



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE ECONOMIA
BACHARELADO EM CIÊNCIAS ECONÔMICAS

LUIZ VICTOR NUNES DA SILVA

**VIOLÊNCIA LETAL E HETEROGENEIDADE MUNICIPAL: Uma análise dos fatores
associados aos homicídios nos municípios do Nordeste brasileiro entre 2013 e 2023**

Recife, Pernambuco

2026

LUIZ VICTOR NUNES DA SILVA

**VIOLÊNCIA LETAL E HETEROGENEIDADE MUNICIPAL: Uma análise dos fatores
associados aos homicídios nos municípios do Nordeste brasileiro entre 2013 e 2023**

Trabalho de Conclusão de Curso apresentado ao curso de Ciências Econômicas da Universidade Federal Rural de Pernambuco, como requisito para obtenção do título de Bacharel em Ciências Econômicas, sob a orientação do Prof. Dr. Diego Firmino Costa da Silva.

Recife

2026

Dados Internacionais de Catalogação na Publicação
Sistema Integrado de Bibliotecas da UFRPE
Bibliotecário(a): Suely Manzi – CRB-4 809

S586v Silva, Luiz Victor Nunes da.
Violência letal e heterogeneidade municipal: uma análise dos fatores associados aos homicídios nos municípios do nordeste brasileiro entre 2013 e 2023 / Luiz Victor Nunes da Silva. – Recife, 2026.
49 f.; il.

Orientador(a): Diego Firmino Costa da Silva.

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal Rural de Pernambuco, Bacharelado em Ciências Econômicas, Recife, BR-PE, 2026.

Inclui referências.

1. Homicídio - Brasil, Nordeste. 2. Crime econômico. 3. Crimes contra a economia popular - Brasil, Nordeste . 4. Modelos para dados em painel 5. Heterogeneidade . I. Silva, Diego Firmino Costa da, orient. II. Título

CDD 330

LUIZ VICTOR NUNES DA SILVA

**VIOLÊNCIA LETAL E HETEROGENEIDADE MUNICIPAL: Uma análise dos fatores
associados aos homicídios nos municípios do Nordeste brasileiro entre 2013 e 2023**

Trabalho de Conclusão de Curso apresentado
ao curso de Ciências Econômicas da
Universidade Federal Rural de Pernambuco,
como requisito para obtenção do título de
Bacharel em Ciências Econômicas, sob a
orientação do Prof. Dr. Diego Firmino Costa
da Silva.

Aprovado em: 08/07/2026

BANCA EXAMINADORA

Prof. Dr. Diego Firmino Costa da Silva (Orientador)
Universidade Federal Rural de Pernambuco

Prof. Dr. Ewerton Felipe de Melo Araújo (Examinador Interno)
Universidade Federal Rural de Pernambuco

Prof.^a Dr.^a Keynis Cândido Souto (Examinadora Interna)
Universidade Federal Rural de Pernambuco

Às vidas que, antes de dados estatísticos, foram
nomes, identidades e trajetórias interrompidas pela
violência.

AGRADECIMENTOS

À minha família, pelo apoio constante e incondicional ao longo desta trajetória. Em especial, aos meus pais, Gustavo e Viviane, cujo incentivo e suporte foram fundamentais para que esta etapa pudesse ser concluída.

Ao Professor Diego Firmino Costa da Silva, pela orientação, disponibilidade, paciência e pelas contribuições que enriqueceram este trabalho e minha formação acadêmica.

À Josy, companheira de graduação e uma das maiores dádivas que a universidade me proporcionou, pela amizade que levarei para o resto da vida, pela parceria construída ao longo desses anos, pelos aprendizados compartilhados, pelas inúmeras experiências vividas dentro e fora da universidade e pelo apoio mútuo que tornou essa caminhada mais leve e significativa.

Ao Pedro, pela presença constante, apoio e compreensão ao longo desta trajetória. Ao Igor, Lucas, João, Luiza e Gey, pelo acolhimento, pelos momentos vividos em conjunto e pelas memórias que permanecem entre as boas lembranças.

Ao Pablo, Victor Hugo (Nikito) e Jamir, pelas conversas e risadas, pelo suporte, pela companhia constante e pelos inúmeros momentos compartilhados ao longo desses anos, que tornaram os dias mais leves e felizes.

À Anna, Erick, Ramon, Stefanny e Thalyta, por fazerem parte da minha história desde diferentes etapas da vida e por permanecerem presentes, mesmo quando o tempo e a distância tornaram os encontros menos frequentes.

Por fim, reconheço o significado pessoal desta conquista. A conclusão deste trabalho representa o encerramento de uma longa etapa de formação, construída entre desafios, incertezas e aprendizados que ultrapassam os limites da vida acadêmica. Acima de tudo, representa a certeza de que nenhuma conquista desta caminhada foi alcançada de forma solitária.

