

# Evolução dos retornos da escolaridade no Brasil\*

Marcos J Ribeiro<sup>†</sup>

FEARP/USP

Fernando Barros Jr<sup>‡</sup>

FEARP/USP

Luciano Nakabashi<sup>§</sup>

FEARP/USP

29 de julho de 2022

## Resumo

Neste artigo utilizamos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) e estimamos os retornos da escolaridade do Brasil entre 1995 e 2015. Para realizar as estimações e corrigir possíveis vieses nas estimativas utilizamos os métodos de [Garen \(1984\)](#) e [Heckman \(1979\)](#). Nossos resultados são robustos e apontam queda nos retornos neste período. Mostramos que essa queda pode ser resultado da expansão da mão de obra qualificada e da qualidade do sistema educacional brasileiro. Também fornecemos evidências de que os trabalhadores que possuem maiores níveis educacionais possuem maiores retornos e a principal causa desse resultado são as características socioeconômicas desses trabalhadores. Além disso, mostramos em um exercício contrafactual que caso os retornos da escolaridade fossem constantes e iguais aos de 1995 o salário hora médio seria 14,95% maior em 2015.

**Palavras Chave:** Retornos da escolaridade, educação, salários.

Área 13 - Economia do Trabalho.

**JEL:** J01, J21, J23, I26.

## Abstract

In this article we use data from the National Household Sample Survey (PNAD) and estimate the returns to schooling in Brazil between 1995 and 2015. To perform the estimates and correct possible biases in the estimates, we used the methods of [Garen \(1984\)](#) and [Heckman \(1979\)](#). Our results are robust and point to a drop in returns in this period. We show that this drop may be the result of the expansion of skilled labor and the quality of the Brazilian education system. We also provide evidence that workers who have higher educational levels have higher returns and the main cause of this result is the socioeconomic characteristics of these workers. In addition, we show in a counterfactual exercise that if the returns to schooling were constant and equal to 1995, the average hourly wage would be 14.95% higher in 2015.

**Keywords:** Returns from schooling, education, wages.

---

\*O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

<sup>†</sup>Endereço: Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, Avenida Bandeirantes, 3900 - Vila Monte Alegre, Ribeirão Preto - SP, 14040-905. Email: [mjribeiro@usp.br](mailto:mjribeiro@usp.br).

<sup>‡</sup>Email: [fabarrosjr@usp.br](mailto:fabarrosjr@usp.br).

<sup>§</sup>Email: [luciano.nakabashi@gmail.com](mailto:luciano.nakabashi@gmail.com).

# 1 Introdução

A teoria econômica preconiza que maiores níveis de escolaridade é elemento chave no processo de desenvolvimento e crescimento econômico, facilita a criação e difusão de tecnologias, aumenta a produtividade e também os salários ([Acemoglu e Angrist, 2000](#); [Wilson e Briscoe, 2004](#); [Perez-Silva e Partridge, 2020](#)). No Brasil, notamos uma aceleração dos investimentos em educação, sendo impulsionado pela Constituição de 1988. Por exemplo, dados do [Banco Mundial](#) mostram que em 1995 os investimentos públicos em educação correspondiam a 4,57% do PIB e passaram para 6,24%, em 2015. Além disso, reformas estruturais relevantes<sup>1</sup>, afetaram a permanência de crianças e jovens na escola ([Menezes Filho, 2003](#)). Sendo assim, tais políticas contribuíram para o aumento significativo na escolaridade média dos brasileiros, a média de anos dos estudo passou de 5,52, em 1995, para 8,54, em 2015<sup>2</sup>.

Com o crescimento da escolaridade ocorreram alterações nos retornos da escolaridade, que são os retornos salariais de cada ano adicional de estudo. Por exemplo, [Nakabashi e Assahide \(2017\)](#) mostram queda relevante do retorno da escolaridade sobre os salários dos trabalhadores brasileiros tanto para jovens (15 a 29 anos) quanto para adultos (25 a 64 anos), entre 1997 e 2012. No período, eles estimam que a queda do retorno da escolaridade foi de 21,2% para jovens e 15% para adultos. [Suliano e Siqueira \(2012\)](#), utilizando dados dos trabalhadores brasileiros entre 25 e 64 anos, também encontram redução nos retornos salariais da educação, entre 2001 e 2006. Já [Montenegro e Patrinos \(2014\)](#) estimaram os retornos de 139 economias, e seus resultados apontam que em 1982 o retorno da escolaridade no Brasil era de 17,3% e em 2012 passou a ser 10,5%, ou seja, uma queda de aproximadamente 39%. Em alguns países com níveis de renda semelhantes ao Brasil os retornos da escolaridade também estão decrescendo ao longo do tempo, por exemplo, [Asadullah e Xiao \(2020\)](#) e [Horie e Iwasaki \(2022\)](#) encontram evidências de queda nos retornos da escolaridade da China, e de países emergentes da Europa, respectivamente. Ambos os estudos apontam o aumento da oferta de mão de obra qualificada como uma das possíveis causas da queda.

No presente artigo estudamos os efeitos da escolaridade no salário dos brasileiros entre 1995 e 2015. O artigo aborda três questões relacionadas: 1) a estimação dos retornos da escolaridade levando em consideração as características socioeconômicas dos trabalhadores (escolaridade, sexo, raça, setor de trabalho, zona de residência); 2) a tendência desses retornos no período e suas possíveis causas; e 3) as possíveis causas do diferencial dos retornos por faixa de escolaridade.

Para responder essas questões utilizamos dados Pesquisa Nacional por Amostra de Domicílios (PNAD) e estimamos os retornos da escolaridade para Brasil utilizando a Equação Minceriana. O método de estimação mais comum é o de Mínimos Quadrados Ordinários (MQO). Porém, podem surgir vieses nas estimativas por MQO devido a endogeneidade da escolha dos anos de estudo e do problema da seleção amostral ([Heckman, 1979](#); [Card, 2001](#); [Blundell et al., 2001](#)). Para contornar

---

<sup>1</sup>A Lei de Diretrizes e Base da Educação Nacional (1996), a criação do Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (FUNDEF) em 1998, a implementação da aprovação automática em várias unidades da federação nas décadas de 1980 e 1990 e o bolsa escola, adotado pelo governo federal em 2001.

<sup>2</sup>Ver Tabela B1 do [Apêndice](#).

para tais problemas, utilizamos o Método de [Garen \(1984\)](#), que é uma alternativa ao método de variáveis instrumentais<sup>3</sup> ([Card, 2001](#)), e controla os vieses de retorno, habilidade e erro de medida, e o Método de [Heckman \(1979\)](#) para controlar o viés de seleção amostral.

No Brasil, desde o início década de 1990, vários estudos empíricos sobre os retornos da escolaridade estimam essa medida no intuito de comparar coortes<sup>4</sup> ([Kassouf, 1997](#); [Sachsida et al., 2004](#); [Marcelo e Wyllie, 2006](#); [Moura, 2008](#); [Pereira et al., 2013](#)), analisar uma região específica ([Van Zaist et al., 2010](#); [Psacharopoulos e Patrinos, 2018](#); [Horie e Iwasaki, 2022](#)) ou analisar apenas um ou poucos anos ([Marcelo e Wyllie, 2006](#); [Van Zaist et al., 2010](#)). Considerando esses estudos, identificamos algumas lacunas que pretendemos sanar com o presente artigo.

A primeira delas é o fato de que os estudos focam em estimar os retornos da escolaridade para alguns anos específicos ou um painel de dados mais curto em relação ao presente estudo ([Sachsida et al., 2004](#); [Nakabashi e Assahide, 2017](#)). A importância de se analisar um período mais longo reside no fato de que no Brasil houve considerável aumento na escolaridade e isso potencialmente afeta a forma como o mercado de trabalho valoriza a mão de obra e os níveis específicos de escolaridade atingidos. Logo, esse artigo contribui para a literatura ao estimar a evolução dos retornos da escolaridade no Brasil em um período mais longo: 1995-2015. Nossos resultados apontam que os retornos da educação no Brasil são decrescentes nesse período. Além disso, fizemos um exercício contrafactual em que mostramos que caso os retornos da escolaridade fossem constantes e iguais aos de 1995, o salário hora médio no Brasil seria 14,9% maior em 2015.

A segunda lacuna diz respeito à análise de possíveis explicações para a evolução dos retornos da escolaridade no Brasil. Uma vez que boa parte dos estudos utilizam dados *cross section*, não é possível avaliar sua evolução, tampouco explicá-la. No presente artigo, discutimos possíveis causas dessa evolução baseada em dois fatos. Primeiro, mostramos que houve aumento da oferta de mão de obra qualificada ao longo dos anos e discutimos como isso pode ter feito os trabalhadores mais qualificados descenderem a escada ocupacional e exercerem profissões em que são super qualificados. Segundo, estimamos os retornos da escolaridade para cada um dos estados brasileiros e comparamos com algumas variáveis que servem de *proxy* para a qualidade da educação<sup>5</sup>, mostramos que nos estados com maior qualidade há maiores retornos. Analisar os fatores que levam às quedas observadas nos retornos da escolaridade no país é relevante para se entender a estagnação da produtividade do trabalho e para ajudar a elaboração de políticas públicas que revertam tal tendência.

A terceira lacuna diz respeito aos fatores que afetam os retornos da escolaridade. Há evidências empíricas que sugerem que há relação causal entre as características socioeconômicas dos trabalhadores e os retornos da escolaridade ([Pereira et al., 2013](#); [Kassouf, 1997](#); [Ottoni, 2017](#); [Psacharopoulos](#)

---

<sup>3</sup>[Card \(2001\)](#) aponta que a escolha do instrumento nem sempre é trivial, e dentro do nosso contexto, em que analisamos um horizonte temporal de 18 anos, variáveis comumente utilizadas como instrumento, por exemplo, a escolaridade da mãe, nem sempre estão disponíveis, e, o uso de variáveis instrumentais pouco correlacionadas com a variável explicativa pode levar a inconsistências nas estimativas ([Bound et al., 1995](#)).

<sup>4</sup>Homens e mulheres, moradores urbanos e rurais, trabalhadores mais e menos qualificados, entre outros

<sup>5</sup>Utilizamos dados do Sistema Nacional de Avaliação da Educação Básica (SAEB), do Índice de Desenvolvimento da Educação Básica (IDEB) e o salário hora médio dos professores como *proxy* para a qualidade do sistema educacional brasileiro.

e Patrinos, 2018). Logo, na especificação da Equação Minceriana que utilizamos permitimos interações entre os anos de estudo e várias dessas características (sexo, raça, setor de atividade, zona de residência). Boa parte dos estudos não levam isso em consideração, ou então, consideram apenas as interações entre anos de estudo, raça e experiência. Isso nos permite verificar se de fato as características socioeconômicas são determinantes dos retornos da escolaridade. Além disso, agrupamos os trabalhadores em quatro grupos de acordo com o nível de escolaridade e estimamos os retornos para cada um<sup>6</sup>. As evidências que encontramos são de que os retornos da escolaridade são maiores para os trabalhadores que possuem mais tempo de estudo, e, em média, a maior parte deles são brancos, moram em regiões urbanas e trabalham no setor de serviços. Por outro lado, os trabalhadores com menores retornos possuem menos tempo de estudo, a maioria são homens não brancos, e boa parte deles residem em áreas rurais e trabalham no setor agrícola.

Por fim, a última lacuna que identificamos está associada a correção dos vieses que surgem na estimação econométrica dos retornos da escolaridade via MQO e impossibilitam a estimação precisa dessa medida. Alguns estudos como o de Montenegro e Patrinos (2014) não utilizam nenhum método para a correção desses vieses. Marcelo e Wyllie (2006), Moura (2008), Van Zaist et al. (2010) e Nakabashi e Assahide (2017) corrigem apenas o viés de seleção amostral. Sendo assim, contribuímos para a literatura ao considerar a existência de vieses ao estimar o retornos da escolaridade no Brasil via Método de Heckman (1979) para corrigir o viés de seleção amostral, e o Método de Garen (1984) para corrigir os demais vieses. Nossas evidências apontam que o viés de seleção amostral viesam as estimativas para cima e os vieses de retorno, habilidade e erro de medida viesam para baixo.

Este estudo pode fornecer informações importantes sobre os incentivos em adquirir graus mais elevados de escolaridade, sobre a eficiência na alocação dos recursos e também nas consequências distributivas. Além disso, a comparação com os retornos de investimentos alternativos pode ajudar os formuladores de políticas educacionais a tomar melhores decisões de investimento.

Além da presente introdução, na Seção 2, apresentamos a especificação econométrica e cada uma dos métodos que utilizamos. Apresentamos, na Seção 3, a fonte dos dados (PNAD) e discutimos os recortes utilizados. Os principais resultados e suas implicações estão na Seção 4. Na seção 5 avaliamos a robustez dos nossos resultados ao testar uma especificação econométrica alternativa a que utilizamos como base. Finalmente, a Seção 6 traz as principais conclusões do artigo.

## 2 Especificações econométricas

Ao estimar os retornos da escolaridade via Mínimos Quadrados Ordinários (MQO) surgem alguns vieses que impedem a estimação correta dessa medida, são eles: viés de habilidade, viés de retorno, viés de erro de medida e viés de seleção amostral (Heckman, 1979; Card, 2001; Blundell et al., 2001). Logo, estimamos os retornos por MQO e utilizamos dois métodos adicionais: o método de Heckman (1979) para corrigir o viés de seleção amostral e o método de Garen (1984) para corrigir os demais.

---

<sup>6</sup>Grupo 1: de 0 a 4 anos de estudo, grupo 2: de 5 a 8, grupo 3: de 9 a 11, e grupo: 4 com 12 ou mais anos de estudo.

