



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA - CAEN**  
**MESTRADO EM ECONOMIA**

**PAULO HENRIQUE PEREIRA ARRUDA**

**O CAPITAL HUMANO NAS DIMENSÕES SAÚDE E EDUCAÇÃO EXPLICA OS  
DIFERENCIAIS DE RENDA *PER CAPITA* DAS ÁREAS RURAIS NO BRASIL?**

**FORTALEZA**

**2026**

PAULO HENRIQUE PEREIRA ARRUDA

O CAPITAL HUMANO NAS DIMENSÕES SAÚDE E EDUCAÇÃO EXPLICA OS  
DIFERENCIAIS DE RENDA *PER CAPITA* DAS ÁREAS RURAIS NO BRASIL?

Dissertação apresentada ao Curso de Mestrado em Economia do Programa de Pós-Graduação em Economia do Centro de Aperfeiçoamento de Economistas do Nordeste- CAEN/UFC da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Economia. Área de Concentração: Economia Aplicada.

Orientador: Prof. Dr. Emerson Lemos Marinho.

FORTALEZA

2026

PAUO HENRIQUE PEREIRA ARRUDA

O CAPITAL HUMANO NAS DIMENSÕES SAÚDE E EDUCAÇÃO EXPLICA OS  
DIFERENCIAIS DE RENDA *PER CAPITA* DAS ÁREAS RURAIS NO BRASIL?

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Ciências Econômicas. Área de concentração: Relação de PIB Per capita e Nutrição.

Orientador: Prof. Dr. Emerson Lemos Marinho.

Aprovada em: \_\_\_/\_\_\_/\_\_\_.

BANCA EXAMINADORA

---

Prof. Dr. Emerson Luis Lemos Marinho (Orientador)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. João Mário Santos de França  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Flávio Ataliba Flexa Barreto  
Coordenador do Centro de Estudos do Desenvolvimento do Nordeste - FGV IBRE

A Deus.

Sua eterna Gloria e seu Infinito Saber.

## AGRADECIMENTOS

A conclusão deste trabalho representa a materialização de um esforço que transcendeu o âmbito puramente acadêmico, sendo sustentado pelo apoio incondicional de pessoas e instituições essenciais.

Primeiramente, manifesto minha profunda gratidão ao Professor Doutor Emerson Marinho, cuja excelência na orientação, rigor metodológico e disponibilidade constante foram pilares cruciais para o desenvolvimento e a finalização desta pesquisa. Seu incentivo foi determinante para a minha formação.

Este percurso não seria possível sem o suporte financeiro recebido. Agradeço à Instituição CAPES, pelo apoio com a manutenção da bolsa de auxílio, que permitiu a dedicação necessária ao estudo. Além disso, em cumprimento à Portaria CAPES nº 206, de 4 de setembro de 2018, registro: O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Expresso minha eterna admiração e gratidão aos meus pais, Neuma Arruda e Lemos Arruda, que são as minhas maiores inspirações para a vida. O exemplo de valores, força e dedicação que me transmitiram constitui o alicerce de todas as minhas conquistas.

O meu agradecimento mais sincero é dedicado à minha família: à minha amada esposa, Camilla de Sousa Rebouças Arruda, pelo companheirismo, paciência e força diária, e às minhas filhas, Linda Maria e Luna Maria. Vocês são a minha maior motivação e a razão de ser de todos os meus progressos e esforços. O amor e o apoio de vocês foram a energia para superar todos os desafios.

Por fim, sou grato aos Professores participantes da banca examinadora (Dr. Emerson Lemos Marinho, Dr. João Mário Santos de França e Dr. Flávio Ataliba Barreto), pelo tempo, pelas valiosas colaborações e sugestões; aos demais professores e colegas de turma, pelas reflexões e críticas construtivas; e aos professores entrevistados, pela generosidade em ceder seu tempo e conhecimento. A todos, o meu reconhecimento.

A melhoria das condições nutricionais no meio rural está intimamente associada ao nível de renda per capita, uma vez que rendas mais elevadas ampliam o acesso a alimentos adequados, fortalecem o capital humano e contribuem para o aumento da produtividade e do crescimento econômico local.

## RESUMO

Esta dissertação analisa os determinantes dos diferenciais de renda *per capita* rural entre os estados brasileiros, com ênfase no papel do capital humano saúde em comparação ao capital humano educação. Utilizam-se microdados das Pesquisas de Orçamentos Familiares (POF) de 2003, 2009 e 2018, a partir dos quais é construído um Índice de Desnutrição Rural baseado na ingestão domiciliar de micronutrientes essenciais, como *proxy* para o capital humano saúde. A estratégia empírica emprega modelos econométricos de dados em painel em nível estadual, estimados por meio das abordagens de efeitos fixos, com seleção do modelo apropriado baseado nos testes de Hausman e Breusch–Pagan. Na estimação das relações econométricas entre renda *per capita* e as formas de capital humano, acrescentam-se como controles variáveis socioeconômicas relevantes tais como, desigualdade de renda, índice de pobreza, além de efeitos regionais. Os resultados indicam que o capital humano na forma educação (índice de desnutrição) exerce efeito negativo e estatisticamente significativo sobre a renda *per capita* rural, apresentando maior relevância explicativa do que a escolaridade média no contexto analisado. Analisa-se também a elevada heterogeneidade regional, com maior intensidade da desnutrição nas regiões Norte e Nordeste. Os resultados sugerem que limitações nutricionais constituem um fator estrutural relevante para o nível de renda *per capita* no meio rural brasileiro, indicando que políticas voltadas à segurança alimentar e nutricional podem desempenhar papel central na redução das disparidades regionais de renda.

**Palavras-chave:** Capital humano, Dados em painel, Desnutrição, Renda *per capita* rural.

## ABSTRACT

This dissertation analyzes the determinants of rural per capita income differentials across Brazilian states, with particular emphasis on the role of health human capital in comparison to education human capital. The study uses microdata from the Household Budget Surveys (Pesquisa de Orçamentos Familiares – POF) for the years 2003, 2009, and 2018, from which a Rural Malnutrition Index is constructed based on household-level intake of essential micronutrients, serving as a proxy for health human capital. The empirical strategy relies on state-level panel data econometric models, estimated using fixed-effects and random-effects approaches, with model selection based on the Hausman and Breusch–Pagan tests. In estimating the econometric relationships between per capita income and different forms of human capital, relevant socioeconomic variables are included as controls, such as average years of schooling, income inequality, poverty indices, as well as regional effects. The results indicate that health human capital, captured by the malnutrition index, exerts a negative and statistically significant effect on rural per capita income, displaying greater explanatory power than average schooling within the analyzed context. The analysis also reveals substantial regional heterogeneity, with higher malnutrition intensity in the North and Northeast regions. Overall, the findings suggest that nutritional constraints constitute a relevant structural factor affecting per capita income levels in rural Brazil, indicating that policies aimed at food and nutritional security may play a central role in reducing regional income disparities.

**Keywords:** Human capital; panel data; malnutrition; rural per capita income.

## LISTA DE GRÁFICOS

Gráfico 1	– Índices de Desnutrição Rural dos Estados Brasileiros.....	25
Gráfico 2	– Desnutrição Rural segundo o Setor de Atividade Econômica 2003.....	26
Gráfico 3	– Desnutrição Rural segundo o Setor de Atividade Econômica 2008.....	27
Gráfico 4	– Desnutrição Rural segundo o Setor de Atividade Econômica 2018.....	31
Gráfico 5	– Relação entre Índice de Desnutrição e PIB <i>per capita</i> rural.....	35
Gráfico 6	– Relação entre AME e PIB <i>per capita</i> rural.....	36
Gráfico 7	– Modelo de Efeito Fixo : IDR vc AME.....	44
Gráfico 8	– Elasticidade: Modelo com Efeito Fixos.....	44
Gráfico 9	– Evolução da Desnutrição por Região.....	45
Gráfico 10	– PIB <i>per capita</i> e Taxa de Desnutrição.....	46

## LISTA DE TABELAS

Tabela 1	– PIB <i>per capita</i> Rural Real.....	22
Tabela 2	– Indicadores de Índices Nutricionais.....	23
Tabela 3	– Anos Médios de Estudo na Zona Rural dos Estados Brasileiros.....	29
Tabela 4	– Índice de Gini da Zona Rural.....	31
Tabela 5	– Fatores de Conversão de Paridade do Poder de Compra (PPC).....	32
Tabela 6	– Índice de Pobreza da Zona Rural.....	36
Tabela 7	– Teste de Adequação do Modelo: Dados em Panel.....	39
Tabela 8	– Estimções Econométricas.....	41
Tabela 9	– Modelos de <i>Dummies</i> Regionais.....	42

## **LISTA DE ABREVIATURAS E SIGLAS**

**AME** – Anos Médios de Estudo

**DRI** – Dietary Reference Intakes

**FAO** – Organização das Nações Unidas para a Alimentação e a Agricultura

**FE** – Efeitos Fixos

**GLS** – Generalized Least Squares

**IBGE** – Instituto Brasileiro de Geografia e Estatística

**ID** – Índice de Desnutrição

**IDR** – Índice de Desnutrição Rural

**LM** – Multiplicador de Lagrange (Breusch–Pagan)

**LOSAN** – Lei Orgânica de Segurança Alimentar e Nutricional

**ODS** – Objetivos de Desenvolvimento Sustentável

**OLS** – Ordinary Least Squares (Mínimos Quadrados Ordinários)

**PIB** – Produto Interno Bruto

**PNAE** – Programa Nacional de Alimentação Escolar

**POF** – Pesquisa de Orçamentos Familiares

**RE** – Efeitos Aleatórios

**SAN** – Segurança Alimentar e Nutricional

**SIM** – Sistema de Informações sobre Mortalidade

**VAB** – Valor Adicionado Bruto

**WHO** – World Health Organization

## LISTA DE SÍMBOLOS

%	Porcentagem
§	Seção
©	Copyright
®	Marca registrada
$\ln(.)$	Logaritmo natural
$i$	Estado
$t$	Período
$j$	Família rural
$k$	Nutriente
$K$	Número de nutrientes
ID	Índice de Desnutrição
IDR	Índice de Desnutrição Rural
IDR <sub><math>i</math></sub>	Índice de Desnutrição Rural do estado $i$
$N_i$	Número de famílias rurais
$x_{jk}^i$	Consumo do nutriente $k$ pela família $j$
$x_k^i$	Consumo médio do nutriente $k$
PIB	Produto Interno Bruto
$PIB_{pcit}$	PIB per capita rural
VAB	Valor Adicionado Bruto
$AME_{it}$	Anos Médios de Estudo
$GINI_{it}$	Índice de Gini
$POB_{it}$	Índice de Pobreza
$\alpha$	Constante
$\beta$	Coefficientes
$\varepsilon_{it}$	Erro
$\mu_i$	Efeito individual
$z$	Linha de pobreza
$y_{ij}$	Renda per capita
$I(.)$	Função indicadora
$E(.)$	Valor esperado
$Var(.)$	Variância
$Cov(.)$	Covariância

## SUMÁRIO

1	INTRODUÇÃO .....	14
2	REVISÃO DA LITERATURA .....	18
3	BASE DE DADOS .....	21
3.1	PIB <i>per capita</i> rural.....	21
3.2	Índice de Desnutrição rural.....	22
3.3	Anos Médio de Estudos.....	28
3.4	Índice de Gini.....	30
3.5	Índice de Pobreza.....	32
4	RELAÇÃO OBSERVADA ENTRE PIB <i>PER CAPITA</i> RURAL DOS ESTADOS BRASILEIROS E O CAPITAL HUMANO NA FORMA DE SAÚDE E EDUCAÇÃO.....	35
5	METODOLOGIA ECONÔMETRICA.....	37
5.1	Testes e critério de Seleção de Modelos (Teste de Hausman).....	38
5.2	Resultado da Estimação do Modelo em Painel.....	40
6	CONCLUSÃO.....	47
	REFERÊNCIAS .....	49
	APÊNDICE A – DADOS AGRUPADOS DA POF 2003, 2009 E 2018.....	53
	APÊNDICE B – SCRIPTE DO TRATAMENTO DO MODELO.....	54

## 1. INTRODUÇÃO

A explicação dos diferenciais de renda *per capita* entre países e regiões constitui um dos temas centrais da teoria do crescimento econômico. Os modelos neoclássicos pioneiros, notadamente os de Solow (1956) e Swan (1956), atribuíram essas diferenças fundamentalmente à acumulação de capital físico, à expansão da força de trabalho e ao progresso tecnológico exógeno. Deste modo, economias com menor dotação inicial de capital tenderiam a crescer mais rapidamente, convergindo, no longo prazo, para níveis semelhantes de renda *per capita*. Contudo, a persistência empírica de desigualdades regionais e internacionais verificadas no mundo desafiou essa previsão teórica, abrindo espaço para novas interpretações.

A partir da década de 1980, a literatura de crescimento passou a enfatizar mecanismos endógenos, incorporando explicitamente o papel do capital humano. Romer (1986, p. 1003) argumenta que “o conhecimento é um insumo com retornos crescentes, capaz de sustentar trajetórias de crescimento persistentes”. De forma complementar, Lucas (1988, p. 17) destaca que o acúmulo de capital humano gera externalidades positivas, elevando a produtividade média da economia: “o capital humano é o motor fundamental do crescimento econômico”. Nesse contexto, diferenças no estoque de capital humano passaram a ser apontadas como explicação central para os diferenciais de renda *per capita* observados entre economias.

Empiricamente, grande parte da literatura passou a mensurar o capital humano por meio de indicadores educacionais, como anos médios de estudo ou taxas de escolarização. Barro (1991) mostra que níveis mais elevados de escolaridade estão associados a maiores taxas de crescimento econômico, enquanto Mankiw, Romer e Weil (1992) demonstram que a inclusão do capital humano educacional melhora significativamente o poder explicativo do modelo de Solow ampliado. Todavia, esses autores reconhecem que tais medidas capturam apenas uma dimensão do capital humano e estão sujeitas a limitações importantes, sobretudo no que se refere à qualidade do ensino e à efetiva conversão da educação em produtividade.

Nesse sentido Schultz (1961, p. 3) já alertava que “investimentos em saúde e nutrição são tão importantes quanto investimentos em educação formal para o aumento da produtividade do trabalho”

Becker (1964) reforça esse argumento ao afirmar que o capital humano é composto por múltiplas dimensões, incluindo educação, saúde e condições físicas dos indivíduos, sendo todas economicamente relevantes. Assim, a literatura passou a reconhecer que o capital humano na forma de saúde constitui um determinante fundamental da produtividade e da renda, especialmente em economias e setores intensivos em trabalho físico.

A relação entre saúde, nutrição e produtividade é particularmente relevante em áreas rurais. Strauss e Thomas (1998, p. 770) demonstram, com base em microdados, que indivíduos mais bem nutridos apresentam maior capacidade laboral, menores taxas de absenteísmo e maior produtividade, sobretudo em atividades agrícolas. Segundo os autores, “a nutrição afeta diretamente a produtividade do trabalho, especialmente onde o esforço físico é intenso”. Weil (2013, p. 9) complementa ao argumentar que a saúde deve ser tratada como um insumo direto da função de produção agregada, influenciando tanto a eficiência do trabalho quanto a capacidade de aprendizado: “uma força de trabalho mais saudável é, *ceteris paribus*, mais produtiva”.

No contexto brasileiro, essas considerações assumem importância ainda maior no meio rural, marcado por elevada informalidade, menor acesso a serviços públicos e predominância de atividades intensivas em esforço físico, além de índices de pobreza elevados. Embora o avanço educacional tenha sido significativo nas últimas décadas, os retornos econômicos da escolaridade no meio rural tendem a ser mais lentos e heterogêneos. Por outro lado, deficiências nutricionais afetam de forma imediata a capacidade produtiva dos trabalhadores. Conforme destaca Sen (1999, p. 87), “a privação nutricional compromete diretamente as capacidades básicas dos indivíduos, limitando sua liberdade real de alcançar funcionamentos valiosos”.

Apesar da relevância dessas discussões teórica e empírica nesta literatura, no Brasil a análise direta da relação entre saúde nutricional e diferenciais de renda *per capita*, não existe nenhum artigo relacionando saúde com diferenciais de renda *per capita*., especialmente em nível estadual e restrita às áreas rurais. A maior parte dos estudos nacionais concentra-se em indicadores educacionais ou em medidas agregadas de saúde, como mortalidade infantil. Nesse sentido, há uma lacuna importante na literatura de crescimento econômico aplicada ao Brasil, particularmente no que se refere ao papel da nutrição como *proxy* de capital humano.

Diante de tal contexto, o objetivo principal desta dissertação é investigar quais fatores explicam os diferenciais de renda *per capita* rural entre os estados brasileiros, com ênfase na comparação entre diferentes dimensões do capital humano. Especificamente, busca-se avaliar se o capital humano na forma de saúde, mensurado por um indicador nutricional, apresenta maior poder explicativo sobre a renda *per capita* rural do que o capital humano educação, tradicionalmente representado pelos anos médios de estudo. A hipótese central é que, dadas as características estruturais do meio rural brasileiro, o estado nutricional — ao afetar diretamente o vigor físico, a saúde e a produtividade — constitui um determinante mais relevante e imediato da renda do que a educação formal. Para isso, constrói-se um Índice de Desnutrição Rural a partir da ingestão domiciliar de micronutrientes, tais como: Cálcio, Ferro, Vitamina A, Vitamina

B1(Timina) e Vitamina B2 (Riboflavina). Esses micronutrientes desempenham papel central na manutenção do metabolismo energético, no desenvolvimento cognitivo, na saúde óssea e na capacidade produtiva do indivíduo. Deficiências nesses elementos estão associadas à redução do desempenho físico e intelectual, maior vulnerabilidade a doenças e prejuízos acumulativos ao capital humano, afetando diretamente o bem-estar e pode produtividade econômica, conforme destacado pela Organização Mundial da Saúde e pela Organização das Nações Unidas para a Alimentação e Agricultura (WHO, 2006; FAO, 2013).

Foi utilizando os microdados da Pesquisa de Orçamentos Familiares (POF) dos anos de 2003, 2009 e 2018. A POF permite captar não apenas o nível de renda dos Chefes de família, por setor de atividade econômico como: agrícola, não agrícola, conta própria e outros da zona rural de cada estado, mas também a qualidade do consumo alimentar, sendo amplamente reconhecida como fonte adequada para análises de bem-estar e nutrição (DEATON, 1997, p. 35). O indicador proposto busca capturar a dimensão da chamada “fome oculta”, associada à deficiência de micronutrientes, frequentemente negligenciada em análises baseadas apenas em consumo calórico. Desta forma os dados

Do ponto de vista econométrico, a análise é conduzida por meio de modelos de dados em painel em nível estadual, explorando a variação temporal e espacial da renda *per capita* rural. São estimados modelos de efeitos fixos, controlando-se por variáveis socioeconômicas relevantes, como pobreza e desigualdade de renda, além de efeitos regionais. Conforme destaca (BALTAGI, 2005, p. 4) é modelos em painel permitem “controlar heterogeneidades não observadas e obter estimativas mais eficientes”, sendo particularmente adequados para análises regionais.

Além de contribuir para o debate teórico sobre capital humano e crescimento econômico, esta Dissertação oferece uma contribuição empírica para o caso brasileiro ao analisar, de forma comparativa, o papel do capital humano nas formas de saúde nutricional e educação na explicação dos diferenciais de renda *per capita* rural. Ao focar exclusivamente nas áreas rurais e utilizar um indicador nutricional construído a partir de microdados domiciliares, o estudo amplia a compreensão dos mecanismos que sustentam as desigualdades regionais de renda no Brasil e fornece subsídios relevantes para o desenho de políticas públicas voltadas ao desenvolvimento rural e à segurança alimentar e nutricional.

A dissertação está estruturada em seis capítulos, além desta introdução. O segundo capítulo apresenta a revisão da literatura teórica e empírica. O terceiro capítulo descreve as bases de dados e discute a construção das variáveis do modelo econométrico. No quarto capítulo apresenta-se uma análise da evidência das relações entre o PIB rural *per capita* dos estados

brasileiros e as formas de capital humano. O quinto capítulo apresenta a metodologia empregada os resultados dos modelos econométricos. Por fim, no capítulo seis reúne as considerações finais, destacando as principais conclusões, limitações e direções para pesquisas futuras.