RESUMO

Este estudo analisa os fatores associados aos homicídios nos municípios do Nordeste brasileiro entre 2013 e 2023. O objetivo é verificar a relação entre a violência letal e variáveis econômicas, educacionais, demográficas e de proteção social, bem como investigar se essas relações diferem de acordo com o porte populacional dos municípios. Para isso, foi utilizada uma base de dados em painel composta por 1.790 municípios nordestinos, analisada por meio de modelos econométricos com efeitos fixos de unidade federativa (UF) e ano. Os resultados indicam associações positivas entre as taxas de homicídios e indicadores relacionados ao dinamismo econômico, às fragilidades no desempenho escolar do ensino médio e à proporção da população com idade entre 15 e 24 anos. Em contrapartida, a cobertura cadastral da população de baixa renda no Cadastro Único apresentou associação negativa com a violência letal. Além disso, as estimações evidenciam que essas associações variam conforme o porte populacional dos municípios, indicando que os fatores associados aos homicídios não se manifestam de forma homogênea no território nordestino. Conclui-se que a violência letal no Nordeste brasileiro constitui um fenômeno marcado por significativa heterogeneidade municipal, de modo que análises agregadas podem ocultar diferenças relevantes na dinâmica dos homicídios e em seus fatores associados.

Palavras-chave: homicídios; economia do crime; Nordeste brasileiro; dados em painel; heterogeneidade municipal.

ABSTRACT

This study analyzes the factors associated with homicides in municipalities in Northeast Brazil between 2013 and 2023. The objective is to examine the relationship between lethal violence and economic, educational, demographic, and social protection variables, as well as to investigate whether these relationships differ according to municipal population size. For this purpose, a panel dataset comprising 1,790 municipalities was analyzed using econometric models with state and year fixed effects. The results indicate positive associations between homicide rates and indicators related to economic dynamism, educational vulnerabilities in upper secondary education, and the proportion of the population aged 15 to 24. In contrast, the coverage rate of low-income individuals registered in the Unified Registry for Social Programs (Cadastro Único) was negatively associated with lethal violence. Furthermore, the estimations show that these associations vary according to municipal population size, indicating that the factors associated with homicides do not manifest uniformly across Northeast Brazil. It is concluded that lethal violence in Northeast Brazil is characterized by substantial municipal heterogeneity, suggesting that aggregate analyses may conceal important differences in the dynamics of homicides and their associated factors.

Keywords: homicides; economics of crime; Northeast Brazil; panel data; municipal heterogeneity.

LISTA DE ILUSTRAÇÕES

Quadro I	Variáveis utilizadas, unidades de medida e fontes	22
Quadro II	Relação esperada entre as dimensões analisadas e a taxa de homicídios	25
Gráfico I	Evolução da taxa média de homicídios por porte populacional nos municípios do Nordeste brasileiro (2013–2023)	36

LISTA DE TABELAS

Tabela I	Composição dos modelos estimados	30
Tabela II	Estatística descritiva das variáveis analisadas, 2013 a 2023	32
Tabela III	Matriz de correlação entre os regressores	35
Tabela IV	Estimações com variáveis <i>dummy</i> de porte populacional	37
Tabela V	Estimações por estrato de faixas populacionais	41

LISTA DE SIGLAS

AEDE	Análise Exploratória de Dados Espaciais
CadÚnico	Cadastro Único para Programas Sociais do Governo Federal
DataSUS	Departamento de Informática do Sistema Único de Saúde
IBGE	Instituto Brasileiro de Geografia e Estatística
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
Ipeadata	Base de Dados do Instituto de Pesquisa Econômica Aplicada
MDS	Ministério do Desenvolvimento Social e Assistência Social, Família e Combate à Fome
MS	Ministério da Saúde
PIB	Produto Interno Bruto
RAIS	Relação Anual de Informações Sociais
SAGICAD	Secretaria de Avaliação, Gestão da Informação e Cadastro Único
SIM	Sistema de Informações sobre Mortalidade
UF	Unidade de Federação
UNODC	Escritório das Nações Unidas sobre Drogas e Crime

SUMÁRIO

1. INTRODUÇÃO	11
2. REVISÃO DE LITERATURA	13
2.1. Conceitos e abordagens sociológicas da criminalidade	13
2.2. O crime sob a perspectiva econômica	14
2.3. Evidências empíricas sobre criminalidade e violência letal	17
3. ASPECTOS METODOLÓGICOS	21
3.1. Abordagem geral da análise	21
3.2. Descrição das variáveis	22
3.3. Especificação econométrica	26
3.4. Procedimentos de diagnóstico e inferência	28
3.5. Estratégia de estimação e composição dos modelos	30
4. RESULTADOS E ANÁLISE	32
4.1. Análise descritiva das variáveis	32
4.2. Matriz de correlação e diagnóstico de multicolinearidade	34
4.3. Heterogeneidade da violência letal por porte populacional	36
4.3.1. <i>Modelo geral e modelo com dummies de porte populacional</i>	37
4.3.2. <i>Modelo geral e modelos por faixas populacionais</i>	41
5. CONSIDERAÇÕES FINAIS	44
6. REFERÊNCIAS	46

