apresentaram uma descrição do método linearização de Taylor, com base em Pessoa e Nascimento Silva e Korn e Graubard.

2.2. ESPECIFICAÇÃO DO MODELO ECONOMÉTRICO

Considerando-se os procedimentos metodológicos descritos anteriormente, o modelo a ser estimado na presente pesquisa é formado pelas equações (6 e 7), que representam a equação de participação e a equação de rendimentos, respectivamente.

$$L_i = \gamma_0 + \gamma_1 Idade_i + \gamma_2 Idade2_i + \gamma_3 Anosest_i + \gamma_4 Anosest_sit_i + \gamma_5 SitcenD_i + \gamma_6 Conjuge_i + \gamma_7 Parente_i + \gamma_8 Filho_i + \gamma_9 Preto_i + \gamma_{10} Amarelo_i + \gamma_{11} Pardo_i + \gamma_{12} Indio_i + \gamma_{13} N_i + \gamma_{14} NE_i + \gamma_{15} CO_i + \gamma_{16} S_i + \gamma_{17} Rndnsal_pcp_i \quad (6)$$

$$W_i = \beta_0 + \beta_1 Idade_i + \beta_2 Idade2_i + \beta_3 Anosest_i + \beta_4 Anosest_sit_i + \beta_5 SitcenD_i + \beta_6 Conjuge_i + \beta_7 Parente_i + \beta_8 Filho_i + \beta_9 Preto_i + \beta_{10} Amarelo_i + \beta_{11} Pardo_i + \beta_{12} Indio_i + \beta_{13} N_i + \beta_{14} NE_i + \beta_{15} CO_i + \beta_{16} S_i + \beta_{17} Sindicato + \beta_{18} lambda \quad (7)$$

Em que as variáveis dependentes são a probabilidade de o indivíduo estar atuando no mercado de trabalho (L_i) e a remuneração salarial de cada indivíduo (W_i). As variáveis explicativas são apresentadas na Tabela 1, onde também constam os seus efeitos esperados sobre as variáveis dependentes, nas equações de participação e de rendimentos.

Tabela 1 – Descrição das variáveis utilizadas e seus efeitos esperados sobre as equações, participação e rendimentos

Variável	Descrição das variáveis	Sinal esperado	
		Participação	Rendimentos
Idade	Idade do indivíduo	+	+
Idade ²	Idade ao quadrado	-	-
Anos_est	Nº de anos de estudos	+	+
Anosest_sit	Nº de anos de estudos x situação censitária	+	+
SitcenD	= 1 se o indivíduo mora na zona urbana	+	+
Conjuge	= 1 se a condição do indivíduo na família é cônjuge	-	-
Parente	= 1 se a condição do indivíduo na família é parente	-	-
Filho	= 1 se a condição do indivíduo na família é filho	-	-
Preto	= 1 se o indivíduo é preto	+	-
Amarelo	= 1 se o indivíduo é amarelo	+	-
Pardo	= 1 se o indivíduo é pardo	+	-
Índio	= 1 se o indivíduo é índio	-	-
Norte	= 1 se o indivíduo mora na região Norte	+	-
Nordeste	= 1 se o indivíduo mora na região Nordeste	+	-
Centro-Oeste	= 1 se o indivíduo mora na região Centro-Oeste	+	+ ou -
Sul	= 1 se o indivíduo mora na região Sul	+	+ ou -
Sindicato	= 1 se o indivíduo é filiado a algum sindicato	n.a.	+
rendnsal_percap	Valor da renda não salarial individual	-	n.a.

Nota: (n.a.) Não aplicável, a variável não faz parte da equação.

Fonte: Elaboração própria.

As variáveis foram selecionadas com base nos trabalhos acerca do capital humano, como aqueles mencionados na revisão de literatura. A ênfase da presente análise, contudo, é sobre o efeito da escolaridade, ou seja, sobre os coeficientes estimados para as variáveis *Anosest* e *Anosest_sit*. Os demais termos foram incluídos para controlar os seus efeitos sobre a participação do indivíduo no mercado de trabalho e sobre a sua remuneração, visando à correta especificação do modelo.

Observa-se que a inclusão da interação entre anos de estudo e situação censitária é uma forma de verificar se existe relação entre os efeitos dessas duas variáveis sobre a renda. Assim, pode-se verificar se as diferenças de salário geradas por diferenças de escolaridade são distintas nos meios urbano e rural. É possível que o efeito de um ano adicional de estudo seja diferente nos meios rural e urbano. Dessa forma, os coeficientes das variáveis *Anosest*, *Anosest_sit* e *SitcenD* devem ser analisados em conjunto.