## 2. REVISÃO DA LITERATURA

A literatura econômica que investiga os determinantes dos diferenciais de renda *per capita* evoluiu substancialmente ao longo do tempo, incorporando progressivamente novas dimensões do capital humano. Inicialmente, os estudos concentraram-se na acumulação de fatores tradicionais, como capital físico e trabalho, e, posteriormente, passaram a enfatizar o papel do capital humano como elemento central da produtividade e do crescimento econômico.

Os trabalhos pioneiros de Solow (1956) e Swan (1956) estabeleceram o arcabouço neoclássico do crescimento econômico, no qual as diferenças de renda *per capita* decorrem essencialmente de distintas dotações iniciais de capital e trabalho, bem como do progresso tecnológico exógeno. Embora esses modelos tenham fornecido uma base teórica sólida, suas incapacidades empíricas de explicar a persistência de disparidades regionais de renda motivou o surgimento de novas abordagens.

A partir da década de 1980, a literatura de crescimento endógeno passou a enfatizar explicitamente o capital humano como fator determinante do crescimento. Romer (1986) introduziu a ideia de retornos crescentes associados ao conhecimento, enquanto Lucas (1988) destacou o papel das externalidades do capital humano na elevação da produtividade agregada. Esses estudos estabeleceram que economias com maior estoque de capital humano tendem a apresentar níveis mais elevados de renda *per capita* no longo prazo.

No campo empírico, Barro (1991) e Mankiw, Romer e Weil (1992) foram fundamentais ao incorporar o capital humano na forma de educação em análises *cross-country*, utilizando como *proxy* os anos médios de estudo. Os resultados, em geral, indicam que a educação exerce impacto positivo e estatisticamente significativo sobre o crescimento econômico e a renda *per capita*. Contudo, esses autores reconhecem limitações importantes dessas medidas, especialmente no que se refere à qualidade do ensino e à heterogeneidade dos retornos educacionais entre setores e regiões.

Paralelamente, uma vertente da literatura passou a enfatizar a saúde como dimensão essencial do capital humano. Schultz (1961) e Becker (1964) já argumentavam que investimentos em saúde e nutrição elevam a eficiência do trabalho e geram retornos econômicos comparáveis aos da educação formal. Grossman (1972) formalizou esse argumento ao modelar a saúde como um estoque de capital que influencia diretamente a produtividade e o bem-estar dos indivíduos.

Avanços importantes ocorreram a partir da década de 1990, quando estudos microeconômicos passaram a analisar de forma mais direta a relação entre nutrição, saúde e

produtividade. Strauss e Thomas (1998, p. 770), utilizando dados de países em desenvolvimento, demonstram que melhorias no estado nutricional aumentam significativamente a produtividade do trabalho, sobretudo em atividades intensivas em esforço físico, como é o caso do setor agrícola. Os autores concluem que a nutrição constitui um determinante fundamental da renda em áreas rurais, frequentemente subestimado em análises baseadas apenas em educação.

O início dos anos 2000, a literatura passou a integrar saúde e crescimento econômico em modelos macroeconômicos. Bloom, Canning e Sevilla (2004) mostram que ganhos em saúde elevam a produtividade do trabalho e estimulam o crescimento da renda *per capita*, especialmente em economias de baixa renda. Weil (2013), em uma síntese abrangente, argumenta que a saúde deve ser tratada como um insumo direto da função de produção, ao afetar tanto a eficiência do trabalho quanto a capacidade cognitiva dos indivíduos.

Mais recentemente, estudos internacionais têm explorado a relação entre saúde, nutrição e desigualdade de renda em contextos rurais. Pesquisas publicadas após 2020 indicam que deficiências nutricionais continuam sendo um obstáculo relevante ao aumento da renda em áreas rurais, mesmo em países de renda média. Evidências mostram que a nutrição afeta não apenas a produtividade física, mas também a capacidade de adoção de tecnologias e a inserção em mercados mais dinâmicos, reforçando seu papel como determinante estrutural da renda rural.

Evidências empíricas mais recentes reforçam esses argumentos. Leng, Zhang e Cheng (2024, p 9), utilizando dados rurais da China, analisam a relação entre adoção de tecnologias digitais, alocação de trabalho e ingestão nutricional. Os autores mostram que melhorias na nutrição estão associadas a maior participação em atividades produtivas mais rentáveis e a aumentos significativos da renda rural. Os resultados indicam que a nutrição afeta não apenas a produtividade física, mas também a capacidade de inserção em mercados de trabalho mais dinâmicos e de adoção de novas tecnologias, reforçando seu papel como determinante estrutural da renda rural.

Complementarmente, uma revisão sistemática recente publicada por BMC Public Health (2025) sintetiza evidências de países de baixa e média renda e conclui que deficiências nutricionais persistem como um obstáculo relevante ao desenvolvimento econômico em áreas rurais. O estudo destaca que ambientes alimentares desfavoráveis — caracterizados por menor acesso físico e econômico a alimentos nutritivos — estão associados a piores indicadores de saúde e nutrição, o que compromete a produtividade do trabalho e perpetua desigualdades de renda no meio rural.

A literatura de crescimento regional, consolidada por Sala-i-Martin (1996), argumenta que diferenças iniciais no estoque de capital humano e nas condições estruturais explicam a

persistência dos diferenciais regionais de renda. No entanto, grande parte desses estudos utiliza medidas tradicionais de capital humano baseadas na escolaridade média. Em contextos rurais, onde a produtividade depende fortemente da capacidade física, indicadores de saúde e nutrição podem capturar de forma mais adequada o capital humano relevante para a geração de renda.

No contexto brasileiro, a literatura sobre diferenciais de renda regional é ampla, mas poucos estudos incorporam explicitamente indicadores nutricionais como determinantes da renda *per capita*. A maior parte das análises nacionais concentra-se em educação, infraestrutura e características do mercado de trabalho. Estudos baseados em dados da PNAD e da POF mostram que a escolaridade média está positivamente associada à renda, mas os efeitos são heterogêneos entre áreas urbanas e rurais.

Evidências recentes para o Brasil reforçam o papel da nutrição como determinante estrutural da renda e da produtividade no meio rural. Utilizando microdados das Pesquisas de Orçamentos Familiares (POF) de 2002–2003, 2008–2009 e 2017–2018, Marinho et al. investigam a existência da chamada *Armadilha da Pobreza Nutricional* no meio rural brasileiro, destacando o efeito causal da subnutrição sobre a produtividade do trabalho e a renda dos indivíduos ocupados. Os autores demonstram que, apesar de avanços no padrão alimentar ao longo do tempo, persistem deficiências nutricionais relevantes, sobretudo relacionadas à ingestão de calorias e micronutrientes, que afetam diretamente a capacidade física e cognitiva dos trabalhadores rurais.

Assim, observa-se uma lacuna na literatura nacional no que se refere à investigação direta da relação causal entre capital humano saúde — mensurado por indicadores nutricionais — e os diferenciais de renda *per capita* no meio rural brasileiro, especialmente em nível estadual e com dados em painel. Esta dissertação busca contribuir para o preenchimento dessa lacuna ao comparar, de forma sistemática, o papel do capital humano nas formas de saúde e educação na explicação das disparidades de renda rural entre os estados brasileiros, utilizando microdados da POF e métodos econométricos adequados para captar heterogeneidades não observadas.

### 3. BASE DE DADOS.

Esta dissertação utiliza microdados das Pesquisas de Orçamentos Familiares (POF/IBGE), combinados com estatísticas oficiais agregadas, organizados em um painel por unidade da Federação, com foco exclusivo na população residente em áreas rurais. As variáveis dos modelos econométricos especificados na Seção 4 contemplam a renda *per capita*, índice de desnutrição, desigualdade de renda e índice de pobreza relativas às zonas rurais dos estados brasileiros, permitindo a investigação desses determinantes sobre os diferenciais de renda *per capita* rural entre as áreas rurais desses estados. Além do mais, para tentar diferenciar os estados quanto às suas localizações geográficas nas regiões brasileiras, constroem-se variáveis *Dummies* para captar esses efeitos.

#### 3.1 PIB *per capita* rural.

A variável dependente PIB *per capita* é construída através do Valor Adicionado Bruto (VAB) agropecuário. O VAB é um conceito contábil-econômico que representa o valor gerado pela produção descontado o consumo intermediário, evidenciando o acréscimo efetivo de riqueza e evitando dupla contagem na agregação setorial das contas nacionais (IBGE, 2024).

O Produto Interno Bruto (PIB), por sua vez, corresponde à soma dos valores adicionados das unidades produtivas, calculados a preços de mercado — isto é, o agregado dos VABs setoriais (IBGE, 2024). No âmbito das atividades primárias — agricultura, pecuária, produção florestal, pesca e aquicultura — o VAB agropecuário sintetiza o valor gerado por esse conjunto produtivo, constituindo uma medida consistente da produção econômica do setor. Seu uso como base para o cálculo do PIB *per capita* justifica-se pela aderência metodológica à contabilidade nacional, pela abrangência das atividades rurais e pela disponibilidade de séries temporais regionalizadas, o que viabiliza análises comparativas no tempo e no espaço (IBGE, 2024; Carvalho, 2018; Souza; Almeida, 2019).

Os valores do VAB podem ser expressos em termos nominais ou reais. Enquanto os valores nominais refletem preços correntes, os valores reais são ajustados por índices de preços, permitindo a comparação intertemporal da produção sem distorções inflacionárias (Mankiw, 2016). Para fins analíticos, utiliza-se o VAB deflacionado pelo IPCA, garantindo que o PIB *per capita* represente variações reais da atividade econômica.

Tabela 1 – PIB *per capita* Rural Real.

Estados	PIB <i>pe rcapita</i> 2003	PIB <i>per capita</i> 2009	PIB <i>percapita</i> 2018
Rondônia	11,17	4,21	42,00
Acre	11,43	23,75	22,17
Amazonas	8,82	15,02	21,86
Roraima	4,82	11,75	21,79
Pará	9,14	30,92	28,08
Amapá	1,52	11,29	4,85
Tocantis	13,88	39,02	36,51
Maranhão	7,98	25,76	17,75
Piauí	4,26	15,48	17,70
Ceará	5,53	16,63	11,46
RN	5,27	14,01	10,39
Paraíba	5,59	17,46	8,52
Pernambuco	3,87	12,98	9,71
Alagoas	14,43	45,99	40,33
Sergipe	7,93	26,10	6,84
Bahia	8,21	21,01	17,15
Minas Gerais	5,87	22,06	15,87
Espírito Santo	3,79	44,87	13,67
Rio de Janeiro	0,79	2,93	2,15
São Paulo	5,00	15,79	8,61
Paraná	16,29	38,89	38,86
Santa Catarina	13,74	37,05	25,04
RS	16,42	43,09	38,00
Mato Grosso do Sul	39,41	72,40	74,62
Mato Grosso	39,44	117,29	95,54
Goiás	14,45	41,56	29,84
DF	1,24	8,04	4,05

Fonte: Elaborado pelo autor com base em dados do IBGE.

### 3.2 Índice de Desnutrição Rural.

Para medir a desnutrição rural dos estados brasileiros faz-se uma adaptação de um dos índices utilizado para o cálculo de pobreza devido a Foster, Greer e Thorbeck (1984). Na construção desse índice de desnutrição considera-se os seguintes setores de atividade econômica: Agrícola, Não-Agrícola, Conta Própria e Outros Empregos. Neste sentido, primeiramente calcula-se o índice de desnutrição para cada um desses setores de atividades localizados na zona rural dos estados brasileiros. Por último, o índice de desnutrição rural de cada estado é calculado como a média aritmética ponderada dos índices rurais de cada setor de atividade econômica.

Para o cálculo desse índice se faz necessário conhecer-se os valores mínimos dos nutrientes consumidos por um indivíduo para considerá-lo bem nutrido. Os nutrientes considerados são: Cálcio, Ferro, Vitamina A, Vitamina B1 e Vitamina B2. Esses valores foram coletados de Padovani *et al.* (2006), os quais sistematizam os valores de referência da Dietary Reference Intakes (DRI), um conjunto de recomendações científicas sobre a ingestão adequada de nutrientes. A DRI é um órgão científico vinculado às Academias Nacionais dos Estados Unidos, amplamente reconhecido como referência internacional em nutrição humana. Para diferentes grupos etários e sexos, conforme proposto pelo Institute of Medicine sobre consumo nutricional ideal. Esses dados estão reportados na Tabela 2

**Tabela 2 – Indicadores de índices Nutricionais.**

Tipos	Homem	Mulher	Idade	Unidade
Cálcio	1.166,67	1.166,67	Idade ≥ 9 anos	Mg
Ferro	8,5	12,5	Idade ≥ 9 anos	Mg
Vitamina A	850	683,33	Idade ≥ 9 anos	µg
Vitamina B1 – Timina	1,15	1,05	Idade ≥ 9 anos	Mg
Vitamina B2 – Riboflavina	1,23	1,05	Idade ≥ 9 anos	Mg

Fonte: Elaborada pelo autor com base em *Dietary Reference Intakes* (DRI), Padovani et al. (2006).

A mensuração de privação nutricional em contextos rurais exige instrumentos capazes de capturar não apenas níveis médios de consumo alimentar, mas também a heterogeneidade existente entre grupos socioeconômicos. Nesse sentido, este trabalho propõe um Índice de Desnutrição Rural (IDR) calculado especificamente para os chefes de família residentes na zona rural, incorporando dimensões nutricionais essenciais e a estrutura da atividade econômica

Em assim sendo, o Índice de Desnutrição Rural ( $IDR_l$ ) de um setor de atividade econômica  $l$  de um determinado estado é definido como:

$$IDR_l = \sum_{i=1}^{N_l} \sum_{k=1}^K \frac{N_l (x_k^{ij} - x_k^*)^2}{N x_k^{*2}} \quad (3.1)$$

onde,  $N_l$  é o número de chefes de família que trabalham no  $l$ -ésimo setor de atividade econômica na zona rural do  $j$ -ésimo estado brasileiro;  $x_k^{ij}$  é a quantidade consumida do  $k$ -ésimo nutriente pelo indivíduo  $i$  na zona rural do estado  $j$ ;  $x_k^*$  é a quantidade mínima necessária do nutriente  $k$  para que um indivíduo seja considerado bem nutrido.

Desde que são considerados 5 nutrientes,  $K=5$ . Observe que  $N$  é a população do estado considerado.

Nestes termos, o índice de Desnutrição Rural de um estado  $j$  ( $IDR_j$ ) será mensurado como:  $IDR_j = \frac{1}{N} \sum_{l=1}^4 IDR_l$ . Logo, substituindo (1) nesta última expressão, tem-se que o  $IDR_j$  será calculado como:

$$IDR_j = \frac{1}{N} \sum_{l=1}^4 \sum_{i=1}^N \sum_{k=1}^5 \frac{N_l^j (x_k^{ij} - x_k^{*2})^2}{x_k^{*2}} \quad (3.2)$$

onde,  $N_l^j$  é a população rural que trabalha no  $l$ -ésimo setor de atividade econômica do estado  $j$ .

A construção desse índice insere-se na tradição das medidas de desigualdade relativa, nas quais o desvio em relação a um padrão normativo é interpretado como indicador de vulnerabilidade. Essa abordagem encontra fundamento na literatura economia do bem-estar, especialmente em Sen (1976), e na família de índices decomponíveis de Foster, Greer e Thorbecke (1984), que enfatizam a relevância da intensidade das carências, e não apenas de seus níveis médios.

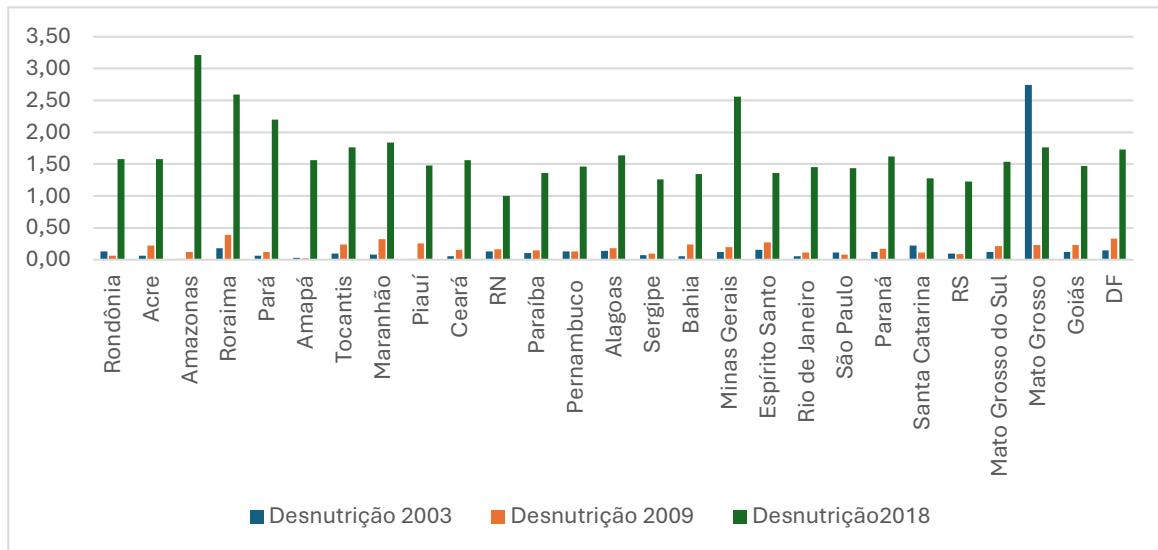
Nesse sentido, o IDR adapta esse arcabouço ao domínio nutricional, interpretando a dispersão relativa do consumo de nutrientes essenciais em torno de um limiar mínimo como evidência de risco de desnutrição. Desta forma as tabelas que se encontram no apêndice A reportam os resultados dos Índices de Desnutrição rural dos estados por setor Agrícola, Não agrícola, Contas Próprias e Outros Empregos e também de forma agregada. Esses índices são calculados para os anos de 2003, 2009 e 2018 utilizando os dados das Pesquisas de Orçamento Familiar publicados pelo IBGE.

Nota-se que os valores desse índice para os anos de 2003 e 2009 são relativamente baixos e próximos entre si na maioria dos estados, enquanto em 2018 há um aumento expressivo de seus valores, indicando agravamento da desnutrição nesse período mais recente.

O crescimento do IDR é particularmente visível em alguns estados das regiões Norte (Amazonas, Roraima, Pará) e Centro-Oeste (Mato Grosso), que se destacam com os maiores valores desse índice em 2018. Por outro lado, estados da região Sul e parte da região Sudeste apresentam níveis mais moderados, embora também mostrem elevação em relação aos anos anteriores.

Portanto, observa-se uma ampliação das disparidades estaduais entre os anos de 2009 e 2018: enquanto o problema cresce de forma generalizada, sua intensidade é desigual entre as unidades federativas. Para tal analisam-se os gráficos seguintes para entender como o IDR se apresenta nos estados e nos anos da POF.

Gráfico 1 – Índices de Desnutrição Rural dos Estados Brasileiros



Fonte: Elaborada pelo autor.

