

DESIGUALDADE DE RENDA E EDUCAÇÃO NAS REGIÕES METROPOLITANAS DO BRASIL: uma análise de dados em painel entre 2013 e 2023

Maurício Almeida Meira ¹

RESUMO:

O objetivo deste artigo é analisar a relação entre desigualdade de renda e educação nas regiões metropolitanas do Brasil, utilizando dados da PNAD entre 2013 e 2023. A pesquisa investiga o impacto da educação sobre a desigualdade de renda, utilizando dados em painel. As variáveis de interesse incluem o Índice de Gini, anos de estudo, taxa de analfabetismo, e renda média habitual, além de variáveis de controle como idade média, proporção de homens e mulheres, proporção de pessoas brancas e pretas, e proporção de pessoas na força de trabalho. Os resultados indicam que a renda média habitual é o principal fator associado ao aumento da desigualdade, enquanto as variáveis educacionais não tiveram efeito significativo. Foi identificada heterocedasticidade e autocorrelação serial, corrigidas com erros-padrão robustos. Conclui-se que a educação, por si só, não é um determinante direto da desigualdade de renda nas regiões metropolitanas brasileiras. Políticas públicas devem ir além da educação, incluindo redistribuição de renda e melhoria na qualidade educacional, para reduzir a desigualdade.

Palavras-chave: Desigualdade de Renda; Educação; Regiões metropolitanas.

1. INTRODUÇÃO

A desigualdade é um problema crônico nas regiões metropolitanas brasileiras, envolvendo variadas questões sociais e econômicas, sobretudo no que tange à disparidades no acesso à educação e saúde de qualidade. O Brasil tem uma tendência de enorme desigualdade na distribuição de renda e elevados níveis de pobreza, devido a traços históricos com uma carga de injustiça social que exclui parte significativa de sua população do acesso a condições básicas de qualidade de vida. As disparidades no acesso à educação, saúde e infraestrutura perpetuam ciclos de pobreza e exclusão social, impactando diretamente a qualidade de vida e o desenvolvimento socioeconômico do país.

Embora a literatura reconheça a importância da educação como fator de redução da desigualdade, a complexa relação entre esses dois elementos exige uma análise mais aprofundada, considerando as particularidades regionais e as dinâmicas temporais. Os estudos

sobre a desigualdade, com o passar dos anos, passou a ser analisado do ponto vista multidimensional, sendo incorporado aspectos da educação, saúde e longevidade, e não somente tendo como base a renda. Um dos exemplos é a inclusão do índice de desenvolvimento humano (IDH) em 1990. Conforme Oliveira (2016), esse indicador promoveu novos caminhos de estudos, principalmente aqueles de impacto direto à sociedade (saúde, educação e renda).

No que concerne à educação, tomando como base os estudos já existentes sobre o tema, percebe-se que a educação têm sido um dos principais fatores que explicam a desigualdade de renda no Brasil. Prova disso, são os estudos de Castro (2006), Souza et al. (2014) e Araújo e Marinho (2015), os quais mostraram que a educação se relaciona com a desigualdade de renda de maneira inversa, uma vez que quanto maior o nível educacional de uma determinada localidade, maior a produtividade do fator trabalho de seus indivíduos e, assim, maior a cota salarial que cada um recebe, melhorando a sua posição social e, assim, fazendo com que haja um aumento na renda média dessa localidade.

Nesse contexto, este artigo se propõe a analisar a relação entre desigualdade de renda e educação nas regiões metropolitanas do Brasil, utilizando uma abordagem de dados em painel que permite investigar a dinâmica da desigualdade ao longo do tempo. A partir de dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) entre 2013 e 2023, o estudo busca responder à seguinte questão: qual o impacto da educação sobre a desigualdade de renda nas regiões metropolitanas brasileiras? A análise econométrica com dados em painel, combinada com a inclusão de variáveis de controle relevantes, visa aprofundar a compreensão sobre os determinantes da desigualdade de renda e fornecer subsídios para a formulação de políticas públicas eficazes de combate à desigualdade e promoção da justiça social.

Para alcançar os objetivos, o artigo foi estruturado em quatro partes, além desta introdução. Inicialmente discute a literatura e evidências recentes sobre o tema. Na sequência versa sobre a metodologia e a definição dos dados utilizados, seguido da análise e discussão dos resultados da estimação do modelo. Por fim, são tecidas as observações finais, a fim de se concluir o presente estudo.

2. REVISÃO TEÓRICA E EVIDÊNCIAS

A desigualdade de renda é um fenômeno multifacetado, influenciado por diversos fatores socioeconômicos. Conforme destacado por Carvalho e Souza (2020), a pesquisa sobre

desigualdade de renda evoluiu ao longo do tempo, expandindo seu escopo para além da distribuição funcional da renda e incorporando aspectos como mensuração, efeitos econômicos e políticos, e causas institucionais.

Dentre os fatores que impactam a desigualdade, a educação tem papel central. Diversos estudos, como os de Souza et al. (2014) e Araújo e Marinho (2015), demonstram a relação inversa entre educação e desigualdade de renda, sugerindo que o aumento do nível educacional contribui para reduzir a disparidade de rendimentos.

Além disso, é importante destacar que, conforme Batista e Mollo (2021), a desigualdade deve ser observada em uma perspectiva multidimensional, de forma a compreender como a desigualdade se comporta frente a alterações em uma determinada dimensão socioeconômica, sobretudo os níveis educacionais. É crucial reconhecer que a desigualdade de renda é um problema complexo, influenciado por uma série de fatores além da educação. Variáveis como idade, sexo, raça, setor de atividade, taxa de ocupação e programas de transferência de renda, também desempenham papel importante na determinação da distribuição de renda. Em sequência, nesta seção será posta, em subseções, as literaturas que relacionam a desigualdade com a educação, e em seguida a relação desse fenômeno com a renda per capita.