A definição dos efeitos esperados das variáveis é oriunda dos resultados de trabalhos anteriores, quais sejam Kassouf (1994), Kassouf (1997), e Hoffman e Kassouf (2005). Especificamente em relação às diferenças nos retornos salariais à educação entre as zonas rural e urbana, espera-se que os indivíduos residentes no meio urbano tenham maior acesso ao mercado de trabalho e auferam retornos salariais mais elevados, conforme resultados apresentados por Kassouf (1997).

No que concerne às diferenças regionais, estatísticas sobre o nível de desemprego nas regiões, calculadas a partir dos dados da PNAD (2008), indicaram que a região Sudeste foi a que apresentou maiores níveis de desemprego, de modo que se espera que os residentes nas demais regiões do país tenham maior probabilidade de participar do mercado de trabalho. Já em relação aos rendimentos, espera-se que nas regiões Norte e Nordeste ocorram retornos salariais à educação menores que no Sudeste, enquanto os retornos nas regiões Sul e Centro-Oeste podem ser maiores ou menores que os da região Sudeste.

Ademais, espera-se que o rendimento oriundo de outras fontes diferentes do salário, representado pela variável *rendnsal_percap*, apresente correlação negativa com a probabilidade de que o indivíduo esteja trabalhando. Reis e Camargo (2007) destacaram que a literatura econômica fornece diversas evidências de que aumentos na renda não proveniente do trabalho oferecem incentivos para os trabalhadores reduzirem a participação na força de trabalho ou as horas trabalhadas. Hurd e Boskin (1984), por exemplo, mostraram evidências empíricas de que aumentos nos benefícios sociais da aposentadoria estão relacionados com queda na participação dos idosos na força de trabalho.

2.3. FONTE E TRATAMENTO DOS DADOS

Os dados utilizados na realização deste estudo foram retirados da PNAD, realizada pelo IBGE, no ano de 2009. A partir desse conjunto de dados, selecionou-se uma amostra de indivíduos do sexo masculino, com idade entre 20 e 65 anos. Essa faixa etária foi escolhida com o intuito de analisar indivíduos que participem do mercado de trabalho e, ainda, que possuam algum período de experiência e nível de escolaridade definido.

Com vistas a preservar o tamanho da amostra e suas informações, não foi aplicado filtro para eliminar os indivíduos que continham *observações faltantes*, uma vez que não houve problema para gerar o modelo com tais lacunas nos dados. A amostra utilizada em todo o país, portanto, consistiu de 113.509 homens, dentre os quais 96.122 estavam trabalhando; 5.386, não; o restante (12.001) não apresentava nenhuma das duas opções e, conseqüentemente, foram desconsiderados no estudo. Em termos percentuais, do total de respostas válidas à pergunta, 94,69% dos homens eram economicamente ativos, estando ou não ocupados.

3. RESULTADOS E DISCUSSÃO

Os resultados estão divididos em três subitens. No primeiro, será apresentada a análise descritiva dos dados. No segundo e no terceiro, serão apresentadas as estimativas da equação de participação no mercado de trabalho e da equação de rendimentos, respectivamente.

3.1. ANÁLISE DESCRITIVA DOS DADOS

A Tabela 2 apresenta os valores médios e os desvios-padrão das variáveis, obtidos para a amostra utilizada nesta pesquisa. Também são apresentadas as médias para o meio urbano e o rural. Todas as médias são incondicionais, ou seja, os valores consideram todos os indivíduos, os que participam do mercado de trabalho e os que não participam.

A variável que mede a renda foi formada com base na remuneração mensal recebida por cada indivíduo como pagamento em dinheiro pelo seu trabalho principal. O número de horas trabalhadas por semana foi multiplicado por quatro, resultando no tempo de serviço mensal. Assim, a renda mensal foi dividida pelo tempo de trabalho obtido por mês, gerando-se um rendimento horário. Em média, o salário por hora trabalhada dos homens brasileiros (*sal_hr*), em 2009, foi de R\$ 7,65. Essa renda média por hora foi de R\$ 8,50 no meio urbano e de R\$ 3,48, no rural.