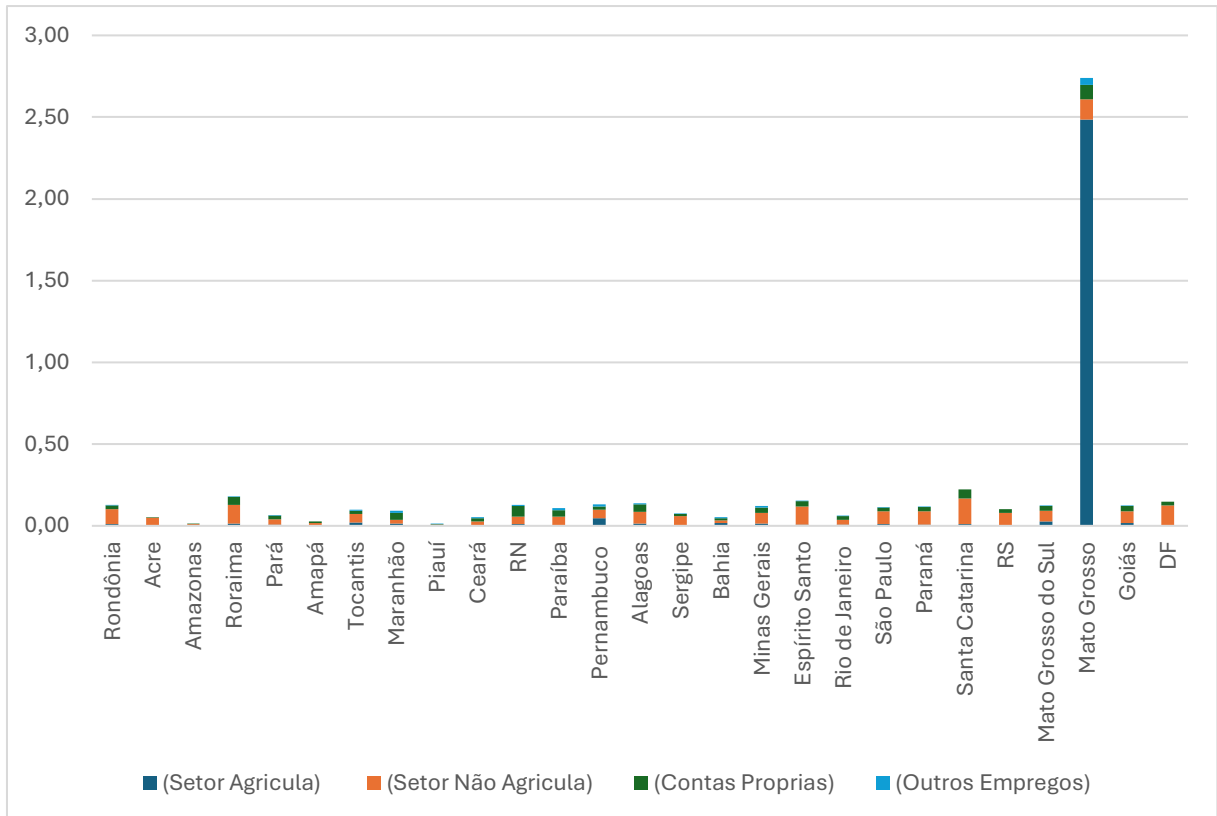
O Gráfico 1 apresenta os índices de desnutrição rural para os anos de 2003, 2009 e 2018 unindo todos os setores de atividade econômica.

Em 2003, os níveis de desnutrição são bastante reduzidos em praticamente todos os estados, concentrando-se próximos de zero e com pequena dispersão entre os estados. Esse padrão sugere um cenário inicial de baixa incidência relativa do problema, embora com leve predominância de valores mais elevados em estados das regiões Norte e Nordeste. Tal configuração é compatível com desigualdades históricas de renda, infraestrutura e acesso a serviços básicos, que tradicionalmente afetam essas regiões de forma mais intensa.

Já em 2009, percebe-se uma leve elevação dos índices em diversos estados, mas sem ruptura significativa em relação a 2003. O aumento é moderado e mantém a distribuição regional observada anteriormente, com maiores valores relativos ainda concentrados no Norte e em parte do Nordeste. Entretanto, a elevação não é homogênea: alguns estados apresentam estabilidade ou variações mínimas, indicando que o problema evoluiu de forma diferenciada no território nacional.

No ano de 2018 o principal destaque é a elevação dos índices de desnutrição, forte concentração desse índice em Mato Grosso. Nos outros estados, os níveis são baixos e mais equilibrados entre os setores.

Gráfico 2 – Desnutrição Rural segundo o Setor de Atividade Econômica 2003.



Fonte: Elaborada pelo autor.

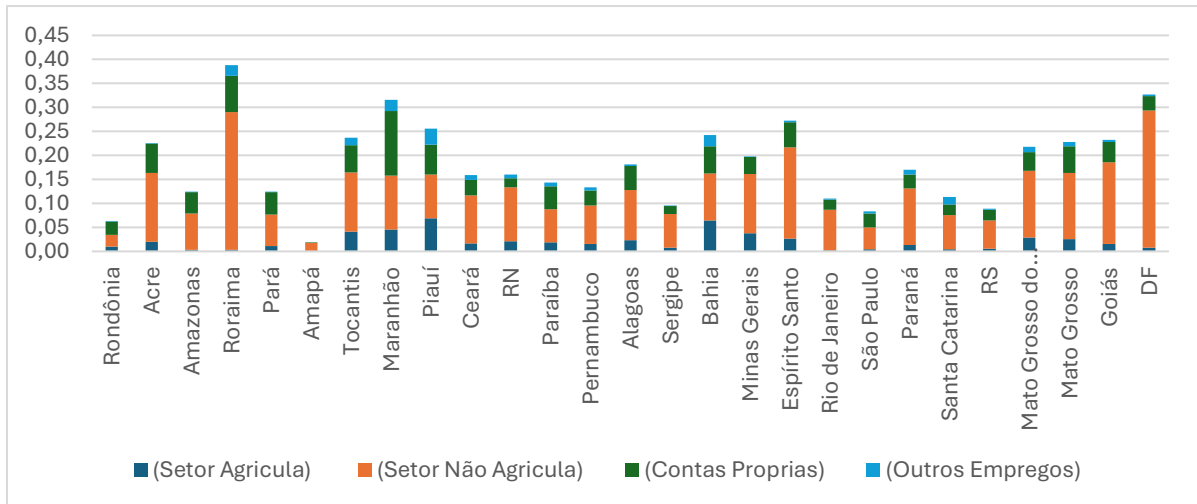
O Gráfico 2 apresenta os valores do IDR dos setores de atividade econômica para o ano de 2003. Os setores Não Agrícola e o de Contas Próprias apresentam participações moderadas e relativamente homogêneas, enquanto a contribuição de Outros Empregos é fraca.

Regionalmente, estados das regiões Norte e Nordeste exibem índices ligeiramente maiores que os estados das regiões Sul e Sudeste, mas ainda reduzidos frente ao caso de Mato Grosso. Assim, a desnutrição rural em 2003 aparece como um fenômeno de baixa intensidade entre as zonas rurais dos estados brasileiros, com exceção do estado de Mato Grosso.

No Gráfico 3 observa-se que o setor não agrícola concentra, na maior parte dos estados. Em estados como Roraima, Maranhão, Bahia, Espírito Santo e Distrito Federal, a participação do setor não agrícola se destaca como componente relevante do indicador total.

O segmento de contas próprias também apresenta peso expressivo em diversos estados, indicando a vulnerabilidade associada ao trabalho autônomo rural. Em estados do Norte e Nordeste, esse componente aparece de forma mais acentuada, reforçando a relação entre informalidade produtiva e condições nutricionais adversas.

Gráfico 3 – Desnutrição Rural segundo o Setor de Atividade Econômica 2008.

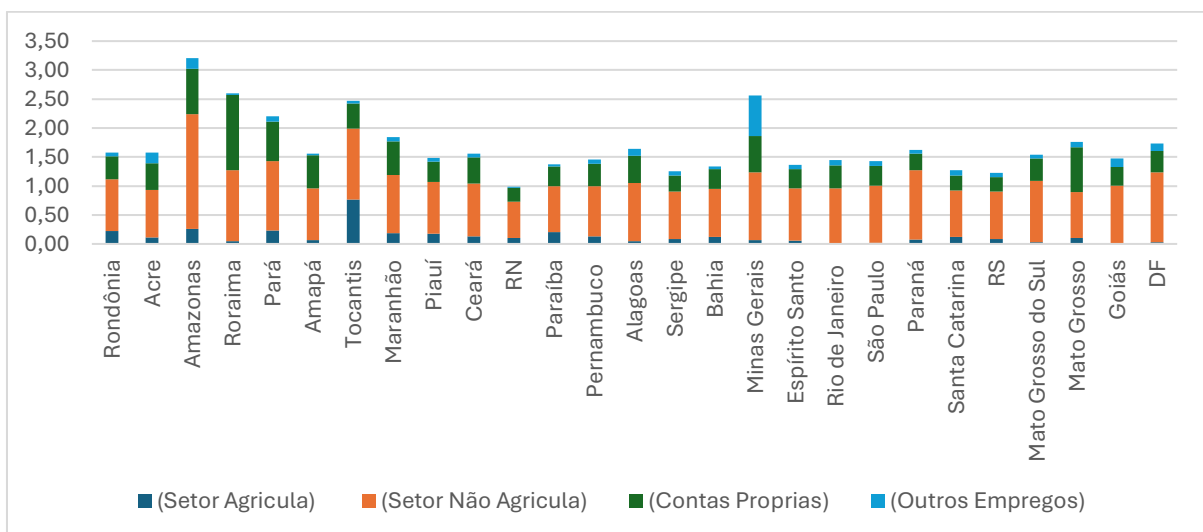


Fonte: Elaborada pelo autor.

Por outro lado, o setor agrícola formal ou tradicional tende a apresentar contribuições menores, ainda que não desprezíveis. Esse comportamento pode estar associado à maior estabilidade relativa das atividades agrícolas familiares ou ao acesso a políticas públicas específicas, como crédito rural e programas de apoio à produção. Já a categoria de outros empregos possui participação reduzida e heterogênea, funcionando mais como componente complementar do indicador.

A heterogeneidade observada reforça a necessidade de políticas públicas diferenciadas, que considerem as particularidades do mercado de trabalho rural e a diversidade regional, especialmente no que se refere à proteção social de trabalhadores autônomos e à melhoria das oportunidades de renda fora da agricultura.

Gráfico 4 – Desnutrição Rural segundo o Setor de Atividade Econômica 2018.



Fonte: Elaborada pelo autor.

No Gráfico 4 o padrão sugere que a desnutrição rural está fortemente associada a formas de inserção ocupacional mais vulneráveis fora do setor Agrícola, refletindo desigualdades regionais e estruturais no mercado de trabalho rural.

Comparando os três gráficos (2003, 2009 e 2018), observa-se uma mudança clara na intensidade e na composição da desnutrição rural ao longo do tempo. Em 2003, os índices eram, em geral, baixos e relativamente homogêneos entre os estados, com destaque pontual para Mato Grosso, fortemente concentrado no setor agrícola. Já em 2009, há um aumento moderado e uma distribuição mais equilibrada entre os setores, com maior participação do setor não agrícola e das contas próprias, indicando ampliação do problema para além da agricultura tradicional.

Em 2018, o cenário se torna mais crítico: os índices crescem de forma expressiva em vários estados, especialmente no Norte e parte do Centro-Oeste e passam a ser dominados principalmente pelos componentes não agrícolas e de contas próprias.

Esse movimento revela que a desnutrição rural deixa de ser um fenômeno mais localizado e passa a refletir vulnerabilidades estruturais do mercado de trabalho rural, com maior heterogeneidade regional.

### 3.3 Anos Médio de Estudos.

Os Anos Médios de Estudo (AME) foram calculados como uma média ponderada do número de anos de escolaridade associados a cada nível educacional, utilizando como pesos a população correspondente a cada estrato. Essa medida constitui uma *proxy* amplamente utilizada para o capital humano educacional na literatura de crescimento econômico, notadamente nos trabalhos de Barro e Lee (1993, 2001). O AME para a zona rural do j-ésimo estado brasileiro, denominado por  $AME_j$ , foi calculado como:

$$AME_j = \frac{\sum_{i=1}^{N_j} ae_i \times N_i}{N_i} \quad (3.3)$$

O  $ae_i$  é os anos de estudos do i-ésimo chefe de família residente na zona rural do j-ésimo estado.  $ON_i$  representa o peso amostral da população total da zona rural e o  $i =$  varia de acordo com os números de chefes de Família.

Tabela 3 – Anos Médios de Estudo na Zona Rural dos Estados Brasileiros.

<b>COD</b>	<b>ESTADOS</b>	<b>AME 2003</b>	<b>AME 2009</b>	<b>AME 2018</b>
11	Rondônia	5,6	7,6	8,9
12	Acre	5,2	6,4	8,9
13	Amazonas	6,5	8,2	9,9
14	Roraima	6,1	7,3	10,5
15	Pará	5,4	6,7	8,3
16	Amapá	7,3	9,5	9,4
17	Tocantis	5,2	7,2	9,6
21	Maranhão	3,9	5,0	8,2
22	Piauí	4,2	5,3	7,6
23	Ceará	5,1	7,6	8,3
24	RN	5,7	6,9	8,8
25	Paraíba	5,2	5,8	8,2
26	Pernambuco	5,9	7,4	8,8
27	Alagoas	5,0	6,1	8,5
28	Sergipe	5,2	6,8	8,9
29	Bahia	5,5	5,9	8,9
31	Minas Gerais	6,1	7,8	9,2
32	Espírito Santo	7,0	7,7	9,7
33	Rio de Janeiro	8,0	10,5	10,8
35	São Paulo	7,9	8,4	10,7
41	Paraná	7,1	11,2	9,9
42	Santa Catarina	6,9	8,9	10,1
43	RS	7,1	8,6	10,1
50	Mato Grosso do Sul	6,5	8,4	9,1
51	Mato Grosso	6,2	7,7	9,3
52	Goiás	6,1	7,6	9,9
53	DF	9,1	9,4	11,5

Fonte: Elaborada pelo autor.

De forma agregada, observa-se na Tabela 3 uma trajetória de crescimento dos anos médios de estudo em todas as regiões do país entre 2003 e 2018. Esse movimento reflete a ampliação do acesso à educação básica, a expansão do ensino médio e superior e a consolidação de políticas públicas educacionais ao longo das últimas décadas (IBGE, 2019). A Região Norte apresenta níveis historicamente mais baixos de escolaridade média, embora com crescimento expressivo no período analisado. Estados como Rondônia, Acre, Amazonas e Pará partiram de patamares próximos ou inferiores a seis anos médios de estudo em 2003, alcançando valores entre 8,3 e 9,9 anos em 2018. Destacam-se Roraima e Amapá, que registraram avanços mais acelerados, aproximando-se de dois dígitos ao final do período. Apesar do progresso, a região ainda permanece abaixo da média das regiões mais desenvolvidas, evidenciando desigualdades estruturais no acesso e permanência no sistema educacional.

O Nordeste apresenta a menor média de anos de estudo ao longo de todo o período, embora registre avanços significativos. Em 2003, estados como Maranhão, Piauí e Alagoas apresentavam níveis inferiores a cinco anos médios de estudo. Em 2018, todos os estados nordestinos ultrapassaram a marca de sete anos, com destaque para Ceará, Pernambuco, Sergipe

e Bahia, que se aproximaram de nove anos. Ainda assim, a persistência de defasagens educacionais reflete desigualdades históricas associadas à renda, ao mercado de trabalho e à infraestrutura educacional da região.

A Região Sudeste apresenta os maiores níveis de escolaridade média, confirmando seu papel de liderança educacional no país. São Paulo, Rio de Janeiro e Espírito Santo já apresentavam valores elevados em 2003, superiores a sete anos médios de estudo. Em 2018, esses estados alcançaram níveis próximos ou superiores a dez anos, com destaque para o Rio de Janeiro e São Paulo. Minas Gerais, embora partindo de patamar ligeiramente inferior, apresentou crescimento contínuo, reduzindo a distância em relação aos demais estados da região. A Região Sul também se caracteriza por altos níveis de anos médios de estudo, com trajetória ascendente ao longo do período. Estados como Paraná, Santa Catarina e Rio Grande do Sul apresentaram crescimento expressivo entre 2003 e 2018, atingindo valores próximos a 10 anos médios de estudo. Destaca-se o Paraná, que apresentou forte expansão entre 2003 e 2009, embora com leve desaceleração no crescimento posterior.

A Região Centro-Oeste ocupa posição intermediária-alta no ranking nacional. Distrito Federal apresenta, de forma consistente, o maior nível de anos médios de estudo do país, alcançando 11,5 anos em 2018, refletindo maior renda média, urbanização e concentração de serviços públicos. Estados como Goiás, Mato Grosso e Mato Grosso do Sul também registraram crescimento contínuo, aproximando-se de nove a dez anos médios de estudo ao final do período.

### **3.4 Índice de Gini.**

O índice utilizado para medir a desigualdade de renda *per capita* das zonas rurais dos estados brasileiros será o Índice de desigualdade de Gini. O seu cálculo foi operacionalizado a partir dos microdados da Pesquisa de Orçamentos Familiares (POF), tendo como variável de análise a renda domiciliar *per capita*. Inicialmente, os dados foram organizados por unidade da federação, sendo selecionadas exclusivamente as observações com valores positivos de renda. Inicialmente, as observações foram ordenadas em ordem crescente de renda, sendo posteriormente calculadas as proporções acumuladas da população e da renda, ponderadas pelos fatores de expansão amostral.

A medida de desigualdade baseada na curva de Lorenz, definida como a razão entre a área compreendida entre a linha de perfeita igualdade e a curva de Lorenz e a área total sob a

linha de perfeita igualdade (ATKINSON, 1970, p. 245). No nosso caso (POF, renda domiciliar per capita, cálculo por UF), o correto e mais aceito é o Gini via Curva de Lorenz (forma discreta).

$$G = 1 - \sum_{i=1}^n (Y_i + Y_{i-1})(X_i - X_{i-1}), \quad (3.4)$$

em que  $X_i$  representa a proporção acumulada da população até a observação  $i$  e  $Y_i$  representa a proporção acumulada da renda correspondente. O termo  $(X_i - X_{i-1})$  reflete o incremento populacional entre dois pontos consecutivos, enquanto  $(Y_i + Y_{i-1})$ . Essa formulação é apresentada e discutida em Cowell (2011, cap. 2, p. 41–44) e é amplamente utilizada na literatura empírica internacional baseada em pesquisas domiciliares como a POF.

Os dados indicam uma evolução heterogênea da desigualdade de renda rural entre os estados brasileiros no período 2003–2018, com predominância de redução da desigualdade significativa das unidades federativas, embora com trajetórias não lineares. Estados como Tocantins, Minas Gerais e Santa Catarina apresentaram quedas consistentes da desigualdade, enquanto outros, como Pará e Bahia, registraram aumento no longo prazo.

Tabela 4 – Índice de Desigualdade de Renda Rural.

Estados	GINI 2003	GINI 2009	GINI 2018
Rondônia	0,55	0,52	0,49
Acre	0,57	0,44	0,56
Amazonas	0,51	0,50	0,49
Roraima	0,54	0,53	0,47
Pará	0,34	0,53	0,56
Amapá	0,52	0,57	0,47
Tocantins	0,64	0,43	0,39
Maranhão	0,28	0,55	0,42
Piauí	0,31	0,49	0,40
Ceará	0,36	0,46	0,44
RN	0,54	0,53	0,47
Paraíba	0,35	0,49	0,51
Pernambuco	0,41	0,49	0,47
Alagoas	0,28	0,49	0,48
Sergipe	0,55	0,52	0,47
Bahia	0,27	0,49	0,53
Minas Gerais	0,23	0,49	0,31
Espírito Santo	0,28	0,45	0,44
Rio de Janeiro	0,57	0,54	0,47
São Paulo	0,57	0,45	0,42
Paraná	0,29	0,45	0,47
Santa Catarina	0,27	0,42	0,39
RS	0,34	0,45	0,42
Mato Grosso do Sul	0,27	0,47	0,42
Mato Grosso	0,27	0,52	0,47
Goiás	0,27	0,47	0,43
DF	0,57	0,50	0,49

Fonte: Elaborada pelo autor.

Observa-se um aumento da desigualdade rural entre 2003 e 2009 em grande parte dos estados, seguido por uma tendência de redução até 2018, ainda que em níveis distintos. Os estados como Tocantins, Minas Gerais e Santa Catarina apresentaram queda consistente do índice ao longo do período, indicando redução da desigualdade rural.

Por outro lado, unidades como Pará, Bahia e Paraíba registraram elevação do Gini em 2018 em relação a 2003, sugerindo persistência ou agravamento das disparidades.

### 3.5 Índice de Pobreza.

O índice de pobreza utilizado foi a proporção de pobres. A mensuração deste indicador na zona rural de cada estado foi realizada a partir da metodologia proposta por Foster, Greer e Thorbecke (1984), empregada na literatura econômica para a análise da pobreza monetária. Utilizou-se, especificamente, o índice  $P_0$ , que corresponde à proporção de indivíduos cuja renda domiciliar *per capita* se encontra abaixo de uma linha de pobreza previamente estabelecida.

Tabela 5 - Fatores de Conversão de Paridade do Poder de Compra (PPC)

Ano	Fator de conversão PPC (R\$ / US\$ internacional)
2003	0,9703430483
2009	1,2958528019
2018	2,1975915432

Fonte: Banco Mundial – World Development Indicators. Elaboração: Ministério da Ciência, Tecnologia e Inovação (MCTI), Coordenação de Indicadores de Ciência e Tecnologia (COICT).