2.1. A relação entre desigualdade e educação

De modo geral, há perspectivas que defendem a escolaridade como o principal fator determinante da desigualdade, por ser responsável por gerar diferenças de produtividade entre os indivíduos, as quais tendem a se manter ao longo de todo o ciclo de vida. Em contrapartida, existem abordagens que destacam o acesso à propriedade, como terra e capital, e as políticas salariais definidas pelo governo como elementos fundamentais para explicar a desigualdade.

Em conformidade com o estudo de Souza et al. (2014), a educação é caracterizada como um investimento, pois possui capacidade de gerar aumento da produtividade do fator trabalho. Eles citam a abordagem da teoria do capital humano, a qual diz que um aumento do nível educacional de uma população pode implicar em ganhos de produtividade do fator trabalho e refletir aumentos, por sua vez, nas rendas médias. Conforme Araújo e Marinho (2015), quanto maior for o grau de produtividade maior será a cota de renda que a pessoa receberá e melhor será sua posição social contribuindo para reduzir as desigualdades.

Tratando-se especificamente do Brasil, baseando-se na literatura já existente, nas evidências empíricas de outros autores e nos testes econométricos realizados, o estudo de Souza et al. (2014), mostrou que existe uma relação inversa entre desigualdade de renda e educação no país, uma vez que um ano adicional de escolaridade no país implica em valorização salarial elevada o que, aliado à escassez de mão de obra qualificada, contribui para concentração de renda pessoal, pois existe uma desigualdade na escolarização da população o que tem como consequência uma menor oferta de mão de obra qualificada fazendo com que haja uma maior dispersão salarial entre os grupos de trabalhadores com diferentes níveis de qualificação. Assim sendo, o estudo corrobora com a abordagem da teoria do capital humano.

De acordo com Araújo e Marinho (2015), de forma geral, os estudos sobre a relação entre desigualdade de renda e educação destacam que no Brasil a desigualdade é, em grande parte, resultado da péssima distribuição educacional existente, tanto em termos pessoais como entre grupos de indivíduos com características similares.

Desse modo, com base na literatura já existente que versa sobre a relação entre educação e desigualdade de renda, percebe-se que essa relação é inversa, uma vez que quanto maior o nível educacional de uma determinada localidade, maior a produtividade do fator trabalho de seus indivíduos e, assim, maior a cota salarial que cada um recebe, melhorando a sua posição social e, assim, fazendo com que haja um aumento na renda média dessa localidade, reduzindo assim a desigualdade de rendimentos entre os indivíduos.

2.2. A relação entre desigualdade e renda per capita

A relação entre desigualdade de renda e renda per capita é complexa e tem sido objeto de debate na literatura econômica. Uma das hipóteses mais conhecidas é a curva de Kuznets (1955), que postula uma relação em forma de "U invertido" entre esses dois indicadores. Segundo essa hipótese, nos estágios iniciais do desenvolvimento econômico, a desigualdade tende a aumentar com o crescimento da renda per capita, mas a partir de certo ponto, a desigualdade começa a diminuir.

No entanto, a validade da hipótese de Kuznets tem sido questionada por diversos estudos empíricos, que encontraram resultados divergentes. Alguns estudos apontam para a existência de uma relação em U invertido, enquanto outros não encontram evidências que confirmem essa hipótese.

É importante destacar que a relação entre desigualdade e renda per capita pode ser influenciada por diversos fatores, como a estrutura produtiva, as políticas públicas e as instituições de cada país ou região. A análise dessa relação requer, portanto, a consideração de diferentes dimensões e a utilização de métodos adequados para controlar por fatores não observados.

Diante da complexidade do tema, este estudo contribui para a literatura ao analisar a relação entre desigualdade de renda e renda per capita nas regiões metropolitanas do Brasil, utilizando uma abordagem de dados em painel que permite controlar por heterogeneidade não observada e investigar a dinâmica da desigualdade ao longo do tempo.

3. METODOLOGIA E DADOS

Este trabalho propõe uma análise da desigualdade de renda nas regiões metropolitanas do Brasil, utilizando dados em painel da Pesquisa Nacional por Amostra de Domicílios (PNAD) entre os anos de 2013 e 2023. Segundo Wooldridge, a abordagem por meio de dados em painel combina elementos de dados de séries temporais e de corte transversal, observando as mesmas unidades ao longo do tempo. Ela permite estudar mudanças dentro das unidades e entre elas. É útil para controlar por características não observadas e estudar efeitos dinâmicos.

Além disso, a estimação foi feita a partir de um modelo de efeitos fixos e também de um modelo de efeitos aleatórios. De acordo com Wooldridge (2006), o método de efeitos fixos é usado para eliminar o viés causado por fatores não observados constantes no tempo. Já o modelo de efeitos aleatórios é baseado na suposição de que o efeito não observado não é correlacionado com as variáveis explicativas. Desse modo, o modelo de efeitos fixos controla a heterogeneidade não observada através da inclusão de um efeito fixo para cada unidade de corte transversal, enquanto o modelo de efeitos aleatórios considera o efeito fixo como parte do termo erro. Posteriormente, seguindo Wooldridge (2006), realizamos o teste de Hausman, que indicou o modelo de efeitos aleatórios como mais eficiente para a análise.

As informações coletadas foram: índice de Gini, renda média habitual, anos de estudo, e taxa de analfabetismo. Além dessas, foram selecionadas as seguintes variáveis de controle: idade média da população, proporção de homes e mulheres na população, proporção de pessoas brancas e pretas, e proporção de pessoas na força de trabalho. A coleta foi feita a partir de dados

disponíveis na Pesquisa Nacional por Amostra de Domicílios (PNAD) para as regiões metropolitanas brasileiras entre os anos de 2013 e 2023.