Tabela 2 – Médias e desvios-padrão das variáveis utilizadas

Variável	Amostra		Meio Urbano	Meio Rural
	Média ou Proporção	Desvio-padrão	Média ou Proporção	
sal_hr	7,649	0,127	8,505	3,481
cupD	0,970	0,001	0,966	0,988
Lnrenda	1,136	0,020	1,482	-0,554
Idade	38,252	0,046	37,992	39,494
Anos_est	8,847	0,036	9,528	5,559
Conjuge	0,095	0,002	0,106	0,038
Parente	0,038	0,001	0,040	0,027
Filho	0,199	0,002	0,201	0,188
Preto	0,079	0,001	0,082	0,063
Amarelo	0,004	0,000	0,005	0,001
Pardo	0,434	0,003	0,412	0,546
Índio	0,002	0,000	0,002	0,002
Norte	0,076	0,003	0,070	0,107
Nordeste	0,261	0,004	0,224	0,446
Centro-oeste	0,076	0,002	0,079	0,062
Sul	0,154	0,003	0,151	0,169
Sindicato	0,196	0,002	0,189	0,235
Rendnsal_percap	26,428	0,868	27,701	20,114

Notas: (*) Diferentemente das variáveis contínuas, no caso das variáveis *dummy*, apresentam-se as proporções de ocorrência de cada categoria.

(**) Calculado por linearização de Taylor.

Fonte: Elaboração própria com base em resultados da pesquisa.

Dada a amostra formada por indivíduos com idade entre 20 e 65 anos, verifica-se que a média de idade é de, aproximadamente, 38 anos. Aproximadamente 83% dos indivíduos residiam na zona urbana, sendo observadas diferenças de destaque entre as características individuais de acordo com a situação censitária do domicílio. Apesar de as idades médias serem próximas, 39,49 e 37,99 anos, na zona rural e urbana, respectivamente, observa-se que, em média, os homens residentes na zona rural estudaram por cerca de cinco anos e meio, enquanto os moradores do meio urbano, por mais de nove anos. Tomando-se a média geral, constata-se que os homens brasileiros possuíam aproximadamente 8,8 anos de estudo. Quanto à participação da força de trabalho, verifica-se que 98,8% e 96,6% dos homens estavam ocupados nos meios rural e urbano, respectivamente.

Em virtude das expressivas distinções constatadas entre o meio urbano e rural, em relação à escolaridade, bem como das diferenças existentes no mercado de trabalho e na qualidade do sistema de ensino, optou-se por comparar o efeito dos anos de estudo entre o meio urbano e o meio rural. Assim, uma *dummy* de situação censitária foi

incluída de forma interativa com o número de anos de estudos em ambas as equações. Essa *dummy* terá valor 1, quando se tratar de zona urbana, e valor 0, zona rural.

Como forma de controlar algumas características específicas dos indivíduos, foram inseridas outras variáveis binárias no modelo, além da situação censitária. Conforme pode ser verificado na Tabela 2, para a região, cuja categoria base é o Sudeste, observa-se que a maior concentração geográfica ocorreu na região Sudeste, seguida pelo Nordeste, Sul, Norte e Centro-Oeste, respectivamente. No que diz respeito à cor ou raça, cuja categoria base é branca, a maior parte dos homens era branca, seguindo-se pelos indivíduos pardos, pretos, amarelos e índios, respectivamente. Finalmente, em relação à condição do indivíduo em seu domicílio, a maioria dos homens era definida na categoria base, chefe de família, seguindo-se pelos filhos, cônjuges e parentes.

Na equação de participação do mercado de trabalho, utilizou-se uma variável para controlar o rendimento oriundo de outras fontes diferentes do salário. Portanto, a variável *Rendnsal_percap* representa o somatório das seguintes fontes de renda: aposentadoria; pensão; aluguel; abono; poupança e doações; calculada para cada indivíduo. O valor agregado dessas fontes de renda para cada indivíduo era, em média, de R\$ 26,43.

Por sua vez, na regressão de rendimentos considerou-se o efeito que a filiação a determinado sindicato pode apresentar sobre os rendimentos do trabalhador. Aproximadamente 19,65% dos indivíduos eram filiados a alguma organização sindical.

3.2. PARTICIPAÇÃO NO MERCADO DE TRABALHO

A estimativa da equação de decisão apresentou os coeficientes descritos na Tabela 3, na qual estão, também, os desvios-padrão obtidos por linearização de Taylor e a estatística MEFF. A variável dependente é binária, terá valor 1 quando o indivíduo participar do mercado de trabalho, e 0, caso contrário.

Os termos referentes à idade, ambos significativos, demonstram que, à medida que a idade se eleva, há tendência inicial de aumento na probabilidade de que o indivíduo esteja ocupado, até uma idade em que essa probabilidade é máxima. A partir desse ponto, o envelhecimento passa a resultar em diminuição na probabilidade de que o homem faça parte do mercado de trabalho.