1. INTRODUÇÃO

A violência letal permanece como um dos mais relevantes problemas sociais da atualidade, produzindo impactos que vão além da perda de vidas humanas. Além de seus efeitos diretos sobre as vítimas e suas famílias, a criminalidade gera consequências que se refletem em diferentes dimensões da vida social e econômica. Nesse sentido, o crime deve ser compreendido não apenas como uma questão de segurança pública, mas também como um fenômeno capaz de afetar negativamente o desenvolvimento econômico e o bem-estar social (LOUREIRO; CARVALHO, 2006). A dimensão desse problema torna-se particularmente evidente quando observados os indicadores de homicídios no Brasil. Embora o país tenha apresentado avanços recentes na redução da violência letal, o fenômeno ainda apresenta elevada magnitude e forte desigualdade territorial. Segundo o Atlas da Violência 2026, o Brasil registrou 42.590 homicídios em 2024, correspondendo a uma taxa de 20,1 mortes por 100 mil habitantes, o menor patamar da série histórica desde 1998. Entretanto, esses avanços não ocorreram de forma homogênea entre as diferentes regiões do país. Em 2024, entre os 20 municípios brasileiros com mais de 100 mil habitantes que apresentaram as maiores taxas de homicídio, 17 localizavam-se na região Nordeste. Essa informação evidencia a relevância da região no cenário nacional da violência letal e reforça a necessidade de compreender os determinantes da violência letal nos municípios nordestinos.

Entre as diferentes abordagens desenvolvidas para compreender a dinâmica criminal, destaca-se a Teoria Econômica do Crime, desenvolvida por Becker (1968). Essa perspectiva parte do pressuposto de que os indivíduos atuam de forma racional, comparando os benefícios esperados da atividade ilícita aos custos potenciais associados à sua prática. Com base nessa contribuição teórica, diversos estudos passaram a investigar empiricamente a relação entre criminalidade e diferentes características socioeconômicas da população, ampliando a compreensão dos determinantes da violência.

No contexto brasileiro, a literatura empírica tem identificado associações entre os homicídios e diferentes fatores econômicos, educacionais e demográficos (OLIVEIRA, 2005; CERQUEIRA; MOURA, 2015). Estudos realizados em distintas escalas territoriais apontam que variáveis relacionadas ao mercado de trabalho, à desigualdade de renda, à escolaridade e à estrutura etária da população podem influenciar a dinâmica da violência. No Nordeste, essas relações também se mostram relevantes, embora as evidências indiquem que a violência letal não se restringe aos grandes centros urbanos, manifestando-se de forma heterogênea no

espaço regional (NÓBREGA JÚNIOR, 2017; PLASSA; PASCHOALINO; SANTOS, 2019; ROCHA; LIMA, 2023). Embora a literatura tenha avançado na identificação dos fatores associados à violência letal e à sua distribuição espacial, ainda são limitados os estudos que analisam os homicídios nos municípios nordestinos considerando diferenças associadas ao porte populacional.

Diante desse contexto, o presente estudo tem como objetivo analisar os fatores associados aos homicídios nos municípios da região Nordeste do Brasil no período de 2013 a 2023. Para isso, utiliza-se uma base de dados em painel composta por informações municipais provenientes de diferentes fontes oficiais, estimada por meio de modelos econométricos com efeitos fixos de UF e ano. Considerando a heterogeneidade dos municípios nordestinos, a estratégia empírica incorpora a estratificação por porte populacional, com o objetivo de verificar se as associações entre os homicídios e os fatores socioeconômicos variam entre grupos de municípios.

Os resultados sugerem que fatores econômicos, educacionais, demográficos e de proteção social estão associados à dinâmica dos homicídios nos municípios nordestinos. Adicionalmente, a estratificação dos municípios por porte populacional revela que a intensidade dessas associações varia entre diferentes grupos de localidades. Nesse sentido, o estudo contribui para a literatura ao evidenciar que os fatores associados aos homicídios se manifestam de forma distinta entre grupos de municípios, podendo auxiliar na formulação de políticas públicas mais adequadas às especificidades dos diferentes contextos municipais.

Além desta introdução, o trabalho está organizado em quatro seções. A segunda seção apresenta o referencial teórico e a revisão da literatura empírica sobre criminalidade e homicídios. A terceira seção descreve os procedimentos metodológicos adotados, as variáveis utilizadas e as fontes de dados empregadas na pesquisa. Na quarta seção, são apresentados e discutidos os resultados das estimações econométricas. Por fim, a quinta seção reúne as considerações finais do estudo, destacando os principais achados, limitações e possibilidades para pesquisas futuras.

2. REVISÃO DE LITERATURA

Esta seção apresenta os fundamentos conceituais, teóricos e empíricos que orientam a análise da violência letal. A discussão parte da delimitação jurídica do crime, avança para abordagens sociológicas e econômicas da criminalidade e, por fim, reúne evidências empíricas sobre os fatores associados aos homicídios, com atenção ao contexto brasileiro e nordestino.

2.1. Conceitos e abordagens sociológicas da criminalidade

A criminalidade constitui um fenômeno complexo que pode ser analisado sob diferentes perspectivas. Sob a ótica jurídica, sua compreensão está associada às condutas definidas como infrações penais por determinado ordenamento legal. No entanto, uma vez que a definição de crime depende das normas estabelecidas por cada sociedade, os comportamentos considerados criminosos podem variar entre países e contextos históricos distintos (UNODC, 2015). No contexto brasileiro, a delimitação legal das infrações penais é estabelecida pelo Decreto-Lei nº 3.914/1941, que adota um critério de separação baseado na gravidade do ato: crimes configuram infrações mais graves, sujeitas a reclusão ou detenção, enquanto contravenções sujeitam-se a sanções mais brandas, como prisão simples ou multa. Contudo, embora a classificação jurídica seja fundamental para delimitar o fenômeno, ela não é suficiente para explicar sua distribuição e intensidade entre diferentes contextos sociais e territoriais. Dessa forma, para compreender a criminalidade, é necessário ir além da norma jurídica e analisar os incentivos socioeconômicos, as restrições e as condições institucionais associadas à ocorrência e à letalidade dos delitos.