Tabela 1 - Variáveis do Modelo

Notação	Variáveis
	Índice de Gini
	Anos de estudo
	Proporção de Analfabetismo
	Renda média habitual
	Idade média da população
	Proporção de homens na população
	Proporção de pessoas brancas
	Proporção de pessoas pretas
	Proporção de pessoas na força de trabalho
	Erro aleatório

Fonte: Elaborada dos autores com base em dados da PNAD

3.1. Procedimentos metodológicos

Para analisar a relação entre desigualdade de renda e educação nas regiões metropolitanas do Brasil, este estudo utiliza um modelo de efeitos aleatórios com dados em painel. A escolha do modelo de efeitos aleatórios se justifica pelos resultados do teste de Hausman, que indicou ser este o modelo mais eficiente.

Para o presente caso, a especificação do modelo de efeitos aleatórios é:

$$\ln GINI_{it} = \beta_0 + \beta_1 \ln ESTUDO_{it} + \beta_2 \ln ANALFABETISMO_{it} + \beta_3 \ln RENDA_{it} + \beta_4 \ln IDADE_{it} + \beta_5 \ln HOMENS_{it} + \beta_6 \ln BRANCAS_{it} + \beta_7 \ln PRETAS_{it} + \beta_8 \ln FORÇA_{it} + a_i + u_{it}$$

onde, $\ln GINI$ é a variável dependente, explicada pelas variáveis independentes $\ln ESTUDO$, $\ln ANALFABETISMO$ e $\ln RENDA$, e as demais variáveis de controle. Os indicadores foram observados em sua forma logarítmica natural (\ln), com o intuito de auferir melhor ajustamento dos dados ao modelo estimado. Outrossim, em consonância com o lecionado por Hoffman (2016), o coeficiente β_0 corresponde ao intercepto, ou seja, a resposta para variável explicada quando as variáveis explicativas se igualam a zero. Ademais, os coeficientes $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$

e β_8 representam a proporção do quanto o índice de gini se altera, caso seja acrescido um ponto no valor da respectiva variável. Por fim, a_i representa o efeito fixo individual para a região metropolitana i e u_{it} é o termo erro idiossincrático.

Para garantir a robustez da inferência estatística, foram realizados testes de diagnóstico para verificar a presença de heterocedasticidade e autocorrelação serial nos resíduos de ambos os modelos, de efeitos fixos e de efeitos aleatórios. Seguindo Wooldridge (2006), a heterocedasticidade foi verificada através do teste de Breusch-Pagan. A presença de heterocedasticidade viola uma das hipóteses do modelo de regressão e pode levar a estimativas ineficientes dos erros padrão. Já a autocorrelação serial foi verificada através do teste de Breusch-Godfrey/Wooldridge para dados em painel. A autocorrelação serial também viola uma das hipóteses do modelo e pode gerar estimativas inconsistentes dos erros padrão.

3.2. Variáveis e Base de dados

Conforme dito anteriormente, a agenda de pesquisa em desigualdade avançou muito ao longo dos anos, com a preocupação na mensuração da mesma. Assim, várias medidas de desigualdade foram sendo construídas. Dentre elas, as principais são: coeficiente de Gini, mais conhecido por índice de Gini, coeficiente de concentração, classe de medidas de Atkinson, índices T de Theil e L de Theil, dentre outros. O presente trabalho tomou como base o índice de Gini como medida de desigualdade, para entender a relação entre a mesma e a educação, sendo esta última representada por: anos de estudos e proporção de analfabetismo.

De acordo com Medeiros (2012) o coeficiente de Gini é uma medida que sintetiza o nível de desigualdade de uma distribuição de renda em um único número, sendo o indicador mais conhecido e usado de desigualdade. Conforme o Atlas do Desenvolvimento Humano no Brasil, o índice de Gini é um instrumento usado para medir o grau de concentração de renda, o qual aponta a diferença entre os rendimentos dos mais pobres e dos mais ricos e, numericamente, varia de 0 a 1, sendo que 0 representa a situação de total igualdade, ou seja, todos têm a mesma renda, e o valor 1 significa completa desigualdade de renda, ou seja, uma só pessoa detém toda a renda do lugar. Para o presente trabalho, o índice de Gini foi calculado a partir da renda domiciliar per capita.

Para analisar o impacto da educação sobre a desigualdade de renda, foram utilizadas as seguintes variáveis:

- **Anos de estudo:** Representa a média de anos de estudo da população com 14 anos ou mais de idade em cada região metropolitana. Essa variável busca capturar o nível de escolaridade da população, que é um importante determinante da produtividade do trabalho e da renda individual.
- **Proporção de analfabetismo:** Representa a proporção de pessoas com 14 anos ou mais de idade que não sabem ler e escrever em cada região metropolitana. O analfabetismo é um indicador de exclusão social e de baixa qualificação da mão de obra, o que pode contribuir para a desigualdade de renda.

Além das variáveis de educação, foram incluídas no modelo as seguintes variáveis de controle:

- **Idade média da população:** A idade média da população pode influenciar a desigualdade de renda, uma vez que faixas etárias mais jovens e mais velhas tendem a ter rendimentos mais baixos.
- **Renda média habitual:** Representa a renda média habitualmente recebida pelas pessoas com 14 anos ou mais de idade em cada região metropolitana. Essa variável busca controlar pelo efeito da renda média na desigualdade, uma vez que regiões com maior renda média podem apresentar menor desigualdade.
- **Proporção de homens e mulheres:** A proporção de homens e mulheres na população pode refletir diferenças na participação no mercado de trabalho e nos níveis de renda entre os sexos.
- **Raça:** A variável raça busca controlar por possíveis diferenças na desigualdade de renda entre grupos raciais, que podem estar associadas a fatores históricos e sociais.
- **Proporção de pessoas na força de trabalho:** A proporção de pessoas na força de trabalho em relação à população total pode influenciar a desigualdade de renda, uma vez que uma maior participação no mercado de trabalho pode estar associada a uma distribuição mais equitativa da renda.