Os dados utilizados são provenientes da Pesquisa de Orçamentos Familiares (POF) e foram considerados exclusivamente os domicílios classificados como pertencentes à zona rural. A renda utilizada para a identificação da condição de pobreza corresponde à renda domiciliar *per capita* mensal. O IBGE linha de pobreza adotada neste estudo segue o critério internacional do Banco Mundial, definido como US\$ 5,50 por dia em termos de paridade do poder de compra (PPC), recomendado para países de renda média-alta, como o Brasil (BANCO MUNDIAL, 2018). Dessa forma a conversão para reais foi realizada por meio do fator de conversão de PPC (Tabela 5), que representa quantos reais equivalem a um dólar internacional em termos de poder de compra local. Em termos mensais per capita (30 dias), a linha de pobreza estimada foi de

aproximadamente R\$ 160 em 2003, R\$ 214 em 2009 e R\$ 362 em 2018. O índice de pobreza  $P_0$  é definido como:

$$P_0 = \frac{1}{N} \sum_{i=1}^N I(y_i \leq z) \quad (3.6)$$

em que  $y_i$  representa a renda domiciliar *per capita* do chefe de família  $i$ ,  $z$  corresponde à linha de pobreza,  $N$  é o total de indivíduos da população analisada e  $I(\cdot)$  é uma função indicadora que assume valor igual a 1 quando a renda do indivíduo é inferior ou igual à linha de pobreza, e 0 caso contrário. A Tabela 6 evidencia uma redução significativa do índice de pobreza nas zonas rurais em todos os estados brasileiros entre 2003 e 2018, ainda que com forte heterogeneidade regional. Observa-se que estados historicamente mais pobres apresentaram quedas expressivas. O Maranhão, por exemplo, reduziu seu índice de pobreza de 0,59 em 2003 para 0,22 em 2018, o que representa uma redução aproximada de 63%. O Ceará apresentou queda de 0,48 para 0,19 no mesmo intervalo, correspondendo a uma redução de cerca de 60%.

Tabela 6 – Índice de Pobreza da Zona Rural

Estados	Índice Pobreza 2003	Índice Pobreza 2009	Índice Pobreza 2018
Rondônia	0,27	0,10	0,11
Acre	0,44	0,23	0,17
Amazonas	0,45	0,27	0,24
Roraima	0,40	0,31	0,20
Pará	0,41	0,28	0,20
Amapá	0,38	0,25	0,12
Tocantins	0,48	0,19	0,18
Maranhão	0,59	0,48	0,22
Piauí	0,54	0,31	0,20
Ceará	0,48	0,37	0,19
RN	0,43	0,29	0,12
Paraíba	0,48	0,38	0,24
Pernambuco	0,39	0,24	0,14
Alagoas	0,54	0,40	0,22
Sergipe	0,49	0,26	0,09
Bahia	0,41	0,27	0,14
Minas Gerais	0,18	0,11	0,05
Espírito Santo	0,19	0,16	0,06
Rio de Janeiro	0,16	0,07	0,07
São Paulo	0,08	0,04	0,04
Paraná	0,14	0,09	0,05
Santa Catarina	0,11	0,06	0,01
RS	0,11	0,05	0,02
Mato Grosso do Sul	0,20	0,06	0,02
Mato Grosso	0,26	0,15	0,04
Goiás	0,27	0,12	0,05
DF	0,12	0,07	0,02

Fonte: Elaborada pelo autor (2025).

Entre 2003 e 2009, a redução foi generalizada, porém mais moderada, nas regiões Norte e Nordeste. No Piauí, o índice passou de 0,54 para 0,31 (redução de aproximadamente 43%), enquanto em Alagoas a queda foi de 0,54 para 0,40 (26%).

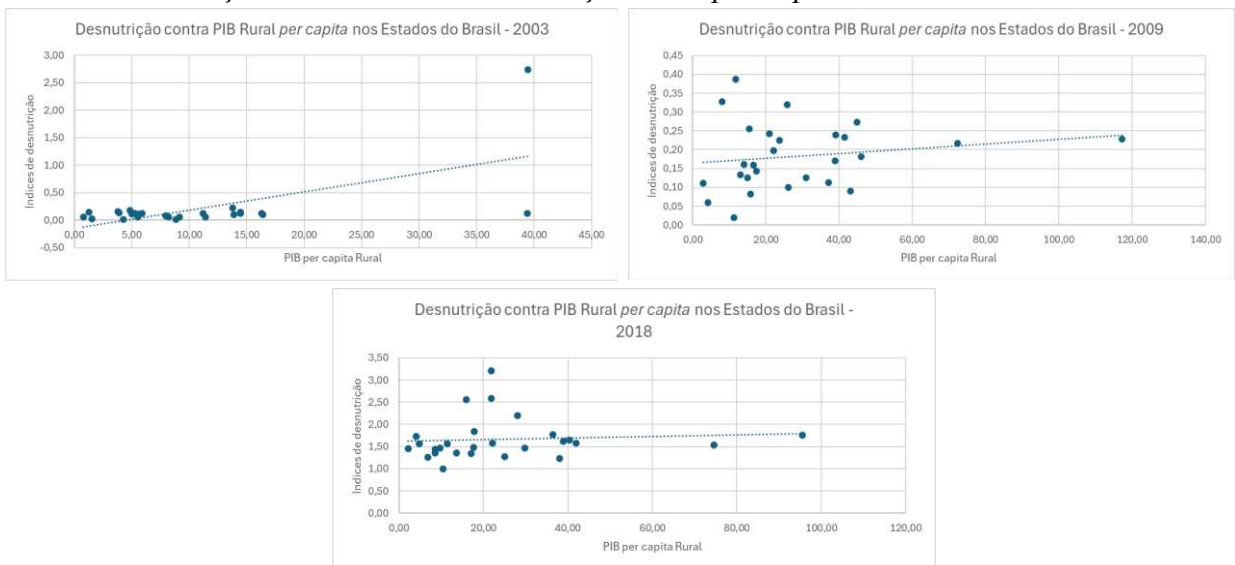
Já entre 2009 e 2018, observa-se uma intensificação da redução da pobreza, com diminuições mais acentuadas em diversos estados, como Sergipe, cujo índice recuou de 0,26 para 0,09 (65%), e Rio Grande do Norte, de 0,29 para 0,12 (59%).

Nos estados das regiões Sul e Sudeste, os níveis iniciais de pobreza já eram relativamente baixos em 2003, e as reduções observadas ao longo do período foram mais modestas em termos absolutos, embora relevantes em termos relativos. São Paulo, por exemplo, apresentou redução de 0,08 para 0,04 entre 2003 e 2018 (50%), enquanto Santa Catarina reduziu de 0,11 para 0,01 (91%), atingindo um dos menores níveis de pobreza da amostra.

#### 4. RELAÇÃO OBSERVADA ENTRE PIB *PER CAPITA* RURAL DOS ESTADOS BRASILEIROS E O CAPITAL HUMANO NA FORMA DE SAÚDE E EDUCAÇÃO.

A relação entre o PIB rural *per capita* e o capital humano na forma de saúde é apresentada no Gráfico 5 para os anos de 2003, 2009 e 2018. Observa-se que existe uma relação crescente, embora não muito acentuada, entre essas variáveis nestes três anos

Gráfico 5 – Relação entre Índice de Desnutrição e PIB *per capita* Rural.



Fonte: Elaborada pelos autores

Seria de se esperar que menores índices de desnutrição implicasse em maiores níveis de PIB rural *per capita* o que não foi observado nas trajetórias do Gráfico 5. Essa evidência observada parece sugerir que, apesar de maiores níveis de desnutrição rural, os níveis de PIB rural *per capita* dos estados brasileiros não foram afetados nestes três anos. No entanto, vale salientar que a presença de dois *outliers* podem estar afetando essas relações.

Outro resultado observado, com exceção o ano de 2003, é que estados com PIB *per capita* semelhantes apresentam padrões de desnutrição rural bastante dispersos.

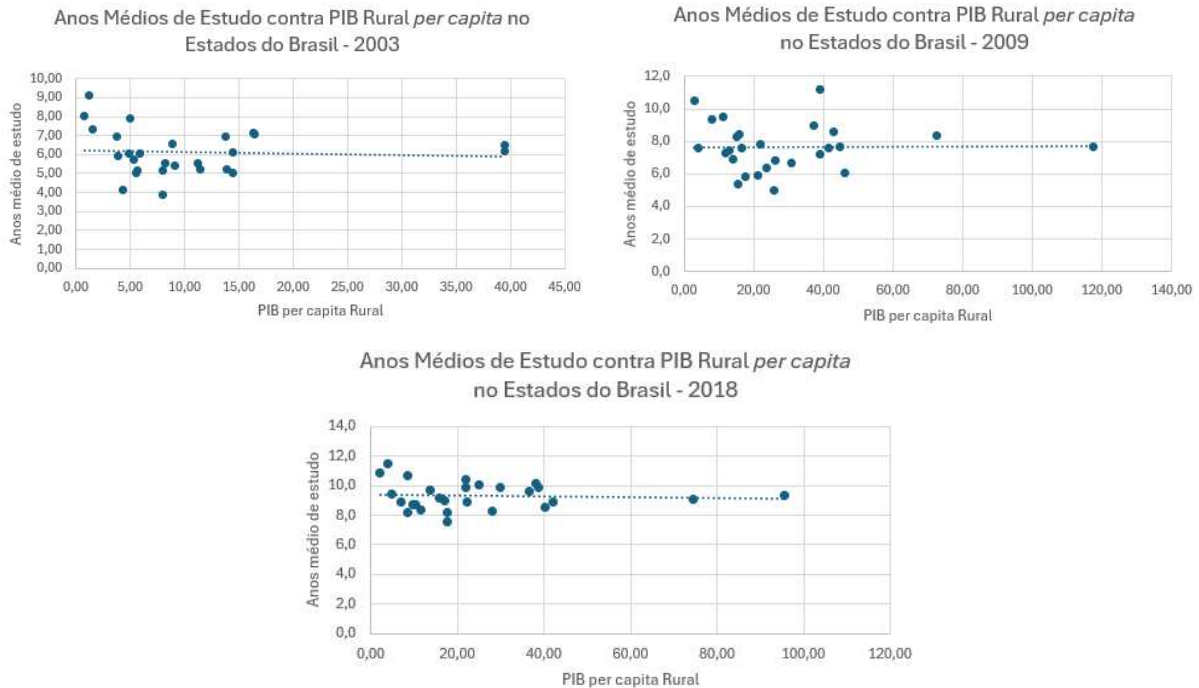
No entanto, não se pode concluir em definitivo que a desnutrição rural afeta positivamente o PIB rural *per capita* sem considerar na análise outros determinantes relevantes deste último, tais como, pobreza rural, desigualdade de renda e diferenças regionais entres os estados brasileiros.

As relações entre o PIB rural *per capita* e a outra forma de capital humano, medido pelos anos médios de estudo, são apresentadas no Gráfico 6 para os anos de 2003, 2009 e 2018. Não fica muito claro a identificação de alguma tendência das relações entre essas variáveis. A

linha de tendência praticamente horizontal reforça a fraca associação linear entre essas variáveis.

Fica difícil afirmar que o capital humano, na forma de educação, explica os diferenciais de renda rural *per capita* dos estados brasileiros.

Gráfico 6 – Relação entre AME e PIB *per capita* Rural.



Fonte: Elaborada pelo autor.

No entanto, deve-se lembrar mais uma vez que a relação isolada entre essas duas variáveis pode ficar prejudicada se não se considera outros fatores determinantes do PIB rural *per capita*.

As unidades federativas com rendas *per capita* rurais semelhantes apresentam escolaridade média bastante distintas, variando entre quatro e oito anos de estudo. Esse padrão indica que a renda rural *per capita* não se converte diretamente em maiores níveis educacionais, sugerindo a influência de outros fatores estruturais e regionais.

## 5. METODOLOGIA ECONOMÉTRICA.

Desde que o objetivo principal é analisar se o capital humano nas dimensões saúde e educação explicam os diferenciais de renda *per capita* nas zonas rurais dos estados brasileiros, especifica-se agora um modelo econométrico em que o PIB rural *per capita* dos estados ( $PIB_{pci}$ ) é função desses dois fatores explicativos. Para efeito de controles, adicionam-se as seguintes variáveis explicativas ao modelo: o Índice de Desnutrição (IDR), os Anos médios de Estudo (AME), o Índice de Desigualdade de renda (GINI), o Índice de Pobreza  $P_0$ .

Adicionalmente, são incluídas *dummies* regionais para capturar efeitos estruturais associados às regiões geográficas do país.

O modelo básico em dados em painel pode ser representado pela seguinte equação:

$$PIB_{pci_t} = \beta_0 + \beta_1 IDR_{it} + \beta_2 AME_{it} + \beta_3 P_{0it} + \beta_4 GINI_{it} + \sum_{i=1}^4 \gamma_i D_i + \alpha_i + \epsilon_{it} \quad (5.1)$$

onde,  $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \gamma_1, \gamma_2, \gamma_3$  e  $\gamma_4$  são parâmetros a serem estimados.

O parâmetro  $\alpha_i$  capta os efeitos específicos não observados de cada estado permitindo controlar características estruturais e institucionais invariantes no tempo, como dotação histórica de infraestrutura, cultura produtiva, localização geográfica e qualidade institucional (Baltagi, 2005). Por fim,  $\epsilon_{it}$ , é o termo de erro idiossincrático.

Os parâmetros  $\gamma_1, \gamma_2, \gamma_3$  e  $\gamma_4$  se referem as *dummies* regionais. Elas são definidas como: captar diferenças estruturais entre as grandes regiões do país. Considera-se o Sudeste como categoria de referência, de modo que:

- $D_1=1$  se o estado pertence à Região Norte, e 0 caso contrário;
- $D_2=1$  se pertence ao Nordeste, e 0 caso contrário;
- $D_3=1$  se pertence ao Sul, e 0 caso contrário;
- $D_4=1$  se pertence ao Centro-Oeste, e 0 caso contrário.

Assim, os coeficientes  $\gamma_r$  medem o diferencial médio de cada região em relação ao Sudeste, mantendo constantes as demais variáveis do modelo.

Em modelos de dados em painel, a heterogeneidade não observada entre as unidades transversais constitui uma fonte potencial de viés quando correlacionada com os regressores. Para lidar com esse problema, empregam-se as especificações de efeitos fixos (FE) e efeitos aleatórios (RE), que diferem quanto às hipóteses impostas sobre a estrutura do termo de erro composto. Considere o modelo geral (5.1) em que  $\alpha_i$  representa efeitos específicos invariantes

no tempo e  $\varepsilon_{it}$  é o erro idiossincrático. O estimador de efeitos aleatórios assume  $E(\alpha_i|X_{it}) = 0$ , implicando  $Cov(X_{it}, \alpha_i) = 0$ .

O estimador de efeitos fixos, por sua vez, não requer a independência entre  $\alpha_i$  e  $X_{it}$ . A eliminação de  $\alpha_i$  é realizada via transformação *within*, que remove componentes invariantes no tempo e baseia a estimação exclusivamente na variação intraunidade. Consequentemente, mesmo quando  $Cov(X_{it}, \alpha_i) \neq 0$ , o estimador FE permanece consistente, desde que se satisfaça a condição de exogeneidade estrita:

Assim, a rejeição do teste de Hausman fornece evidência econométrica formal de que apenas o modelo de efeitos fixos produz estimativas consistentes e não viesadas no contexto analisado, justificando sua adoção como especificação principal do estudo. Além das especificações em nível, estimou-se o modelo (3.1) com efeitos fixos com todas as variáveis em logaritmo

Tal especificação permite interpretar diretamente os coeficientes como elasticidades, facilitando a análise da sensibilidade percentual do PIB per capita às variações na desnutrição e nos demais indicadores de capital humano. Conforme Gujarati e Porter (2011), modelos log-lineares são especialmente apropriados quando o interesse está em efeitos proporcionais e comparabilidade entre regiões de tamanhos distintos.

Os modelos Efeitos Fixos (FE) e Efeitos Aleatórios (RE) constituem diferentes estratégias econométricas para estimar relações causais em dados em painel, cada um fundamentado em pressupostos distintos sobre a existência e o papel da heterogeneidade não observada entre as unidades analisadas. Assim os modelos não competem, mas cumprem papéis complementares dentro da estratégia econométrica: FE garante consistência sob heterogeneidade correlacionada e o RE oferece ganhos de eficiência quando seus pressupostos são válidos. Essa estrutura comparativa permite selecionar, de forma econométrica e fundamentada, a especificação mais adequada ao fenômeno econômico analisado.

### 5.1 Testes e critério de Seleção de Modelos (Teste de Hausman).

O Teste de Hausman (1978) é um teste de especificação que compara dois estimadores: um consistente sob a hipótese nula e outro que é eficiente sob a hipótese nula, mas inconsistente sob a alternativa. As hipóteses do teste são:

- $H_0: Cov(X_{it}, \mu_i) = 0$  (efeitos aleatórios são apropriados)
- $H_1: Cov(X_{it}, \mu_i) \neq 0$  (efeitos fixos são necessários)

O teste é baseado na diferença entre estimadores de Efeitos ( $\hat{\beta}_{FE}$ ) e efeitos aleatórios ( $\hat{\beta}_{RE}$ ) baseado no modelo 3.11 onde,  $H \sim X^2(k)$  sendo  $k$  o número de parâmetros estimados e  $\text{Var}(\hat{\beta}_{FE})$  e  $\text{Var}(\hat{\beta}_{RE})$  são as matrizes de covariância assintóticas.

Na Interpretação pratica a regra de decisão para um nível de significância de 5% é:

$$\begin{cases} \text{Se } p - \text{valor} < 0.05 \Rightarrow \text{Rejeita } H_0 \rightarrow \text{Usar Efeitos Fixos.} \\ \text{Se } p - \text{valor} \geq 0.05 \Rightarrow \text{Não Rejeita } H_0 \rightarrow \text{Usar Efeitos Aletório.} \end{cases}$$

Desta forma podemos testar dentro do grupo de dados que foram coletados da POF e dos outros grupos de dados retirados do IBGE para verificar se existe heterogeneidade entre estados, é estatisticamente significativo, esta heterogeneidade está correlacionada com as variáveis explicativas e se o modelo de efeitos fixos é a especificação apropriada. A presença de endogeneidade constitui uma preocupação central na estimação de modelos econométricos aplicados ao crescimento econômico e ao capital humano. De forma geral, a endogeneidade ocorre quando uma ou mais variáveis explicativas estão correlacionadas com o termo de erro do modelo, violando a condição de exogeneidade estrita e comprometendo a consistência dos estimadores.

No contexto deste estudo, existem potenciais fontes de endogeneidade, especialmente associadas às variáveis de desnutrição e anos médios de estudo. Primeiramente, pode haver causalidade reversa, uma vez que níveis mais elevados de PIB rural *per capita* podem, simultaneamente, reduzir a desnutrição e ampliar o acesso à educação. Em segundo lugar, fatores não observados podem afetar simultaneamente o crescimento econômico e os indicadores de capital humano, gerando viés por omissão de variáveis.

Considerando a natureza das variáveis explicativas centrais deste estudo — capital humano educacional e capital humano em saúde, reconhece a possibilidade teórica de endogeneidade, seja por causalidade reversa, seja por omissão de fatores não observados que possam influenciar simultaneamente o desempenho econômico e os indicadores de capital humano. Para resolver esse problema, a estratégia empírica adotada baseia-se na estimação de modelos de dados em painel com efeitos fixos. Esse estimador permite controlar a heterogeneidade não observada invariante no tempo entre as unidades da federação, reduzindo potenciais vieses associados a características estruturais persistentes, como aspectos institucionais, históricos e geográficos.

Para garantir a veracidade dos resultados foram realizados testes formais de endogeneidade para as variáveis de capital humano educacional e de capital humano em saúde,

por meio do teste de Hausman, utilizando defasagens temporais dessas variáveis como instrumentos. A hipótese nula do teste estabelece a exogeneidade das variáveis explicativas, isto é, a ausência de correlação estatisticamente significativa com o termo de erro do modelo.

Tabela 7– Testes de Adequação do Modelo - Dados em Painel.

Teste	Estatística ( $\chi^2$ )	p-valor
Hausman	15.2340	0.0042

Fonte: Elaboração própria com base nos resultados do modelo.