Se a definição jurídica se concentra na tipificação legal, o pensamento científico buscou compreender as raízes do comportamento criminal. A partir desse deslocamento, a criminalidade passa a ser analisada não apenas como infração à lei, mas como um fenômeno social complexo, associado a condições sociais, econômicas, institucionais e territoriais. Entre as abordagens sociológicas que contribuíram para essa compreensão, destaca-se a Escola de Chicago, especialmente a partir da Teoria da Desorganização Social, formulada por Shaw e McKay (1942). Ao transferir a ênfase analítica do indivíduo para a dinâmica do ambiente urbano, o modelo passou a interpretar a criminalidade não como uma patologia individual, mas como um sintoma do enfraquecimento das instituições comunitárias, principalmente em áreas marcadas por rápida urbanização, instabilidade residencial e reduzida capacidade de controle social. Essa abordagem, ainda que formulada na primeira metade do século XX,

permanece relevante por destacar como características do ambiente social podem influenciar a distribuição da criminalidade entre diferentes territórios.

Outra contribuição importante é a Teoria do Controle Social, desenvolvida por Travis Hirschi. Em *Causes of Delinquency* (1969), o autor desloca a pergunta central da criminologia: em vez de questionar apenas por que os indivíduos cometem crimes, busca compreender o que os impede de delinquir. Para Hirschi, os vínculos sociais funcionam como mecanismos de contenção da conduta desviante, sendo compostos por elementos como apego, compromisso, envolvimento em atividades convencionais e crenças compartilhadas. Quando esses vínculos se enfraquecem, aumentam as chances de envolvimento com práticas delituosas. Dessa forma, a teoria destaca a importância dos mecanismos de integração social e do controle comunitário para a compreensão da criminalidade.

A Teoria da Anomia, proposta por Robert Merton (1938), oferece uma perspectiva complementar para compreender a criminalidade a partir das desigualdades presentes na estrutura social. Para Merton, a criminalidade pode surgir da tensão estrutural entre metas culturais amplamente valorizadas, como ascensão econômica e status social, e a distribuição desigual dos meios legítimos para alcançá-las. Nesse sentido, o acesso restrito ou de baixa qualidade à educação e ao mercado de trabalho pode ampliar as pressões sociais associadas ao envolvimento em práticas ilícitas, principalmente em contextos de oportunidades limitadas. Dessa forma, a teoria contribui para compreender como desigualdades no acesso a oportunidades legítimas podem influenciar decisões individuais, dialogando com interpretações posteriores baseadas em incentivos e custos de oportunidade¹.

2.2. O crime sob a perspectiva econômica

Embora essas abordagens sociológicas tenham contribuído para compreender a criminalidade a partir das dinâmicas sociais, territoriais e dos mecanismos de integração comunitária, a economia do crime passou a interpretar o comportamento delituoso como resultado de escolhas realizadas diante de incentivos, restrições e oportunidades. Nesse contexto, os estudos empíricos de Belton Fleisher (1963; 1966) figuram entre as primeiras contribuições a aproximar a criminalidade de fatores econômicos, ao relacionar a delinquência juvenil ao desemprego e à desigualdade de renda. Ao destacar que a alocação do

¹ O custo de oportunidade corresponde àquilo de que se abre mão ao realizar determinada escolha (MANKIW, 2021). No contexto da economia do crime, refere-se aos benefícios associados à atividade lícita que são renunciados diante da opção pela atividade ilícita.

tempo entre atividades lícitas e ilícitas depende das oportunidades disponíveis, Fleisher abriu caminho para a interpretação do crime como objeto de análise econômica.

A consolidação da economia do crime como campo de estudo é frequentemente associada ao trabalho de Gary Becker em *Crime and Punishment: An Economic Approach* (1968), em que o autor propõe que o ato criminoso pode ser compreendido como resultado de uma decisão racional diante de custos e benefícios esperados. Nessa abordagem, o indivíduo avalia a viabilidade da atividade ilícita comparando os ganhos monetários ou psicológicos decorrentes do crime com os custos associados à probabilidade de punição e à severidade da pena. Como essa decisão ocorre em um ambiente de incerteza, o agente j pondera os possíveis resultados antes de agir e escolhe a alternativa que maximiza sua função de utilidade esperada² (EU_j), representada pela equação:

$$EU_j = p_j U_j(Y_j - f_j) + (1 - p_j) U_j(Y_j) \quad (1)$$

Em que:

- p_j : probabilidade de punição;
- Y_j : ganho decorrente do delito;
- f_j : punição convertida em valor monetário;
- U_j : função de utilidade do indivíduo j .