A inclusão dessas variáveis de controle visa isolar o efeito da educação sobre a desigualdade de renda, controlando por outros fatores que podem influenciar a distribuição de renda nas regiões metropolitanas.

3.3. Ferramenta de análise

A análise de dados deste estudo foi conduzida utilizando a linguagem de programação R, em um fluxo de trabalho que combinou os ambientes Google Colab e Jupyter Notebook (via Anaconda) para otimizar as etapas de coleta e análise. A coleta de dados foi realizada por meio de web scraping dos microdados da Pesquisa Nacional por Amostra de Domicílios (PNAD) Contínua para o período de 2013 a 2023, usando as linguagens R e SQL. Para esta tarefa, foi empregado o ambiente Google Colab, aproveitando seus recursos de nuvem para lidar com o grande volume de dados. O processo foi automatizado com o pacote PNADcIBGE, que gerencia a complexidade de acesso e download dos arquivos diretamente dos repositórios do IBGE. O pacote googledrive foi utilizado para armazenar os dados brutos coletados em nuvem, garantindo a integridade e a portabilidade dos dados. Posteriormente, a limpeza, o tratamento e toda a análise econométrica foram desenvolvidos no ambiente Jupyter Notebook, escolhido por sua interatividade e capacidade de organização da pesquisa. Para a manipulação e agregação dos dados, foi utilizado o pacote dplyr. A estimação dos modelos de dados em painel, incluindo efeitos fixos e aleatórios, foi realizada com o pacote plm. Os testes de diagnóstico, como o Teste de Hausman e os testes para heterocedasticidade e autocorrelação, foram conduzidos com o pacote lmtest, e as correções de erros-padrão robustos (Driscoll-Kraay) foram aplicadas com o auxílio do pacote sandwich. Finalmente, a formatação das tabelas de resultados para apresentação no artigo foi feita com os pacotes stargazer e kableExtra.

4. ANÁLISE E DISCUSSÃO DOS RESULTADOS

Nesta seção, apresentamos os resultados da análise econométrica, organizados em quatro subseções. Primeiramente, são discutidas as estatísticas descritivas das variáveis utilizadas

(Seção 4.1.1). Em seguida, são apresentados os resultados do modelo de efeitos fixos robusto (Seção 4.1.2) e do Teste de Hausman utilizado para a escolha do modelo adequado (Seção 4.1.3). Depois, são apresentados os resultados do modelo de efeitos aleatórios robusto (Seção 4.1.4). Por fim, é feita uma análise socioeconômica para os resultados do modelo escolhido (Seção 4.2).

4.1 ANÁLISE DESCRITIVA DOS RESULTADOS ECONÔMETRICOS

Esta seção aborda os resultados da análise econométrica. Serão apresentadas as estatísticas descritivas, os resultados do modelo de efeitos fixos robusto, do Teste de Hausman e os resultados do modelo de efeitos aleatórios robusto.

4.1.1 Estatísticas Descritivas

A Tabela 2 apresenta as estatísticas descritivas das variáveis analisadas neste estudo, calculadas a partir dos dados da PNAD para as regiões metropolitanas do Brasil entre os anos de 2013 e 2023.

Tabela 2 - Estatística descritiva das variáveis analisadas

Variáveis	Média	Desvio Padrão	Valor Mínimo	Valor Máximo
GINI	0,5837	0,0324	0,48	0,66
ESTUDO	11,7926	3,8003	0,00	16,00
ANALFABETISMO	0,0217	0,0185	0,00	0,09
RENDIA	2636,7909	4061,3754	2,00	300.000,00

Fonte: Elaborada dos autores com base em dados da PNAD

Como podemos observar, o Índice de Gini médio para as regiões metropolitanas é de 0,5837, com um desvio padrão de 0,0324. O valor mínimo observado para o Índice de Gini foi de 0,48, enquanto o valor máximo foi de 0,66. Esses valores indicam que há variação na desigualdade de renda entre as regiões metropolitanas, mas que a desigualdade é, em geral, alta.

A média de anos de estudo da população é de 11,7926 anos, com um desvio padrão de 3,8003 anos. O valor mínimo observado para os anos de estudo foi de 0 anos, o que pode indicar a presença de pessoas que nunca frequentaram a escola, enquanto o valor máximo foi de 16 anos, o que sugere que algumas pessoas possuem nível de escolaridade elevado.

A taxa de analfabetismo média é de 0,0217, com um desvio padrão de 0,0185. O valor mínimo observado para a taxa de analfabetismo foi de 0%, enquanto o valor máximo foi de 9%. Esses valores indicam que a taxa de analfabetismo é relativamente baixa nas regiões metropolitanas, mas que ainda há variação entre as regiões.

A renda média habitual da população é de R\$ 2.636,79, com um desvio padrão de R\$ 4.061,38. O valor mínimo observado para a renda foi de R\$ 2,00, enquanto o valor máximo foi de R\$ 300.000,00. A alta variabilidade na renda, evidenciada pelo desvio padrão elevado, sugere que há grande disparidade nos rendimentos entre os indivíduos nas regiões metropolitanas.