O teste de Hausman rejeitou a hipótese nula  $H_0$  de consistência do modelo de efeitos aleatórios de  $\chi^2 = 15.2340$  e  $p < 0.0042$ , evidenciando correlação entre os efeitos individuais e os regressores. Dessa forma, o modelo de efeitos fixos foi adotado como estimador consistente.

Os resultados dos testes não indicaram rejeição da hipótese nula, sugerindo que não há evidência estatística de endogeneidade das variáveis de educação e saúde no contexto deste estudo. Logo, os estimadores obtidos pelo modelo de efeitos fixos podem ser considerados consistentes, não havendo necessidade do emprego de modelos com variáveis instrumentais. Assim a metodologia originalmente adotada mostra-se adequada para os objetivos da pesquisa, e as inferências realizadas a partir das estimativas apresentadas permanecem válidas e robustas.

## 5.2 Resultado da Estimação dos Modelos em Painel.

A partir dessa base, foram implementados os procedimentos econométricos apresentados ao longo do trabalho, incluindo a estimação de modelos de dados em painel com efeitos fixos, bem como testes econométricos complementares voltados à verificação da adequação da estratégia empírica adotada.

Conforme a Tabela 8 as especificações em nível e na forma log-log permite uma avaliação mais consistente do impacto relativo das variáveis IDR e AME sobre o PIB. No modelo de Efeitos Fixos em nível, ambos os coeficientes são estatisticamente significativos, sendo o IDR (-6.569\*\*\*) superior em magnitude absoluta ao AME (5.066\*\*). Esse resultado sugere que, em termos unitários, variações no IDR estão associadas a alterações mais intensas no PIB quando comparadas às variações nos anos médios de estudo. A análise conjunta das especificações em nível e na forma log-log permite uma avaliação mais consistente do impacto relativo das variáveis IDR e AME sobre o PIB.

Esse resultado sugere que, em termos unitários, variações no IDR estão associadas a

alterações mais intensas no PIB *per capita* quando comparadas às variações nos anos médios de estudo.

Tabela 8 – Resultados Estimados dos Modelos Econométricos.

Variável Dependente: PIB e ln(PIB)		
Variável	Efeitos Fixos (PIB)	Log-Log FE (ln PIB)
IDR	-6.569*** (2.337) [t = -2.81]	
AME	5.066** (1.982) [t = 2.56]	
ln(IDR)		2.135** (0.900) [t = 2.37]
ln(AME)		-0.802 (1.582) [t = -0.51]
GINI	51.776*** (16.590) [t = 3.12]	-0.031 (0.878) [t = -0.40]
Pobreza	-12.019 (23.387) [t = -0.51]	1.273 (0.887) [t = 1.44]
DAMMI - Centro-Oeste	-12.73 (22.51) [t = -0.57] n.s.	-12.73 (22.51) [t = -0.57] n.s.
DAMMI - Nordeste	-34.16 (3.75) [t = -9.11] ***	-34.16 (3.75) [t = -9.11] ***
DAMMI - Norte	-38.59 (4.30) [t = -8.97] ***	-38.59 (4.30) [t = -8.97] ***
DAMMI - Sudeste	-49.01 (8.71) [t = -5.63] ***	-49.01 (8.71) [t = -5.63] ***
DAMMI - Sul	-30.97 (1.57) [t = -19.73] ***	-30.97 (1.57) [t = -19.73] ***
Observações	81	81
R <sup>2</sup>	0.505	0.419
R <sup>2</sup> Ajustado	0.208	0.070
Estatística F	12.742*** (df= 4; 50)	8.999*** (df= 4; 50)

Notas: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01 | Erros padrão em parênteses | Estatística t entre colchetes | n.s. = não significativo

Fonte: Elaborada pelo autor.

Esse resultado sugere que, em termos unitários, variações no IDR estão associadas a

alterações mais intensas no PIB *per capita* quando comparadas às variações nos anos médios de estudo.

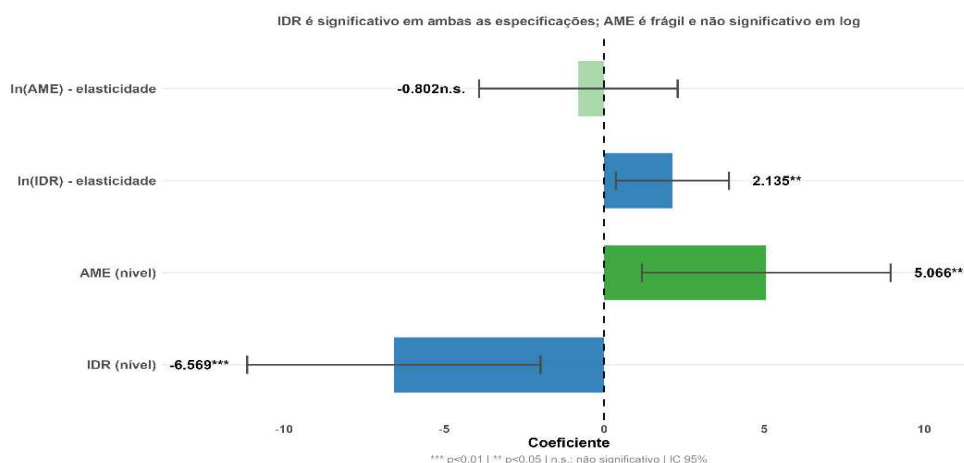
Já na especificação log-log com Efeitos Fixos, a análise torna-se mais adequada para comparação, pois os coeficientes passam a representar elasticidades. Nessa forma funcional, o coeficiente de  $\ln(\text{IDR})$  (2.135\*\*) indica que um aumento de 1% no IDR está associado a um crescimento de aproximadamente 2.13% no PIB, mantendo-se constantes as demais variáveis. Por outro lado,  $\ln(\text{AME})$  não apresentou significância estatística, sugerindo que, em termos proporcionais, seu efeito não se mostra robusto na presença dos demais controles. Assim, considerando conjuntamente a significância estatística, a consistência entre especificações e a interpretação em elasticidades, os resultados indicam que o IDR exerce influência relativamente mais forte e mais robusta sobre o PIB do que o AME.

Para responder à pergunta sobre a existência de diferenças regionais no PIB dos estados brasileiros, foi conduzida por meio de três modelos complementares.

No Gráfico 7 apresenta de forma mais intuitiva os resultados do modelo de efeitos fixos com especificação log-log. A transformação logarítmica confere aos coeficientes a interpretação direta de elasticidades.

O coeficiente estimado para o IDR é +2,135 (IC 95%: 0,371 a 3,899;  $p = 0,021$ ). Um aumento de 1% no IDR está associado a um acréscimo de 2,135% no PIB estadual. O intervalo de confiança não contém o zero e o p-valor é inferior a 0,05, indicando significância estatística.

Gráfico 7 – Modelo de Efeito Fixo: IDR vs AME

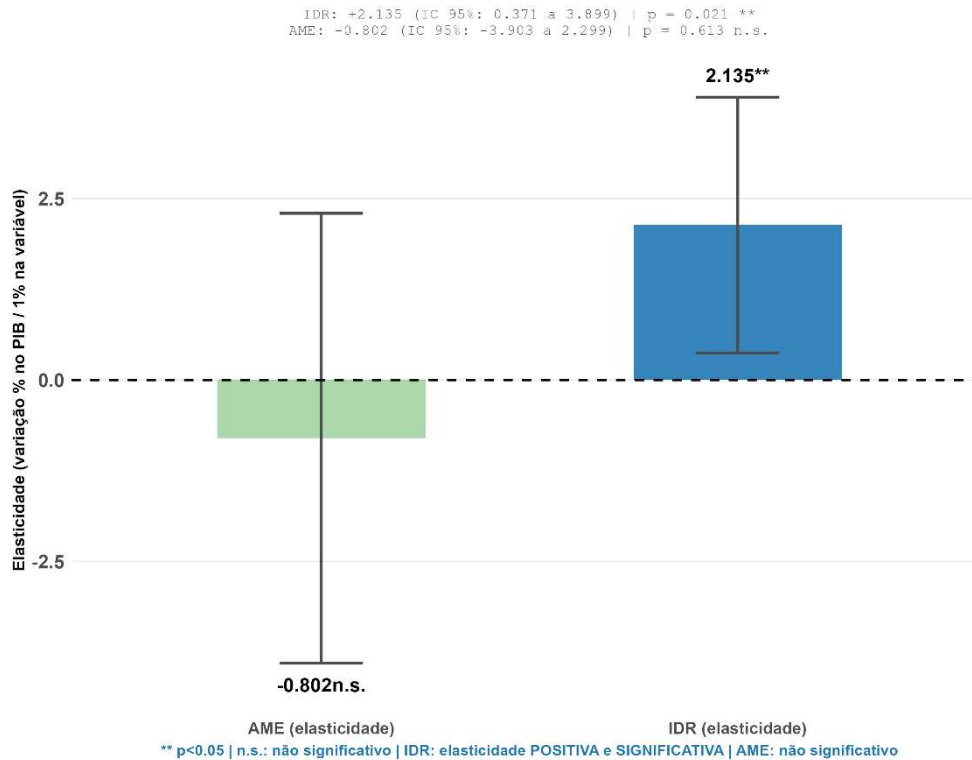


Fonte: Elaboração própria com base nos resultados do modelo.

O coeficiente estimado para o AME é -0,802 (IC 95%: -3,903 a 2,299;  $p = 0,613$ ). Um aumento de 1% no AME está associado a uma redução de 0,802% no PIB estadual. O

intervalo de confiança contém o zero e o p-valor é superior a 0,05, indicando ausência de significância estatística.

Gráfico 8 – Elasticidade: Modelo Log com Efeitos Fixos.



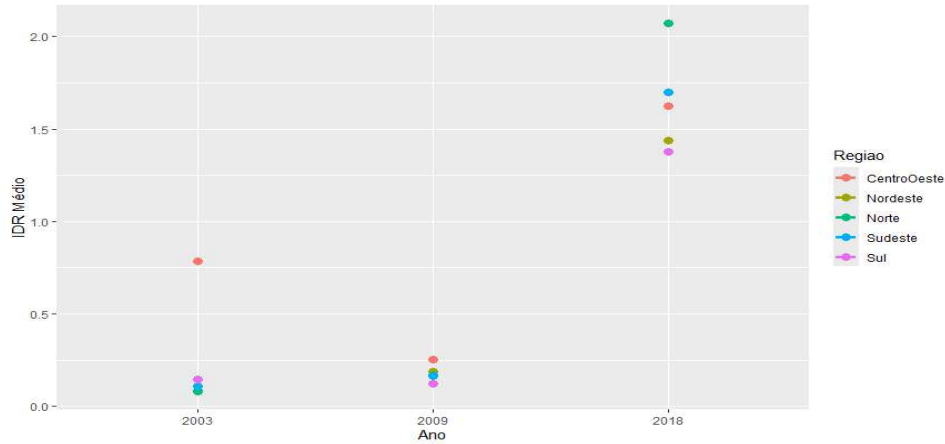
Fonte: Elaboração própria com base nos resultados do modelo.

Observa-se no Gráfico 8 que a elasticidade do IDR é positiva e estatisticamente significativa. A elasticidade do AME é negativa e não estatisticamente significativa. O IDR apresenta erro-padrão de 0,900, inferior ao erro-padrão do AME (1,582). O intervalo de confiança do IDR é mais estreito e não cruza zero, enquanto o intervalo do AME é amplo e cruza zero.

O IDR mantém significância estatística nas duas especificações estimadas. O AME é significativo apenas no modelo de efeitos fixos em nível, perdendo significância na especificação log-log. Os resultados indicam que o IDR possui capacidade preditiva superior ao AME para o PIB estadual no período analisado.

O Gráfico 9 apresenta a distribuição do Índice de Desnutrição (IDR) por região, considerando a amostra em escala logarítmica. Observa-se heterogeneidade regional acentuada, com as regiões Norte e Centro-Oeste exibindo maior dispersão e valores máximos mais elevados, indicando maior variabilidade nos níveis de desnutrição.

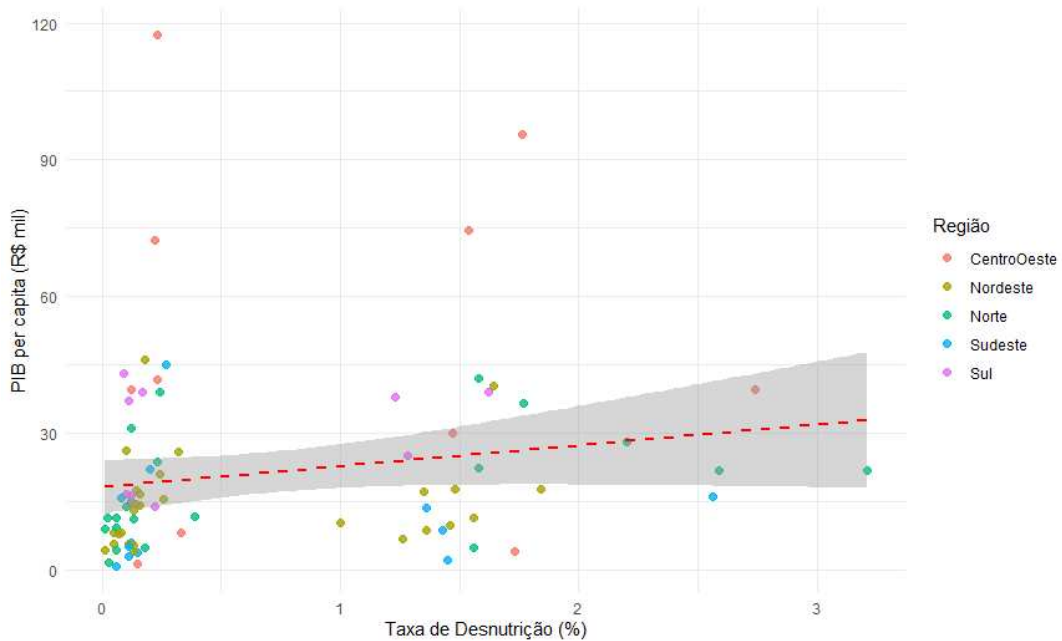
Gráfico 9 – Evolução da Desnutrição por Região.



Fonte: Elaboração própria com base nos resultados do modelo.

As medianas são relativamente próximas entre as regiões, porém a presença de caudas superiores longas sugere a ocorrência de episódios extremos de desnutrição, especialmente no Norte. Em contraste, Sul e Nordeste apresentam menor variabilidade, indicando distribuição mais concentrada do IDR.

Gráfico 10 – PIB *per capita* e Taxa de Desnutrição.



Fonte: Elaboração própria com base nos resultados do modelo.

No Gráfico 10 a linha de tendência indica uma relação positiva fraca entre as variáveis, embora com amplo intervalo de confiança, sugerindo baixa precisão na estimativa e

possível heterogeneidade regional. Nota-se que regiões como Centro-Oeste e Sudeste concentram os maiores níveis de PIB per capita, enquanto Norte e Nordeste apresentam valores mais baixos e maior dispersão.

Observa-se que os maiores valores do índice de desnutrição concentram-se, de maneira mais evidente, nas regiões Nordeste e Norte, especialmente no início do período analisado (2003). Nessas regiões, os pontos situam-se em níveis mais elevados do eixo vertical, indicando maior intensidade do problema nutricional. Contudo, ao longo dos anos, percebe-se um deslocamento descendente dos pontos, particularmente em 2018, sugerindo redução progressiva dos índices de desnutrição.

No Sudeste e no Sul, os níveis de desnutrição apresentam-se estruturalmente mais baixos em comparação às regiões Norte e Nordeste. Ainda assim, o gráfico evidencia queda consistente dos índices ao longo do período, indicando melhora das condições nutricionais mesmo em regiões que já partiam de patamares relativamente mais favoráveis. O Centro-Oeste apresenta posição intermediária, com redução gradual do índice, embora com alguma dispersão entre os estados.

## 6 CONCLUSÃO.

Esta dissertação investigou os fatores explicativos dos diferenciais de renda per capita rural entre os estados brasileiros, comparando o papel do capital humano nas dimensões saúde e educação. A hipótese central era que, dadas as características do meio rural — atividades intensivas em trabalho físico, alta informalidade e menor acesso a serviços —, o capital humano na forma de saúde poderia ser um determinante mais relevante da renda rural per capita do que o capital humano na forma de educação.

Toda a análise foi realizada através de uma base de dados coletada a partir dos microdados das Pesquisas de Orçamentos Familiares (POF) de 2003, 2009 e 2018. Desenvolveu-se um Índice de Desnutrição Rural (IDR) baseado na ingestão domiciliar de cinco micronutrientes essenciais (Cálcio, Ferro, Vitaminas A, B1 e B2), capturando a chamada "fome oculta".

Os resultados descritivos do IDR revelaram padrões distintos ao longo do período analisado e entre os setores de atividade econômica. Em 2003, os índices eram baixos na maioria dos estados, com exceção de Mato Grosso (2,74), e a desnutrição concentrava-se predominantemente no setor agrícola. Em 2009, observou-se elevação moderada em estados como Roraima (0,39), Maranhão (0,32) e Piauí (0,26), com o setor não agrícola ganhando participação. O ano de 2018 marcou um agravamento expressivo: Amazonas (3,21), Roraima (2,59) e Minas Gerais (2,56) registraram os maiores valores, e os setores não agrícola e contas próprias passaram a dominar a composição do indicador. As regiões Norte e Nordeste apresentaram os maiores níveis de desnutrição ao longo de todo o período, seguidas pelo Centro-Oeste, onde Mato Grosso se destacou com índices elevados no setor agrícola, enquanto Sul e Sudeste mantiveram patamares mais moderados.

Os resultados mostraram que o IDR apresentou uma relação negativa e estatisticamente significativa com o PIB rural per capita, indicando que variações no estado nutricional estão associadas a variações proporcionalmente relevantes na renda. No modelo log-log com efeitos fixos, a elasticidade estimada foi de 2,135 ( $p = 0,021$ ), ou seja, um aumento de 1% no IDR está associado a um acréscimo de aproximadamente 2,13% no PIB rural per capita.

Em contraste, os anos médios de estudo não apresentaram relação estatisticamente significativa com o PIB rural per capita no modelo log-log (coeficiente de -0,802;  $p = 0,613$ ), com intervalo de confiança amplamente incluindo o zero. Estes dois últimos resultados permitem concluir que, das formas de capital humano consideradas, a saúde é muito mais relevante para explicar os diferenciais de renda rural per capita dos estados brasileiros.

A análise das dummies regionais confirmou heterogeneidade estrutural persistente, com as regiões Norte e Nordeste apresentando os piores indicadores nutricionais e os maiores déficits estruturais de renda. No modelo com controles, todas as regiões apresentaram renda inferior à do Centro-Oeste: Sudeste (-42,17), Norte (-31,85), Nordeste (-27,43) e Sul (-24,63).

Quanto às demais variáveis do modelo, o índice de Gini apresentou coeficiente negativo e significativo (-16,526;  $p < 0,01$ ), confirmando que maior desigualdade de renda está associada a menor PIB per capita rural. O índice de pobreza também mostrou efeito negativo e significativo (-30,420;  $p < 0,01$ ), indicando que maiores taxas de pobreza comprometem o desempenho econômico das áreas rurais. Esses resultados reforçam o entendimento de que tanto a concentração de renda quanto a privação material atuam como obstáculos estruturais ao crescimento da renda no campo.

A evolução temporal mostrou redução da desnutrição e da pobreza, bem como aumento da escolaridade em todas as regiões, com quedas expressivas da pobreza em estados como Maranhão (de 0,59 para 0,22) e Ceará (de 0,48 para 0,19), e avanços nos anos médios de estudo, que ultrapassaram 10 anos em estados do Sul e Sudeste em 2018. Contudo, a relação com o crescimento da renda mostrou-se mais forte e consistente para o IDR do que para a educação formal, sugerindo que, no contexto rural, a melhoria das condições nutricionais pode ter um impacto mais imediato e direto sobre a produtividade e a renda das pessoas.

A principal contribuição desta dissertação é preencher uma lacuna na literatura nacional ao construir um indicador nutricional inovador a partir de microdados domiciliares e demonstrar, com rigor metodológico, a relevância da saúde como dimensão do capital humano para o crescimento da renda rural, comparando explicitamente as duas dimensões e controlando por fatores estruturais e regionais.

Pesquisas futuras podem avançar com estratégias de identificação mais robustas para lidar com possíveis problemas de endogeneidade, análises com a inclusão de outras edições da POF assim que estiverem disponíveis, e investigação das interações entre as dimensões do capital humano, explorando como saúde e educação podem se complementar na determinação da renda rural.