A utilidade esperada do crime depende, portanto, da comparação entre dois cenários possíveis: ser punido e sofrer a punição correspondente, ou não ser punido e usufruir integralmente do ganho Y_j . Assim, quanto maior a probabilidade de punição ou a severidade da pena, menor tende a ser a utilidade esperada da atividade criminosa. A partir dessa estrutura de decisão individual, Becker (1968) amplia a análise para a oferta de delitos, argumentando que o número de crimes cometidos responde às variações nos custos esperados do crime e nas oportunidades disponíveis ao agente. Nessa formulação, u_j representa o conjunto de fatores que influenciam a oferta de delitos, incluindo características econômicas, sociais e individuais do agente. Assim, a oferta de delitos pode ser representada como:

$$O_j = O_j(p_j, f_j, u_j) \quad (2)$$

Essa formulação é relevante para este estudo, pois permite interpretar variáveis socioeconômicas como indicadores associados ao custo de oportunidade do crime. A

² A função de utilidade expressa a relação entre determinada condição ou resultado e o nível de satisfação obtido pelo indivíduo (MANKIW, 2021). A utilidade esperada, por sua vez, considera os possíveis resultados de uma escolha sob incerteza e suas respectivas probabilidades.

expansão do emprego formal pode ampliar os retornos esperados da atividade lícita e, consequentemente, reduzir a atratividade relativa da conduta delituosa. De modo semelhante, níveis mais elevados de escolaridade podem ampliar expectativas de rendimento futuro, tornando a permanência na legalidade mais vantajosa. Por outro lado, contextos de baixa inserção produtiva, fragilidade educacional e maior vulnerabilidade social podem reduzir o custo de oportunidade da entrada em atividades ilícitas.

Isaac Ehrlich, em *Participation in Illegitimate Activities: A Theoretical and Empirical Investigation* (1973), complementa a análise de Becker ao incorporar de forma mais explícita variáveis estruturais ao cálculo racional do indivíduo. Segundo o autor, a decisão de delinquir está associada à comparação entre os retornos esperados das atividades legais e ilegais, sendo influenciada pela distribuição de renda, pelo nível educacional e pelo acesso a oportunidades legais. Embora sua análise tenha sido desenvolvida principalmente para compreender crimes contra a propriedade, sua contribuição é relevante para este estudo por destacar como desigualdade, escolaridade e inserção no mercado de trabalho podem afetar os incentivos à atividade ilícita.

Ao tratar os retornos no mercado legal (l) e ilegal (i) como alternativas concorrentes na alocação do tempo, Ehrlich (1973) formaliza a decisão de delinquir a partir da comparação entre os ganhos esperados das atividades lícitas e ilícitas. De forma simplificada, o retorno líquido esperado da infração pode ser representado como:

$$E[\pi_i] = w_i - c_i - p_i f_i \quad (3)$$

Onde:

- w_i : ganho esperado com a infração;
- c_i : custos diretos da ação criminosa;
- $p_i f_i$: custo esperado da punição.

A decisão de delinquir, contudo, não ocorre simplesmente porque o retorno líquido da atividade ilícita é positivo ($E[\pi_i] > 0$). Para Ehrlich (1973), o indivíduo compara esse retorno com aquilo que poderia obter em atividades legais. Assim, a entrada na criminalidade depende de que o retorno líquido esperado da infração supere o custo de oportunidade associado à atividade lícita, acrescido de elementos como aversão ao risco e custos morais:

$$E[\pi_i] > w_l + \rho \quad (4)$$

Sendo ρ o somatório da aversão ao risco e dos custos morais, e w_l o retorno esperado no mercado legal, a teoria sugere que a legalidade se torna relativamente mais atrativa quando o retorno líquido esperado da infração $E[\pi_i]$ não supera o seu custo de oportunidade ($w_l + \rho$),

mesmo que a infração apresente retorno líquido positivo. Com isso, o autor reforça a visão de Becker ao demonstrar que o aumento do retorno esperado no mercado legal, impulsionado por maior nível educacional e acesso a oportunidades, pode contribuir para reduzir a atratividade relativa do crime.

Block e Heineke (1975) refinam essa lógica ao argumentar que a atividade ilegal não deve ser compreendida como um substituto perfeito para o trabalho lícito. No artigo *A Labor Theoretic Analysis of the Criminal Choice*, os autores incorporam à função de utilidade variáveis como risco, esforço, estigma social e custos morais, indicando que a oferta de crimes não depende apenas do retorno líquido esperado. Dessa forma, fatores como escolaridade, inserção no mercado formal e vínculos institucionais podem influenciar não apenas os ganhos esperados da legalidade, mas também os custos subjetivos associados à entrada em trajetórias criminais.

Zhang (1997), em *The Effects of Welfare Programs on Criminal Behaviour: A Theoretical and Empirical Analysis*, amplia a análise econômica do crime ao incorporar os programas de assistência social à restrição orçamentária do indivíduo. Para o autor, a provisão estatal de bem-estar pode criar um piso mínimo de utilidade na legalidade, elevando o custo de oportunidade da entrada em atividades ilícitas. Além disso, a possibilidade de perda do acesso a benefícios ou mecanismos de proteção social em caso de punição aumenta o custo esperado da conduta criminosa. Assim, políticas de assistência social podem reduzir a atratividade relativa do crime ao tornar a permanência na legalidade mais vantajosa, especialmente para indivíduos em situação de vulnerabilidade.