Quanto à escolaridade, o efeito de anos adicionais de estudo sobre a probabilidade de que o indivíduo esteja ocupado é positivo somente para a zona urbana. Todavia, no meio rural, o efeito de anos adicionais de estudo sobre a probabilidade de que os homens consigam fazer parte da força de trabalho é contrário, de modo que um maior número de anos dedicados aos estudos afasta o indivíduo da força de trabalho. Uma possível explicação para essa constatação reside na diferença de exigências acerca da

qualificação do trabalhador. No meio rural, há maior número de atividades que requerem baixos níveis de qualificação, inclusive de funções que não requerem nenhum nível mínimo de estudo. Adicionalmente, observa-se, no meio urbano, maior nível de exigência de qualificação, treinamento e escolaridade. Trata-se de fatores que podem levar à exclusão de trabalhadores mais qualificados na zona rural, bem como à redução dos trabalhadores menos qualificados na zona urbana.

Alternativamente, é possível que os indivíduos mais qualificados no meio rural sejam mais seletivos em relação aos empregos. O nível médio de escolaridade é menor no meio rural. Dessa forma, os trabalhadores mais bem qualificados são mais escassos e, portanto, são mais seletivos e escolhem empregos melhores.

Já em relação à localização do domicílio, nota-se que os residentes em área urbana apresentam menor probabilidade de estar ocupado, resultado coerente com as expectativas apresentadas na análise descritiva dos dados. A região Norte não apresentou coeficiente significativo; com isso, não há diferença relevante em relação aos moradores da região Sudeste (categoria excluída). Os moradores da região Sul e Centro-Oeste têm probabilidade maior de estarem trabalhando do que os do Sudeste. Já os que residem no Nordeste têm probabilidade menor do que os da região Sudeste.

Tabela 3 – Coeficientes estimados para a equação de participação no mercado de trabalho e efeito do Plano Amostral (MEFF) – 2008

Variável	Coefficiente	Desvio-padrão	Significância	MEFF
Constante	1,6131	0,1103	0,0000	–
Idade	0,0427	0,0046	0,0000	1,1932
Idade ²	-0,0004	0,0001	0,0000	1,1720
anos_est_sit	0,0463	0,0071	0,0000	1,2113
anos_est	-0,0336	0,0070	0,0000	1,2202
sitcenD	-0,9162	0,0636	0,0000	1,3625
Conjuge	-0,1266	0,0287	0,0000	1,4258
Parente	-0,4027	0,0343	0,0000	1,2865
Filho	-0,4879	0,0206	0,0000	1,3379
Preto	-0,0895	0,0300	0,0030	1,3924
Amarelo	-0,0213	0,1326	0,8720	1,2370
Pardo	-0,0566	0,0192	0,0030	1,4589
Índio	-0,0132	0,1344	0,9220	1,0507
Norte	-0,0259	0,0270	0,3370	1,2248
Nordeste	-0,0757	0,0226	0,0010	1,6116
Centro-Oeste	0,0869	0,0285	0,0020	1,3303
Sul	0,1235	0,0272	0,0000	1,3890
rendnsal_percap	-0,0001	0,0000	0,0070	1,3944

Fonte: Elaboração própria com base nos resultados da pesquisa.

Em relação à condição do homem em seu domicílio, a categoria base dos chefes de família é aquela em que os indivíduos têm maior probabilidade de estarem ocupados. No que diz respeito à discriminação do mercado de trabalho com base na raça, comparando-se com os brancos, os homens que se declararam pretos e pardos apresentavam as menores chances de estarem empregados. A probabilidade de os indivíduos que se declararam amarelos ou índios estarem no mercado de trabalho não era estatisticamente distinta dessa probabilidade para os homens brancos.

A variável representativa da renda proveniente de fontes não salariais mostrou-se significativa e com o sinal esperado, para o efeito sobre a ocupação do indivíduo. Consta-se que, quanto maiores esses rendimentos, menores são as chances de que o homem participe do mercado de trabalho.

Por fim, em todos os casos a estatística MEFF calculada é maior que a unidade. Isso significa que a descon sideração do plano amostral (pesos, estratificação e conglomeração) levaria à subestimação da verdadeira variância. A título de demonstração, a estimação desse modelo descon siderando o plano amostral está incluída no anexo. Podem ser observadas as diferenças nos betas estimados e também nos desvios-padrão e nas significâncias, conseqüentemente. A comparação com a estimativa descon siderando o plano amostral, apresentada em anexo, mostra que não houve divergência de sinais entre os coeficientes gerados com as duas formas de estimação. Contudo, houve diferença de magnitude. Como exemplo, o coeficiente de anos de estudos na equação de rendimentos calculado com base na amostragem complexa foi 12% menor do que aquele obtido com amostragem simples.

