

Métodos Multivariados de Análise de Dados*

Draft

Alberson da Silva Miranda

21 de janeiro de 2025

*Código disponível em https://github.com/albersonmiranda/analise_multivariada.

SUMÁRIO

1	INTRODUÇÃO	3
2	REFERENCIAL TEÓRICO	4
3	METODOLOGIA	6
4	RESULTADOS PRELIMINARES	8
5	REFERÊNCIAS	9

1 INTRODUÇÃO

[...]

2 REFERENCIAL TEÓRICO

A primeira abordagem quantitativa, baseada em modelagem estatística, para evidenciar o poder de explicação da escolaridade sobre a renda do trabalhador foi realizada no trabalho seminal de [Mincer \(1974\)](#). O autor propôs um modelo de capital humano, onde a escolaridade é um dos principais determinantes da renda – especificamente, do logaritmo da renda. A especificação que ficou conhecida como *equação de Mincer* é a seguinte:

$$\ln Y = \beta_0 + \beta_1 S + \beta_2 E + \beta_3 S^2 + \beta_4 E^2 + \varepsilon \quad (2.1)$$

em que Y é a renda, S é a escolaridade (medida em anos), E é o tempo de experiência (estimada como a idade subtraída dos anos de escolaridade e da constante 6, considerada a idade para início da formação) e ε é o erro aleatório. O modelo de Mincer foi amplamente utilizado e adaptado em diversos estudos subsequentes, tornando-se uma referência na literatura sobre capital humano.

A partir de dados do censo norte-americano, ele estimou diversas especificações de modelos estatísticos, concluindo que aqueles que utilizaram a escolaridade e tempo de experiência como variáveis dependentes foram capazes de explicar até 33% da variação da renda do trabalhador¹, com o coeficiente de retorno médio de escolaridade (β_1 na Equação 2.1) de até 0.16, o que significa um incremento médio de 17,4% na renda por ano de escolaridade².

Desde então, diversos estudos têm sido realizados para estimar os efeitos da escolaridade sobre a renda do trabalhador, cada um estendendo ou aplicando o modelo de Mincer em diferentes contextos. [Psa-charopoulos e Patrinos * \(2004\)](#) estimam, em média, cerca de 10% de incremento de renda para cada ano de estudo. [Colclough, Kingdon e Patrinos \(2010\)](#) mostram, a partir de dados de 34 países, que até a década de 1990 os retornos a cada nível de escolaridade eram decrescentes, mas que, a partir de então, o retorno a cada ano extra de estudo aumenta a cada incremento no nível de escolaridade. Neste mesmo trabalho, os autores demonstram que os retornos em todos os níveis de escolaridade (primário, secundário e terciário) estão reduzindo ao longo das décadas, mas que a redução é mais acentuada no nível primário.

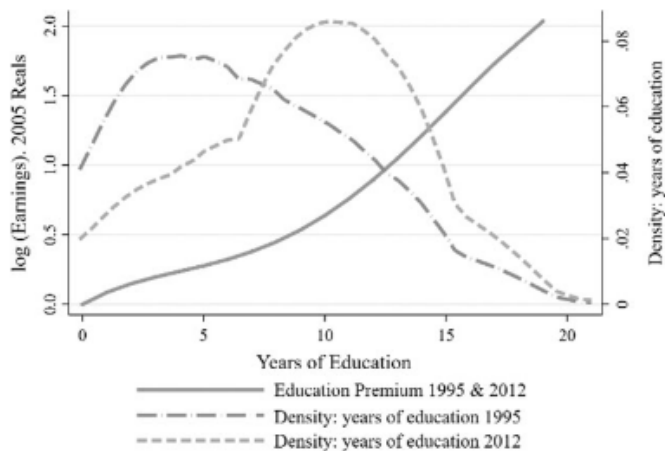
[Ferreira, Firpo e Messina \(2022\)](#) analisam a desigualdade salarial no Brasil no período entre 1995-2012 e mostram que os dois principais fatores que aumentam a desigualdade são 1) a maior disparidade de salários entre diferentes setores da economia, e; 2) o chamado “paradoxo do progresso”, o efeito intensificador da desigualdade quando há aumento da educação da população em uma sociedade em

¹Medidos pelo coeficiente de determinação R^2 .

²Como o modelo utiliza o log da renda, para se computar o efeito em moeda deve-se realizar a operação inversa, resultando em $e^{0.16} = 1.1735$.

que os retornos à educação são convexos, ou seja, aumentam exponencialmente a cada aumento do nível educacional (Figura 2.1). Esse fenômeno é paradoxal no sentido de que se espera que a educação reduza a desigualdade, mas, em sociedades em que os retornos associados aos níveis superiores de educação são muito altos em relação aos menores, a escolaridade tem efeito contrário, de intensificação das desigualdades.

Figura 2.1: Paradoxo do progresso no Brasil.



Fonte: Ferreira et al. (2022).

Nesse sentido, [Altonji, Blom e Meghir \(2012\)](#) aponta que a escolha do curso superior é de grande impacto na determinação da renda, destacando que a diferença no retorno médio entre alguns cursos superiores, como engenharia elétrica e pedagogia, é quase tão grande quanto a diferença média entre indivíduos com ensino médio e ensino superior. [Ophem e Mazza \(2024\)](#) reforçam essa ideia, mostrando que a escolha do curso superior é um dos principais elementos não apenas na renda inicial quanto também na progressão salarial ao longo da carreira.