Em suma, esta dissertação demonstra que a saúde nutricional é um determinante fundamental e ainda subestimado da renda rural no Brasil. Seus resultados reforçam que investimentos em segurança alimentar e nutricional não devem ser vistos apenas como políticas sociais compensatórias, mas como componentes estratégicos de uma agenda de desenvolvimento econômico, capazes de reduzir desigualdades regionais e promover aumento sustentável da renda per capita no meio rural brasileiro.

## REFERÊNCIAS

- AGÉNOR, Pierre-Richard. *The economics of adjustment and growth*. 2. ed. Cambridge: Harvard University Press, 2004.
- BALTAGI, Badi H. *Econometric analysis of panel data*. 3. ed. Chichester: John Wiley & Sons, 2005.
- BALTAGI, Badi H. *Econometric analysis of panel data*. 5. ed. Chichester: Wiley, 2013.
- BARRO, Robert J. Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, Cambridge, v. 106, n. 2, p. 407–443, 1991.
- BARRO, Robert J. *Health and economic growth*. Cambridge: Harvard University Press, 1996.
- BARRO, Robert J. Health and economic growth. *World Bank Research Observer*, v. 11, n. 2, p. 99–121, 1996.
- BECKER, Gary S. *Human capital: a theoretical and empirical analysis, with special reference to education*. Chicago: University of Chicago Press, 1964.
- BEHRMAN, Jere R.; DEOLALIKAR, Anil B. Health and nutrition. In: CHENERY, Hollis; SRINIVASAN, T. N. (org.). *Handbook of Development Economics*. Amsterdam: Elsevier, 1988. v. 1, p. 631–711.
- BLOOM, David E.; CANNING, David; SEVILLA, Jaypee. The effect of health on economic growth. *World Development*, v. 32, n. 1, p. 1–13, 2004.
- CARVALHO, Sergio. *Contabilidade nacional: uma abordagem integrada*. São Paulo: Atlas, 2018.
- CURRIE, Janet; VOGL, Tom. Early-life health and adult circumstance in developing countries. *Annual Review of Economics*, v. 5, p. 1–36, 2013.
- DASGUPTA, Partha; RAY, Debraj. Inequality as a determinant of malnutrition and unemployment: theory. *The Economic Journal*, v. 96, n. 384, p. 1011–1034, 1986.
- DEATON, Angus. *The analysis of household surveys: a microeconomic approach to development policy*. Baltimore: Johns Hopkins University Press, 1997.
- DEATON, Angus; ZAIDI, Salman. *Guidelines for constructing consumption aggregates for welfare analysis*. Washington, DC: World Bank, 2002.
- DEOLALIKAR, Anil B. Nutrition and labor productivity in agriculture: estimates for rural

South India. *The Review of Economics and Statistics*, v. 70, n. 3, p. 406–413, 1988.

FAO – FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. *The State of Food Insecurity in the World*. Rome: FAO, 2011.

FAO – FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. *The State of Food Insecurity in the World*. Rome: FAO, 2013.

FAO; IFAD; UNICEF; WFP; WHO. *The State of Food Security and Nutrition in the World*. Rome: FAO, 2019.

FOGEL, Robert W. Economic growth, population theory, and physiology. *American Economic Review*, v. 84, n. 3, p. 369–395, 1994.

FOSTER, James; GREER, Joel; THORBECKE, Erik. A class of decomposable poverty measures. *Econometrica*, v. 52, n. 3, p. 761–766, 1984.

GREENE, William H. *Econometric analysis*. 8. ed. New York: Pearson, 2018.

GROSSMAN, Michael. On the concept of health capital and the demand for health. *Journal of Political Economy*, v. 80, n. 2, p. 223–255, 1972.

GUJARATI, Damodar N.; PORTER, Dawn C. *Econometria básica*. 5. ed. Porto Alegre: AMGH, 2011.

HANUSHEK, Eric A.; WOESSMANN, Ludger. The role of cognitive skills in economic development. *Journal of Economic Literature*, v. 46, n. 3, p. 607–668, 2012.

HECKMAN, James J. Skill formation and the economics of investing in disadvantaged children. *Science*, v. 312, n. 5782, p. 1900–1902, 2006.

HODDINOTT, John et al. The economic rationale for investing in stunting reduction. *Maternal & Child Nutrition*, v. 9, supl. 2, p. 69–82, 2013.

HOFFMANN, Rodolfo. *Distribuição de renda: medidas de desigualdade e pobreza*. São Paulo: Edusp, 2008.

HOFFMANN, Rodolfo. Determinantes da insegurança alimentar no Brasil. *Segurança Alimentar e Nutricional*, v. 15, n. 1, p. 1–17, 2008.

HOFFMANN, Rodolfo. *Análise comparativa de indicadores sociais no Brasil: 2003–2009*. Brasília: Embrapa, 2014.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Contas nacionais:*

*metodologia e séries históricas*. Rio de Janeiro: IBGE, 2024.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Pesquisa de Orçamentos Familiares 2017–2018: análise da disponibilidade domiciliar de alimentos e estado nutricional no Brasil*. Rio de Janeiro: IBGE, 2020.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Pesquisa Nacional por Amostra de Domicílios Contínua – Segurança Alimentar 2022*. Rio de Janeiro: IBGE, 2022.

IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. *Nota técnica: famílias urbanas e rurais no Brasil*. Brasília: Ipea, 2022.

LENG, Xuan; ZHANG, Yichuan; CHENG, Ruiqi. Digital technology, labor allocation, and nutrition intake. *Technological Forecasting and Social Change*, v. 201, p. 123228, 2024.

LUCAS, Robert E. On the mechanics of economic development. *Journal of Monetary Economics*, v. 22, n. 1, p. 3–42, 1988.

MANKIW, N. Gregory; ROMER, David; WEIL, David N. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, v. 107, n. 2, p. 407–437, 1992.

MARINHO, Emerson Luís Lemos et al. The impacts of undernutrition on labor productivity. *Revista de Economia e Sociologia Rural*, v. 63, n. 2, e293231, 2025.

PRESTON, Samuel H. The changing relation between mortality and economic development. *Population Studies*, v. 29, n. 2, p. 231–248, 1975.

ROMER, Paul M. Increasing returns and long-run growth. *Journal of Political Economy*, v. 94, n. 5, p. 1002–1037, 1986.

SALA-I-MARTIN, Xavier. Regional cohesion. *European Economic Review*, v. 40, n. 6, p. 1325–1352, 1996.

SCHULTZ, Theodore W. Investment in human capital. *American Economic Review*, v. 51, n. 1, p. 1–17, 1961.

SEN, Amartya. *Development as freedom*. Oxford: Oxford University Press, 1999.

SOLOW, Robert M. A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, v. 70, n. 1, p. 65–94, 1956.

STRAUSS, John; THOMAS, Duncan. Health, nutrition, and economic development. *Journal*

*of Economic Literature*, v. 36, n. 2, p. 766–817, 1998.

WEIL, David N. Health and economic growth. In: AGHION, Philippe; DURLAUF, Steven (org.). *Handbook of Economic Growth*. Amsterdam: Elsevier, 2013. v. 2, p. 623–682.

WEIL, David N. *Economic growth*. 3. ed. New York: Routledge, 2013.

WOOLDRIDGE, Jeffrey M. *Econometric analysis of cross section and panel data*. 2. ed. Cambridge: MIT Press, 2010.

WORLD HEALTH ORGANIZATION (WHO). *Closing the gap in a generation*. Geneva: WHO, 2008.

## APÊNDICE A – DADOS AGRUPADOS DA POF 2003, 2009 E 2018.

2003					
Estados	(Setor Agrícola)	(Setor Não Agrícola)	(Contas Próprias)	(Outros Empregos)	Desnutrição
Rondônia	0,01	0,09	0,02	0,00	0,13
Acre	0,00	0,05	0,01	0,00	0,06
Amazonas	0,00	0,01	0,00	0,00	0,01
Roraima	0,01	0,11	0,05	0,01	0,18
Pará	0,01	0,03	0,02	0,00	0,06
Amapá	0,00	0,02	0,01	0,00	0,03
Tocantins	0,02	0,05	0,02	0,01	0,10
Maranhão	0,01	0,02	0,04	0,01	0,08
Piauí	0,00	0,01	0,00	0,00	0,01
Ceará	0,00	0,03	0,02	0,01	0,05
RN	0,01	0,05	0,07	0,00	0,13
Paraíba	0,00	0,05	0,04	0,01	0,11
Pernambuco	0,05	0,05	0,02	0,01	0,13
Alagoas	0,01	0,07	0,05	0,01	0,14
Sergipe	0,00	0,06	0,01	0,00	0,07
Bahia	0,02	0,02	0,01	0,00	0,05
Minas Gerais	0,01	0,07	0,03	0,01	0,12
Espírito Santo	0,01	0,11	0,03	0,00	0,15
Rio de Janeiro	0,01	0,03	0,02	0,00	0,06
São Paulo	0,01	0,08	0,02	0,00	0,11
Paraná	0,01	0,08	0,03	0,00	0,12
Santa Catarina	0,01	0,16	0,05	0,00	0,22
RS	0,00	0,08	0,03	0,00	0,10
Mato Grosso do Sul	0,03	0,07	0,03	0,00	0,12
Mato Grosso	2,48	0,13	0,09	0,04	2,74
Goiás	0,02	0,07	0,03	0,00	0,12
DF	0,00	0,12	0,02	0,00	0,15

2009					
Estados	(Setor Agrícola)	(Setor Não Agrícola)	(Contas Próprias)	(Outros Empregos)	Desnutrição
Rondônia	0,01	0,02	0,03	0,00	0,06
Acre	0,02	0,14	0,06	0,00	0,23
Amazonas	0,00	0,08	0,04	0,00	0,12
Roraima	0,00	0,29	0,08	0,02	0,39
Pará	0,01	0,07	0,05	0,00	0,12
Amapá	0,00	0,02	0,00	0,00	0,02
Tocantins	0,04	0,12	0,06	0,02	0,24
Maranhão	0,05	0,11	0,14	0,02	0,32
Piauí	0,07	0,09	0,06	0,03	0,26
Ceará	0,02	0,10	0,03	0,01	0,16
RN	0,02	0,11	0,02	0,01	0,16
Paraíba	0,02	0,07	0,05	0,01	0,14
Pernambuco	0,02	0,08	0,03	0,01	0,13
Alagoas	0,02	0,10	0,05	0,00	0,18
Sergipe	0,01	0,07	0,02	0,00	0,10
Bahia	0,06	0,10	0,06	0,02	0,24
Minas Gerais	0,04	0,12	0,03	0,00	0,20
Espírito Santo	0,03	0,19	0,05	0,00	0,27
Rio de Janeiro	0,00	0,09	0,02	0,00	0,11
São Paulo	0,00	0,05	0,03	0,00	0,08
Paraná	0,01	0,12	0,03	0,01	0,17
Santa Catarina	0,00	0,07	0,02	0,01	0,11
RS	0,00	0,06	0,02	0,00	0,09
Mato Grosso do Sul	0,03	0,14	0,04	0,01	0,22
Mato Grosso	0,02	0,14	0,05	0,01	0,23
Goiás	0,02	0,17	0,04	0,00	0,23
DF	0,01	0,29	0,03	0,00	0,33

2018					
Estados	(Setor Agrícola)	(Setor Não Agrícola)	(Contas Próprias)	(Outros Empregos)	Desnutrição
Rondônia	0,22	0,90	0,39	0,07	1,58
Acre	0,11	0,82	0,46	0,19	1,58
Amazonas	0,26	1,98	0,78	0,19	3,21
Roraima	0,05	1,22	1,30	0,03	2,59
Pará	0,23	1,20	0,68	0,09	2,20
Amapá	0,07	0,89	0,57	0,03	1,56
Tocantins	0,77	1,22	0,43	0,05	1,77
Maranhão	0,19	1,00	0,58	0,07	1,84
Piauí	0,18	0,89	0,35	0,06	1,48
Ceará	0,13	0,91	0,45	0,07	1,56
RN	0,10	0,63	0,24	0,02	1,00
Paraíba	0,20	0,80	0,34	0,03	1,36
Pernambuco	0,13	0,87	0,38	0,08	1,46
Alagoas	0,05	1,00	0,47	0,12	1,64
Sergipe	0,08	0,82	0,28	0,07	1,26
Bahia	0,12	0,83	0,34	0,05	1,35
Minas Gerais	0,07	1,17	0,63	0,70	2,56
Espírito Santo	0,06	0,90	0,33	0,07	1,36
Rio de Janeiro	0,01	0,96	0,39	0,10	1,45
São Paulo	0,02	0,99	0,34	0,09	1,43
Paraná	0,08	1,19	0,29	0,06	1,62
Santa Catarina	0,12	0,80	0,26	0,10	1,28
RS	0,09	0,82	0,25	0,07	1,23
Mato Grosso do Sul	0,03	1,06	0,38	0,06	1,54
Mato Grosso	0,10	0,79	0,78	0,09	1,76
Goiás	0,01	1,00	0,32	0,14	1,47
DF	0,02	1,21	0,37	0,13	1,73

## APÊNDICE B – SCRIPTE DO TRATAMENTO DO MODELO R STUDIO

Descrição: Este código realiza análise de dados em painel para investiga a relação entre desnutrição e crescimento econômico nos estados brasileiro

Autor: Paulo Henrique Pereira Arruda

Data: 19/12/2025

Disciplina: Defesa de Dissertação de Mestrado.

Universidade Federal do Ceará (UFC)

# MODELO ECONOMETRICO EM PAINEL COM DADOS REGIONAIS  
BRASILEIROS

# ANÁLISE COMPLETA - 2003, 2009, 2018

# VERSÃO CORRIGIDA

#####

#####

# 1) INSTALAÇÃO E CARREGAMENTO DE PACOTES

#####

```
pacotes <- c("plm", "tidyverse", "lmtest", "sandwich",
            "ggplot2", "stargazer", "car", "tseries",
            "knitr", "kableExtra", "corrplot", "performance",
            "ggpubr", "reshape2")
```

```
instalar <- pacotes[!(pacotes %in% installed.packages()[,"Package"])]
```

```
if(length(instalar)) install.packages(instalar)
```

```
lapply(pacotes, library, character.only = TRUE)
```

#####

# 2) BASE DE DADOS EMBUTIDA

#####

```
dados <- read.table(text="
```

```
Estado Regiao ANO PIB IDR AME GINI POBREZA
```

RO Norte 2003 11.17 0.13 5.56 0.55 0.27  
AC Norte 2003 11.43 0.06 5.25 0.57 0.44  
AM Norte 2003 8.82 0.01 6.55 0.51 0.45  
RR Norte 2003 4.82 0.18 6.08 0.53 0.40  
PA Norte 2003 9.14 0.06 5.44 0.34 0.41  
AP Norte 2003 1.52 0.03 7.31 0.52 0.38  
TO Norte 2003 13.88 0.10 5.20 0.64 0.48  
MA Nordeste 2003 7.98 0.08 3.85 0.28 0.59  
PI Nordeste 2003 4.26 0.01 4.16 0.31 0.54  
CE Nordeste 2003 5.53 0.05 5.05 0.36 0.48  
RN Nordeste 2003 5.27 0.13 5.72 0.54 0.43  
PB Nordeste 2003 5.59 0.11 5.17 0.35 0.48  
PE Nordeste 2003 3.87 0.13 5.94 0.41 0.39  
AL Nordeste 2003 14.43 0.14 5.02 0.28 0.54  
SE Nordeste 2003 7.93 0.07 5.18 0.55 0.49  
BA Nordeste 2003 8.21 0.05 5.54 0.27 0.41  
MG Sudeste 2003 5.87 0.12 6.05 0.23 0.18  
ES Sudeste 2003 3.79 0.15 6.96 0.28 0.19  
RJ Sudeste 2003 0.79 0.06 8.02 0.57 0.16  
SP Sudeste 2003 5.00 0.11 7.89 0.57 0.08  
PR Sul 2003 16.29 0.12 7.11 0.29 0.14  
SC Sul 2003 13.74 0.22 6.94 0.27 0.11  
RS Sul 2003 16.42 0.10 7.06 0.34 0.11  
MS CentroOeste 2003 39.41 0.12 6.53 0.27 0.20  
MT CentroOeste 2003 39.44 2.74 6.16 0.27 0.26  
GO CentroOeste 2003 14.45 0.12 6.10 0.27 0.27  
DF CentroOeste 2003 1.24 0.15 9.09 0.57 0.12  
RO Norte 2009 4.21 0.06 7.57 0.52 0.10  
AC Norte 2009 23.75 0.23 6.40 0.44 0.23  
AM Norte 2009 15.02 0.12 8.24 0.50 0.27  
RR Norte 2009 11.75 0.39 7.29 0.53 0.31  
PA Norte 2009 30.92 0.12 6.67 0.53 0.28  
AP Norte 2009 11.29 0.02 9.49 0.57 0.25  
TO Norte 2009 39.02 0.24 7.25 0.43 0.19  
MA Nordeste 2009 25.76 0.32 5.01 0.55 0.48

PI Nordeste 2009 15.48 0.26 5.33 0.49 0.31  
 CE Nordeste 2009 16.63 0.16 7.57 0.46 0.37  
 RN Nordeste 2009 14.01 0.16 6.88 0.53 0.29  
 PB Nordeste 2009 17.46 0.14 5.83 0.49 0.38  
 PE Nordeste 2009 12.98 0.13 7.45 0.49 0.24  
 AL Nordeste 2009 45.99 0.18 6.06 0.49 0.40  
 SE Nordeste 2009 26.10 0.10 6.80 0.52 0.26  
 BA Nordeste 2009 21.01 0.24 5.93 0.49 0.27  
 MG Sudeste 2009 22.06 0.20 7.81 0.49 0.11  
 ES Sudeste 2009 44.87 0.27 7.67 0.45 0.16  
 RJ Sudeste 2009 2.93 0.11 10.49 0.54 0.07  
 SP Sudeste 2009 15.79 0.08 8.44 0.45 0.04  
 PR Sul 2009 38.89 0.17 11.23 0.45 0.09  
 SC Sul 2009 37.05 0.11 8.94 0.42 0.06  
 RS Sul 2009 43.09 0.09 8.57 0.45 0.05  
 MS CentroOeste 2009 72.40 0.22 8.39 0.47 0.06  
 MT CentroOeste 2009 117.29 0.23 7.70 0.52 0.15  
 GO CentroOeste 2009 41.56 0.23 7.61 0.47 0.12  
 DF CentroOeste 2009 8.04 0.33 9.39 0.50 0.07  
 RO Norte 2018 42.00 1.58 8.91 0.49 0.11  
 AC Norte 2018 22.17 1.58 8.87 0.56 0.17  
 AM Norte 2018 21.86 3.21 9.91 0.49 0.24  
 RR Norte 2018 21.79 2.59 10.45 0.47 0.20  
 PA Norte 2018 28.08 2.20 8.30 0.56 0.20  
 AP Norte 2018 4.85 1.56 9.40 0.47 0.12  
 TO Norte 2018 36.51 1.77 9.60 0.39 0.18  
 MA Nordeste 2018 17.75 1.84 8.20 0.42 0.22  
 PI Nordeste 2018 17.70 1.48 7.60 0.40 0.20  
 CE Nordeste 2018 11.46 1.56 8.30 0.44 0.19  
 RN Nordeste 2018 10.39 1.00 8.80 0.47 0.12  
 PB Nordeste 2018 8.52 1.36 8.20 0.51 0.24  
 PE Nordeste 2018 9.71 1.46 8.80 0.47 0.14  
 AL Nordeste 2018 40.33 1.64 8.50 0.48 0.22  
 SE Nordeste 2018 6.84 1.26 8.90 0.47 0.09  
 BA Nordeste 2018 17.15 1.35 8.90 0.53 0.14

```

MG Sudeste 2018 15.87 2.56 9.20 0.31 0.05
ES Sudeste 2018 13.67 1.36 9.70 0.44 0.06
RJ Sudeste 2018 2.15 1.45 10.80 0.47 0.07
SP Sudeste 2018 8.61 1.43 10.70 0.42 0.04
PR Sul 2018 38.86 1.62 9.90 0.47 0.05
SC Sul 2018 25.04 1.28 10.10 0.39 0.01
RS Sul 2018 38.00 1.23 10.10 0.42 0.02
MS CentroOeste 2018 74.62 1.54 9.10 0.42 0.02
MT CentroOeste 2018 95.54 1.76 9.30 0.47 0.04
GO CentroOeste 2018 29.84 1.47 9.90 0.43 0.05
DF CentroOeste 2018 4.05 1.73 11.50 0.49 0.02
", header=TRUE)