3.3. EQUAÇÃO DE RENDIMENTOS

Os coeficientes estimados para a regressão de rendimentos estão apresentados na Tabela 4. Inicialmente, ratifica-se a necessidade do processo de seleção amostral, visto que a variável *lambda* (equação 2), que é o inverso da razão de Mills, é significativa. Portanto, a inclusão dessa variável na equação de rendimentos é necessária para a correção do viés de seletividade amostral. Adicionalmente, o fato de o *lambda* possuir sinal positivo indica que os fatores não observados que aumentam a probabilidade de participação do mercado de trabalho também aumentam os rendimentos salariais.

Nesse caso é possível concluir que o uso de métodos tradicionais de estimação do logaritmo das taxas salariais horárias para toda a amostra, utilizando somente a amostra de homens que trabalham, resultaria, em valor absoluto, em uma estimativa negativamente viesada do verdadeiro efeito.

Tabela 4 – Coeficientes estimados para a equação de rendimentos e efeito do Plano Amostral (MEFF)

Variável	Coeficiente	Desvio-padrão	Significância	MEFF
Constante	-3,2259	0,1711	0,0000	–
Idade	0,1442	0,0070	0,0000	2,3737
Idade ²	-0,0017	0,0001	0,0000	2,5468
anos_est_sit	0,0271	0,0132	0,0400	7,0293
anos_est	0,0972	0,0130	0,0000	7,7271
sitcenD	1,1991	0,1161	0,0000	9,6040
Conjuge	-0,1708	0,0275	0,0000	1,2233
Parente	-0,5000	0,0551	0,0000	1,9391
Filho	-0,8938	0,0377	0,0000	2,9033
Preto	-0,0949	0,0300	0,0020	1,1299
Amarelo	0,0598	0,1207	0,6210	0,8970
Pardo	-0,1349	0,0241	0,0000	2,0172
Índio	-0,4770	0,3033	0,1160	4,5358
Norte	-0,1684	0,0541	0,0020	4,3145
Nordeste	-0,7075	0,0445	0,0000	5,0854
Centro-Oeste	0,0508	0,0370	0,1700	2,1188
Sul	-0,1084	0,0294	0,0000	1,6265
Sindicato	0,1384	0,0263	0,0000	1,9598
<i>Lambda</i>	0,1925	0,0224	0,0000	–

Fonte: Elaboração própria com base nos resultados da pesquisa.

Em relação ao efeito da idade, a partir dos 20 anos, nota-se tendência de aumento na remuneração do indivíduo até um ponto máximo, a partir do qual o rendimento apresenta trajetória descendente, comportamento coerente com o esperado inicialmente.

No que diz respeito à localização regional, o nível salarial observado na região Centro-Oeste não foi estatisticamente diferente do nível salarial da região Sudeste. As regiões Nordeste, Sul e Norte apresentaram níveis salariais menores que os da região Sudeste. Ademais, conforme esperado *a priori*, o fato de o trabalhador ser filiado a determinado sindicato é um fator que contribui, positivamente, para elevar o seu nível salarial.

A equação de participação mostrou que os residentes na zona rural possuíam maior probabilidade de estar no mercado de trabalho. Contudo, as estimativas da equação de rendimentos mostram que a remuneração dos trabalhadores rurais é estatisticamente inferior à dos moradores das zonas urbanas. Os sinais positivos do coeficiente da interação entre anos de estudos e situação censitária e do coeficiente de situação censitária mostram que os empregados do meio urbano recebem remuneração maior do que a dos empregados do meio rural, bem como também recebem mais por cada ano

adicional de estudo. Nesse ponto, é interessante destacar que a expressiva superioridade do tempo médio de escolaridade na zona urbana é um possível aspecto relevante para explicar a distinção salarial.

Nesse contexto, em relação ao efeito da escolaridade, foco do presente estudo, observa-se que há efeito positivo de anos adicionais de estudo sobre os rendimentos dos indivíduos. Para os trabalhadores do meio urbano, em média, a cada ano a mais de estudo tem-se um aumento⁸ aproximado de 31,83% nos salários recebidos por hora. Por sua vez, o trabalhador da zona rural que possui um ano a mais de estudo tem um salário horário superior em cerca de 9,72%. Esses efeitos são condicionais, ou seja, representam o ganho salarial decorrente de um ano adicional de estudo para os indivíduos que estejam atuando no mercado de trabalho.