4.1.2 Resultados do Modelo de Efeitos Fixos robustos

Nesta seção, são discutidos os resultados do Modelo de Efeitos Fixos estimado com correção para heterocedasticidade e autocorrelação serial utilizando erros-padrão robustos pelo método de Driscoll-Kraay. A decisão por essa abordagem decorre de uma estimação inicial, na qual foram detectados problemas de heterocedasticidade e autocorrelação serial. A correção robusta foi implementada para garantir a confiabilidade dos coeficientes e das inferências estatísticas. A tabela 3 apresenta os resultados estimados para o modelo de efeitos fixos robustos:

Tabela 3: Resultados do Modelo de Efeitos Fixos robustos

Variável	Coefficiente	Erro-Padrão Robusto	Valor t	p-valor	Significância
ln_media_anos_estudo	-0.02005367	0.15712564	-0.1276	0.89857	Não signif.
ln_proporcao_analfabetismo	0.00742058	0.00999088	0.7427	0.45850	Não signif.
ln_media_rendimento_habitual	0.09271453	0.04584541	2.0223	0.04445	* (p<0.05)
media_idade	0.00060319	0.00756470	0.0797	0.93652	Não signif.
proporcao_homens	0.25797330	0.20569949	1.2541	0.21124	Não signif.
proporcao_raca_branca	0.02975746	0.07622662	0.3904	0.69666	Não signif.
proporcao_raca_preta	0.10549773	0.10324334	1.0218	0.30807	Não signif.

Observações: 231

R² Ajustado: 0.194

Estatística F: 11.743*** (df = 7; 203) | Significância Global: p<0.01

Elaborada pelos autores utilizando os dados da PNAD

O modelo apresentou um R² ajustado de 0.194, indicando que aproximadamente 19,4% da variação do Índice de Gini entre as regiões metropolitanas analisadas é explicada pelas variáveis independentes incluídas. A estatística F (11.743; p<0.01) sugere que o

modelo, como um todo, é estatisticamente significativo ao nível de 1%, confirmando que as variáveis incluídas têm poder explicativo conjunto.

A variável Log da Renda Média Habitual apresenta um coeficiente positivo e significativo ($\beta = 0.093$; $p < 0.05$), sugerindo que um aumento de 1% na renda média habitual está associado a um aumento de aproximadamente 0.093% no Índice de Gini. O resultado é relevante porque destaca que o crescimento da renda média não está necessariamente acompanhado por uma redução na desigualdade, possivelmente indicando que os ganhos não são distribuídos de forma equitativa entre as camadas sociais.

Com relação a variável Log da Média de Anos de Estudo, seu coeficiente estimado é negativo (-0.020), mas estatisticamente não significativo ($p = 0.899$). Isso indica que, mesmo com um efeito esperado negativo (mais anos de estudo levando a menor desigualdade), os dados não suportam tal relação no contexto analisado. No que concerne a variável Log da Proporção de Analfabetismo, embora o coeficiente seja positivo (0.007), ele não é estatisticamente significativo ($p = 0.459$). A ausência de significância sugere que a proporção de analfabetos, isoladamente, não está fortemente associada à variação do Índice de Gini no modelo robusto.

Já as variáveis Proporção de Homens e Idade Média, têm coeficientes positivos (0.258 e 0.001, respectivamente), mas não são estatisticamente significativas. Isso indica que, no contexto das regiões metropolitanas, a proporção de homens e a idade média não explica, de maneira robusta, as variações na desigualdade de renda. Por fim, para as variáveis Proporção de Raça Branca e Proporção de Raça Preta, os coeficientes são positivos (0.030 e 0.105, respectivamente), mas também não apresentam significância estatística. Isso sugere que, embora a composição racial possa ser relevante para outros fenômenos sociais, seu impacto direto sobre a desigualdade de renda não é evidenciado nos dados após a correção dos problemas de heterocedasticidade e autocorrelação.

Os resultados sugerem que a renda média habitual é a principal variável explicativa associada ao Índice de Gini nas regiões metropolitanas brasileiras. O impacto positivo e significativo dessa variável reforça a importância de se discutir a distribuição dos ganhos de renda, uma vez que o crescimento econômico, por si só, não implica em redução das desigualdades.

Por outro lado, variáveis relacionadas à educação (como média de anos de estudo e analfabetismo) e composição demográfica (idade, gênero e raça) não se mostraram

estatisticamente significativas no modelo robusto, possivelmente devido à presença de efeitos não observados ou à necessidade de considerar interações mais complexas entre essas variáveis.

Por fim, o modelo robusto (Driscoll-Kraay) aprimorou as inferências estatísticas ao corrigir problemas de heterocedasticidade e autocorrelação serial detectados na estimação inicial. Desse modo, renda média habitual se destaca como o principal fator associado ao aumento do Índice de Gini, enquanto demais variáveis, embora relevantes do ponto de vista teórico, não apresentaram significância estatística nos dados.

4.1.3 Teste Hausman para escolha do modelo

O Teste de Hausman é utilizado para determinar a escolha entre o modelo de efeitos fixos e o modelo de efeitos aleatórios. De acordo com Wooldridge (2006), o teste verifica se os efeitos individuais não observados estão correlacionados com as variáveis explicativas. A hipótese nula do teste assume que não há correlação entre e as variáveis explicativas, ou seja, o modelo de efeitos aleatórios é consistente e mais eficiente. Caso a hipótese nula seja rejeitada, o modelo de efeitos fixos deve ser preferido, pois fornece estimativas consistentes mesmo na presença de correlação. A tabela 4 mostra os resultados do teste:

Tabela 4: Resultados do Teste de Hausman

Teste	Estatística Qui-Quadrado (χ^2)	Graus de Liberdade (df)	p-valor	Hipótese Nula	Interpretação
Hausman Test	9.7297	7	0.2044	O modelo de Efeitos Aleatórios é consistente	Não rejeita $H_0 \rightarrow$ Efeitos Aleatórios preferido

Elaborada pelos autores utilizando os dados da PNAD

Para este trabalho, o Teste de Hausman resultou em uma estatística qui-quadrado de 9.7297, com 7 graus de liberdade e um p-valor de 0.2044. Como o p-valor é maior que 0.05, não rejeitamos a hipótese nula de que o modelo de efeitos aleatórios é consistente. Isso significa que, com base no teste, não há evidências estatísticas de que os efeitos individuais estejam correlacionados com as variáveis explicativas. Portanto, o modelo de efeitos

aleatórios é o mais apropriado para analisar os dados em questão, sendo mais eficiente do ponto de vista econométrico.