## 2.1 MQO

Estimamos a seguinte especificação da Equação Minceriana:

$$\begin{aligned} \ln Y_i = & \alpha_0 + \beta_0 S_i + \beta_1 Exp_i + \beta_2 Exp_i^2 + \beta_3 Ra\breve{c}a_i + \beta_4 Urb_i + \beta_5 Agr_i + \\ & \beta_6 Serv_i + \beta_7 Sexo_i + \beta_8 (Exp_i \times S_i) + \beta_9 (Ra\breve{c}a_i \times S_i) + \beta_{10} (Urb_i \times S_i) + \\ & \beta_{11} (Agr_i \times S_i) + \beta_{12} (Serv_i \times S_i) + \beta_{13} (Sexo_i \times S_i) + \gamma \mathbf{X} + \epsilon_i, \end{aligned} \quad (1)$$

em que:  $\ln Y_i$  é o logaritmo natural do salário hora mensal do indivíduo  $i$ ;  $S_i$  são os anos de estudo;  $Exp_i$  é a experiência no mercado de trabalho<sup>7</sup>; As demais variáveis são *dummies*, à saber:  $Ra\breve{c}a_i$  toma o valor 1 se é branco e 0 caso seja não branco;  $Urb_i$  é igual a 1 caso more em áreas urbanas e 0 caso contrário;  $Agr_i$  e  $Serv_i$  indicam o setor de atividade no qual o indivíduo trabalha, logo,  $Agr_i$  toma o valor 1 caso esteja na agricultura e  $Serv_i$  é igual a 1 caso esteja no setor de serviços;  $Sexo_i$  é igual a 1 caso seja do sexo masculino e 0 caso contrário;  $\mathbf{X}$  é um conjunto de variáveis de controle que inclui *dummies* indicando a região do Brasil em que o indivíduo reside (Sudeste, Nordeste, Centro Oeste e Sul) e se é sindicalizado; por fim,  $\epsilon_i$  é um termo de erro aleatório.

As variáveis anos de estudo, experiência e experiência ao quadrado são provenientes dos modelos de Mincer (1958) e Mincer (1974). As *dummies* de raça, zona de residência (rural, urbana), sexo, setor de ocupação, região de residência e sindicato são variáveis que captam as heterogeneidades dos indivíduos e são bastante comuns na literatura empírica, algumas delas podem ser vistas nos estudos de Sachsida et al. (2004), Marcelo e Wyllie (2006), Suliano e Siqueira (2012), Nakabashi e Assahide (2017), Brotherhood et al. (2019), dentre outros.

O que diferencia nossa especificação, expressa na Equação (1), de outras pesquisas é que adicionamos interações entre os anos de estudo e algumas das variáveis (Experiência, Raça, Urb, Agr, Serv, Sexo). Há algumas evidências empíricas mostrando que há diferenças nos retornos da escolaridade entre moradores urbanos e rurais (Pereira et al., 2013; Kassouf, 1997), entre pessoas do sexo masculino e feminino (Zhang et al., 2005; Psacharopoulos e Patrinos, 2018), entre pessoas de diferentes raças (Garen, 1984) e entre os setores de atividade da economia (Ottoni, 2017), e essas interações nos permitem verificar se de fato as características socioeconômicas dos trabalhadores influenciam os retornos da escolaridade. Na seção 5 testamos uma especificação alternativa para verificar a robustez dos resultados.

Ao derivarmos a Equação (1) em relação a  $S_i$  temos o retorno marginal da escolaridade:

$$\begin{aligned} \frac{\partial \ln Y_i}{\partial S_i} = & \beta_0 + \beta_8 Exp_i + \beta_9 Ra\breve{c}a_i + \beta_{10} Urb_i + \\ & \beta_{11} Agr_i + \beta_{12} Serv_i + \beta_{13} Sexo_i, \end{aligned} \quad (2)$$

Essa medida fornece o aumento percentual médio no salário hora para cada ano adicional de estudo.

<sup>7</sup>Seguimos Heckman et al. (2003) e calculamos a experiência como: idade - anos de estudo - 6.

## 2.2 Método de Garen

Devido ao problema de endogeneidade causado pela correlação entre  $\epsilon_i$  e  $S_i$ , o procedimento MQO produz estimativas viesadas dos retornos da escolaridade. Os três principais vieses são: viés de habilidade, retorno e viés de erro de medida (Blundell et al., 2001; Card, 2001). Para corrigí-los utilizamos o método proposto por Garen (1984).

Esse método é bastante comum na literatura sobre os retornos da escolaridade e pode ser visto nos estudos de Sachsida et al. (2004) e Suliano e Siqueira (2012). A ideia principal é estimar, no primeiro estágio, a variável endógena (anos de estudo) em função de alguns instrumentos, e no segundo estágio usar os erros desta estimação como regressor adicional na especificação principal (Equação (1)). Nessa abordagem a escolaridade é vista como uma decisão endógena e no primeiro estágio estimamos a equação de escolaridade onde as variáveis explicativas são variáveis que afetam a decisão de estudar. Logo, no primeiro estágio estimamos a seguinte equação:

$$S_i = \kappa_0 + \kappa_1 Filhos_i + \kappa_2 Raça_i + \kappa_3 Mãe_i + \kappa_4 Urb_i + \eta_i, \quad (3)$$

Onde:  $S_i$  representa os anos de escolaridade do indivíduo  $i$ ;  $Filhos_i$  toma o valor 1 caso tenha filhos menores de 14 anos e 0 caso não tenha;  $Raça_i$  toma o valor 1 se é branco e 0 caso seja não branco;  $Mãe_i$  assume o valor 1 caso a mãe esteja viva e 0 caso contrário;  $Urb_i$  tem valor 1 caso more em áreas urbanas e 0 caso contrário. Incluímos também *dummies* que indicam a região do Brasil onde o indivíduo reside<sup>8</sup>; por fim,  $\eta_i$  é o termo de erro aleatório.

Variáveis como raça e região onde reside podem ser vistas nos estudos de Garen (1984) e Sachsida et al. (2004). Por outro lado, a variável Filhos se justifica pelo fato de que as decisões de fertilidade e educação são interdependentes e pais pobres tendem a ter mais filhos e consequentemente investir menos em educação (De La Croix e Doepke, 2003). A zona de residência (rural ou urbana) também influencia nas decisões de estudar, uma vez que, indivíduos que moram na zona rural podem ter pior saúde e nutrição e maiores dificuldades de acesso a escola, quando comparados aos residentes urbanos, o que por sua vez, pode fazer com que não frequentem a escola (Gollin et al., 2014). Já a mãe do indivíduo, estando viva, pode influenciar na educação provendo os meios para que ele estude, sobretudo na infância<sup>9</sup>.

No segundo estágio, utilizamos o termo de erro da Equação (3) e estimamos a Equação (1) acrescida do seguinte termo:

$$\beta_{14}\hat{\eta}_i + \beta_{15}(\hat{\eta}_i \times S_i). \quad (4)$$

Note que aqui estamos assumindo que o termo de erro da Equação (1) pode ser capturado usando a relação linear  $\epsilon_i = \beta_{14}\hat{\eta}_i + \beta_{15}(\hat{\eta}_i \times S_i) + \omega_i$ , e que  $\mathbb{E}(\omega_i|S_i, \hat{\eta}_i, \hat{\eta}_i S_i) = 0$ . Logo, ao incluir a expressão (4) na Equação (1) estamos produzindo um novo termo de erro  $\omega_i$  que não está correlacionado com  $S_i$ , que por sua vez resolve o problema da endogeneidade.

<sup>8</sup>Por conveniência não indicamos na Equação (3).

<sup>9</sup>Suliano e Siqueira (2012) também utilizam uma *dummy* que indica se a mãe do indivíduo está ou não viva.



## 2.3 Método de Heckman

Em algumas pesquisas a Equação de Mincer é estimada utilizando-se somente dados dos indivíduos que estão no mercado de trabalho e cujo os dados de salários estão disponíveis na base de dados utilizada. Ao proceder dessa forma a amostra está sendo selecionada de forma não aleatória causando viés de seleção amostral. Uma das fontes desse viés pode ser o fato de que os indivíduos tenham um determinado salário reserva que afetam suas decisões de entrar ou não no mercado de trabalho. Logo, isso deve ser levado em consideração ao se estimar os retornos da escolaridade. Para corrigir esse viés o método de Heckman (1979) é o mais adotado<sup>10</sup>.

Esse método também é estimado em dois estágios, no primeiro, estimamos a equação de seleção, que deve conter variáveis que afetam a decisão dos indivíduos de entrar no mercado de trabalho, por exemplo, anos de estudo, experiência de trabalho e *dummies* que indicam a raça, zona de residência, sexo, se tem ou não filhos. Essa equação é estimada no intuito de calcular a razão inversa de Mills. A equação de seleção que estimamos é a seguinte:

$$D_i = \psi_0 + \psi_1 Ra\tilde{c}a_i + \psi_2 Urb_i + \psi_3 Sexo_i + \psi_4 Filhos_i + \psi_5 S_i + \zeta_i, \quad (5)$$

Onde:  $D_i$  é uma *dummy* que toma o valor 1 caso o indivíduo esteja empregado e 0 caso contrário;  $Ra\tilde{c}a_i$  toma o valor 1 se é branco e 0 caso seja não branco;  $Urb_i$  toma o valor 1 quando o indivíduo mora na área urbana e 0 caso contrário;  $Sexo_i$  tem valor 1 se é do sexo masculino e 0 caso contrário;  $Filhos_i$  toma o valor 1 se tem filhos menores de 14 anos e zero caso contrário;  $S_i$  são os anos de estudo. Adicionamos também *dummies* que indicam a região do Brasil que o indivíduo reside; por fim,  $\zeta_i$  é um termo de erro aleatório.

As variáveis de sexo, raça e filhos são comuns na literatura e podem influenciar na probabilidade de participação no mercado de trabalho. Por exemplo, mulheres que possuem filhos podem ter salário reserva maior que os demais participantes do mercado de trabalho. Segundo Monte et al. (2011) o salário reserva das mulheres é quase o dobro do dos homens em todas as regiões do Brasil e o dos indivíduos não brancos é superior ao dos brancos. Já variáveis de região e a zona de residência (urbano ou rural) servem para captar as heterogeneidades do mercado de trabalho brasileiro.

A Equação (5) é estimada utilizando-se o modelo Probit e a razão inversa de Mills é calculada da seguinte forma:

$$\lambda_i = \frac{\phi(\hat{D}_i)}{\Phi(\hat{D}_i)}, \quad (6)$$

Sendo que  $\phi(\cdot)$  é a função densidade de probabilidade e  $\Phi(\cdot)$  é a função de distribuição acumulada da distribuição normal padrão. A Equação (1) omite a razão inversa de Mills como regressor de modo que o viés de seleção amostral é resultado da omissão de variável explicativa. Logo, no segundo estágio do método de Heckman estimamos a Equação (1) acrescida da razão inversa de Mills:

---

<sup>10</sup>Marcelo e Wyllie (2006), Moura (2008), Nakabashi e Assahide (2017) também empregaram esse método.

$$\beta_{14}\lambda_i, \tag{7}$$

Consequentemente, isso produzirá estimativas dos retornos da escolaridade sem o viés de seleção amostral.

### 3 Base de dados

Para estimar a evolução dos retornos da escolaridade no Brasil utilizamos dados da Pesquisa Nacional de Amostras por Domicílios (PNAD) de 1995 a 2015. A PNAD é uma pesquisa amostral complexa, logo, envolve estratificação, conglomeração e probabilidades desiguais de seleção. Segundo [Moura \(2008\)](#) a não consideração do desenho amostral da PNAD ao estimar a variância, pode produzir testes de hipóteses e intervalos de confiança viesados, o que por sua vez, viesas as estimativas dos retornos da escolaridade<sup>11</sup>. Logo, seguimos [Moura \(2008\)](#) e consideramos o desenho amostral da PNAD em nossa análise.

Além disso, fizemos alguns ajustes nos dados para filtrar as informações relevantes e estimar corretamente os retornos da escolaridade. Restringimos a análise a trabalhadores entre 25 e 64 anos, que não estudam, e não são militares ou servidores públicos. Também excluímos os indivíduos que se declararam indígenas e amarelos e indivíduos que possuíam um ou mais dados faltantes nas variáveis utilizadas. Destaca-se que, em dois métodos econométricos descritos na Seção 2 (MQO e Garen) utilizamos dados somente dos indivíduos que estão empregados, no entanto, para implementar o método de Heckman utilizamos também dados dos indivíduos desempregados.

Salientamos, na seção anterior, que nossa variável dependente da Equação de Mincer (Equação 1) é o logaritmo do salário hora mensal e dois fatos sobre isso merecem destaque. Primeiro, calculamos essa medida dividindo o rendimento mensal em reais recebido no trabalho principal, pela quantidade de horas trabalhadas semanalmente multiplicada por 4,2<sup>12</sup>. Segundo, excluímos os trabalhadores que ganham mais de 100 salários mínimos<sup>13</sup>.

Para comparar os retornos da escolaridade que estimamos, a nível estadual, com medidas de qualidade da educação, utilizamos dados do Sistema Nacional de Avaliação da Educação Básica (SAEB), do Índice de Desenvolvimento da Educação Básica (IDEB) e escolaridade média dos professores. Os dois primeiros são fornecidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), o último coletamos na PNAD.

O IDEB é calculado a partir dos dados de aprovação escolar, obtidos no Censo Escolar, e das médias de desempenho em matemática e português do SAEB que é aplicado aos estudantes ao final

<sup>11</sup>Ver [Lumley \(2004\)](#).

<sup>12</sup>O salário reportado na PNAD é mensal, e as horas trabalhadas, são semanais. Logo, devemos fazer o ajuste para que ambas as variáveis estejam na mesma unidade temporal. Sendo assim, supomos que há 252 dias úteis no ano e 21 no mês, além disso supomos também que o indivíduo trabalha 5 dias da semana. Uma vez que na PNAD é reportado as horas trabalhadas na semana então dividimos 21 por 5 para obter o número de “semanas” trabalhadas no mês, que é 4,2.

<sup>13</sup>Ao fazer esse ajuste consideramos o salário mínimo em reais de 2010 do respectivo ano em análise.



de cada ciclo de ensino (4° e 8° ano do ensino fundamental e 3° ano do ensino médio). No âmbito internacional a pontuação em testes de avaliação de desempenho escolar também vem sendo empregada como medida de qualidade da educação, sobretudo para entender a relação entre capital humano e crescimento econômico. Uma das mais utilizadas é a pontuação no Programa Internacional de Avaliação de Alunos (PISA). [Hanushek e Kimko \(2000\)](#) enfatizam que essa medida é um componente primário do capital humano e melhora significativamente a capacidade de explicar o crescimento econômico.