Algumas considerações relevantes acerca da diferença verificada entre os efeitos devem ser realizadas. Primeiramente, a qualidade da educação não está sendo controlada na regressão. É natural que indivíduos com o mesmo tempo de estudo tenham adquirido níveis de conhecimento bastante diferentes, em função de diferenças na qualidade do ensino. Há diversos fatores, como qualificação e interesse dos professores, qualidade do material didático, infraestrutura e ambiente escolar e domiciliar, entre outros, que interferem na resposta dos indivíduos ao ensino. Na maioria das vezes, esses são aspectos que representam dificuldades e limitações à melhoria da qualidade educacional na zona rural.

Ademais, no meio rural é comum que as crianças dividam o seu tempo entre estudo e trabalho, o que prejudica o seu rendimento e reflete na sua qualificação profissional. Portanto, o diferencial no impacto dos anos de estudo nas taxas salariais reflete, em parte, uma significativa diferença estrutural.

Há diferenças notórias, também, no mercado de trabalho e nos salários, entre a zona urbana e a rural. Em geral, a mão de obra desqualificada tem a remuneração definida com base no salário mínimo e não há variações expressivas em torno desse valor. Dessa forma, os trabalhadores do meio rural têm o salário mais rígido. Já no meio urbano há maior variedade e complexidade de atividades e, portanto, os salários variam mais, além de atingir valores mais elevados. Ademais, os custos de vida inferiores na zona rural permitem que os trabalhadores sobrevivam com menores níveis de renda.

Observa-se que as características inerentes aos trabalhadores, como habilidade e inteligência, não estão sendo consideradas. Essas podem ser úteis para explicar parte do

⁸ O efeito de um ano adicional de estudos no meio urbano foi obtido pelo comando `nlcom`, no Stata, que faz combinação não linear de estimadores.

diferencial no retorno salarial à educação. Contudo, a importância do viés de variável omitida referente à habilidade dos indivíduos foi avaliada por Sachsida *et al.* (2004). A partir da estimação de um pseudopainel montado com as edições da PNAD relativas aos anos de 1992 a 1999, esses autores destacaram que o efeito de variável omitida sobre o retorno em escolaridade é nulo ou pouco perceptível.

Dados esses resultados, observa-se que os indivíduos residentes no meio urbano possuem maiores incentivos à realização de investimentos que envolvem os estudos. No meio rural, não foi identificada resposta positiva à escolaridade em termos de probabilidade de participar do mercado de trabalho, além de a resposta salarial aos estudos ter sido expressivamente inferior à do meio urbano. Com isso, os homens do meio rural com maior nível de escolaridade não têm estímulo para permanecerem nesse setor, sendo levados a procurar empregos nas cidades, em busca de maior compensação pelas suas qualificações adquiridas.

As taxas de retorno de 9,72% e 31,83% são úteis em uma comparação com possíveis investimentos alternativos, como forma de direcionar os recursos para as opções mais eficientes. Ademais, deve-se atentar para o fato de que não estão sendo mensurados efeitos sociais que advêm da elevação da escolaridade, o que é capaz de elevar ainda mais a resposta positiva à educação.

Os resultados são coerentes com os encontrados por Kassouf (1997), apesar da diferença na consideração de amostra complexa. Essa autora identificou superioridade na resposta salarial à educação no meio urbano, tanto para homens quanto para mulheres. Em termos de valores, contudo, os retornos calculados na presente pesquisa encontram-se em patamares inferiores aos encontrados nos trabalhos mencionados na revisão de literatura. Contudo, os coeficientes estimados não podem ser comparados, em função da utilização de diferentes variáveis explicativas e da consideração da amostra complexa.

Nesse contexto torna-se relevante destacar as diferenças entre os resultados desta pesquisa com os que se formariam a partir da desconsideração do plano amostral complexo. A partir dos valores estimados do efeito plano amostral (MEFF), constata-se que, para quase todas as variáveis, caso não fosse considerado o desenho complexo da amostra, as variâncias estariam subestimadas. Portanto, a realização de testes estatísticos poderia levar à avaliação de determinado coeficiente como significativo, equivocadamente. Os maiores equívocos envolveriam a *dummy* de situação censitária, a variável interativa de anos de estudo com situação censitária, a *dummy* para a região Nordeste e o coeficiente da escolaridade. Com isso, ratifica-se que a incorporação do desenho amostral complexo é necessária para a correta avaliação dos efeitos da escolaridade sobre os rendimentos.

4. CONCLUSÃO

Primeiramente, como aspecto relevante desta pesquisa destaca-se a importância de consideração do desenho amostral complexo para gerar estimativas no ponto e erros-padrões calculados corretamente. A expressividade dos valores da estatística MEF, para grande parte das variáveis, foi um resultado que chamou atenção, o que indica um caminho que, necessariamente, deve ser seguido pelos novos estudos baseados no banco de dados da PNAD. A comparação das estimativas com e sem consideração da amostragem complexa ratificou essa importância.