Em síntese, a teoria econômica do crime permite interpretar a criminalidade como resultado da interação entre incentivos individuais, oportunidades legais, custos esperados da punição e condições institucionais. Embora o modelo de Becker (1968) represente o ponto de partida dessa abordagem, as contribuições posteriores mostram que a decisão de delinquir não depende apenas do cálculo entre ganho e punição, mas também de fatores estruturais, como desigualdade, acesso ao mercado de trabalho, escolaridade, proteção social e condições de vulnerabilidade. Esses elementos fornecem a base teórica para interpretar a violência letal como um fenômeno influenciado por características socioeconômicas, demográficas e institucionais presentes nos municípios.

2.3. Evidências empíricas sobre criminalidade e violência letal

A pluralidade de recortes espaciais e temporais da literatura empírica sobre criminalidade e violência letal no Brasil oferece um importante ponto de partida para observar

como a violência se manifesta em distintos contextos regionais, especialmente em um país marcado por grande diversidade social, econômica e territorial.

Cerqueira e Moura (2015), por exemplo, analisam o efeito das oportunidades no mercado de trabalho sobre as taxas de homicídios nos municípios brasileiros entre 1980 e 2010. Por meio de um painel com efeitos fixos municipais e temporais, os autores encontraram efeito positivo da taxa de desemprego dos homens sobre os homicídios, especialmente entre os jovens, com maior intensidade para o grupo de 25 a 29 anos. Em contraste, a renda do trabalho por hora trabalhada apresentou efeito estatisticamente nulo sobre a taxa de homicídios. Oliveira (2005), por sua vez, investiga as causas da criminalidade nas cidades brasileiras e sua relação com o porte populacional, utilizando dados municipais em painel para a década de 1990. Os resultados indicam que a criminalidade tende a ser mais elevada em municípios de maior porte populacional e destacam a desigualdade de renda, a pobreza e fatores relacionados à educação como elementos relevantes para a compreensão do fenômeno. Em conjunto, esses estudos indicam que, em nível nacional, a criminalidade está associada a condições que variam entre os municípios, como oportunidades no mercado de trabalho, porte populacional, pobreza e desigualdade.

Essa leitura ganha maior complexidade quando se volta para a região Nordeste, onde a violência letal se manifesta em um contexto marcado por desigualdades históricas e diversidade territorial. Dessa forma, estudos voltados especificamente ao Nordeste brasileiro são importantes para compreender a dinâmica da violência letal no contexto regional.

Nóbrega Júnior (2017), ao analisar a dinâmica dos homicídios nos estados nordestinos entre 2000 e 2013, destaca a associação entre fatores socioeconômicos e violência letal. Por meio da correlação de Pearson, o autor identificou relação positiva entre desemprego e homicídios, sugerindo que a escassez de oportunidades de trabalho pode contribuir para o aumento da violência. Em contrapartida, variáveis como renda domiciliar *per capita*, gastos públicos em segurança e número de prisões apresentaram associações mais fracas com os homicídios, enquanto o efetivo policial mostrou associação positiva de magnitude moderada.

Para além dos fatores socioeconômicos, parte da literatura aplicada ao Nordeste também tem investigado a relação entre segurança pública e violência letal. Nogueira et al. (2019), ao analisarem a eficiência dos gastos em segurança nos estados nordestinos entre 2008 e 2012 por meio da Análise Envoltória de Dados (DEA), identificaram elevado grau de ineficiência na gestão desses recursos. De forma complementar, Batista (2021), utilizando

correlações para os estados da região no período de 2014 a 2020, não encontrou associações consistentes entre a ampliação dos gastos policiais e a redução dos crimes violentos. Em conjunto, esses resultados sugerem que o aumento dos recursos destinados à segurança pública, por si só, não garante reduções nos indicadores de violência, reforçando a importância da eficiência na alocação dos gastos e da efetividade das políticas adotadas.

Ao deslocar o foco para a escala municipal, outra vertente da literatura evidencia que a violência letal nordestina também possui uma dimensão espacial relevante. Nessa perspectiva, estudos que empregam técnicas de análise espacial buscam identificar padrões de dependência espacial e concentração geográfica, contribuindo para a compreensão da distribuição territorial da criminalidade letal. Plassa, Paschoalino e Santos (2019) analisaram a distribuição espacial e os determinantes socioeconômicos das taxas de homicídios nos municípios nordestinos entre 2010 e 2014. Os autores identificaram forte dependência espacial, com agrupamentos de municípios de alta violência concentrados especialmente em áreas litorâneas. Além disso, apontaram a desigualdade de renda como um dos principais fatores associados às taxas de homicídios na região. Rocha e Lima (2023) analisaram a evolução e a dinâmica territorial das taxas de homicídios nos municípios nordestinos em 2000, 2010 e 2020. Os resultados evidenciaram um processo de convergência das taxas de homicídios entre municípios de diferentes portes, uma vez que os municípios menores apresentaram crescimento proporcionalmente mais intenso da violência letal ao longo do período. Esse comportamento sugere que a criminalidade violenta na região não se restringe às capitais ou aos grandes centros urbanos, alcançando também municípios de menor porte. Tais estudos evidenciam que a violência letal no Nordeste possui forte dimensão territorial, marcada por dependência espacial, concentração geográfica e heterogeneidade entre municípios. Dessa forma, a análise da criminalidade na região exige atenção não apenas aos fatores socioeconômicos, mas também à forma como a violência se distribui no espaço e às diferenças observadas entre municípios de distintos portes populacionais.