4.1.4 Resultados do Modelo de Efeitos Aleatórios robustos

Nesta seção, analisamos os resultados do Modelo de Efeitos Aleatórios estimado com correção para heterocedasticidade e autocorrelação serial utilizando erros-padrão robustos pelo método de Driscoll-Kraay. A aplicação dessa correção foi necessária após a estimação inicial, na qual os testes de diagnóstico revelaram a presença de heterocedasticidade (Teste de Breusch-Pagan) e autocorrelação serial (Teste de Wooldridge). A violação dessas hipóteses poderia levar a erros-padrão subestimados e inferências estatísticas incorretas, tornando imprescindível a utilização de erros robustos. A tabela 5 apresenta os resultados estimados para o modelo de efeitos aleatórios robustos:

Tabela 5: Resultados do Modelo de Efeitos Aleatórios robustos

Variável	Coefficiente	Erro-Padrão Robusto	Valor t	p-valor	Significância
Intercept (Constante)	-1.2918	0.3209	-4.0257	7.785e-05	*** (p<0.01)
ln_media_anos_estudo	-0.0747	0.1489	-0.5015	0.6165	Não signif.
ln_proporcao_analfabetismo	0.0147	0.0100	1.4791	0.1405	Não signif.
ln_media_rendimento_habitual	0.0982	0.0578	1.6973	0.0910	* (p<0.10)
media_idade	0.0030	0.0083	0.3596	0.7195	Não signif.
proporcao_homens	0.1864	0.1827	1.0203	0.3087	Não signif.
proporcao_raca_branca	-0.0027	0.0267	-0.0999	0.9205	Não signif.
proporcao_raca_preta	0.0942	0.0746	1.2626	0.2080	Não signif.

Observações: 231

R² Ajustado: 0.2611

Significância Global (F-Statistic): p<0.01

Elaborada pelos autores utilizando os dados da PNAD

O modelo apresentou um R^2 ajustado de 0.2611, indicando que cerca de 26,1% da variação no logaritmo do Índice de Gini é explicada pelas variáveis independentes incluídas no modelo. A estatística F apresentou significância global ao nível de 1%, confirmando que as variáveis, em conjunto, possuem poder explicativo sobre a desigualdade de renda medida pelo Índice de Gini.

O coeficiente estimado para a constante é -1.2918 ($p < 0.01$), sendo altamente significativo. Esse valor representa o nível base do logaritmo do Índice de Gini na ausência das variáveis explicativas. Com relação a variável Log da Renda Média Habitual, foi positivo e marginalmente significativo (0.0982, $p < 0.10$), indicando que um aumento de 1% na renda média habitual está associado a um aumento de aproximadamente 0.098% no Índice de Gini. Esse resultado é consistente com a literatura econômica, que destaca que o crescimento da renda média, sem políticas redistributivas, pode ampliar a desigualdade, uma vez que os ganhos tendem a se concentrar em determinados grupos.

No que concerne ao Log da Média de Anos de Estudo foi negativo (-0.0747) e não significativo ($p = 0.6165$). Embora a teoria sugira que um aumento nos anos de estudo contribui para a redução da desigualdade ao elevar a produtividade e os rendimentos, a falta de significância pode indicar limitações estruturais, como disparidades na qualidade da educação ou outras variáveis omitidas. Já o Log da Proporção de Analfabetismo foi positivo (0.0147), mas não significativo ($p = 0.1405$). O sinal positivo está alinhado com a teoria econômica, uma vez que maiores taxas de analfabetismo refletem baixa qualificação da força de trabalho, contribuindo para maior desigualdade. Contudo, a falta de significância sugere que o efeito direto dessa variável sobre o Índice de Gini não foi robustamente capturado.

Com relação as Variáveis de Controle (Idade Média, Proporção de Homens, Raça Branca e Raça Preta), a Idade Média apresentou um coeficiente muito próximo de zero (0.0030) e não significativo ($p = 0.7195$), sugerindo que não há efeito relevante dessa variável sobre a desigualdade no modelo. Já a Proporção de Homens e Proporção de Raça Preta apresentaram coeficientes positivos, mas também não significativos (0.1864 e 0.0942, respectivamente). Por fim, a Proporção de Raça Branca apresentou um coeficiente negativo (-0.0027), mas sem significância estatística ($p = 0.9205$). Esses resultados indicam que, individualmente, as variáveis de controle demográficas não possuem efeitos estatisticamente robustos sobre a desigualdade de renda no contexto analisado.

4.2 ANÁLISE SOCIECONÔMICA DOS RESULTADOS DO MODELO ESCOLHIDO

A análise econométrica realizada permitiu compreender a relação entre educação e desigualdade de renda nas regiões metropolitanas do Brasil. Com base nos resultados do modelo de efeitos aleatórios robusto, o Índice de Gini foi impactado de forma significativa apenas pela renda média habitual, enquanto as variáveis educacionais e demográficas não mostraram efeitos relevantes. A partir disso, podemos refletir sobre as implicações socioeconômicas desses resultados, especialmente em relação à interação entre educação, desigualdade e a distribuição da renda.