Especificamente em relação aos resultados, verificou-se que as questões de cor, condição do indivíduo na família e localização geográfica interferem na formação dos padrões salariais. Já em relação à escolaridade, os resultados corroboraram os de outras pesquisas, como a de Kassouf (1994), Loureiro e Galvão (2001), nas quais os retornos salariais dos trabalhadores do meio urbano foram superiores aos do meio rural. Os retornos salariais a anos adicionais de estudos também foram superiores no meio urbano.

A diferença entre os retornos salariais mostra que há uma lacuna entre o ensino no meio rural e no meio urbano. Indivíduos com as mesmas características de idade, cor, região e escolaridade possuem retornos salariais a um ano adicional de estudo diferentes, pelo fato de morarem no meio urbano ou rural. A baixa qualificação da mão de obra rural resulta em rigidez salarial, dado que os valores das remunerações situam-se em torno do salário mínimo. Entretanto, a qualificação dos profissionais do meio urbano resulta em salários significativamente superiores e variáveis. Isso mostra que as produtividades dos dois são distintas e, como a produtividade reflete os conhecimentos e a qualificação, constata-se uma diferença entre a qualidade do ensino.

Dado retorno salarial à educação inferior no meio rural, possivelmente, a tendência de os trabalhadores mais qualificados migrarem do campo para as cidades permanece, e os estímulos aos estudos no meio rural tendem a se manter inferiores aos do meio urbano. Portanto, essa diferença somente se alteraria com alguma intervenção, a qual poderia ser por meio de melhoria da qualidade do ensino no meio rural, o que elevaria a produtividade dos trabalhadores e, conseqüentemente, sua remuneração.

Diante do contexto de acelerada evolução tecnológica e da tendência à produção de bens e serviços de maior valor agregado, as empresas tendem a buscar profissionais mais qualificados, em detrimento dos de limitados conhecimentos. Em geral, são requeridos conhecimentos específicos para que o trabalhador tenha a capacidade requerida para lidar com essas novas tecnologias, o que torna a educação ainda mais importante.

Por fim, observa-se que aumentos na remuneração do trabalhador rural, gerados por melhores condições do ensino, ajudariam a reduzir a desigualdade de renda entre o meio urbano e o rural e, conseqüentemente, fortaleceria o processo de queda da desigualdade da distribuição de renda no Brasil.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ANGRIST, J.D.; PISCHKE, J-S. *Mostly harmless econometric: An empiricist's companion*. Princeton, New Jersey: Princeton University Press, 2009.
- BECKER, G. S. Investment in human capital: A theoretical analysis. *Journal of Political Economy*, v. 70, n. 5, p. 9-49, jan., 1962.
- CARVALHO, A. P.; NÉRI, M. C.; NASCIMENTO SILVA, D. B. Diferenciais de salários por raça e gênero no Brasil: aplicação dos procedimentos de Oaxaca e Heckman em pesquisas amostrais complexas. In: Encontro Nacional de Estudos Populacionais, ABEP, 15, Caxambu, setembro, 2006.
- FARIA, A. L. C. *Aplicação do teste de elegibilidade multidimensional na definição do público alvo beneficiário de políticas públicas*. Dissertação de Mestrado, Escola Nacional de Ciências Estatísticas, Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro, 2006.
- GAREN, J. The Returns to schooling: a selectivity bias approach with a continuous choice variable. *Econometrica*, v. 52, n. 5, p. 1.199-1.218, set., 1984.
- GUIMARÃES, P. W. *Variação de renda familiar, desigualdade e pobreza no Brasil*. Tese de Doutorado, Departamento de Economia Aplicada, Universidade Federal de Viçosa, Viçosa, 2007.
- HECKMAN, J. J. Sample selection bias as a specification error. *Econometrica*, v. 45, n. 1, p. 153-161, jan., 1979.
- HOFFMAN, R.; KASSOUF, A. L. Deriving conditional and unconditional marginal effects in log earnings equation estimated by Heckman's procedure. *Applied Economics*, v. 37, p. 1.303-1.311, 2005.
- HURD, M. D.; BOSKIN, M.J. The effect of social security on retirement in the early 1970s. *The Quarterly Journal of Economics*, v. 99, n. 4, p. 767-790, 1984.
- KASSOUF, A. L. Retornos à escolaridade e ao treinamento nos setores urbano e rural do Brasil. *Revista de Economia e Sociologia Rural*, v. 35, n. 2, p. 59-76, 1997.
- KASSOUF, A. L. The wage rate estimation using the Heckman procedure. *Revista de Econometria*, Rio de Janeiro, v. 14, n. 1, p. 89-107, abr./set., 1994.
- LAM, D.; LEVISON, D. Idade, experiência, escolaridade e diferenciais de renda: Estados Unidos e Brasil. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 20, n. 2, p. 219-256, ago., 1990.
- LANGONI, C. G. Investimento em educação no Brasil: um comentário. *Pesquisa e Planejamento*, Rio de Janeiro, v. 1, n. 2, p. 381-392, dez., 1971.