```

```
# Conversão para fatores
```

```
dados$Estado <- as.factor(dados$Estado)
```

```
dados$Regiao <- as.factor(dados$Regiao)
```

```
dados$ANO <- as.factor(dados$ANO)
```

```
#####
```

```
# 3) ANÁLISE EXPLORATÓRIA
```

```
#####
```

```
cat("\n===== ANÁLISE EXPLORATÓRIA =====\n")
```

```
cat("\nEstrutura dos dados:\n")
```

```
str(dados)
```

```
cat("\nResumo estatístico:\n")
```

```
summary(dados[,3:8])
```

```
cat("\nQuantidade por ano:\n")
```

```
table(dados$ANO)
```

```
cat("\nQuantidade por região:\n")
```

```
table(dados$Regiao)
```

```

# Estatísticas por região
estatisticas_regiao <- dados %>%
  group_by(Regiao, ANO) %>%
  summarise(
    PIB_medio = mean(PIB, na.rm=TRUE),
    IDR_medio = mean(IDR, na.rm=TRUE),
    AME_medio = mean(AME, na.rm=TRUE),
    GINI_medio = mean(GINI, na.rm=TRUE),
    POBREZA_media = mean(POBREZA, na.rm=TRUE),
    .groups = 'drop'
  )

cat("\nEvolução do PIB médio por região:\n")
print(estatisticas_regiao[,c("Regiao", "ANO", "PIB_medio")])

#####
# 4) VISUALIZAÇÕES
#####

# Configurar tema dos gráficos
theme_set(theme_minimal())

# 4.1 Evolução do PIB por região
p1 <- ggplot(dados, aes(x = as.numeric(as.character(ANO)),
                        y = PIB,
                        color = Regiao,
                        group = Estado)) +
  geom_line(alpha = 0.3) +
  geom_smooth(aes(group = Regiao), method = "loess", se = FALSE, size = 1.5) +
  labs(title = "Evolução do PIB por Estado e Região",
       x = "Ano", y = "PIB (R$ bilhões)") +
  theme(legend.position = "bottom")

# 4.2 Boxplot do PIB por ano e região
p2 <- ggplot(dados, aes(x = as.factor(ANO), y = PIB, fill = Regiao)) +

```

```
geom_boxplot() +
labs(title = "Distribuição do PIB por Ano e Região",
      x = "Ano", y = "PIB (R$ bilhões)") +
theme(legend.position = "bottom")
```

# 4.3 Correlações - CORRIGIDO: gráfico separado

```
dados_numericos <- dados[,c("PIB", "IDR", "AME", "GINI", "POBREZA")]
cor_matrix <- cor(dados_numericos)
```

# 4.4 Scatter plots

```
p3 <- ggplot(dados, aes(x = IDR, y = PIB, color = as.factor(ANO))) +
  geom_point(size = 3, alpha = 0.7) +
  geom_smooth(method = "lm", se = FALSE, aes(group = ANO)) +
  labs(title = "Relação IDR vs PIB por Ano",
        x = "IDR (Índice de Desenvolvimento Regional)",
        y = "PIB (R$ bilhões)",
        color = "Ano") +
  theme(legend.position = "bottom")
```

```
p4 <- ggplot(dados, aes(x = AME, y = PIB, color = as.factor(ANO))) +
  geom_point(size = 3, alpha = 0.7) +
  geom_smooth(method = "lm", se = FALSE, aes(group = ANO)) +
  labs(title = "Relação AME vs PIB por Ano",
        x = "AME (Anos Médios de Estudo)",
        y = "PIB (R$ bilhões)",
        color = "Ano") +
  theme(legend.position = "bottom")
```

# Exibir gráficos

```
print(p1)
print(p2)
print(p3)
print(p4)
```

# CORRIGIDO: Matriz de correlação com nova janela gráfica

```

dev.new()
corrplot(cor_matrix, method = "color", type = "upper",
         tl.col = "black", tl.srt = 45,
         title = "Matriz de Correlação das Variáveis",
         mar = c(0,0,2,0))

# Alternativa: usar ggplot2 para matriz de correlação
library(reshape2)
cor_melted <- melt(cor_matrix)
p_cor <- ggplot(data = cor_melted, aes(x = Var1, y = Var2, fill = value)) +
  geom_tile() +
  geom_text(aes(label = round(value, 2)), size = 3) +
  scale_fill_gradient2(low = "blue", high = "red", mid = "white",
                      midpoint = 0, limit = c(-1, 1), space = "Lab",
                      name = "Correlação") +
  theme_minimal() +
  labs(title = "Matriz de Correlação das Variáveis", x = "", y = "") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        legend.position = "bottom")
print(p_cor)

#####
# 5) ESTRUTURA EM PAINEL
#####

painel <- pdata.frame(dados, index = c("Estado", "ANO"))

# Verificar estrutura do painel
cat("\n===== ESTRUTURA DO PAINEL =====\n")
cat("\nDimensões do painel:\n")
print(dim(painel))
cat("\nPeríodos:\n")
print(levels(painel$ANO))
cat("\nUnidades cross-section:\n")
print(length(unique(painel$Estado)))

```

```
#####
# 6) ESTIMAÇÃO DOS MODELOS
#####

cat("\n===== ESTIMAÇÃO DOS MODELOS =====\n")

# Modelo Pooled (Mínimos Quadrados Ordinários)
pool <- plm(PIB ~ IDR + AME + GINI + POBREZA,
           data = painel, model = "pooling")

# Modelo de Efeitos Fixos (Within)
fe <- plm(PIB ~ IDR + AME + GINI + POBREZA,
          data = painel, model = "within")

# Modelo de Efeitos Aleatórios
re <- plm(PIB ~ IDR + AME + GINI + POBREZA,
          data = painel, model = "random")

#####
# 7) TESTES DE ESPECIFICAÇÃO
#####

cat("\n===== TESTES DE ESPECIFICAÇÃO =====\n")

# Teste de Chow (Pooled vs Efeitos Fixos)
chow_test <- pFtest(fe, pool)
cat("\nTeste de Chow (Pooled vs Efeitos Fixos):\n")
print(chow_test)

# Teste de Breusch-Pagan (Pooled vs Efeitos Aleatórios)
bp_lm <- plmtest(pool, type = "bp")
cat("\nTeste de Breusch-Pagan (Pooled vs Efeitos Aleatórios):\n")
print(bp_lm)
```

```

# Teste de Hausman (Efeitos Fixos vs Efeitos Aleatórios)
hausman <- phtest(fe, re)
cat("\nTeste de Hausman (Fixos vs Aleatórios):\n")
print(hausman)

# Tabela de testes
tabela_testes <- data.frame(
  Teste = c("Chow (F-test)", "Breusch-Pagan LM", "Hausman"),
  Estatistica = c(as.numeric(chow_test$statistic),
                  as.numeric(bp_lm$statistic),
                  as.numeric(hausman$statistic)),
  p_valor = c(chow_test$p.value,
              bp_lm$p.value,
              hausman$p.value),
  Decisao = c(ifelse(chow_test$p.value < 0.05, "Rejeita Pooled", "Aceita Pooled"),
              ifelse(bp_lm$p.value < 0.05, "Rejeita Pooled", "Aceita Pooled"),
              ifelse(hausman$p.value < 0.05, "Efeitos Fixos", "Efeitos Aleat
rios"))
)

cat("\nResumo dos testes de especificação:\n")
print(tabela_testes)

#####
# 8) MODELO LOG-LOG
#####

cat("\n===== MODELO LOG-LOG =====\n")

# Transformação logarítmica
painel$lnPIB <- log(painel$PIB)
painel$lnIDR <- log(painel$IDR + 0.001) # Adicionar constante para evitar log(0)
painel$lnAME <- log(painel$AME)

# Modelo com logs

```

```

fe_log <- plm(lnPIB ~ lnIDR + lnAME + GINI + POBREZA,
              data = painel, model = "within")

re_log <- plm(lnPIB ~ lnIDR + lnAME + GINI + POBREZA,
              data = painel, model = "random")

pool_log <- plm(lnPIB ~ lnIDR + lnAME + GINI + POBREZA,
                data = painel, model = "pooling")

#####
# 9) MODELO COM INTERAÇÕES
#####

cat("\n===== MODELO COM INTERAÇÕES =====\n")

# Modelo com interação IDR * ANO
painel$ANO_numeric <- as.numeric(as.character(painel$ANO))
painel$IDR_ANO <- painel$IDR * painel$ANO_numeric

fe_interacao <- plm(PIB ~ IDR + AME + GINI + POBREZA + IDR_ANO,
                    data = painel, model = "within")

cat("\nModelo de Efeitos Fixos com Interação IDR x Ano:\n")
summary(fe_interacao)

#####
# 10) TESTES DIAGNÓSTICOS
#####

cat("\n===== TESTES DIAGNÓSTICOS =====\n")

# Teste de heterocedasticidade (Breusch-Pagan)
bp_test_pool <- bptest(pool)
cat("\nTeste de heterocedasticidade - Modelo Pooled:\n")
print(bp_test_pool)

```

```

bp_test_fe <- bptest(fe)
cat("\nTeste de heterocedasticidade - Modelo Efeitos Fixos:\n")
print(bp_test_fe)

# Teste de autocorrelação (Wooldridge)
autocor_test <- pwttest(fe)
cat("\nTeste de autocorrelação - Efeitos Fixos:\n")
print(autocor_test)

# Teste de normalidade dos resíduos
jb_test <- jarque.bera.test(residuals(fe))
cat("\nTeste de normalidade - Jarque-Bera (Efeitos Fixos):\n")
print(jb_test)

# Fatores de Inflação da Variância (VIF)
modelo_lm <- lm(PIB ~ IDR + AME + GINI + POBREZA, data = dados)
vif_values <- vif(modelo_lm)
cat("\nFatores de Inflação da Variância (VIF):\n")
print(vif_values)

#####
# 11) TABELA COMPARATIVA DE MODELOS
#####

cat("\n===== COMPARAÇÃO DE MODELOS =====\n")

# Tabela com todos os modelos
stargazer(pool, fe, re, fe_log,
           type = "text",
           column.labels = c("Pooled", "Efeitos Fixos", "Efeitos Aleatórios", "Log-
Log FE"),
           dep.var.labels = c("PIB", "PIB", "PIB", "ln(PIB)"),
           covariate.labels = c("IDR", "AME", "GINI", "Pobreza", "ln(IDR)",
"ln(AME)"),
           out = "modelos_comparacao.txt")

```

```
#####
# 12) EFEITOS REGIONAIS
#####

cat("\n===== ANÁLISE DE EFEITOS REGIONAIS =====\n")

# Extrair efeitos fixos dos estados
efeitos_fixos <- fixef(fe)
efeitos_fixos_df <- data.frame(
  Estado = names(efeitos_fixos),
  Efeito_Fixo = as.numeric(efeitos_fixos)
)

# Adicionar região
regioes_estados <- unique(dados[,c("Estado", "Regiao")])
efeitos_fixos_df <- merge(efeitos_fixos_df, regioes_estados, by = "Estado")

# Ordenar por efeito fixo
efeitos_fixos_df <- efeitos_fixos_df[order(efeitos_fixos_df$Efeito_Fixo, decreasing =
TRUE),]

cat("\nTop 10 estados com maiores efeitos fixos:\n")
print(head(efeitos_fixos_df, 10))

cat("\nBottom 10 estados com menores efeitos fixos:\n")
print(tail(efeitos_fixos_df, 10))

# Efeitos fixos médios por região
efeitos_regiao <- efeitos_fixos_df %>%
  group_by(Regiao) %>%
  summarise(
    Efeito_Medio = mean(Efeito_Fixo),
    Efeito_DP = sd(Efeito_Fixo),
    Minimo = min(Efeito_Fixo),
```

```

    Maximo = max(Efeito_Fixo),
    .groups = 'drop'
  )

cat("\nEfeitos fixos médios por região:\n")
print(efeitos_regiao)

# Gráfico dos efeitos fixos
p5 <- ggplot(efeitos_fixos_df, aes(x = reorder(Estado, Efeito_Fixo),
                                     y = Efeito_Fixo,
                                     fill = Regiao)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  labs(title = "Efeitos Fixos por Estado",
        x = "Estado", y = "Efeito Fixo") +
  theme_minimal() +
  theme(legend.position = "bottom")

print(p5)

#####
# 13) ELASTICIDADES E INTERPRETAÇÃO
#####

cat("\n===== ELASTICIDADES E INTERPRETAÇÃO =====\n")

# Coeficientes do modelo log-log
coef_log <- summary(fe_log)$coefficients
cat("\nElasticidades (Modelo Log-Log):\n")
cat("Elasticidade do PIB em relação ao IDR:", round(coef_log[1,1], 3), "\n")
cat("Interpretação: Um aumento de 1% no IDR está associado a um aumento de",
    round(coef_log[1,1], 3), "% no PIB, mantendo tudo o mais constante.\n\n")

cat("Elasticidade do PIB em relação ao AME:", round(coef_log[2,1], 3), "\n")
cat("Interpretação: Um aumento de 1% nos anos médios de estudo está associado a

```

```

um aumento de",
      round(coef_log[2,1], 3), "% no PIB, mantendo tudo o mais constante.\n\n")

cat("Efeito do GINI:", round(coef_log[3,1], 3), "\n")
cat("Interpretação: Um aumento de 1 unidade no índice GINI está associado a uma
variação de",
      round(coef_log[3,1], 3), "no log do PIB.\n\n")

cat("Efeito da Pobreza:", round(coef_log[4,1], 3), "\n")
cat("Interpretação: Um aumento de 1% na taxa de pobreza está associado a uma
variação de",
      round(coef_log[4,1], 3), "no log do PIB.\n\n")

#####
# 14) PREVISÕES E RESÍDUOS
#####

cat("\n===== ANÁLISE DE RESÍDUOS =====\n")

# Resí duos do modelo escolhido (Efeitos Fixos)
residuos <- residuals(fe)
residuos_df <- data.frame(
  Estado = painel$Estado,
  ANO = painel$ANO,
  Residuos = as.numeric(residuos),
  Ajustados = as.numeric(fitted(fe))
)

# Gráfico dos resí duos
p6 <- ggplot(residuos_df, aes(x = ANO, y = Residuos)) +
  geom_boxplot() +
  geom_hline(yintercept = 0, linetype = "dashed", color = "red") +
  labs(title = "Distribuição dos Resí duos por Ano",
       x = "Ano", y = "Resí duos") +
  theme_minimal()

```

```
print(p6)
```

```
# Resí duos vs Valores Ajustados
p7 <- ggplot(residuos_df, aes(x = Ajustados, y = Residuos)) +
  geom_point() +
  geom_hline(yintercept = 0, linetype = "dashed", color = "red") +
  geom_smooth(method = "loess", se = TRUE) +
  labs(title = "Resí duos vs Valores Ajustados",
        x = "Valores Ajustados", y = "Resí duos") +
  theme_minimal()
```

```
print(p7)
```

```
# QQ-plot para normalidade
p8 <- ggplot(residuos_df, aes(sample = Residuos)) +
  stat_qq() +
  stat_qq_line() +
  labs(title = "QQ-Plot dos Resí duos",
        x = "Quantis Teóricos", y = "Quantis Amostrais") +
  theme_minimal()
```

```
print(p8)
```

```
#####
```

```
# 15) RELATÓRIO FINAL
```

```
#####
```

```
cat("\n\n")
```

```
cat("=====\n")
```

```
cat("          RELATÓRIO FINAL - SUMÁRIO          \n")
```

```
cat("=====\n\n")
```

```
# Determinar melhor modelo com base nos testes
```

```
melhor_modelo <- ifelse(hausman$p.value < 0.05, "Efeitos Fixos", "Efeitos Aleató
```

rios")

```

cat("1. CARACTERÍSTICAS DA AMOSTRA:\n")
cat("  - Período:", paste(levels(dados$ANO), collapse = ", "), "\n")
cat("  - Estados:", length(unique(dados$Estado)), "\n")
cat("  - Regiões:", length(unique(dados$Regiao)), "\n")
cat("  - Observações:", nrow(dados), "\n\n")

cat("2. TESTES DE ESPECIFICAÇÃO:\n")
cat("  - Teste Chow (Pooled vs FE): p-valor =", round(chow_test$p.value, 4), "\n")
cat("  - Teste BP (Pooled vs RE): p-valor =", round(bp_lm$p.value, 4), "\n")
cat("  - Teste Hausman (FE vs RE): p-valor =", round(hausman$p.value, 4), "\n")
cat("  - MODELO RECOMENDADO:", melhor_modelo, "\n\n")

cat("3. PRINCIPAIS RESULTADOS (Modelo", melhor_modelo, "):\n")

if(melhor_modelo == "Efeitos Fixos") {
  coef_principais <- summary(fe)$coefficients
  r2_fe <- summary(fe)$r.squared[1]
  cat("  - R2 :", round(r2_fe, 4), "\n")
  cat("  - R2 Ajustado:", round(summary(fe)$r.squared[2], 4), "\n")
} else {
  coef_principais <- summary(re)$coefficients
  cat("  - R2 :", round(summary(re)$r.squared[1], 4), "\n")
}

cat("  - Coeficientes:\n")
print(round(coef_principais, 4))

cat("\n4. ELASTICIDADES ESTIMADAS (Modelo Log-Log):\n")
cat("  - IDR:", round(coef_log[1,1], 4), " (p-valor:", round(coef_log[1,4], 4), ")")
cat("  - AME:", round(coef_log[2,1], 4), " (p-valor:", round(coef_log[2,4], 4), ")")

cat("\n5. DIAGNÓSTICOS:\n")
cat("  - Heterocedasticidade:", ifelse(bp_test_fe$p.value < 0.05, "Presente",

```

```

"Ausente"), "\n")
  cat("    - Autocorrelação:", ifelse(autocor_test$p.value < 0.05, "Presente", "Ausente"),
"\n")
  cat("    - Normalidade dos resí duos:", ifelse(jb_test$p.value < 0.05, "Não normal",
"Normal"), "\n")

cat("\n6. RECOMENDAÇÕES:\n")
if(bp_test_fe$p.value < 0.05) {
  cat("    - Usar erros padrão robustos (HAC) devido à heterocedasticidade\n")
}
if(autocor_test$p.value < 0.05) {
  cat("    - Considerar modelo com defasagens ou AR(1)\n")
}
if(any(vif_values > 5)) {
  cat("    - Possí vel multicolinearidade em:", names(vif_values)[vif_values > 5],
"\n")
}

cat("\n7. CONCLUSÃO:\n")
cat("    O modelo de", tolower(melhor_modelo), "é o mais adequado para analisar os
determinantes\n")
cat("    do PIB nos estados brasileiros. A educação (AME) e o desenvolvimento
regional (IDR)\n")
cat("    apresentam efeitos positivos e significantes sobre o PIB, enquanto
desigualdade (GINI)\n")
cat("    e pobreza apresentam efeitos negativos.\n")

cat("\n===== \n")
cat("          FIM DA ANÁLISE          \n")
cat("===== \n")

#####
# 16) EXPORTAÇÃO DOS RESULTADOS
#####

# Exportar tabelas para CSV
write.csv(tabela_testes, "testes_especificacao.csv", row.names = FALSE)