## 4 Resultados

Nesta seção apresentamos os resultados e suas principais implicações. Na subseção [4.1](#) discutimos as particularidades das Equações Mincerianas e de cada método que estimamos (MQO, Heckman e Garen). Em seguida, na subseção [4.2](#) mostramos que houve queda nos retornos da educação brasileiro entre 1995 e 2015. Também mostramos que trabalhadores com maiores níveis de escolaridade possuem maiores retornos, e na subseção [4.3](#) discutimos o principal motivo disso. Na subseção [4.4](#) argumentamos que a expansão da mão de obra qualificada e a qualidade da educação podem ser determinantes dos retornos da educação no Brasil. Por fim, na subseção [4.5](#) fazemos um exercício contrafactual para verificar como o salário hora médio do brasileiro evoluiria caso os retornos da escolaridade fossem constantes e iguais aos de 1995.

### 4.1 Estimações econométricas

As regressões que estimamos com cada um dos métodos (MQO, Heckman e Garen) e para cada ano de 1995 a 2015 podem ser vistas no Apêndice [A](#). Para facilitar o entendimento primeiro discutimos os aspectos gerais das nossas estimações e depois as particularidades de cada um dos métodos que empregamos para estimar os retornos da escolaridade.

De modo geral, a maioria dos parâmetros estimados tem o sinal esperado. As variáveis experiência e anos de estudo possuem sinais positivos e a experiência ao quadrado sinal negativo e são relevantes para explicar variações nos salários, como prevê o modelo de [Mincer \(1974\)](#). O setor de atividade (agricultura e serviços) também mostrou ser relevante na determinação de salários. De acordo com nossos resultados os trabalhadores da agricultura possuem salários menores quando comparados trabalhadores da indústria e do setor de serviços.

Já a variável que indica se o trabalhador é ou não sindicalizado apresenta efeito positivo nos salários dos trabalhadores. No entanto, esse efeito vem diminuindo ao longo do tempo, indicando perda da importância dos sindicatos na determinação dos salários dos trabalhadores. Resultado semelhante pode ser visto em [Sachsida et al. \(2004\)](#), [Suliano e Siqueira \(2012\)](#) e [Nakabashi e Assahide \(2017\)](#). Nossos resultados também apontam outros fatos bastante comuns na literatura empírica, por exemplo, o salário dos homens é maior que o das mulheres, e, o salário dos trabalhadores urbanos é maior que dos rurais. [Kassouf \(1997\)](#), [Crespo e Reis \(2009\)](#), [Pereira et al. \(2013\)](#) e [Nakabashi e](#)

Assahide (2017) mostram evidências similares.

Outro padrão que encontramos nas estimações é o fato de que os trabalhadores da região nordeste possuem salários menores que os das demais regiões, fato que pode ser observado nos estudos de Sachsida et al. (2004) e Crespo e Reis (2009). Por outro lado, a variável raça não apresenta um padrão bem definido no sinal, alternando entre positivo e negativo.

No que diz respeito as particularidades dos métodos que empregamos, temos nas Tabelas A3 e A4 as estimações do segundo estágio do Modelo de Heckman, onde corrigimos o viés de seleção amostral<sup>14</sup>. Podemos ver que os coeficientes da variável  $\lambda$  que representa a razão inversa de Mills é significativo em 16 das 19 regressões estimadas, logo, nessas regressões,  $\lambda$  é uma variável relevante para controlar o viés de seleção amostral. Nota-se também, que em quase todas as regressões esse coeficiente é positivo, indicando correlação positiva entre as variáveis não observáveis das equações de seleção e das equações salariais. Isso quer dizer que algo não observado que aumenta as chances de trabalhar tende a aumentar o salário.

Quando comparamos as médias dos retornos estimados com e sem correção do viés de seleção amostral (Tabela B2) notamos que as estimativas por MQO são de fato viesadas, para cima. Ou seja, a não consideração do viés de seleção amostral faz com que as estimativas dos retornos da escolaridade sejam maiores que de fato são.

Nas Tabelas A5 e A6 reportamos as estimativas método de Garen. Observa-se que os coeficientes do termo de erro da equação de escolaridade,  $\hat{\eta}$ , e da interação desse termo com os anos de escolaridade,  $\hat{\eta} : S$ , são positivos e significativos na maioria das regressões. Nos artigos de Garen (1984) e Sachsida et al. (2004) eles encontram coeficientes negativos para a variável  $\hat{\eta}$ . Essa diferença em relação a nossas estimações pode ser atribuída às diferenças na especificação da equação de escolaridade, que destacamos ser devido a indisponibilidade de dados para o período que analisamos. Outro fato importante do método de Garen é que as estimativas dos retornos da escolaridade são maiores que as de MQO (Tabela B2). Logo, nossas evidências mostram que os vieses de habilidade, retorno, e erro de medida estão viesando os retornos da educação para baixo.

## 4.2 Retornos da escolaridade

Na Figura 1 apresentamos a evolução dos retornos da escolaridade no Brasil entre 1995 e 2015 estimada via Mínimos Quadrados Ordinários (MQO), método de Garen e Heckman<sup>15</sup>. Nota-se claramente que em todos os métodos os retornos apresentaram queda. No caso do MQO a queda foi de 23,3%, e nos métodos de Garen e Heckman 17 % e 33% respectivamente. Isso quer dizer que anos adicionais de estudo estão resultando em acréscimos cada vez menores ao salário dos trabalhadores brasileiros ao longo do tempo.

Uma vez que estamos interessados na evolução dos retornos da escolaridade, utilizamos o Teste

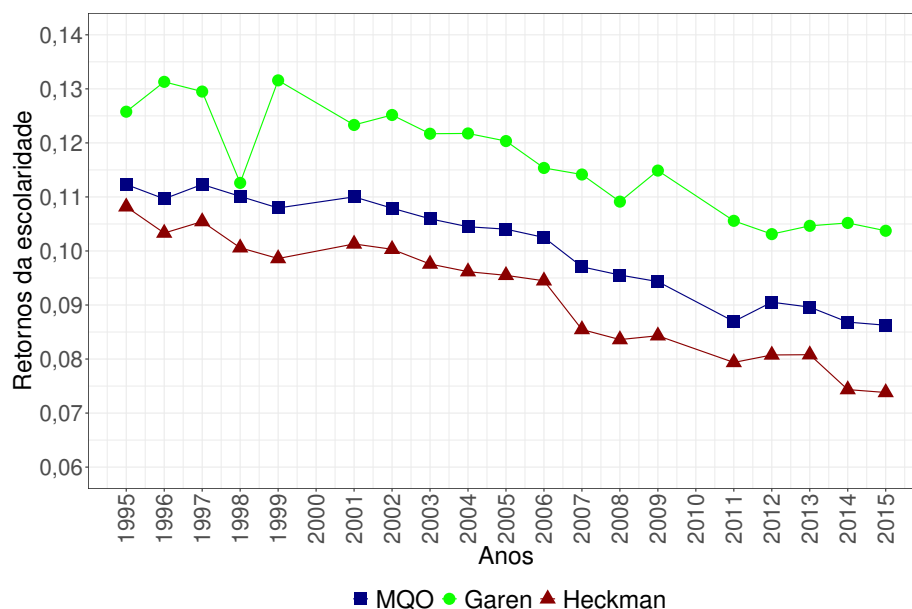
---

<sup>14</sup>As estimações das equações de seleção, e das equações de escolaridade, primeiro estágio do Método de Heckman e Garen, respectivamente, podem ser vistas no [Apêndice Online](#).

<sup>15</sup>Destaca-se que, para calcular os retornos da escolaridade substituímos na Equação (2) as médias amostrais das variáveis Exp, Raça, Urb, Agr, Serv e Sexo. E no caso do método de Garen utilizamos também a média de  $\hat{\eta}$ .

de Wald para verificar se de fato os parâmetros associados aos retornos da escolaridade (Equação (2)) diferem entre 1995 e 2015. Os resultados podem ser vistos no Apêndice C e rejeitam a hipótese nula de que são iguais, nos três métodos.

**Figura 1:** Evolução dos retornos da escolaridade no Brasil - 1995:2015



Fonte: Elaborado pelos autores.

Alguns estudos encontraram resultados semelhantes, por exemplo, [Moura \(2008\)](#) ao estimar os retornos da escolaridade no Brasil entre 1992 a 2004 encontrou evidências de que estes estão em tendência de queda, exceto para trabalhadores que possuem nível superior. Já [Montenegro e Patrinos \(2014\)](#) encontraram queda de aproximadamente 39% entre 1982 e 2012. No entanto, essa tendência de queda não está restrita apenas ao Brasil, como pode ser visto nos estudos de [Asadullah e Xiao \(2020\)](#) e [Horie e Iwasaki \(2022\)](#) que analisaram o caso chinês e de países emergentes da Europa respectivamente.

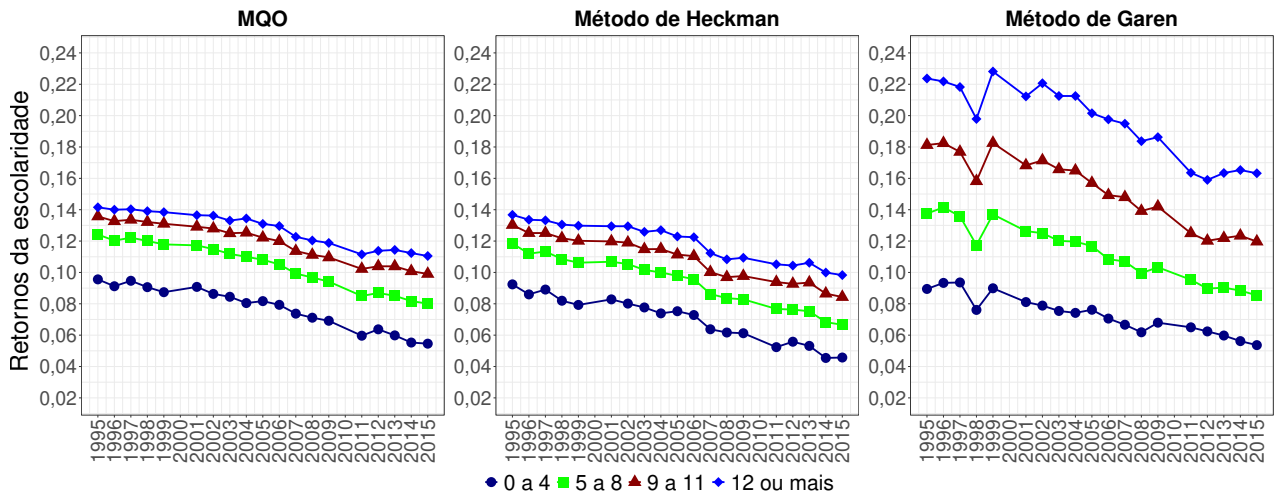
Há também evidências empíricas que mostram heterogeneidade no retorno da escolaridade dado o nível de escolaridade atingido, ou seja, indivíduos com mais anos de escolaridade tendem a ter maiores retornos de escolaridade e remuneração do trabalho, sobretudo se tiverem o ensino superior completo ([Garen, 1984](#); [Jaeger e Page, 1996](#); [Sachsida et al., 2004](#); [Crespo e Reis, 2009](#)). Logo, neste estudo, agrupamos os trabalhadores de acordo com o nível escolaridade atingido e calculamos os retornos da escolaridade para cada um deles. O primeiro grupo é formado por indivíduos que possuem entre 0 e 4 anos de estudo, o segundo 5 a 8, o terceiro 9 a 11 e o quarto mais de 12 anos de estudo. O primeiro e segundo grupos compreendem os dois primeiros ciclos do ensino fundamental, o terceiro abrange o ensino médio e o quarto o ensino superior.

Para facilitar a análise dos retornos da escolaridade no Brasil por nível de escolaridade atingido seguimos [Garen \(1984\)](#) e [Sachsida et al. \(2004\)](#), logo, o que fazemos é reportar o resultado da Equação (2) para cada um dos quatro grupos de indivíduos classificados de acordo com os anos de estudo.

O procedimento é simples, basta substituir na Equação (2) as médias amostrais, condicionada ao grupo, das variáveis Exp, Raça, Urb, Agr, Serv e Sexo. No caso do método de Garen utilizamos também a média condicional da variável  $\hat{\eta}_i$ .

Os resultados podem ser vistos na Figura 2, e notamos que os retornos da escolaridade estão em tendência de queda para todos os grupos<sup>16</sup>. Além disso, um ano adicional de estudo, em média, tem maior impacto positivo nos rendimentos dos indivíduos mais escolarizados. Esse padrão não é exclusivo no Brasil, por exemplo, Colclough et al. (2010) apontam em seu estudo que houve uma mudança no padrão mundial dos retornos da escolaridade que antes era côncavo e passou a ser convexo, ou seja, a educação adicional tem um impacto proporcional muito maior nos salários de trabalhadores com níveis educacionais mais altos que nos mais baixos. Nas próximas subseções discutimos as possíveis causas da diferença nos retornos dos quatro grupos, e, da tendência de queda.

**Figura 2:** Evolução dos retornos da escolaridade no Brasil por níveis de escolaridade - 1995:2015



Fonte: Elaborado pelos autores.

### 4.3 Diferença nos retornos da escolaridade por nível educacional

Na Tabela 1 apresentamos a média das estimações dos retornos da escolaridade, entre 1995 e 2015, por nível de escolaridade. Como já havíamos enfatizado, a medida que o nível educacional aumenta o retorno da escolaridade aumenta em todos os métodos. Chama atenção também, o fato de que no método de Garen a média dos retornos para quem tem mais de 12 anos de estudo é 2,6 vezes maior do que a média de quem tem entre 0 e 4 anos de estudo. Além disso, verificamos que o retorno da escolaridade pelo método de Garen ultrapassa o retorno estimado por MQO a partir da faixa de 5 a 8 anos de estudo, isso é um indício de que há uma mudança no viés gerado pela endogeneidade na decisão pelo nível de escolaridade. Logo, dado a baixa expectativa de retorno, os indivíduos que estão abaixo dos 5 anos de estudo podem optar por não investirem em educação.

<sup>16</sup>Disponibilizamos esses resultados no formato de Tabela no [Apêndice Online](#).