- LEAL, C. I. S.; WERLANG, S. R. C. Retornos em educação no Brasil: 1976/1989. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 21, n. 3, p. 559-574, dez., 1991.
- LEITE, P. G.; SILVA, D. B. N. Análise da situação ocupacional de crianças e adolescentes nas regiões Sudeste e Nordeste do Brasil utilizando informações da PNAD 1999. *Revista Brasileira de Estudos de População*, v. 19, n. 2, jul./dez., 2002.
- LOUREIRO, P. R. A.; GALRÃO, F. G. Discriminação no mercado de trabalho: uma análise dos setores rural e urbano no Brasil. *Economia Aplicada*, v. 5, n. 3, p. 519-545, 2001.
- MINCER, J. Investment in human capital and personal income distribution. *Journal of Political Economy*, v. 66, n. 4, p. 281-302, jan., 1958.
- SILVA, P. L. N.; PESSOA, D. G. C.; LILA, M. F. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. *Ciência e Saúde Coletiva*, v. 7, n. 4, p. 659-670, 2002.
- PESSOA, D. G. C.; SILVA, P. L. N. *Análise de dados amostrais complexos*. São Paulo: Associação Brasileira de Estatística, 1998.
- REIS, M. C.; CAMARGO, J. M. Impactos de aposentadorias e pensões sobre a educação e a participação dos jovens na força de trabalho. *Pesquisa e Planejamento Econômico*, v. 37, n. 2, p. 221-246, 2007.
- RESENDE, M.; WYLLIE, R. Retornos para educação no Brasil: Evidências empíricas adicionais. *Economia Aplicada*, v. 10, n. 3, p. 349-365, jul./set., 2006.
- ROSENZWEIG, M. R. Why Are There Returns to Schooling? *The American Economic Review*, v. 85, n. 2, p. 153-158, 1995.
- SACHSIDA, A; LOUREIRO, P. R. A; MENDONÇA, M. J. C. Um estudo sobre retorno em escolaridade no Brasil. *Revista Brasileira de Economia*, v. 58, n. 2, p. 249-265, abr./jun., 2004.

ANEXO**RESULTADO DO MODELO ESTIMADO DESCONSIDERANDO O PLANO AMOSTRAL****Tabela A1 – Coeficientes estimados para a equação de participação no mercado de trabalho desconsiderando o Plano Amostral**

Variável	Coefficiente	Desvio-padrão
Constante	1,5566	0,1023
Idade	0,0505	0,0044
Idade ²	-0,0005	0,0001
anos_est	-0,0196	0,0069
anos_est_sit	0,0313	0,0071
sitcenD	-0,8791	0,0570
Conjuge	-0,0689	0,0259
Parente	-0,3830	0,0325
Filho	-0,5068	0,0190
Preto	-0,1518	0,0263
Amarelo	-0,0108	0,1114
Pardo	-0,0487	0,0169
Índio	-0,1506	0,1248
Norte	0,0276	0,0261
Nordeste	-0,1556	0,0188
Centro-Oeste	0,0084	0,0260
Sul	0,0913	0,0252
rendnsal_percap	-0,0001	0,00003

Fonte: Elaboração própria com base nos resultados da pesquisa.

Tabela A2 – Coeficientes estimados para a equação de rendimentos desconsiderando o plano amostral

Variável	Coeficiente	Desvio padrão
Constante	-3,4082	0,0997
Idade	0,1399	0,0046
Idade ²	-0,0017	0,0001
anos_est	0,1104	0,0049
anos_est_sit	0,0160	0,0052
sitcenD	1,3656	0,0377
Conjuge	-0,1404	0,0250
Parente	-0,5406	0,0406
Filho	-0,9770	0,0227
Preto	-0,1045	0,0289
Pardo	-0,1284	0,0173
Amarelo	0,0782	0,1103
Índio	-0,5696	0,1307
Norte	-0,2218	0,0259
Nordeste	-0,6201	0,0202
Centro-Oeste	0,1154	0,0269
Sul	-0,0421	0,0238
Sindicato	0,2022	0,0191
<i>Lambda</i>	0,1939	0,0316

Fonte: Elaboração própria com base nos resultados da pesquisa.