Primeiramente, o coeficiente positivo da renda média habitual (0.0982, $p < 0.10$) indica que o aumento da renda média, embora possa gerar benefícios econômicos gerais, também está associado a aumento da desigualdade. Esse resultado está em linha com a teoria econômica, que sugere que os ganhos de renda podem se concentrar em setores ou grupos sociais específicos, ampliando as disparidades de rendimento. Em um contexto de desigualdade estrutural, como o observado no Brasil, a renda média pode ser um reflexo da falta de redistribuição, uma vez que os setores mais altos da renda tendem a se beneficiar mais do crescimento econômico, em detrimento dos mais pobres.

Por outro lado, a ausência de significância estatística das variáveis relacionadas à educação, como anos de estudo e taxa de analfabetismo, sugere que, neste estudo específico, a educação formal não é um fator determinante direto para a redução da desigualdade de renda nas regiões metropolitanas analisadas. Embora a teoria sugira que a expansão educacional seja um caminho crucial para a equidade, os resultados indicam que as disparidades educacionais regionais e a qualidade da educação podem ser fatores limitantes que não são adequadamente capturados pelas variáveis de anos de estudo e analfabetismo. Assim, um aumento quantitativo na escolaridade pode não ser suficiente para reduzir a desigualdade sem uma melhoria na qualidade e acesso à educação de qualidade para todas as camadas sociais.

Adicionalmente, a variabilidade na renda observada (com um desvio padrão elevado) reforça a ideia de que há uma grande disparidade de oportunidades nas regiões metropolitanas do Brasil. A alta disparidade de rendimentos, com valores extremos entre os mais pobres e mais ricos, sugere que, para promover uma verdadeira redução da desigualdade, não basta apenas focar no aumento da renda média. Medidas complementares, como políticas de redistribuição da

renda e acesso equitativo à educação e serviços públicos, são essenciais para reduzir essas disparidades. Isso reforça a necessidade de políticas públicas focadas em justiça social, capazes de distribuir os benefícios do crescimento de forma mais equitativa.

Finalmente, os resultados não significativos das variáveis relacionadas a gênero e raça indicam que, no modelo analisado, essas variáveis não têm efeito direto sobre a desigualdade de renda, mas isso não significa que não sejam fatores importantes na dinâmica socioeconômica. De fato, estudos mais aprofundados poderiam explorar a interseccionalidade dessas variáveis, ou seja, como a combinação de fatores como raça, gênero e classe social pode influenciar a desigualdade de maneira mais complexa. As políticas públicas, portanto, precisam ser sensíveis a essas nuances e buscar ações que combinem a educação de qualidade, acesso a melhores empregos e diminuição da discriminação racial e de gênero.

Em resumo, embora os resultados econométricos indiquem que a renda média habitual seja o principal fator associado ao aumento da desigualdade de renda, as variáveis educacionais e demográficas não apresentaram o efeito esperado. Isso sugere que as políticas públicas no Brasil devem ir além do aumento da escolaridade e abordar questões estruturais mais amplas, como a qualidade da educação, acesso a oportunidades e a distribuição equitativa da renda. O estudo aponta, assim, para a importância de políticas que integrem crescimento econômico inclusivo com estratégias de redistribuição e melhoria das condições educacionais para todos os cidadãos, com foco especialmente nas regiões mais vulneráveis

5. CONSIDERAÇÕES FINAIS

Este estudo buscou analisar a relação entre educação e desigualdade de renda nas regiões metropolitanas do Brasil, utilizando uma abordagem de dados em painel com a Pesquisa Nacional por Amostra de Domicílios (PNAD), abrangendo o período de 2013 a 2023. Os resultados evidenciam que, embora a educação seja um fator relevante para o desenvolvimento socioeconômico, a renda média foi o principal determinante associado ao aumento da desigualdade de renda, confirmando que o crescimento econômico, por si só, pode não ser suficiente para reduzir as disparidades sociais. A análise também apontou que as variáveis educacionais e demográficas, como anos de estudo e taxa de analfabetismo, não apresentaram

um impacto direto sobre a desigualdade nas regiões metropolitanas, sugerindo que fatores estruturais, como a qualidade da educação e o acesso desigual a oportunidades, devem ser mais bem abordados em futuras políticas públicas.

Os resultados do modelo de efeitos aleatórios indicaram que, apesar da significância da renda média, as demais variáveis analisadas, como anos de estudo, analfabetismo, gênero e raça, não tiveram um impacto direto significativo sobre o Índice de Gini. Isso sugere que as disparidades educacionais e sociais não estão sendo adequadamente capturadas pelas variáveis utilizadas, o que pode ser resultado de fatores não observados ou da necessidade de uma abordagem mais aprofundada que considere as complexidades dessas variáveis em um contexto de desigualdade histórica e estrutural. Além disso, a grande variabilidade na renda observada, com disparidades extremas entre os mais ricos e os mais pobres, reforça a ideia de que a desigualdade nas regiões metropolitanas brasileiras é multifacetada e exige ações que transcendam a simples expansão educacional.

Em termos de implicações políticas, os resultados sugerem que o Brasil precisa adotar políticas públicas mais eficazes, não só para promover o crescimento econômico, mas para garantir que os benefícios desse crescimento sejam distribuídos de forma mais equitativa. Para isso, é crucial implementar políticas redistributivas, melhorar a qualidade da educação e garantir acesso igualitário a oportunidades de emprego e capacitação. A combinação dessas estratégias pode, a longo prazo, ajudar a reduzir a desigualdade de renda e promover uma sociedade mais justa e igualitária. O estudo também destaca a importância de explorar a interseccionalidade das variáveis sociais, como gênero e raça, para entender melhor os mecanismos que perpetuam a desigualdade, indicando que a luta pela igualdade social precisa ser mais inclusiva e abrangente.