**Tabela 1:** Média dos retornos da escolaridade condicionada ao nível de escolaridade - 1995:2015

Escolaridade	MQO	Heckman	Garen
0 a 4	0,077	0,070	0,073
5 a 8	0,104	0,094	0,113
9 a 11	0,119	0,108	0,152
12 ou mais	0,128	0,119	0,196

Fonte: Elaborado pelos autores.

Na Tabela 2 reportamos as médias das variáveis utilizadas para calcular os retornos da escolaridade, por nível de escolaridade atingido. Nota-se que, em média, os trabalhadores com mais tempo de estudo possuem menos experiência de trabalho, a maior parcela deles são brancos, moram em regiões urbanas e trabalham no setor de serviços. Por outro lado, os indivíduos com menos tempo de estudo possuem mais experiência de trabalho, a maioria são homens não brancos, e uma parcela significativa desses trabalhadores residem em áreas rurais e trabalham no setor agrícola. Essas heterogeneidades entre os diferentes grupos de trabalhadores são a principal causa das diferenças nos retornos da escolaridade entre eles.

**Tabela 2:** Médias das variáveis utilizadas para calcular os retornos da escolaridade - 1995:2015

Escolaridade	Exp	Raça	Urb	Agr	Serv	Sexo
0 a 4	35,54	37,81%	66,04%	37,08%	40,39%	63,24%
5 a 8	25,82	47,64%	86,94%	11,97%	60,11%	61,40%
9 a 11	19,49	52,76%	94,37%	4,09%	73,40%	55,40%
12 ou mais	18,53	74,18%	97,35%	2,29%	81,45%	51,78%

Fonte: Elaborado pelos autores.

Notas: **Exp**: tempo de experiência no mercado de trabalho. As demais variáveis são *dummies*, à saber: **Raça**: toma o valor 1 caso o indivíduo seja branco e 0 caso contrário, **Urb**: tem valor 1 caso more em áreas urbanas e 0 caso contrário; **Agr**: tem valor 1 caso trabalhe no setor agrícola e 0 caso contrário; **Serv**: toma o valor 1 caso trabalhe no setor de serviços e 0 caso contrário; **Sexo**: tem valor 1 caso seja do sexo masculino e 0 caso contrário.

Algumas pesquisas apresentam evidências semelhantes a essas. Por exemplo, [Pereira et al. \(2013\)](#) ao analisarem os retornos da escolaridade no meio urbano e rural brasileiro concluíram que esses retornos são maiores no meio urbano. Eles argumentam que isso resulta da lacuna de qualidade na educação que há entre ambas as zonas de residência, e isso leva a menor produtividade e remuneração no meio rural. Já [Ottoni \(2017\)](#) estima os retornos da escolaridade para os setores de serviços modernos, tradicionais, indústria, agropecuária e administração pública no Brasil. Ele encontra evidências de que os retornos da escolaridade são maiores no setor de serviços modernos<sup>17</sup>. Há também evidências mostrando que os retornos da escolaridade das mulheres são maiores que dos homens, por exemplo, [Psacharopoulos e Patrinos \(2018\)](#).

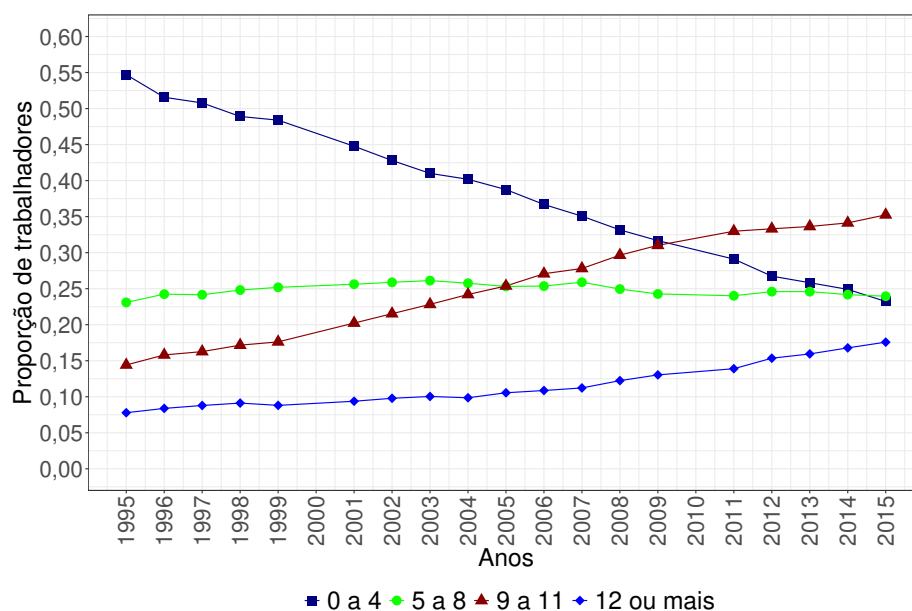
<sup>17</sup>[Ottoni \(2017\)](#) interpreta o salário dos trabalhadores como uma medida de produtividade e refere-se ao retorno da escolaridade como sendo uma taxa de conversão de anos de estudo em produtividade. Em essência, taxa de conversão e retornos da escolaridade são a mesma coisa, tanto é que o procedimento que ele utiliza para estimar essa medida é semelhante ao que utilizamos aqui, ou seja, a Equação Minceriana.

## 4.4 Retornos decrescentes

### 4.4.1 Aumento da mão de obra qualificada

Um ponto importante que merece destaque é o fato de que houve um aumento significativo no nível de escolaridade dos brasileiros entre 1995 e 2015. Como podemos na Figura 3 a proporção de trabalhadores de 0 a 4 anos de estudo caiu de 54% da população para 22,8% entre 1995 e 2015, por outro lado, a proporção de trabalhadores com 12 anos ou mais de estudos aumentou de 7,7% para 17,6% no mesmo período. Essa rápida expansão educacional pode ter ocorrido a uma taxa que excedeu a da criação de empregos. Logo, dado o aumento na oferta de mão de obra qualificada a demanda pode não ter acompanhado, o que por sua vez, pode ter reduzido os retornos da escolaridade no Brasil.

**Figura 3:** Evolução da proporção de trabalhadores por nível de escolaridade no Brasil - 1995:2015



Fonte: Elaborado pelos autores.

Nesse contexto, a redução pode acontecer por duas vias. A primeira é o fato de que os empregadores podem estar ofertando salários menores aos trabalhadores dado o excesso de mão de obra disponível. [Garcia \(2021\)](#), por exemplo, observaram aumento na oferta relativa de trabalhadores qualificados no Brasil entre 1981 e 2014, segundo eles, esse fato é a chave para explicar o declínio acentuado na prêmio por qualificação dos trabalhadores.

A segunda via está associada ao fato de que os trabalhadores mais qualificados podem estar descendo a escada ocupacional e indo para profissões que são menos intensivas em habilidades, o que por sua vez, também faz com que os trabalhadores menos qualificados desçam ainda mais a escada ocupacional. Isso resulta em descompasso educacional uma vez que o trabalhador tem um nível educacional maior que o requerido na profissão que exerce. Consequentemente, esse descompasso pode gerar má alocação de talentos uma vez que a realocação desses trabalhadores poderia aumentar



a produtividade da economia.

Marioni (2021) analisou os efeitos do descompasso educacional nos salários dos trabalhadores brasileiros e uma das conclusões é de que um quarto dos trabalhadores formais possuem qualificação maior que a necessária para exercer sua profissão, ou seja, são trabalhadores supereducados. Esses trabalhadores supereducados ganham significativamente menos que os trabalhadores com mesmo nível de qualificação e emprego compatível com suas habilidades e isso resulta em queda nos retornos da educação. Mais evidências sobre o descompasso educacional no Brasil podem ser vistos em Esteves (2009), Reis (2017) e Annegues e Souza (2020)<sup>18</sup>.

#### 4.4.2 Qualidade do sistema educacional

Outro fato importante que afeta os retornos da escolaridade é a qualidade do sistema educacional. Há sistemas educacionais eficazes e capazes de desenvolver nos alunos várias habilidades, enquanto outros são inúteis e não contribuem para a formação de capital humano. Nesse sentido, a qualidade do ensino pode servir para os empregadores como um sinal de alta ou baixa produtividade. Logo, trabalhadores que estudaram em melhores escolas são melhor remunerados e conseqüentemente possuem maiores retornos.

Card e Krueger (1992), por exemplo, estimaram os retornos da escolaridade com dados de homens americanos nascidos entre 1920 e 1949, e um dos resultados encontrados mostra que homens que foram educados em escolas com alta qualidade possuíam retornos maiores. Card e Krueger (1992) utilizam a proporção professor aluno e o salário relativo dos professores como medidas de qualidade do ensino.

Já Brotherhood et al. (2019) interpretam os retornos da escolaridade como sendo uma medida de qualidade da educação. A ideia principal é de que os retornos da escolaridade são provenientes do valor que as forças do mercado atribuem a educação dos trabalhadores, logo, *ceteris paribus*, as diferenças nos retornos da escolaridade entre indivíduos que estudam em diferentes estados podem ser atribuídas a qualidade da educação recebida. Após estimar os retornos para os estados brasileiros Brotherhood et al. (2019) comparam os resultados com algumas medidas de qualidade da educação a nível estadual, por exemplo, o gasto médio com aluno na educação primária e as notas do Sistema Nacional de Avaliação da Educação Básica (SAEB), e encontram forte correlação positiva entre essas medidas.

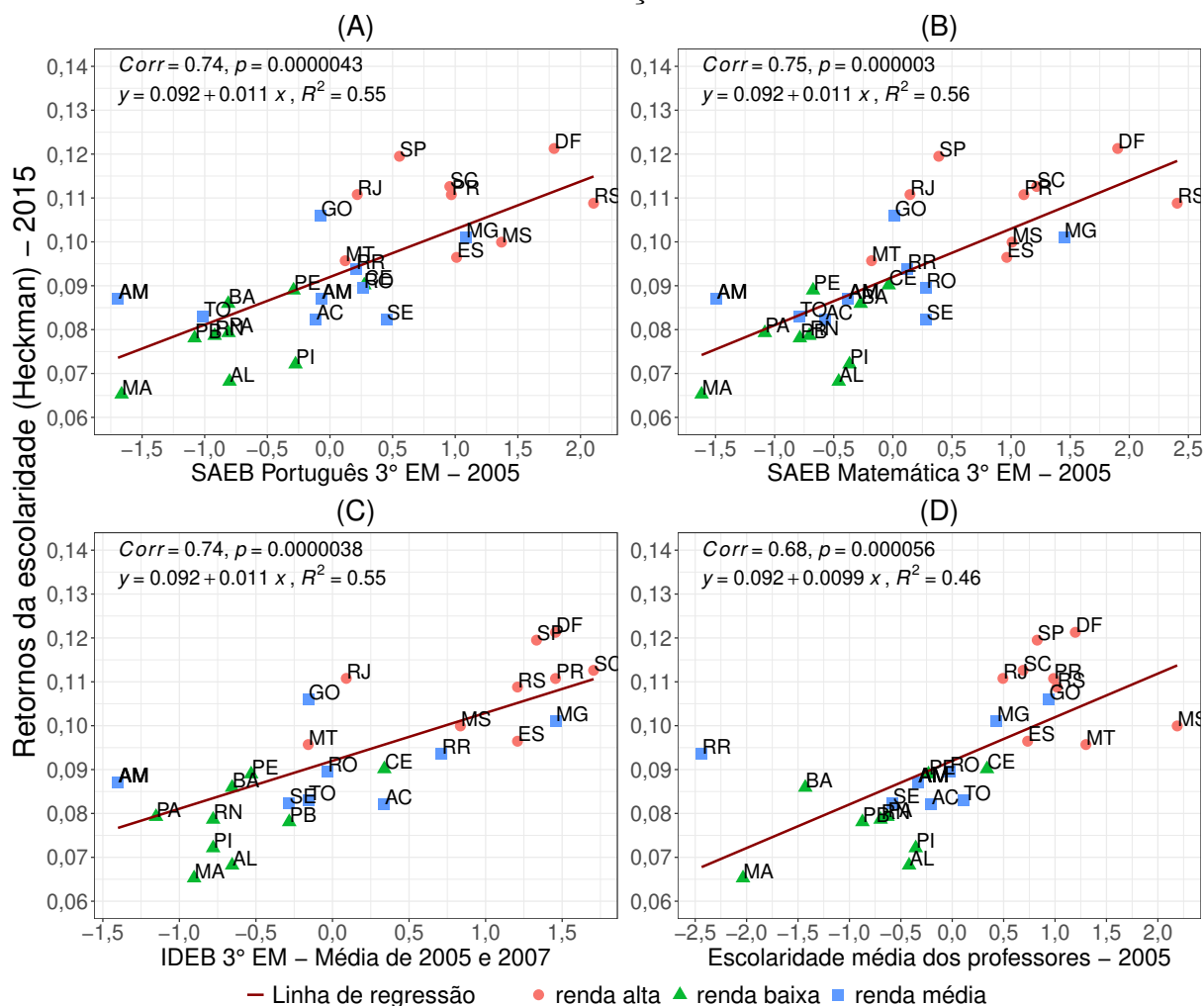
Seguimos Brotherhood et al. (2019) e estimamos os retornos da escolaridade para os 26 estados brasileiros mais o Distrito Federal<sup>19</sup> e comparamos com algumas medidas que servem como *proxy* para a qualidade da educação no Brasil a nível estadual, por exemplo, as notas do SAEB nos exames de matemática e português, a pontuação do Índice de Desenvolvimento da Educação Básica (IDEB) e a escolaridade média dos professores. Para fazer uma comparação adequada em termos de tempo

<sup>18</sup>Esses fatos também foram observados na economia americana por Beaudry et al. (2016).

<sup>19</sup>Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE), Tocantins (TO).

utilizamos dados de 2005 das medidas de qualidade da educação e estimamos os retornos da escolaridade de 2015 com uma subamostra de indivíduos entre 26 e 30 anos. Provavelmente os indivíduos nesta faixa etária estavam estudando no 3º ano do ensino médio (3º EM) em 2005.

**Figura 4:** Comparação entre os retornos da escolaridade e medidas de qualidade da educação



Fonte: Elaborado pelos autores com dados do INEP e PNAD.