3 METODOLOGIA

Neste trabalho, de ordem quantitativa, utilizo os dados da Relação Anual de Informações Sociais (Rais) de 2006 e 2022 para estimar os efeitos da escolaridade sobre a renda dos trabalhadores do estado do Espírito Santo. Para tanto, utilizo o *data lake* tratado e disponibilizado gratuitamente pelo projeto Base dos Dados (CAVALCANTE, 2022). O acesso, manipulação dos dados e a análise foram realizados com o *software* R (R CORE TEAM, 2024) e o repositório com todo o código realizado aqui está disponível publicamente e pode ser reproduzido em sua totalidade¹.

As variáveis de interesse extraídas da Rais foram:

1. renda média nominal naquele ano
2. ciclo de escolaridade
3. idade
4. raça/cor
5. sexo

Após selecionadas, apliquei condições às variáveis para obter amostra completa, ou seja, sem valores faltantes, e coerente. Essas condições estão resumidas na tabela a seguir. Elas implicam na restrição às entradas com renda média positiva não nula; na exclusão de entradas sem quaisquer dos campos escolaridade, raça/cor ou sexo preenchidos.

¹https://github.com/albersonmiranda/analise_multivariada.

Tabela 3.1: Possíveis valores para as variáveis selecionadas da Rais

Variável	Valores
Sigla UF	ES
Renda Média Nominal	Númericos, não negativos
Ciclo de Escolaridade	Analfabeto, Ensino Fundamental (I/II, completo/incompleto), Ensino Médio (completo/incompleto), Ensino Superior (completo/incompleto), Mestrado ou Doutorado
Idade	Sem restrições
Raça/Cor	Branco, Preto, Pardo, Indígena ou Amarelo
Sexo	Masculino ou Feminino

Fonte: Elaborado pelo Autor (2025)

Além das condições de interesse do pesquisador, é necessário atentar que a Rais trata do mercado de trabalho formal, o que exclui trabalhadores informais e profissionais autônomos. Portanto, o presente trabalho mira estimar as relações escolaridade-renda no mercado de trabalho formal do Espírito Santo, destacando o impacto de substratos marginalizados da sociedade na determinação da renda do trabalhador.

Seguindo [Mincer \(1974\)](#), mas adaptando à nossa Lei de Diretrizes e Bases da Educação Nacional (LDB), faremos uma aproximação para o tempo de experiência de cada indivíduo a partir da seguinte fórmula:

$$\text{exp} = \text{idade} - \text{anos de formação} - 6 \quad (3.1)$$

Em que supõe-se:

- Idade de 6 anos para ingresso no ensino fundamental
- 5 anos para o fundamental I
- 4 anos para o fundamental II
- 3 anos para o ensino médio
- 4 anos para o superior
- 2 anos para o mestrado
- 4 anos par ao doutorado

Assim, por exemplo, para um indivíduo de 30 anos com mestrado, temos $30 - 18 - 6 = 6$ anos de experiência.

4 RESULTADOS PRELIMINARES

[...]

5 REFERÊNCIAS

ALTONJI, J. G.; BLOM, E.; MEGHIR, C. Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers. **Annual Review of Economics**, v. 4, n. 1, p. 185–223, 1 set. 2012. ISSN 1941-1383, 1941-1391. DOI: [10.1146/annurev-economics-080511-110908](https://doi.org/10.1146/annurev-economics-080511-110908). Disponível em: <https://www.annualreviews.org/doi/10.1146/annurev-economics-080511-110908>. Acesso em: 1 mai. 2024. Citado na p. 5.

CAVALCANTE, P. **basedosdados: 'Base Dos Dados' R Client**. [S.l.: s.n.], 2022. Disponível em: <https://CRAN.R-project.org/package=basedosdados>. Citado na p. 6.

COLCLOUGH, C.; KINGDON, G.; PATRINOS, H. The Changing Pattern of Wage Returns to Education and its Implications. **Development Policy Review**, v. 28, n. 6, p. 733–747, 2010. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-7679.2010.00507.x>. ISSN 1467-7679. DOI: [10.1111/j.1467-7679.2010.00507.x](https://doi.org/10.1111/j.1467-7679.2010.00507.x). Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-7679.2010.00507.x>. Acesso em: 27 abr. 2024. Citado na p. 4.

FERREIRA, F. H. G.; FIRPO, S. P.; MESSINA, J. Labor Market Experience and Falling Earnings Inequality in Brazil: 1995–2012. **The World Bank Economic Review**, v. 36, n. 1, p. 37–67, 2 fev. 2022. ISSN 0258-6770. DOI: [10.1093/wber/lhab005](https://doi.org/10.1093/wber/lhab005). Disponível em: <https://doi.org/10.1093/wber/lhab005>. Acesso em: 27 abr. 2024. Citado na p. 4.

MINCER, J. **Schooling, experience, and earnings**. New York: Columbia University Press, 1974. 152 p. (Human behavior and social institutions, 2). ISBN 978-0-87014-265-9. Citado nas pp. 4, 7.

OPHEM, H. van; MAZZA, J. Educational choice, initial wage and wage growth. **Empirical Economics**, 21 mar. 2024. ISSN 1435-8921. DOI: [10.1007/s00181-024-02580-5](https://doi.org/10.1007/s00181-024-02580-5). Disponível em: <https://doi.org/10.1007/s00181-024-02580-5>. Acesso em: 27 abr. 2024. Citado na p. 5.

PSACHAROPOULOS, G.; PATRINOS *, H. A. Returns to investment in education: a further update. **Education Economics**, v. 12, n. 2, p. 111–134, 1 ago. 2004. ISSN 0964-5292, 1469-5782. DOI: [10.1080/0964529042000239140](https://doi.org/10.1080/0964529042000239140). Disponível em: <https://www.tandfonline.com/doi/full/10.1080/0964529042000239140>. Acesso em: 17 out. 2024. Citado na p. 4.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2024. Disponível em: <https://www.R-project.org/>. Citado na p. 6.