Além das abordagens espaciais e socioeconômicas, parte da literatura também tem buscado avaliar em que medida os modelos tradicionais da economia do crime são capazes de explicar a trajetória da violência letal nordestina. Nessa perspectiva, Jorge (2018) propõe uma abordagem exploratória para compreender a dinâmica dos homicídios no Nordeste entre 2000 e 2015, replicando exercícios desenvolvidos por Levitt (2004) e Cerqueira (2010). Ao considerar variáveis sociais, econômicas e de dissuasão, o autor identifica uma expressiva discrepância entre a previsão do modelo e a trajetória observada da violência letal: enquanto

os determinantes utilizados indicariam uma redução de aproximadamente 70% nas taxas de homicídio, a região registrou aumento de 71% no período. Esse resultado sugere que a dinâmica nordestina apresenta especificidades não plenamente captadas pelos modelos tradicionais da economia do crime, reforçando a necessidade de análises empíricas voltadas às particularidades regionais.

Diante dessas contribuições, observa-se que a literatura tem avançado na compreensão da criminalidade letal no Nordeste a partir de diferentes abordagens, incluindo análises socioeconômicas, estudos sobre segurança pública, análises espaciais e aplicações da economia do crime. Ainda assim, permanecem oportunidades para investigações que utilizem dados municipais em painel, abrangendo um período mais recente e incorporando diferentes dimensões associadas à violência letal. Nesse sentido, o presente estudo contribui ao analisar os fatores associados aos homicídios nos municípios do Nordeste brasileiro entre 2013 e 2023, considerando indicadores representativos das dimensões econômica, do mercado de trabalho, educacional, demográfica e de proteção social. Adicionalmente, busca-se verificar se esses fatores apresentam comportamentos distintos entre municípios de diferentes portes populacionais.

3. ASPECTOS METODOLÓGICOS

Esta seção apresenta os procedimentos metodológicos adotados para analisar os fatores associados aos homicídios nos municípios do Nordeste brasileiro. Inicialmente, descreve-se a estrutura da base de dados em painel e as variáveis selecionadas. Em seguida, são apresentados a especificação econométrica dos modelos e os procedimentos empregados para avaliar a adequação da inferência estatística.

3.1. Abordagem geral da análise

O estudo de fenômenos sociais complexos, como a criminalidade letal, requer uma estrutura de dados capaz de capturar tanto as diferenças territoriais quanto as mudanças observadas ao longo dos anos. Na literatura econométrica, a dimensão temporal é captada pelas séries temporais, que consistem em observações de uma mesma unidade em períodos sucessivos, permitindo analisar tendências e mudanças ao longo do tempo, embora não permitam, isoladamente, a comparação entre diferentes unidades de análise. Por outro lado, a dimensão territorial pode ser observada por meio de dados em corte transversal, que reúnem observações de diferentes unidades em um único período, possibilitando comparações entre elas, mas sem acompanhar suas trajetórias evolutivas (GUJARATI; PORTER, 2011). A estrutura de dados em painel, por sua vez, integra essas dimensões ao acompanhar diferentes unidades de análise ao longo de múltiplos períodos, o que permite observar, simultaneamente, as diferenças entre as unidades e as mudanças ocorridas ao longo do tempo. Conforme Gujarati e Porter (2011) e Baltagi (2005), essa organização amplia a quantidade de informação disponível, aumenta a variabilidade dos dados e possibilita o controle de características não observadas associadas às unidades ou aos períodos considerados.

Dessa forma, a unidade de análise do presente estudo corresponde aos municípios da Região Nordeste, abrangendo os estados de Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia, observados anualmente entre 2013 e 2023. Cada observação corresponde a uma combinação município-ano, representada pelo par it , em que i identifica o município e t o período. Essa estrutura permite analisar a relação entre a taxa de homicídios e indicadores socioeconômicos, educacionais e demográficos, considerando tanto a heterogeneidade municipal da região quanto as mudanças ocorridas ao longo do

período. Por fim, a base inicial compreende 1.793 municípios nordestinos³ ao longo de 11 anos, totalizando 19.723 observações município-ano. Após a exclusão das observações com valores ausentes⁴ nas variáveis utilizadas nos modelos econométricos, a amostra efetivamente empregada nas estimações passou a contar com 1.790 municípios e 19.669 observações. Dessa forma, a base utilizada nas estimações caracteriza-se, nos termos de Gujarati e Porter (2011), como um painel curto, uma vez que o número de unidades observadas (N) é substancialmente superior ao número de períodos (T), e como um painel desbalanceado, devido à exclusão de observações com valores ausentes nas variáveis utilizadas nos modelos econométricos.

3.2. Descrição das variáveis

A base empírica do estudo foi construída a partir de dados secundários extraídos de diferentes plataformas governamentais e posteriormente organizados em um painel anual. Conforme Marconi e Lakatos (2003), o uso de dados secundários permite mobilizar informações previamente produzidas e sistematizadas por instituições especializadas, sendo especialmente útil em análises de fenômenos sociais em larga escala. A seleção das variáveis explicativas foi orientada pelo referencial teórico discutido anteriormente e pela disponibilidade de informações em escala municipal e com periodicidade anual. Assim, as dimensões teóricas abordadas neste estudo foram operacionalizadas por meio de indicadores *proxy*, entendidos como variáveis observáveis utilizadas para representá-las empiricamente. O Quadro I apresenta as variáveis utilizadas no estudo, suas descrições, unidades de medida e respectivas fontes.