Notas: Corr: Correlação de Pearson;  $p$ : Valor  $p$  da correlação. As medidas de qualidade da educação foram padronizadas.

Na Figura 4 reportamos os resultados e notamos alguns fatos relevantes<sup>20</sup>. Primeiro, tanto as medidas de qualidade da educação quanto os retornos da escolaridade são bastante heterogêneos entre os estados. Segundo, os estados com maiores níveis de renda possuem maiores retornos da escolaridade<sup>21</sup>. Em média, o retorno dos estados de renda alta é 0,108, contra 0,09 e 0,079 dos estados com renda média e baixa, respectivamente. Terceiro, há forte correlação positiva entre os

<sup>20</sup>Por conveniência, reportamos a comparação da qualidade da educação somente com os retornos da escolaridade calculados via Método de Heckman. Quando comparamos os retornos calculados com os demais métodos e as medidas de qualidade notamos que apresentam comportamento similar ao observado na Figura 4.

<sup>21</sup>Classificamos os 26 estados brasileiros mais o Distrito Federal usando dados do PIB per capita de 2015. Os nove primeiros são considerados renda alta, os nove do meio são renda média e os nove últimos são renda baixa.

retornos da escolaridade que estimamos para 2015 e as medidas de qualidade da educação de 2005. Ou seja, no estados onde a qualidade da educação era alta (baixa) em 2005 os retornos são maiores (menores) em 2015.

## 4.5 Exercício contrafactual

Há uma pergunta adicional que queremos responder, que é: Como o salário hora médio evoluiria caso os retornos da escolaridade fossem constantes e iguais aos de 1995? Responder à essa pergunta vai nos dizer quão maior o salário hora seria caso as forças de mercado (demanda e oferta por trabalho) e a qualidade da educação fossem constantes. Para responder a essas perguntas fizemos o seguinte exercício contrafactual:

1. Estimamos a Equação (1) e utilizamos os coeficientes e a média das variáveis explicativas para calcular a média do logaritmo do salário hora de 1995 a 2015. Essa estimação é a referência.
2. Recalculamos a média do logaritmo do salário hora de modo similar ao passo 1, porém, fixamos os parâmetros e a média das variáveis associadas ao retorno da escolaridade como sendo de 1995, ou seja, recalculamos o salário com os retornos da escolaridade de 1995, conforme a Equação (8).

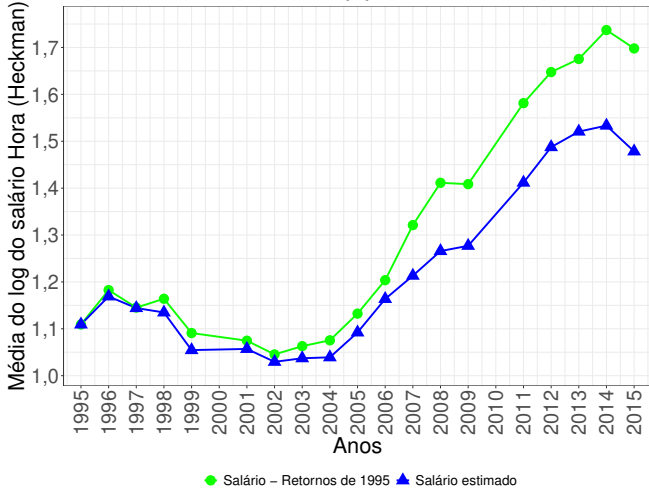
$$\begin{aligned} \ln Y^t = & \alpha_0^t + \beta_0^{1995} \bar{S}^t + \beta_1^t \bar{Exp}^t + \beta_2^t (\bar{Exp}^t)^2 + \beta_3^t \bar{Raça}^t + \beta_4^t \bar{Urb}^t + \beta_5^t \bar{Agr}^t + \\ & \beta_6^t \bar{Serv}^t + \beta_7^t \bar{Sexo}^t + \beta_8^{1995} (\bar{Exp}^{1995} \times \bar{S}^{1995}) + \beta_9^{1995} (\bar{Raça}^{1995} \times \bar{S}^{1995}) + \\ & \beta_{10}^{1995} (\bar{Urb}^{1995} \times \bar{S}^{1995}) + \beta_{11}^{1995} (\bar{Agr}^{1995} \times \bar{S}^{1995}) + \beta_{12}^{1995} (\bar{Serv}^{1995} \times \bar{S}^{1995}) + \\ & \beta_{13}^{1995} (\bar{Sexo}^{1995} \times \bar{S}^{1995}) + \gamma^t \bar{\mathbf{X}}^t, \text{ onde } t \in [1995, 1996, \dots, 2015], \end{aligned} \quad (8)$$

Onde  $\ln Y$  é o logaritmo natural do salário hora mensal;  $S$  são os anos de estudo;  $Exp$  é a experiência no mercado de trabalho; As demais variáveis são *dummies*, a saber: **Raça** toma o valor 1 se é branco e 0 caso seja não branco; **Urb** é igual a 1 caso more em áreas urbanas e 0 caso contrário; **Agr** e **Serv** indicam o setor de atividade no qual o indivíduo trabalha, logo, *Agr* toma o valor 1 caso esteja na agricultura e *Serv* é igual a 1 caso esteja no setor de serviços; **Sexo** é igual a 1 caso seja do sexo masculino e 0 caso contrário;  $\mathbf{X}$  é um conjunto de variáveis de controle que inclui *dummies* indicando a região do Brasil em que o indivíduo reside (Sudeste, Nordeste, Centro Oeste e Sul) e se é sindicalizado; por fim,  $\epsilon$  é um termo de erro aleatório.

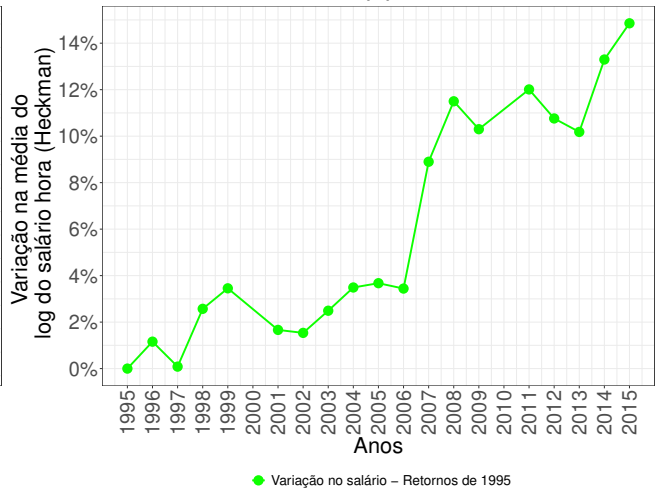
Na Figura 5 (A) reportamos a evolução do salário (referência), e o salário calculado com o retorno da escolaridade de 1995. Na Figura 5 (B) mostramos a variação percentual entre essas duas medidas. Notamos que, caso o retorno da escolaridade fosse constante e igual ao de 1995, o salário hora médio seria maior que de fato é. Considerando 2015 seria 14,9% maior<sup>22</sup>.

<sup>22</sup>Por conveniência reportamos apenas os resultados do Método de Heckman. Destaca-se que os resultados do MQO e Método de Garen também apontam salários maiores caso o retorno da escolaridade fosse constante e igual ao de 1995, considerando 2015 seria 17,6% e 3,62% maior, respectivamente.

**Figura 5: Exercício contrafactual - 1995:2015**  
(A)



(B)



Fonte: Elaborado pelos autores.

## 5 Análise de robustez

Para testar a robustez dos nossos resultados estimamos uma especificação da Equação Minceriana semelhante a utilizada no estudo de [Garen \(1984\)](#), [Sachsida et al. \(2004\)](#) e [Suliano e Siqueira \(2012\)](#), ou seja:

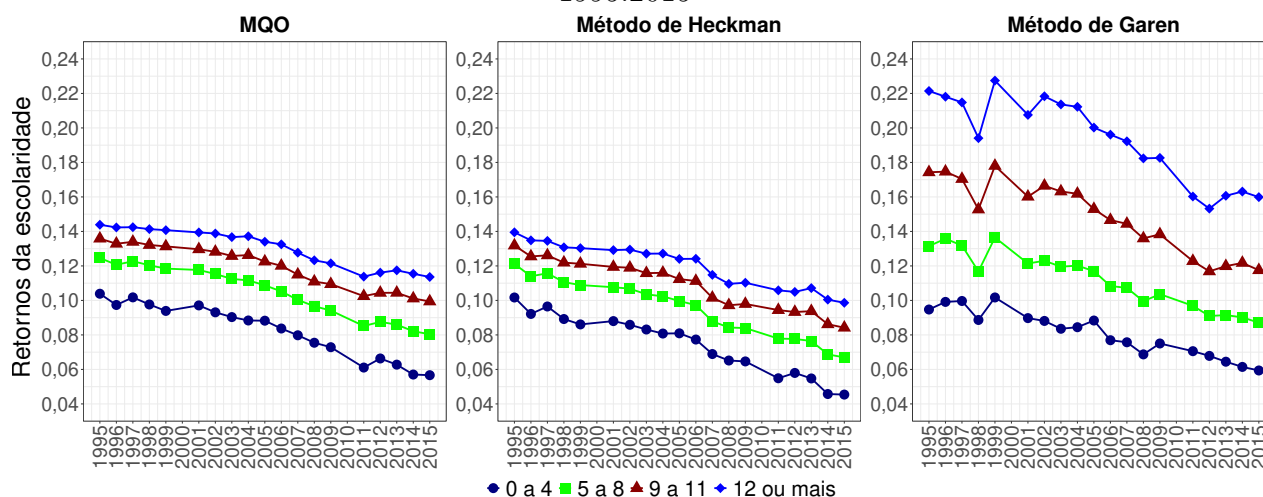
$$\begin{aligned} \ln Y_i = & \alpha_0 + \beta_0 S_i + \beta_1 Exp_i + \beta_2 Exp_i^2 + \beta_3 Ra\grave{c}a_i + \\ & \beta_4 Urb_i + \beta_5 Agr_i + \beta_6 Serv_i + \beta_7 Sexo + \beta_8 (Exp_i \times S_i) + \\ & \beta_9 (Ra\grave{c}a_i \times S_i) + \beta_{10} (Ra\grave{c}a_i \times Exp_i \times S_i) + \gamma \mathbf{X} + \epsilon_i, \end{aligned} \quad (9)$$

Sendo que  $\ln Y_i$  é o logaritmo natural do salário hora mensal do indivíduo  $i$ ;  $S_i$  são os anos de estudo;  $Exp_i$  é a experiência no mercado de trabalho; As demais variáveis são *dummies*, a saber:  $Ra\grave{c}a_i$  toma o valor 1 se é branco e 0 caso seja não branco;  $Urb_i$  é igual a 1 caso more em áreas urbanas e 0 caso contrário;  $Agr_i$  e  $Serv_i$  indicam o setor de atividade no qual o indivíduo trabalha, logo,  $Agr_i$  toma o valor 1 caso esteja na agricultura e  $Serv_i$  é igual a 1 caso esteja no setor de serviços;  $Sexo_i$  é igual a 1 caso seja do sexo masculino e 0 caso contrário;  $\mathbf{X}$  é um conjunto de variáveis de controle que inclui *dummies* indicando a região do Brasil em que o indivíduo reside (Sudeste, Nordeste, Centro Oeste e Sul) e se é sindicalizado; Por fim,  $\epsilon_i$  é um termo de erro aleatório.

A principal diferença entre a especificação expressa na Equação (1) para esta é que aqui há somente três interações da variável anos de estudo. Logo, o retorno marginal da educação pode ser escrito como:

$$\frac{\partial \ln Y_i}{\partial S_i} = \beta_0 + \beta_8 Exp_i + \beta_9 Ra\grave{c}a_i + \beta_{10} (Ra\grave{c}a_i \times Exp_i), \quad (10)$$

**Figura 6:** Evolução dos retornos da escolaridade no Brasil por níveis de escolaridade - 1995:2015



Fonte: Elaborado pelos autores.

Na Figura 6 vemos que o resultado é bastante semelhante ao da Figura 2, ou seja, os retornos da escolaridade são maiores para indivíduos com maior nível de escolaridade e são decrescentes. As estimações podem ser vistas no [Apêndice Online](#) desse artigo.

## 6 Conclusão

No presente artigo utilizamos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) e estimamos os retornos da escolaridade do Brasil entre 1995 e 2015. Para fazer as estimações utilizamos o método de Mínimos Quadrados Ordinários (MQO) e dois métodos adicionais que corrigem alguns vieses que podem surgir, são eles: Método de [Garen \(1984\)](#) que corrige os vieses de habilidade, retorno e erro de medida, e Método de [Heckman \(1979\)](#) que corrige o viés de seleção amostral.

Nossos resultados são robustos, e os três métodos que utilizamos são unânimes em apontar queda nos retornos da escolaridade brasileiro no período analisado. Além disso, testamos uma especificação diferente da especificação que utilizamos como base e a tendência de queda nos retornos se manteve. Também mostramos evidências de que os vieses de retorno, habilidade e erro de medida viesam as estimativas dos retornos da escolaridade para baixo, por outro lado, o viés de seleção amostral viesam as estimativas para cima.

Argumentamos que uma das possíveis causas da queda nos retornos da escolaridade é a expansão da mão de obra qualificada no Brasil. Essa expansão pode fazer com que os empregadores ofereçam salários menores aos trabalhadores dado o excesso de mão de obra qualificada. Além disso, os trabalhadores qualificados podem estar descendo a escada ocupacional e ocupando profissões as quais são superqualificados, consequentemente recebendo uma remuneração incompatível com seu nível de ensino.

Outra possível causa da queda nos retornos é a qualidade do sistema educacional brasileiro. Para

verificar como essas variáveis se relacionam estimamos os retornos da escolaridade para os estados brasileiros com dados de 2015, e comparamos com quatro medidas de qualidade da educação de 2005. Nossas evidências mostram que os estados que possuíam maior qualidade da educação em 2005 são os que possuem maiores retornos da educação em 2015.