REFERÊNCIAS

ARAÚJO, Jair Andrade de; MARINHO, Emerson. **ESTUDO SOBRE A DESIGUALDADE DE RENDA E SEUS DETERMINANTES NO BRASIL**. Revista de Políticas Públicas, [s. l.], v. 19, ed. 2, 2015. Disponível em: <https://www.redalyc.org/>. Acesso em: 23 Setembro 2024.

ATLAS BRASIL. **Atlas do Desenvolvimento Humano no Brasil**. Consulta. Disponível em: <http://www.atlasbrasil.org.br/>. Acesso em: 04 Novembro 2024.

BARROS, Ricardo Paes de. *et al.* **A Estabilidade Inaceitável: Desigualdade e Pobreza no Brasil**. IPEA. Rio de Janeiro, jun, 2001.

BARROS, Ricardo Paes de. *et al.* **Desigualdade e Pobreza no Brasil: retrato de uma estabilidade inaceitável**. RBCS. Vol. 15, n 42. fev. 2000.

BATISTA, Henrique Rogê; MOLLO, Maria de Lourdes Rollemberg. **A Questão da Desigualdade Multidimensional: discutindo a construção de um indicador**. Revista de Economia Contemporânea.vol.25, n.1, p. 1-30. Rio de Janeiro, 2021.

BERNI, Duilio de Avila; *et al.* (2002). **A Desigualdade Económica do Rio Grande do Sul: Primeiras Investigações sobre a Curva de Kuznets**. 1º Encontro de Economia Gaúcha (EEG). Porto Alegre, 2002.

CARVALHO, André Roncaglia de; SOUZA, Luciana Rosa de. **A evolução conceitual da desigualdade e da pobreza no pensamento econômico**. Revista de Economia Política, v. 41, ed. 2, 2021.

CASTRO, Rafael Santos. **Efeitos da Desigualdade de Renda Sobre o Crescimento Econômico no Brasil: Uma análise não linear**. Ciências Econômicas. UFRGS, Porto Alegre, 2006

FIGUEIREDO, Allisianne Krystina Saraiva de; et al. (2016). **Análise Espacial do Desenvolvimento e das Desigualdades no Território Sudoeste Baiano**. Editora Unijuí. n. 44. 2016.

GUJARATI, D. N.; PORTER, D. C. **Econometria básica**. 5ª. ed., Porto Alegre: AMGH Editora Ltda, 2011.

HOFFMANN, Rodolfo. **Análise de Regressão: uma introdução à econometria**. 5. ed. Piracicaba: O Autor, 2016.

IBGE. **Pesquisa Nacional por Amostra de Domicílios – PNAD**. Microdados. Disponível em: <https://www.ibge.gov.br/>. Acesso em: 5 ago. 2018.

KUZNETS, Simon. **Economic Growth and Income Inequality**. The American Economic Review, v. XLV, n. 01. MARCH, 1955.

MEDEIROS, Marcelo. **Medidas de Desigualdade e Pobreza**. 1. ed. [S. l.]: UnB, 2012. Disponível em: <https://econpolrg.files.wordpress.com/>. Acesso em: 15 Outubro 2024.

MUNIZ, Sergio Tadeu Gonçalves. **Crescimento pró-pobre e políticas de distribuição de renda e de redução da pobreza**. EDUTFPR. Curitiba, 2018.

OLIVEIRA, José Luis Santos de. **Relação entre gastos sociais e o IDH-M: uma análise econométrica para os municípios paranaenses**. Ciências Econômicas. UFPR, Curitiba, jul. 2016.

OSORIO, Rafael Guerreiro; SOUZA, Pedro H. G. Ferreira de. **Evolução da pobreza extrema e da desigualdade de renda na Bahia: 1995 a 2009**. Texto para Discussão, No. 1696, Instituto de Pesquisa Econômica Aplicada (IPEA). Brasília, 2012.

RIPSA. **Taxa de Analfabetismo**. Rede Interagencial de Informações para a Saúde (RIPSA), [S. l.], S/D. Disponível em: <http://www.ripsa.org.br/>. Acesso em: 12 Novembro 2024.

SALVATO, Márcio Antônio. *et al.* **Crescimento e desigualdade: Evidências da curva de Kuznets para os municípios de Minas Gerais – 1991/2000**. Ibmec, Minas Gerais. Working Paper, n.33. ago, 2006.

SANTOS, Moisés Pais dos. *et al.* **Instruções de Renda e Desenvolvimento Econômico: Análise da hipótese de Kuznets para os estados brasileiros no período 1992-2010**. Revista Brasileira de Estudos Regionais e Urbanos (RBERU). Vol. 11, n. 2, pp. 251-271. abr, 2017.

SOUZA, Maria Iná de Abreu. *et al.* **Relação Entre a Desigualdade e Educação no Brasil: uma estimativa de dados em painel**. Ciências Econômicas. UESC, Ilhéus, out. 2014.

STOCK, James H.; WATSON, Mark W. **Introduction to Econometrics**. 3. ed. Hoboken: Pearson Education, 2015.

TAQUES, Fernando Henrique; MAZZUTTI, Caio Cícero de Toledo Piza da Costa. **Qual a Relação Entre Desigualdade de Renda e Nível de Renda Per Capita? Testando a Hipótese de de Kuznets para as Unidade Federativas Brasileira**. Planejamento e Políticas Públicas,ppp. n. 35. jul./dez. 2010.

WOOLDRIDGE, Jeffre M. **Introdução à Econometria: Uma Abordagem Moderna**. São Paulo: Pioneira Thompson Learning, 2006.