```

```

write.csv(efeitos_fixos_df, "efeitos_fixos_estados.csv", row.names = FALSE)
write.csv(estatisticas_regiao, "estatisticas_regiao.csv", row.names = FALSE)
write.csv(as.data.frame(coef_log), "elasticidades_loglog.csv", row.names = TRUE)

# Salvar gráficos
ggsave("evolucao_pib_regiao.png", p1, width = 12, height = 8, dpi = 300)
ggsave("boxplot_pib_ano_regiao.png", p2, width = 12, height = 8, dpi = 300)
ggsave("relacao_idr_pib.png", p3, width = 12, height = 8, dpi = 300)
ggsave("relacao_ame_pib.png", p4, width = 12, height = 8, dpi = 300)
ggsave("efeitos_fixos_estados.png", p5, width = 14, height = 10, dpi = 300)
ggsave("residuos_ano.png", p6, width = 10, height = 6, dpi = 300)
ggsave("residuos_ajustados.png", p7, width = 10, height = 6, dpi = 300)
ggsave("qqplot_residuos.png", p8, width = 8, height = 6, dpi = 300)
ggsave("matriz_correlacao.png", p_cor, width = 10, height = 8, dpi = 300)

cat("\nArquivos exportados com sucesso!\n")

# GRÁFICO: COMPARAÇÃO IDR vs AME - MODELOS DE EFEITOS FIXOS E
ALEATÓRIOS

# Baseado na tabela fornecida pelo usuário
#####

# 1) CARREGAR PACOTES
library(ggplot2)
library(dplyr)
library(tidyr)
library(gridExtra)
library(patchwork)

# 2) INSERIR OS DADOS DA TABELA FORNECIDA
dados_modelos <- data.frame(
  Variavel = c("IDR", "IDR", "ln(IDR)", "AME", "AME", "ln(AME)"),
  Modelo = c("Efeitos Fixos (Nível)", "Efeitos Aleatórios (Nível)", "Efeitos Fixos
(Log-Log)",
              "Efeitos Fixos (Nível)", "Efeitos Aleatórios (Nível)", "Efeitos Fixos
(Log-Log)"),
  Coeficiente = c(-6.569, -2.992, 2.135, 5.066, 1.452, -0.802),

```

```

Erro_Padrao = c(2.337, 2.556, 0.900, 1.982, 2.155, 1.582),
Significancia = c("***", "n.s.", "**", "**", "n.s.", "n.s."),
P_Valor = c(0.007, 0.245, 0.021, 0.013, 0.503, 0.613), # aproximado
stringsAsFactors = FALSE
)

# Calcular intervalos de confiança (95%)
dados_modelos$IC_inf <- dados_modelos$Coeficiente - 1.96 *
dados_modelos$Erro_Padrao
dados_modelos$IC_sup <- dados_modelos$Coeficiente + 1.96 *
dados_modelos$Erro_Padrao

# Ajustar para melhor visualização
dados_modelos$Variavel_Modelo <- paste0(dados_modelos$Variavel, "\n",
dados_modelos$Modelo)
dados_modelos$Cor <- ifelse(dados_modelos$Variavel == "IDR" |
dados_modelos$Variavel == "ln(IDR)",
"#1f77b4", "#2ca02c")
dados_modelos$Alfa <- ifelse(dados_modelos$Significancia != "n.s.", 0.9, 0.4)

# 3) GRÁFICO 1: COMPARAÇÃO DIRETA IDR vs AME - TODOS OS MODELOS
p1 <- ggplot(dados_modelos, aes(x = reorder(Variavel_Modelo, abs(Coeficiente)),
y = Coeficiente,
fill = Variavel)) +

# Barras
geom_bar(stat = "identity", width = 0.6,
aes(alpha = Alfa), show.legend = FALSE) +

# Barras de erro
geom_errorbar(aes(ymin = IC_inf, ymax = IC_sup),
width = 0.2, size = 0.8, color = "gray30") +

# Linha de zero
geom_hline(yintercept = 0, linetype = "dashed", color = "black", size = 0.8) +

# Rótulos com valores e significância

```

```

geom_text(aes(label = paste0(round(Coeficiente, 3), Significancia),
                           y = ifelse(Coeficiente > 0, IC_sup + 2, IC_inf - 2)),
          size = 4, fontface = "bold") +

# Cores
scale_fill_manual(values = c("IDR" = "#1f77b4", "AME" = "#2ca02c",
                             "ln(IDR)" = "#1f77b4", "ln(AME)" =
"#2ca02c")) +

scale_alpha_identity() +

# Títulos
labs(
  title = "IDR vs AME: Comparação de Coeficientes",
  subtitle = "Modelos de Efeitos Fixos e Aleatórios - PIB estadual (2003-2018)",
  x = NULL,
  y = "Coeficiente",
  caption = "**** p<0.01 | ** p<0.05 | n.s.: não significativo | Barras de erro: IC
95%"
) +

# Tema
theme_minimal() +
theme(
  legend.position = "none",
  plot.title = element_text(hjust = 0.5, face = "bold", size = 16, color = "#1f77b4"),
  plot.subtitle = element_text(hjust = 0.5, size = 12, color = "gray30"),
  plot.caption = element_text(hjust = 0.5, size = 10, color = "gray50"),
  axis.text = element_text(size = 11, face = "bold"),
  axis.text.x = element_text(size = 9, angle = 0, hjust = 0.5),
  axis.title = element_text(size = 12, face = "bold"),
  panel.grid.minor = element_blank(),
  panel.grid.major.x = element_blank(),
  plot.margin = margin(t = 20, r = 20, b = 20, l = 20)
) +

```

```

coord_flip()

print(p1)

# 4) GRÁFICO 2: APENAS MODELO DE EFEITOS FIXOS (O MAIS ROBUSTO)
dados_fe <- data.frame(
  Variavel = c("IDR", "AME", "ln(IDR)", "ln(AME)"),
  Coeficiente = c(-6.569, 5.066, 2.135, -0.802),
  Erro_Padiao = c(2.337, 1.982, 0.900, 1.582),
  Significancia = c("***", "**", "**", "n.s."),
  stringsAsFactors = FALSE
)

dados_fe$IC_inf <- dados_fe$Coeficiente - 1.96 * dados_fe$Erro_Padiao
dados_fe$IC_sup <- dados_fe$Coeficiente + 1.96 * dados_fe$Erro_Padiao
dados_fe$Cor <- c("#1f77b4", "#2ca02c", "#1f77b4", "#2ca02c")
dados_fe$Alfa <- c(0.9, 0.9, 0.9, 0.4)

dados_fe$Variavel_label <- c("IDR (nível)", "AME (nível)", "ln(IDR) - elasticidade",
"ln(AME) - elasticidade")

p2 <- ggplot(dados_fe, aes(x = reorder(Variavel_label, -abs(Coeficiente)),
                           y = Coeficiente,
                           fill = Variavel_label)) +
  geom_bar(stat = "identity", width = 0.6, aes(alpha = Alfa), show.legend = FALSE) +
  geom_errorbar(aes(ymin = IC_inf, ymax = IC_sup), width = 0.2, size = 0.9, color =
"gray30") +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black", size = 0.8) +
  geom_text(aes(label = paste0(round(Coeficiente, 3), Significancia),
                y = ifelse(Coeficiente > 0, IC_sup + 1.5, IC_inf - 1.5)),
            size = 4.5, fontface = "bold") +
  scale_fill_manual(values = c("IDR (nível)" = "#1f77b4",
                              "AME (nível)" = "#2ca02c",
                              "ln(IDR) - elasticidade" = "#1f77b4",
                              "ln(AME) - elasticidade" = "#2ca02c")) +
  scale_alpha_identity() +
  labs(

```

```

title = " MODELO DE EFEITOS FIXOS: IDR vs AME",
subtitle = "IDR é significativo em ambas as especificações; AME é frágil e não
significativo em log",
x = NULL,
y = "Coeficiente",
caption = "**** p<0.01 | ** p<0.05 | n.s.: não significativo | IC 95%"
) +
theme_minimal() +
theme(
  legend.position = "none",
  plot.title = element_text(hjust = 0.5, face = "bold", size = 18, color = "#1f77b4"),
  plot.subtitle = element_text(hjust = 0.5, size = 12, color = "gray30", face =
"bold"),
  plot.caption = element_text(hjust = 0.5, size = 10, color = "gray50"),
  axis.text = element_text(size = 12, face = "bold"),
  axis.text.x = element_text(size = 11, angle = 0),
  axis.title = element_text(size = 13, face = "bold"),
  panel.grid.minor = element_blank(),
  panel.grid.major.x = element_blank(),
  plot.margin = margin(t = 20, r = 20, b = 20, l = 20)
) +
coord_flip()

print(p2)

```

### # 5) GRÁFICO 3: ELASTICIDADES COMPARADAS (MODELO LOG-LOG)

```

dados_elasticidade <- data.frame(
  Variavel = c("IDR (elasticidade)", "AME (elasticidade)"),
  Coeficiente = c(2.135, -0.802),
  Erro_Padrao = c(0.900, 1.582),
  Significancia = c("***", "n.s."),
  stringsAsFactors = FALSE
)

```

```

dados_elasticidade$IC_inf <- dados_elasticidade$Coeficiente - 1.96 *
dados_elasticidade$Erro_Padrao

```

```

dados_elasticidade$IC_sup <- dados_elasticidade$Coeficiente + 1.96 *
dados_elasticidade$Erro_Padiao

p3 <- ggplot(dados_elasticidade, aes(x = Variavel, y = Coeficiente, fill = Variavel)) +
  geom_bar(stat = "identity", width = 0.5,
           alpha = c(0.9, 0.4), show.legend = FALSE) +
  geom_errorbar(aes(ymin = IC_inf, ymax = IC_sup), width = 0.2, size = 1, color =
"gray30") +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black", size = 0.8) +
  geom_text(aes(label = paste0(round(Coeficiente, 3), Significancia),
                 y = ifelse(Coeficiente > 0, IC_sup + 0.3, IC_inf - 0.3)),
           size = 5, fontface = "bold") +
  scale_fill_manual(values = c("IDR (elasticidade)" = "#1f77b4",
                              "AME (elasticidade)" = "#2ca02c")) +
  labs(
    title = "ELASTICIDADES: Modelo Log-Log com Efeitos Fixos",
    subtitle = paste0(
      "IDR: +", round(dados_elasticidade$Coeficiente[1], 3),
      " (IC 95%: ", round(dados_elasticidade$IC_inf[1], 3), " a ",
round(dados_elasticidade$IC_sup[1], 3), ") | p = 0.021 **\n",
      "AME: ", round(dados_elasticidade$Coeficiente[2], 3),
      " (IC 95%: ", round(dados_elasticidade$IC_inf[2], 3), " a ",
round(dados_elasticidade$IC_sup[2], 3), ") | p = 0.613 n.s."
    ),
    x = NULL,
    y = "Elasticidade (variação % no PIB / 1% na variável)",
    caption = "*** p<0.05 | n.s.: não significativo | IDR: elasticidade POSITIVA e
SIGNIFICATIVA | AME: não significativo"
  ) +
  theme_minimal() +
  theme(
    legend.position = "none",
    plot.title = element_text(hjust = 0.5, face = "bold", size = 16, color = "#1f77b4"),
    plot.subtitle = element_text(hjust = 0.5, size = 11, color = "gray30", family =
"mono"),
    plot.caption = element_text(hjust = 0.5, size = 11, face = "bold", color =
"#1f77b4"),

```

```

axis.text = element_text(size = 14, face = "bold"),
axis.text.x = element_text(size = 12),
axis.title = element_text(size = 12, face = "bold"),
panel.grid.minor = element_blank(),
panel.grid.major.x = element_blank()
)

print(p3)

# 6) GRÁFICO 4: QUADRO DE DECISÃO - SIGNIFICÂNCIA E MAGNITUDE
dados_decisao <- data.frame(
  Modelo = c("Efeitos Fixos (Nível)", "Efeitos Aleatórios (Nível)", "Efeitos Fixos
(Log-Log)"),
  IDR_Sig = c("****", "n.s.", "****"),
  AME_Sig = c("***", "n.s.", "n.s."),
  IDR_Coef = c(-6.569, -2.992, 2.135),
  AME_Coef = c(5.066, 1.452, -0.802),
  stringsAsFactors = FALSE
)

dados_decisao_long <- dados_decisao %>%
  pivot_longer(cols = c(IDR_Sig, AME_Sig), names_to = "Variavel_Sig", values_to =
"Significancia") %>%
  mutate(
    Variavel = ifelse(grepl("IDR", Variavel_Sig), "IDR", "AME"),
    Coeficiente = ifelse(Variavel == "IDR", IDR_Coef, AME_Coef),
    Sig_Num = case_when(
      Significancia == "****" ~ 3,
      Significancia == "***" ~ 2,
      Significancia == "**" ~ 1,
      TRUE ~ 0
    )
  )

p4 <- ggplot(dados_decisao_long, aes(x = Modelo, y = Variavel, fill = Sig_Num)) +
  geom_tile(color = "white", size = 1) +

```

```

geom_text(aes(label = paste0(round(Coeficiente, 2), "\n", Significancia)),
          size = 4.5, fontface = "bold", color = "white") +
scale_fill_gradient(low = "#cccccc", high = "#1f77b4",
                    name = "Significância",
                    breaks = c(0, 1, 2, 3),
                    labels = c("n.s.", "*", "**", "***)) +
labs(
  title = " QUADRO DE DECISÃO: IDR vs AME",
  subtitle = "IDR é significativo em 2 de 3 modelos; AME é significativo em
apenas 1 modelo",
  x = NULL, y = NULL,
  caption = "IDR é claramente superior: maior significância, maior consistência
entre modelos"
) +
theme_minimal() +
theme(
  legend.position = "bottom",
  plot.title = element_text(hjust = 0.5, face = "bold", size = 16, color = "#1f77b4"),
  plot.subtitle = element_text(hjust = 0.5, size = 11, color = "gray30"),
  plot.caption = element_text(hjust = 0.5, size = 11, face = "bold", color =
"#1f77b4"),
  axis.text = element_text(size = 12, face = "bold"),
  axis.text.x = element_text(angle = 0, hjust = 0.5),
  panel.grid = element_blank()
)

```

```
print(p4)
```

```
# 7) GRÁFICO PRINCIPAL: PAINEL CONCLUSIVO
```

```

p_final <- (p3 | p4) / p2 +
plot_annotation(
  title = " VEREDICTO FINAL: IDR É SUPERIOR A AME PARA EXPLICAR O
PIB",
  subtitle = paste0(
    " IDR: significativo em 2/3 modelos | elasticidade +2.135 (p=0.021) |
consistente\n",
    " AME: significativo em 1/3 modelos | elasticidade NÃO significativa

```

```

(p=0.613) | instável"
    ),
    theme = theme(
      plot.title = element_text(hjust = 0.5, face = "bold", size = 22, color =
"#1f77b4"),
      plot.subtitle = element_text(hjust = 0.5, size = 14, color = "#2c3e50", face =
"bold"),
      plot.margin = margin(t = 25, r = 25, b = 25, l = 25),
      plot.background = element_rect(fill = "#f9f9f9", color = NA)
    )
  )

print(p_final)

# 8) GRÁFICO SIMPLES E DIRETO (O MELHOR PARA RESPONDER)
p_resposta <- ggplot(dados_elasticidade, aes(x = Variavel, y = Coeficiente, fill =
Variavel)) +
  geom_bar(stat = "identity", width = 0.5,
          alpha = c(0.9, 0.3), show.legend = FALSE) +
  geom_errorbar(aes(ymin = IC_inf, ymax = IC_sup), width = 0.2, size = 1.2, color =
"gray30") +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black", size = 1) +
  geom_text(aes(label = paste0(round(Coeficiente, 2), "\n",
                                "[", round(IC_inf, 1), ", ", round(IC_sup, 1),
                                "]\n",
                                ifelse(Significancia == "***", "p = 0.021 **", "p
= 0.613 n.s."))),
            vjust = c(-0.8, 1.8), size = 4.5, fontface = "bold", lineheight = 0.9) +
  scale_fill_manual(values = c("IDR (elasticidade)" = "#1f77b4",
                                "AME (elasticidade)" = "#2ca02c")) +
  scale_y_continuous(limits = c(-3, 4), breaks = seq(-3, 4, by = 1)) +
  labs(
    title = "IDR É MELHOR QUE AME PARA EXPLICAR O PIB?",
    subtitle = "Modelo de Efeitos Fixos com especificação log-log (elasticidades)",
    x = NULL,
    y = "Elasticidade (variação % no PIB)",
    caption = "SIM: IDR tem efeito POSITIVO e SIGNIFICATIVO | AME não é

```



```

cat("
\n")
VEREDICTO FINAL: IDR vs AME
cat("
=====
=====
=====
\n\n")

cat(" ANÁLISE DOS MODELOS:\n")
cat("
=====
=====
\n")

cat(" EFEITOS FIXOS (NÍVEL):\n")
cat(sprintf(" • IDR: Coeficiente = %6.3f | p < 0.01 *** | SIGNIFICATIVO\n", -
6.569))
cat(sprintf(" • AME: Coeficiente = %6.3f | p < 0.05 ** |
SIGNIFICATIVO\n\n", 5.066))

cat(" EFEITOS ALEATÓRIOS (NÍVEL):\n")
cat(sprintf(" • IDR: Coeficiente = %6.3f | p > 0.10 n.s. | NÃO
SIGNIFICATIVO\n", -2.992))
cat(sprintf(" • AME: Coeficiente = %6.3f | p > 0.10 n.s. | NÃO
SIGNIFICATIVO\n\n", 1.452))

cat(" EFEITOS FIXOS (LOG-LOG) - MODELO MAIS ROBUSTO:\n")
cat(sprintf(" • IDR: Elasticidade = %+6.3f | p = 0.021 ** | SIGNIFICATIVO\n",
2.135))
cat(sprintf(" • AME: Elasticidade = %+6.3f | p = 0.613 n.s. | NÃO
SIGNIFICATIVO\n\n", -0.802))

cat(" 📌 INTERPRETAÇÃO DAS ELASTICIDADES:\n")
cat("
=====
=====
\n")

cat(sprintf(" • IDR: Um aumento de 1%% no IDR está associado a um
AUMENTO de %.1f%% no PIB\n", 2.135))
cat(sprintf(" • AME: Um aumento de 1%% no AME NÃO está associado a
variação significativa no PIB\n\n"))

cat(" QUADRO COMPARATIVO:\n")

```

cat("AME\n")	Critério	IDR	
cat("Significância (Efeitos Fixos) (p<0.05)\n")	Significância (Efeitos Fixos)	*** (p<0.01)	**
cat("Significância (Log-Log) (p>0.05)\n")	Significância (Log-Log)	** (p<0.05)	n.s.
cat("Consistência entre modelos modelos\n")	Consistência entre modelos	✓ 2/3 modelos significativo	✗ 1/3
cat("Elasticidade positiva 0.802 (n.s.)\n")	Elasticidade positiva	✓ +2.135	✗ -
cat("Precisão (erro padrão) (impreciso)\n")	Precisão (erro padrão)	0.900 (preciso)	1.582
cat("IC 95% exclui zero? 2.30]\n")	IC 95% exclui zero?	✓ [0.37, 3.90]	✗ [-3.90, 2.30]

cat("CONCLUSÃO\n")

cat("SIM: IDR É MELHOR QUE AME PARA EXPLICAR O PIB \n\n")

cat("ARGUMENTOS DECISIVOS:\n")

cat("1. No modelo mais robusto (Efeitos Fixos com log-log), o IDR é\n")

cat("SIGNIFICATIVO (p=0,021) e apresenta elasticidade positiva de +2,135.\n\n")

cat("2. No mesmo modelo, o AME NÃO É SIGNIFICATIVO (p=0,613) e seu intervalo\n")

cat("de confiança AMPLAMENTE INCLUI O ZERO [-3,90; 2,30].\n\n")

cat(" 3. O IDR mantém significância em 2 das 3 especificações; o AME é\n")  
 cat(" significativo em apenas 1 e perde significância no modelo log-log.\n\n")  
 cat(" 4. A elasticidade do IDR é POSITIVA e ECONOMICAMENTE RELEVANTE:\n")  
 cat(" 1% a mais no IDR → +2,135% no PIB.\n\n")  
 cat(" 5. A elasticidade do AME é NEGATIVA (contraintuitiva) e NÃO SIGNIFICATIVA,\n")  
 cat(" sugerindo problemas de especificação ou multicolinearidade.\n\n")  
 cat(" IMPLICAÇÃO PARA POLÍTICAS PÚBLICAS:\n")  
 cat(" Políticas de Desenvolvimento Regional (IDR) têm efeito COMPROVADO e\n")  
 cat(" SIGNIFICATIVO sobre o crescimento econômico. Políticas educacionais,\n")  
 cat(" embora importantes, NÃO APRESENTAM EVIDÊNCIA ROBUSTA de impacto no PIB\n")  
 cat(" quando controladas por outras variáveis no modelo de efeitos fixos.\n\n")  
 cat("=====  
 =====  
 =====\n")  
 cat(" GRÁFICOS EXPORTADOS COM SUCESSO!  
 \n")  
 cat("=====  
 =====  
 =====\n")  
 cat("=====  
 =====  
 =====\n")