Para captar as heterogeneidades nos retornos da escolaridade, condicionado ao nível de escolaridade, agrupamos os trabalhadores de acordo com a escolaridade e calculamos os retornos para cada um desses grupos. Nossas evidências mostram que os trabalhadores com maiores níveis de escolaridade possuem maiores retornos, e a principal causa desse resultado são as características socioeconômicas desses trabalhadores. Boa parte dos trabalhadores com menos tempo de estudo são não brancos, residem em áreas rurais, trabalham na agricultura, são do sexo masculino e possuem mais experiência de trabalho. Por outro lado, a maior parcela dos trabalhadores com mais tempo de estudo são brancos, residem em áreas urbanas, trabalham no setor de serviços e possuem menos tempo de experiência de trabalho.

Destaca-se que tanto o aumento da mão de obra qualificada quanto a baixa qualidade do sistema educacional brasileiro podem ter contribuído para que os retornos sejam decrescentes no período que analisamos. Embora essa seja uma explicação factível, não pretendemos esgotar o assunto, logo, pesquisas futuras podem fornecer informações importantes sobre outros fatores que possivelmente afetam os retornos da escolaridade no Brasil, por exemplo, o comércio internacional e o investimento estrangeiro direto.

## Referências

- Acemoglu, D. e Angrist, J. (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER macroeconomics annual*, 15:9–59.
- Annegues, A. C. e Souza, W. P. S. d. F. (2020). Retorno salarial do overeducation: Viés de seleção ou penalização ao excesso de escolaridade? *Revista Brasileira de Economia*, 74:119–138.
- Asadullah, M. N. e Xiao, S. (2020). The changing pattern of wage returns to education in post-reform china. *Structural Change and Economic Dynamics*, 53:137–148.
- Beaudry, P., Green, D. A., e Sand, B. M. (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics*, 34(S1):S199–S247.
- Blundell, R., Dearden, L., e Sianesi, B. (2001). Estimating the returns to education: Models, methods and results.
- Bound, J., Jaeger, D. A., e Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American statistical association*, 90(430):443–450.



- Brotherhood, L. M., Ferreira, P. C., e Santos, C. (2019). Education quality and returns to schooling: Evidence from migrants in Brazil. *Economic Development and Cultural Change*, 67(3):439–459.
- Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160.
- Card, D. e Krueger, A. B. (1992). Does school quality matter? returns to education and the characteristics of public schools in the United States. *Journal of political Economy*, 100(1):1–40.
- Colclough, C., Kingdon, G., e Patrinos, H. (2010). The changing pattern of wage returns to education and its implications. *Development Policy Review*, 28(6):733–747.
- Crespo, A. e Reis, M. C. (2009). Sheepskin effects and the relationship between earnings and education: Analyzing their evolution over time in Brazil. *Revista Brasileira de Economia*, 63:209–231.
- De La Croix, D. e Doepke, M. (2003). Inequality and growth: Why differential fertility matters. *American Economic Review*, 93(4):1091–1113.
- Esteves, L. A. (2009). Incompatibilidade escolaridade-ocupação e salários: Evidências de uma empresa industrial brasileira. *Revista Brasileira de Economia*, 63:77–90.
- Garcia, P. (2021). Skill prices and compositional effects on the declining wage inequality in Latin America: Evidence from Brazil. *Revista Brasileira de Economia*, 75(2).
- Garen, J. (1984). The returns to schooling: A selectivity bias approach with a continuous choice variable. *Econometrica: Journal of the Econometric Society*, pages 1199–1218.
- Gollin, D., Lagakos, D., e Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2):939–993.
- Hanushek, E. A. e Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American economic review*, 90(5):1184–1208.
- Heckman, J., Tobias, J. L., e Vytlačil, E. (2003). Simple estimators for treatment parameters in a latent-variable framework. *Review of Economics and Statistics*, 85(3):748–755.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161.
- Horie, N. e Iwasaki, I. (2022). Returns to schooling in european emerging markets: A meta-analysis. *Education Economics*.
- Jaeger, D. A. e Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *The review of economics and statistics*, pages 733–740.

- Kassouf, A. L. (1997). Retornos à escolaridade e ao treinamento nos setores urbano e rural do Brasil. *Revista de Economia e Sociologia Rural*, 35(2):59–76.
- Lumley, T. (2004). Analysis of complex survey samples. *Journal of statistical software*, 9(1):1–19.
- Marcelo, R. e Wyllie, R. (2006). Retornos para educação no Brasil: Evidências empíricas adicionais. *Economia aplicada*, 10:349–365.
- Marioni, L. d. S. (2021). Overeducation in the labour market: Evidence from Brazil. *Education Economics*, 29(1):53–72.
- Menezes Filho, N. A. (2003). *A evolução recente da educação no Brasil*. PhD thesis, Universidade de São Paulo.
- Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of political economy*, 66(4):281–302.
- Mincer, J. (1974). Schooling, experience, and earnings. *Human Behavior & Social Institutions*, (2).
- Monte, P. A. d., Ramalho, H. M. d. B., e Pereira, M. d. L. (2011). O salário de reserva e a oferta de trabalho: Evidências para o Brasil. *Economia Aplicada*, 15(4):613–639.
- Montenegro, C. E. e Patrinos, H. A. (2014). Comparable estimates of returns to schooling around the world. *World Bank policy research working paper*, (7020).
- Moura, R. L. d. (2008). Testando as hipóteses do modelo de mincer para o Brasil. *Revista brasileira de economia*, 62:407–449.
- Nakabashi, L. e Assahide, L. (2017). Estimando o retorno da escolaridade dos jovens por classe de renda: 1997-2012. *Pesquisa e Planejamento Econômico-PPE, Brasília*, 47(3):137–83.
- Ottoni, B. (2017). *Educação setores de atividade e produtividade*. Elsevier.
- Pereira, V. d. F., Lima, J. E. d., Lima, J. R. F. d., Braga, M. J., e Mendonça, T. G. d. (2013). Avaliação dos retornos à escolaridade para trabalhadores do sexo masculino no Brasil. *Revista de Economia Contemporânea*, 17:153–176.
- Perez-Silva, R. e Partridge, M. D. (2020). Concentration of human capital, externalities and the wage gap in us metro areas. *Regional Studies*, 54(11):1564–1573.
- Psacharopoulos, G. e Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature. *Education Economics*, 26(5):445–458.
- Reis, M. C. (2017). Educational mismatch and labor earnings in Brazil. *International Journal of Manpower*.

- Sachsida, A., Loureiro, P. R. A., e Mendonça, M. J. C. d. (2004). Um estudo sobre retorno em escolaridade no Brasil. *Revista Brasileira de Economia*, 58:249–265.
- Suliano, D. C. e Siqueira, M. L. (2012). Retornos da educação no Brasil em âmbito regional considerando um ambiente de menor desigualdade. *Economia Aplicada*, 16(1):137–165.
- Van Zaist, J. K., Nakabashi, L., e Salvato, M. A. (2010). Retornos privados da escolaridade no Paraná. *Revista EconomiA*.
- Wilson, R. A. e Briscoe, G. (2004). The impact of human capital on economic growth: A review. *Impact of Education and Training (Third report on vocational training research in Europe: background report)*, pages 9–70.
- Zhang, J., Zhao, Y., Park, A., e Song, X. (2005). Economic returns to schooling in urban china, 1988 to 2001. *Journal of comparative economics*, 33(4):730–752.

# Apêndice A Regressões estimadas

**Tabela A1:** Estimações do MQO, 1995-2004

	Variável Dependente								
	Logaritmo do salário hora								
	1995	1996	1997	1998	1999	2001	2002	2003	2004
Constante	-0,84***	-0,89***	-0,85***	-0,86***	-0,93***	-0,86***	-0,95***	-0,96***	-0,92***
S	0,12***	0,12***	0,11***	0,13***	0,12***	0,11***	0,11***	0,11***	0,11***
Raça	0,06***	0,04***	0,04***	0,02**	0,02**	-0,003	-0,01*	-0,01	-0,01
Urb	0,10***	0,12***	0,08***	0,10***	0,06***	0,05***	0,03**	0,02	-0,03**
Sexo	0,41***	0,41***	0,40***	0,40***	0,39***	0,36***	0,37***	0,39***	0,42***
Agr	-0,47***	-0,47***	-0,52***	-0,48***	-0,45***	-0,45***	-0,40***	-0,37***	-0,36***
Serv	-0,09***	-0,05***	-0,07***	-0,06***	-0,06***	-0,02*	-0,02	0,001	-0,02*
Exp	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***
Exp <sup>2</sup>	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***
SE	0,14***	0,19***	0,20***	0,22***	0,20***	0,16***	0,15***	0,17***	0,07***
NE	-0,20***	-0,16***	-0,19***	-0,14***	-0,15***	-0,20***	-0,20***	-0,20***	-0,31***
CO	0,03*	0,09***	0,12***	0,14***	0,12***	0,14***	0,15***	0,18***	0,09***
Sul	0,003	0,03*	0,06***	0,07***	0,04***	0,04***	0,08***	0,10***	0,02**
Sind	0,21***	0,22***	0,22***	0,22***	0,21***	0,20***	0,21***	0,21***	0,19***
S:Raça	0,02***	0,02***	0,02***	0,02***	0,03***	0,03***	0,03***	0,03***	0,03***
S:Urb	0,02***	0,02***	0,02***	0,01***	0,02***	0,02***	0,03***	0,02***	0,03***
S:Sexo	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,004***	0,004***	0,001
S:Agr	-0,03***	-0,02***	-0,01***	-0,03***	-0,02***	-0,02***	-0,02***	-0,02***	-0,01***
S:Serv	0,01***	0,01***	0,01***	0,005***	0,01***	0,004***	0,01***	0,004***	0,01***
S:Exp	-0,001***	-0,002***	-0,001***	-0,001***	-0,002***	-0,001***	-0,002***	-0,001***	-0,002***
Observações	90.224	86.401	92.475	91.400	94.398	101.485	105.616	105.111	112.696
R <sup>2</sup>	0,52	0,50	0,52	0,52	0,50	0,48	0,47	0,46	0,46
R <sup>2</sup> Ajustado	0,52	0,50	0,51	0,51	0,50	0,48	0,47	0,46	0,46
EPR	15,70	16,35	15,88	15,66	15,43	15,68	15,57	15,75	15,44
Estatística F	5.053,04***	4.544,71***	5.167,47***	5.107,92***	4.983,37***	4.926,99***	5.015,69***	4.767,78***	5.013,15***

Fonte: Elaborado pelos autores.

Notas: **S**: Anos de estudo; **Raça**: *dummy* com valor um caso o indivíduo seja branco e zero caso contrário, **Urb**: *dummy* com valor um caso more em áreas urbanas e zero caso contrário; **Sexo**: *dummy* com valor um caso seja do sexo masculino e zero caso contrário; **Agr**: *dummy* com valor um caso trabalhe no setor agrícola e zero caso contrário; **Serv**: *dummy* com valor um caso trabalhe no setor de serviços e zero caso contrário; **Exp**: tempo de experiência no mercado de trabalho; **Exp<sup>2</sup>**: experiência ao quadrado; **SE**, **NE**, **CO**, e **Sul** são *dummies* que tomam o valor um caso more na regiões Sudeste, Nordeste, Centro Oeste ou Sul, respectivamente, e zero caso contrário; **Sind**: *dummy* com valor um caso o indivíduo seja sindicalizado. As demais variáveis são interações entre os anos de estudo e as *dummies*; **EPR**: Erro padrão do resíduo. Asterisco simples (\*), duplo (\*\*) e triplo (\*\*\*) indicam significância a 10%, 5% e 1% respectivamente.

**Tabela A2:** Estimacões do MQO, 2005-2015

	Variável Dependente									
	Logaritmo do salário hora									
	2005	2006	2007	2008	2009	2011	2012	2013	2014	2015
Constante	-0,85***	-0,82***	-0,75***	-0,68***	-0,68***	-0,68***	-0,56***	-0,59***	-0,53***	-0,64***
S	0,12***	0,11***	0,11***	0,11***	0,11***	0,11***	0,11***	0,12***	0,12***	0,11***
Raça	-0,04***	-0,05***	-0,04***	-0,05***	-0,03***	-0,01	-0,04***	-0,05***	-0,06***	-0,07***
Urb	0,03***	0,01	0,03***	0,02*	0,04***	0,13***	0,11***	0,11***	0,08***	0,09***
Sexo	0,41***	0,42***	0,44***	0,46***	0,46***	0,45***	0,46***	0,49***	0,49***	0,46***
Agr	-0,35***	-0,41***	-0,37***	-0,41***	-0,44***	-0,53***	-0,58***	-0,61***	-0,60***	-0,55***
Serv	0,01	0,02*	0,04***	0,01	-0,02*	-0,08***	-0,07***	-0,04***	-0,04***	-0,02*
Exp	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***
Exp <sup>2</sup>	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***
SE	0,10***	0,10***	0,10***	0,11***	0,10***	0,16***	0,17***	0,14***	0,16***	0,21***
NE	-0,27***	-0,26***	-0,29***	-0,26***	-0,27***	-0,20***	-0,21***	-0,24***	-0,23***	-0,19***
CO	0,13***	0,14***	0,14***	0,14***	0,15***	0,19***	0,24***	0,22***	0,21***	0,26***
Sul	0,06***	0,06***	0,06***	0,09***	0,10***	0,11***	0,14***	0,12***	0,16***	0,20***
Sind	0,18***	0,17***	0,15***	0,16***	0,15***	0,12***	0,11***	0,12***	0,08***	0,08***
S:Raça	0,03***	0,03***	0,03***	0,03***	0,02***	0,02***	0,02***	0,02***	0,02***	0,02***
S:Urb	0,02***	0,02***	0,02***	0,02***	0,02***	0,01***	0,01***	0,02***	0,02***	0,02***
S:Sexo	0,0001	-0,001	-0,004***	-0,004***	-0,005***	-0,01***	-0,01***	-0,01***	-0,01***	-0,01***
S:Agr	-0,02***	-0,01***	-0,02***	-0,01***	-0,01***	-0,003	0,003	0,002	0,001	0,002
S:Serv	0,01***	0,01***	0,001	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***
S:Exp	-0,001***	-0,002***	-0,001***	-0,001***	-0,002***	-0,002***	-0,002***	-0,002***	-0,002***	-0,002***
Observações	117.663	120.197	118.829	119.713	122.558	109.955	113.613	113.523	117.927	113.405
R <sup>2</sup>	0,45	0,45	0,43	0,44	0,43	0,39	0,40	0,39	0,40	0,37
R <sup>2</sup> Ajustado	0,45	0,45	0,43	0,44	0,43	0,39	0,40	0,39	0,40	0,37
EPR	15,46	15,44	15,97	15,79	15,47	17,68	17,33	17,98	17,23	17,92
Estatística F	5.159,72***	5.214,15***	4.745,35***	4.864,98***	4.947,58***	3.627,45***	3.926,50***	3.823,00***	4.118,18***	3.573,21***

Fonte: Elaborado pelos autores.

Notas: **S**: Anos de estudo; **Raça**: *dummy* com valor um caso o indivíduo seja branco e zero caso contrário; **Urb**: *dummy* com valor um caso more em áreas urbanas e zero caso contrário; **Sexo**: *dummy* com valor um caso seja do sexo masculino e zero caso contrário; **Agr**: *dummy* com valor um caso trabalhe no setor agrícola e zero caso contrário; **Serv**: *dummy* com valor um caso trabalhe no setor de serviços e zero caso contrário; **Exp**: tempo de experiência no mercado de trabalho; **Exp<sup>2</sup>**: experiência ao quadrado; **SE**, **NE**, **CO**, e **Sul** são *dummies* que tomam o valor um caso more na regiões Sudeste, Nordeste, Centro Oeste ou Sul, respectivamente, e zero caso contrário; **Sind**: *dummy* com valor um caso o indivíduo seja sindicalizado. As demais variáveis são interações entre os anos de estudo e as *dummies*; **EPR**: Erro padrão do resíduo. Asterisco simples (\*), duplo (\*\*) e triplo (\*\*\*) indicam significância a 10%, 5% e 1% respectivamente.

**Tabela A3:** Estimações do Modelo de Heckman, 1995-2004

	Variável Dependente								
	Logaritmo do salário hora								
	1995	1996	1997	1998	1999	2001	2002	2003	2004
Constante	-0,88***	-0,90***	-0,91***	-0,99***	-0,96***	-0,99***	-1,04***	-1,12***	-1,06***
S	0,10***	0,10***	0,08***	0,09***	0,09***	0,08***	0,08***	0,08***	0,08***
Raça	0,03***	0,01	0,02	0,02	-0,01	-0,01	-0,01	0,03***	0,03***
Urb	0,22***	0,27***	0,17***	0,08**	0,17***	0,05*	0,02	-0,18***	-0,21***
Sexo	0,38***	0,36***	0,38***	0,42***	0,35***	0,37***	0,39***	0,47***	0,53***
Agr	-0,37***	-0,36***	-0,40***	-0,37***	-0,34***	-0,34***	-0,33***	-0,29***	-0,31***
Serv	0,03***	0,09***	0,07***	0,08***	0,08***	0,11***	0,07***	0,09***	0,05***
Exp	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***	0,06***
Exp <sup>2</sup>	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***
SE	0,14***	0,20***	0,20***	0,21***	0,20***	0,14***	0,13***	0,09***	0,004
NE	-0,20***	-0,14***	-0,18***	-0,14***	-0,14***	-0,21***	-0,21***	-0,26***	-0,38***
CO	0,02	0,11***	0,11***	0,13***	0,11***	0,12***	0,13***	0,15***	0,06***
Sul	-0,002	0,03**	0,07***	0,09***	0,06***	0,06***	0,09***	0,11***	0,05***
Sind	0,27***	0,28***	0,29***	0,30***	0,29***	0,27***	0,26***	0,27***	0,25***
S:Raça	0,02***	0,03***	0,02***	0,02***	0,02***	0,03***	0,03***	0,03***	0,03***
S:Urb	0,02***	0,01***	0,02***	0,01***	0,01***	0,02***	0,02***	0,02***	0,02***
S:Sexo	0,02***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***
S:Agr	-0,01***	0,0001	0,01***	0,001	0,01***	0,01***	0,01***	0,01**	0,01***
S:Serv	0,03***	0,03***	0,04***	0,04***	0,04***	0,03***	0,03***	0,03***	0,04***
S:Exp	-0,001***	-0,001***	-0,001***	-0,001***	-0,002***	-0,001***	-0,001***	-0,001***	-0,002***
$\lambda$	-1,50***	-1,59***	-0,87***	0,03	-1,00***	-0,09	0,08	1,48***	1,58***
Observações	94.652	91.582	98.585	98.328	102.336	109.473	113.642	113.782	120.968
R <sup>2</sup>	0,47	0,46	0,47	0,46	0,45	0,44	0,44	0,43	0,43
R <sup>2</sup> Ajustado	0,47	0,46	0,47	0,46	0,45	0,44	0,44	0,43	0,43
EPR	16,41	17,10	16,68	16,61	16,22	16,35	16,17	16,38	16,00
Estatística F	4.263,28***	3.881,74***	4.367,14***	4.234,42***	4.244,59***	4.337,68***	4.429,69***	4.223,06***	4.504,22***

Fonte: Elaborado pelos autores.

Notas: **S**: Anos de estudo; **Raça**: *dummy* com valor um caso o indivíduo seja branco e zero caso contrário; **Urb**: *dummy* com valor um caso more em áreas urbanas e zero caso contrário; **Sexo**: *dummy* com valor um caso seja do sexo masculino e zero caso contrário; **Agr**: *dummy* com valor um caso trabalhe no setor agrícola e zero caso contrário; **Serv**: *dummy* com valor um caso trabalhe no setor de serviços e zero caso contrário; **Exp**: tempo de experiência no mercado de trabalho; **Exp<sup>2</sup>**: experiência ao quadrado; **SE**, **NE**, **CO**, e **Sul** são *dummies* que tomam o valor um caso more na regiões Sudeste, Nordeste, Centro Oeste ou Sul, respectivamente, e zero caso contrário; **Sind**: *dummy* com valor um caso o indivíduo seja sindicalizado;  **$\lambda$** : razão inversa de Mills. As demais variáveis são interações entre os anos de estudo e as *dummies*; **EPR**: Erro padrão do resíduo. Asterisco simples (\*), duplo (\*\*) e triplo (\*\*\*) indicam significância a 10%, 5% e 1% respectivamente.



**Tabela A4:** Estimações do Modelo de Heckman, 2005-2015

	Variável Dependente									
	Logaritmo do salário hora									
	2005	2006	2007	2008	2009	2011	2012	2013	2014	2015
Constante	-0,99***	-1,07***	-1,03***	-0,95***	-1,06***	-1,05***	-0,94***	-1,08***	-0,95***	-1,18***
S	0,08***	0,08***	0,07***	0,07***	0,08***	0,09***	0,08***	0,08***	0,08***	0,07***
Raça	0,002	0,004	0,01	0,004	0,03***	0,02**	0,01	0,003	-0,02*	0,01
Urb	-0,14***	-0,22***	-0,22***	-0,18***	-0,19***	-0,06***	-0,09***	-0,12***	-0,15***	-0,32***
Sexo	0,51***	0,56***	0,62***	0,64***	0,67***	0,64***	0,66***	0,74***	0,70***	0,73***
Agr	-0,29***	-0,34***	-0,31***	-0,35***	-0,36***	-0,43***	-0,50***	-0,52***	-0,53***	-0,43***
Serv	0,08***	0,10***	0,10***	0,06***	0,07***	0,04***	0,01	0,05***	0,03***	0,11***
Exp	0,06***	0,06***	0,06***	0,06***	0,06***	0,07***	0,06***	0,06***	0,06***	0,06***
Exp <sup>2</sup>	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***
SE	0,05***	0,03***	0,08***	0,11***	0,12***	0,19***	0,22***	0,21***	0,20***	0,15***
NE	-0,33***	-0,34***	-0,36***	-0,32***	-0,32***	-0,25***	-0,29***	-0,28***	-0,26***	-0,30***
CO	0,09***	0,09***	0,12***	0,12***	0,19***	0,28***	0,34***	0,33***	0,35***	0,40***
Sul	0,09***	0,09***	0,15***	0,17***	0,19***	0,27***	0,29***	0,35***	0,38***	0,37***
Sind	0,24***	0,23***	0,21***	0,21***	0,21***	0,18***	0,16***	0,17***	0,14***	0,15***
S:Raça	0,03***	0,03***	0,03***	0,03***	0,03***	0,02***	0,02***	0,03***	0,03***	0,03***
S:Urb	0,02***	0,02***	0,02***	0,01***	0,02***	0,01***	0,01***	0,01***	0,01***	0,01***
S:Sexo	0,01***	0,01***	0,01***	0,01***	0,003***	0,002**	0,001	-0,003**	-0,0001	-0,0004
S:Agr	0,004**	0,01***	0,004*	0,01***	0,01***	0,01***	0,02***	0,02***	0,02***	0,03***
S:Serv	0,03***	0,03***	0,03***	0,03***	0,03***	0,03***	0,03***	0,03***	0,04***	0,04***
S:Exp	-0,001***	-0,001***	-0,001***	-0,001***	-0,001***	-0,002***	-0,002***	-0,002***	-0,002***	-0,002***
$\lambda$	1,38***	2,23***	2,60***	2,72***	2,49***	2,26***	3,03***	3,39***	3,45***	4,09***
Observações	126.504	128.325	127.173	126.851	131.335	116.225	119.545	119.914	124.459	122.666
R <sup>2</sup>	0,42	0,42	0,40	0,40	0,40	0,36	0,37	0,36	0,36	0,34
R <sup>2</sup> Ajustado	0,42	0,42	0,40	0,40	0,40	0,36	0,37	0,36	0,36	0,34
EPR	16,08	16,14	16,75	16,59	16,45	18,65	18,36	19,12	18,54	19,53
Estatística F	4.615,64***	4.644,84***	4.211,02***	4.298,97***	4.374,64***	3.268,70***	3.481,79***	3.395,63***	3.567,36***	3.169,07***

Fonte: Elaborado pelos autores.

Notas: **S**: Anos de estudo; **Raça**: *dummy* com valor um caso o indivíduo seja branco e zero caso contrário; **Urb**: *dummy* com valor um caso more em áreas urbanas e zero caso contrário; **Sexo**: *dummy* com valor um caso seja do sexo masculino e zero caso contrário; **Agr**: *dummy* com valor um caso trabalhe no setor agrícola e zero caso contrário; **Serv**: *dummy* com valor um caso trabalhe no setor de serviços e zero caso contrário; **Exp**: tempo de experiência no mercado de trabalho; **Exp<sup>2</sup>**: experiência ao quadrado; **SE**, **NE**, **CO**, e **Sul** são *dummies* que tomam o valor um caso more na regiões Sudeste, Nordeste, Centro Oeste ou Sul, respectivamente, e zero caso contrário; **Sind**: *dummy* com valor um caso o indivíduo seja sindicalizado;  **$\lambda$** : razão inversa de Mills. As demais variáveis são interações entre os anos de estudo e as *dummies*; **EPR**: Erro padrão do resíduo. Asterisco simples (\*), duplo (\*\*) e triplo (\*\*\*) indicam significância a 10%, 5% e 1% respectivamente.

**Tabela A5:** Estimações do modelo de Garen, 1995-2004

	Variável Dependente								
	Logaritmo do salário hora								
	1995	1996	1997	1998	1999	2001	2002	2003	2004
Constante	-0,14***	-0,15***	-0,08*	0,05	-0,05	-0,05	-0,08	-0,20***	-0,10**
S	0,02***	0,03***	0,02**	0,01	0,02**	0,01*	0,02**	0,04***	0,03***
Raça	0,04*	-0,01	-0,002	0,04*	-0,03	-0,03	-0,08***	-0,05*	-0,07***
Urb	-0,08***	-0,13***	-0,12***	-0,15***	-0,22***	-0,16***	-0,19***	-0,22***	-0,28***
Sexo	0,37***	0,39***	0,36***	0,36***	0,36***	0,32***	0,33***	0,40***	0,47***
Agr	-0,38***	-0,47***	-0,52***	-0,55***	-0,47***	-0,40***	-0,36***	-0,32***	-0,35***
Serv	0,01	0,01	-0,03	-0,01	0,002	0,07***	0,12***	0,21***	0,17***
$\eta$	0,05***	0,06***	0,05***	0,05***	0,04***	0,04***	0,02***	0,02***	0,02***
Exp	0,03***	0,03***	0,03***	0,02***	0,03***	0,02***	0,02***	0,02***	0,02***
Exp <sup>2</sup>	-0,0004***	-0,0004***	-0,0004***	-0,0003***	-0,0004***	-0,0003***	-0,0003***	-0,0003***	-0,0003***
SE	0,13***	0,17***	0,18***	0,21***	0,17***	0,15***	0,14***	0,16***	0,06***
NE	-0,20***	-0,16***	-0,19***	-0,14***	-0,15***	-0,19***	-0,19***	-0,19***	-0,29***
CO	0,02	0,07***	0,11***	0,13***	0,10***	0,13***	0,14***	0,17***	0,09***
Sul	-0,01	0,01	0,05***	0,06***	0,02	0,04***	0,07***	0,09***	0,02*
Sind	0,22***	0,23***	0,23***	0,23***	0,22***	0,21***	0,22***	0,22***	0,20***
S:Raça	0,02***	0,03***	0,03***	0,02***	0,03***	0,03***	0,03***	0,03***	0,03***
S:Urb	0,07***	0,07***	0,07***	0,07***	0,07***	0,07***	0,06***	0,06***	0,07***
S:Sexo	0,02***	0,01***	0,02***	0,01***	0,01***	0,01***	0,01***	0,002	-0,01***
S:Agr	-0,04***	-0,01***	-0,01	-0,004	-0,01**	-0,02***	-0,01***	-0,02***	-0,004
S:Serv	-0,01**	-0,001	0,002	-0,001	-0,001	-0,01**	-0,01***	-0,02***	-0,02***
S: $\eta$	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***
S:Exp	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***
Observações	90.224	86.401	92.475	91.400	94.398	101.485	105.616	105.111	112.696
R <sup>2</sup>	0,51	0,50	0,51	0,51	0,50	0,48	0,47	0,46	0,45
R <sup>2</sup> Ajustado	0,51	0,50	0,51	0,51	0,50	0,48	0,47	0,46	0,45
EPR	15,74	16,40	15,91	15,72	15,47	15,72	15,61	15,78	15,50
Estatística F	4.534,50***	4.061,48***	4.635,12***	4.551,19***	4.455,37***	4.411,10***	4.488,15***	4.273,74***	4.468,02***

Fonte: Elaborado pelos autores.

Notas: **S**: Anos de estudo; **Raça**: *dummy* com valor um caso o indivíduo seja branco e zero caso contrário, **Urb**: *dummy* com valor um caso more em áreas urbanas e zero caso contrário; **Sexo**: *dummy* com valor um caso seja do sexo masculino e zero caso contrário; **Agr**: *dummy* com valor um caso trabalhe no setor agrícola e zero caso contrário; **Serv**: *dummy* com valor um caso trabalhe no setor de serviços e zero caso contrário;  **$\eta$** : resíduo da equação de escolaridade (equação 3); **Exp**: tempo de experiência no mercado de trabalho; **Exp<sup>2</sup>**: experiência ao quadrado; **SE**, **NE**, **CO**, e **Sul** são *dummies* que tomam o valor um caso more na regiões Sudeste, Nordeste, Centro Oeste ou Sul, respectivamente, e zero caso contrário; **Sind**: *dummy* com valor um caso o indivíduo seja sindicalizado. As demais variáveis são interações entre os anos de estudo e as *dummies*; **EPR**: Erro padrão do resíduo. Asterisco simples (\*), duplo (\*\*) e triplo (\*\*\*) indicam significância a 10%, 5% e 1% respectivamente.

**Tabela A6:** Estimações do Modelo de Garen, 2005-2015

	Variável Dependente									
	Logaritmo do salário hora									
	2005	2006	2007	2008	2009	2011	2012	2013	2014	2015
Constante	-0,04	-0,01	0,11**	0,16***	0,25***	0,37***	0,43***	0,40***	0,45***	0,45***
S	0,04***	0,04***	0,03***	0,03***	0,03***	0,03***	0,03***	0,04***	0,05***	0,03***
Raça	-0,12***	-0,19***	-0,12***	-0,11***	-0,10***	-0,12***	-0,11***	-0,17***	-0,17***	-0,16***
Urb	-0,22***	-0,18***	-0,26***	-0,26***	-0,32***	-0,22***	-0,16***	-0,15***	-0,22***	-0,25***
Sexo	0,46***	0,47***	0,50***	0,53***	0,59***	0,61***	0,60***	0,63***	0,68***	0,68***
Agr	-0,28***	-0,44***	-0,46***	-0,49***	-0,59***	-0,73***	-0,79***	-0,79***	-0,82***	-0,78***
Serv	0,21***	0,24***	0,25***	0,23***	0,17***	0,15***	0,11***	0,19***	0,18***	0,19***
$\eta$	0,02***	0,01***	0,01*	0,004*	-0,001	0,001	0,0000	-0,01**	-0,01***	-0,02***
Exp	0,02***	0,02***	0,02***	0,02***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***
Exp <sup>2</sup>	-0,0003***	-0,0003***	-0,0003***	-0,0003***	-0,0002***	-0,0002***	-0,0002***	-0,0002***	-0,0001***	-0,0002***
SE	0,09***	0,10***	0,09***	0,10***	0,10***	0,14***	0,16***	0,13***	0,14***	0,20***
NE	-0,26***	-0,24***	-0,27***	-0,24***	-0,24***	-0,19***	-0,19***	-0,23***	-0,21***	-0,18***
CO	0,13***	0,14***	0,13***	0,14***	0,15***	0,18***	0,24***	0,21***	0,21***	0,26***
Sul	0,06***	0,06***	0,06***	0,08***	0,10***	0,10***	0,14***	0,12***	0,15***	0,20***
Sind	0,19***	0,18***	0,17***	0,17***	0,16***	0,13***	0,12***	0,13***	0,10***	0,09***
S:Raça	0,03***	0,04***	0,03***	0,03***	0,03***	0,03***	0,03***	0,03***	0,03***	0,03***
S:Urb	0,06***	0,05***	0,06***	0,06***	0,07***	0,05***	0,04***	0,04***	0,05***	0,05***
S:Sexo	-0,01***	-0,01***	-0,01***	-0,01***	-0,02***	-0,03***	-0,03***	-0,03***	-0,03***	-0,04***
S:Agr	-0,02***	0,002	0,01	0,01**	0,02***	0,03***	0,04***	0,03***	0,04***	0,04***
S:Serv	-0,02***	-0,02***	-0,03***	-0,02***	-0,02***	-0,02***	-0,01***	-0,02***	-0,02***	-0,02***
S: $\eta$	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***	0,01***
S:Exp	0,002***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***
Observações	117.663	120.197	118.829	119.713	122.558	109.955	113.613	113.523	117.927	113.405
R <sup>2</sup>	0,45	0,45	0,43	0,43	0,43	0,38	0,39	0,39	0,39	0,37
R <sup>2</sup> Ajustado	0,45	0,45	0,43	0,43	0,43	0,38	0,39	0,39	0,39	0,37
EPR	15,50	15,48	16,01	15,83	15,51	17,74	17,38	18,04	17,29	17,96
Estatística F	4.618,32***	4.660,37***	4.246,59***	4.339,33***	4.430,00***	3.227,86***	3.501,55***	3.402,52***	3.658,80***	3.192,06***

Fonte: Elaborado pelos autores.

Notas: **S**: Anos de estudo; **Raça**: *dummy* com valor um caso o indivíduo seja branco e zero caso contrário, **Urb**: *dummy* com valor um caso more em áreas urbanas e zero caso contrário; **Sexo**: *dummy* com valor um caso seja do sexo masculino e zero caso contrário; **Agr**: *dummy* com valor um caso trabalhe no setor agrícola e zero caso contrário; **Serv**: *dummy* com valor um caso trabalhe no setor de serviços e zero caso contrário;  **$\eta$** : resíduo da equação de escolaridade (equação 3); **Exp**: tempo de experiência no mercado de trabalho; **Exp<sup>2</sup>**: experiência ao quadrado; **SE**, **NE**, **CO**, e **Sul** são *dummies* que tomam o valor um caso more na regiões Sudeste, Nordeste, Centro Oeste ou Sul, respectivamente, e zero caso contrário; **Sind**: *dummy* com valor um caso o indivíduo seja sindicalizado. As demais variáveis são interações entre os anos de estudo e as *dummies*; **EPR**: Erro padrão do resíduo. Asterisco simples (\*), duplo (\*\*) e triplo (\*\*\*) indicam significância a 10%, 5% e 1% respectivamente.

## Apêndice B Estatísticas descritivas

**Tabela B1:** Média das variáveis utilizadas na pesquisa

Anos	Y	S	Exp	Fam	Filhos	Raça	Sind	Urb	Sexo	Mãe	Agr	Ind	Serv
1995	6,60	5,52	27,32	4,09	59,35	53,41	17,94	78,98	60,28	68,45	22,30	22,01	55,69
1996	7,14	5,78	27,20	4,07	57,39	54,08	17,79	79,74	60,79	68,53	21,07	21,97	56,96
1997	7,02	5,84	27,15	4,00	56,83	53,20	17,16	79,38	60,78	68,82	21,04	21,99	56,97
1998	7,01	5,99	27,23	3,93	55,38	52,89	16,98	79,65	60,54	68,94	19,95	22,22	57,83
1999	6,33	6,03	27,32	3,90	54,29	52,83	16,64	79,26	59,89	68,73	20,84	21,34	57,82
2001	6,19	6,30	27,00	3,79	52,52	51,33	16,63	83,47	60,02	69,86	18,50	21,79	59,71
2002	5,85	6,48	27,02	3,78	51,04	50,83	17,22	83,62	59,54	70,04	18,81	23,69	57,50
2003	5,96	6,62	26,94	3,73	49,73	49,67	17,92	83,23	59,47	70,30	19,10	23,13	57,76
2004	5,84	6,71	26,84	3,69	49,15	48,82	17,99	81,71	59,11	70,43	19,70	23,17	57,13
2005	6,07	6,88	26,73	3,65	48,00	47,40	18,53	81,59	58,82	70,75	19,31	23,31	57,38
2006	6,39	7,09	26,63	3,60	47,33	47,06	18,67	82,04	58,48	70,98	18,33	23,51	58,17
2007	6,80	7,22	26,60	3,60	45,75	46,25	17,58	82,44	58,65	70,82	17,56	23,96	58,49
2008	6,66	7,44	26,56	3,55	43,92	45,52	18,11	82,85	58,32	70,69	16,95	24,60	58,45
2009	6,87	7,62	26,43	3,49	43,28	45,22	17,58	83,18	58,16	71,59	16,48	24,11	59,41
2011	8,76	7,83	26,43	3,46	41,73	44,97	16,84	84,55	59,04	71,65	15,84	23,67	60,50
2012	9,11	8,12	26,28	3,42	40,62	43,45	16,51	84,95	58,68	72,09	14,60	24,31	61,09
2013	10,25	8,22	26,45	3,37	39,51	43,06	15,62	84,54	58,61	71,81	14,61	24,57	60,81
2014	9,44	8,34	26,58	3,32	38,60	42,48	16,14	84,96	57,94	71,54	14,26	24,17	61,57
2015	10,33	8,54	26,59	3,29	37,76	41,88	18,55	84,70	58,31	71,93	13,93	23,48	62,60

Elaborado pelos autores.

Notas: **Y**: Salário hora médio; **S**: Média dos anos de estudo; **Exp**: Experiência; **Fam**: Número médio de moradores nos domicílios; **Filhos**: % de indivíduos com filhos menores de 14 anos; **Raça**: % de indivíduos brancos; **Sind**: % de trabalhadores sindicalizados; **Urb**: % de Moradores na área urbana; **Sexo**: % de Homens; **Mãe**: % de indivíduos que tem a mãe viva; **Agr**: % de trabalhadores da agricultura; **Ind**: % de trabalhadores da indústria; **Serv**: % de trabalhadores dos serviços.

**Tabela B2:** Estatísticas descritivas dos retornos da escolaridade

Escolaridade	Modelo	Mínimo	Q25	Mediana	Q75	Máximo	Média	Desvio p.
0 a 4	Garen	0,0536	0,0637	0,0742	0,0800	0,0936	0,0733	0,0123
0 a 4	Heckman	0,0455	0,0585	0,0739	0,0810	0,0924	0,0700	0,0147
0 a 4	MQO	0,0546	0,0664	0,0805	0,0890	0,0956	0,0773	0,0137
5 a 8	Garen	0,0854	0,0974	0,1168	0,1257	0,1415	0,1129	0,0183
5 a 8	Heckman	0,0666	0,0798	0,0978	0,1065	0,1185	0,0938	0,0162
5 a 8	MQO	0,0801	0,0905	0,1081	0,1176	0,1242	0,1043	0,0151
9 a 11	Garen	0,1197	0,1321	0,1570	0,1699	0,1825	0,1525	0,0226
9 a 11	Heckman	0,0843	0,0954	0,1114	0,1200	0,1303	0,1083	0,0144
9 a 11	MQO	0,0991	0,1068	0,1222	0,1301	0,1357	0,1189	0,0127
12 ou mais	Garen	0,1590	0,1745	0,1979	0,2154	0,2281	0,1961	0,0238
12 ou mais	Heckman	0,0983	0,1073	0,1229	0,1296	0,1367	0,1192	0,0128
12 ou mais	MQO	0,1105	0,1166	0,1309	0,1375	0,1415	0,1276	0,0113

Fonte: Elaborado pelos autores.

## Apêndice C Teste de Wald

Queremos testar se os parâmetros da Equação (1) associados aos retornos da escolaridade estimados com dados 1995 são iguais aos de 2015 em cada respectivo método que utilizamos. Implementar tal teste é um meio formal de verificar se de fato os retornos da escolaridade diferem entre os anos de 1995 e 2015. Um dos meios de se fazer isso é utilizando o Teste de Wald, que por sua vez, requer a matriz de variância-covariância dos parâmetros. Se estimarmos separadamente a Equação (1) para 1995 e 2015 testar a igualdade dos parâmetros dependeria da suposição de que esses parâmetros têm covariância zero.

Logo, nossa estratégia consiste em estimar ambas equações em conjunto. Ou seja, empilhamos  $\ln Y_{i,1995}$  e  $\ln Y_{i,2015}$  um em cima do outro e fazemos algo semelhante com as matrizes de variáveis explicativas  $\mathbf{X}_{1995}$  e  $\mathbf{X}_{2015}$ . Logo, o sistema a ser estimado é o seguinte:

$$\begin{pmatrix} \ln Y_{1995} \\ \ln Y_{2015} \end{pmatrix} = \begin{pmatrix} \theta_{1995} \\ \theta_{2015} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{1995} & \mathbf{0}_{1995} \\ \mathbf{0}_{2015} & \mathbf{X}_{2015} \end{pmatrix} + \begin{pmatrix} \epsilon_{1995} \\ \epsilon_{2015} \end{pmatrix}, \quad (11)$$

onde  $\mathbf{0}_{1995}$  ( $\mathbf{0}_{2015}$ ) é uma matriz de zeros com o mesmo número de linhas de  $\mathbf{X}_{1995}$  ( $\mathbf{X}_{2015}$ ) e o mesmo número de colunas de  $\mathbf{X}_{2015}$  ( $\mathbf{X}_{1995}$ ) e  $\epsilon_{1995}$  e  $\epsilon_{2015}$  são os termos de erro. Além disso, adicionamos uma *dummy* que toma o valor 1 caso o ano seja 2015 no intuito de capturar os diferentes interceptos das equações de salários estimadas para 1995 e 2015. Ao proceder dessa maneira, conseguimos calcular a matriz de variância-covariância dos parâmetros, que por sua vez, permite testar, via Teste de Wald, a igualdade de dois ou mais coeficientes.

A hipótese nula do teste é de que os coeficientes associados a estimação dos retornos da escolaridade de 1995 e 2015 são iguais,  $H_0 : \beta^{2015} = \beta^{1995}$ . Os resultados podem ser vistos na Tabela C1,

e, em todos os três métodos rejeita-se a hipótese nula.

**Tabela C1:** Teste de Wald

Método	Estatística F	Valor P
MQO	1136,97	0,0
Heckman	1298,62	0,0
Garen	273,35	0,0

Elaborado pelos autores.