Quadro I: Variáveis utilizadas, unidades de medida e fontes

Variável	Descrição	Unidade	Fonte
Taxa de homicídios por 100 mil hab.	Óbitos por agressão a cada 100 mil habitantes	Taxa por 100 mil habitantes	SIM/DataSUS; IBGE
Log PIB real <i>per capita</i>	PIB municipal real <i>per capita</i> em logaritmo natural	Log natural	IBGE; Ipeadata

³ Fernando de Noronha foi removido da amostra por se tratar de uma unidade administrativa autônoma do Governo do Estado de Pernambuco.

⁴ Ao todo, foram excluídas 54 observações município-ano: 52 por ausência simultânea nas taxas de reprovação e abandono no ensino médio e 2 por ausência na taxa de emprego formal. Entre as perdas educacionais, destacam-se os municípios alagoanos de Flexeiras, Jequiá da Praia e Roteiro, excluídos integralmente da amostra por não apresentarem informações de reprovação e abandono escolar em nenhum dos anos analisados.

Emprego formal	Vínculos formais ativos em relação à população estimada	Percentual (%)	RAIS; IBGE
Cobertura do CadÚnico	População cadastrada no CadÚnico em relação à população estimada	Percentual (%)	MDS/SAGICAD; IBGE
Reprovação escolar	Taxa de reprovação do ensino médio	Percentual (%)	INEP
Abandono escolar	Taxa de abandono do ensino médio	Percentual (%)	INEP
Jovens 15-24	População de 15 a 24 anos em relação à população total estimada	Percentual (%)	MS

Fonte: Elaboração própria, a partir das bases indicadas no quadro.

A dimensão da criminalidade letal é representada pela variável dependente, mensurada pela taxa de homicídios por 100 mil habitantes, construída a partir do número absoluto de óbitos decorrentes de agressões⁵ registrados no Sistema de Informações sobre Mortalidade (SIM/DataSUS). A conversão dos valores absolutos em taxas foi adotada para permitir a comparação entre municípios de diferentes portes populacionais, tendo como referência a população municipal disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) para o período de 2013 a 2023. Essa mesma base populacional também foi empregada na construção dos demais indicadores proporcionais utilizados nas estimações.

Como indicador da dimensão econômica, utilizou-se o Produto Interno Bruto (PIB) municipal, disponibilizado pelo IBGE. O indicador foi deflacionado para preços constantes de 2023, com base nos fatores de deflação disponibilizados pelo Ipeadata. Posteriormente, o indicador foi transformado em termos *per capita* e convertido em logaritmo natural. Essa transformação foi adotada para reduzir discrepâncias de escala entre os municípios, suavizar a influência de valores extremos e facilitar a interpretação dos coeficientes estimados⁶.

⁵ Os óbitos decorrentes de agressões correspondem aos códigos X85-Y09 da CID-10, conforme a classificação adotada pelo SIM/DataSUS.

⁶ Modelos em que a variável dependente é mantida em nível e uma variável explicativa é expressa em logaritmo natural são denominados nível-log. O coeficiente indica, aproximadamente, a variação em unidades da variável dependente associada a um aumento de 1% na variável explicativa, mantidas constantes as demais variáveis (GUJARATI; PORTER, 2011).

Para mensurar a dimensão mercado de trabalho, foram utilizados os dados da Relação Anual de Informações Sociais (RAIS), considerando os vínculos ativos em 31 de dezembro de cada ano. O indicador foi construído como proporção da população municipal estimada, permitindo aproximar o grau de inserção formal no mercado de trabalho em nível municipal. No entanto, por mensurar apenas os vínculos registrados, o indicador não capta a totalidade da ocupação, especialmente em um contexto regional marcado por elevada informalidade no mercado de trabalho⁷. Ainda assim, o emprego formal permanece relevante para a análise por representar uma dimensão associada às oportunidades no mercado formal de trabalho, em diálogo com a abordagem econômica do crime.

Para captar a dimensão da proteção social, utilizou-se a taxa de cobertura do Cadastro Único para Programas Sociais do Governo Federal (CadÚnico)⁸, utilizada como indicador da cobertura cadastral da população de baixa renda nos municípios. Os dados foram extraídos da Secretaria de Avaliação, Gestão da Informação e Cadastro Único (SAGICAD), secretaria vinculada ao Ministério do Desenvolvimento e Assistência Social, Família e Combate à Fome (MDS). O indicador foi construído a partir da proporção entre a população cadastrada e a população municipal estimada.

No bloco educacional, foram utilizadas duas medidas associadas às condições educacionais do ensino médio: as taxas de reprovação e abandono, extraídas do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Enquanto a taxa de reprovação representa a proporção de estudantes que não obtiveram aprovação ao final do período letivo, a taxa de abandono indica a parcela de alunos que interromperam sua trajetória escolar durante o ano.

Quanto à composição demográfica, foi incorporada ao modelo a proporção de jovens de 15 a 24 anos em relação à população total municipal, calculada a partir das informações de população por faixa etária disponibilizadas pelo Ministério da Saúde (MS). Essa variável foi utilizada como controle da composição etária dos municípios, considerando que a literatura aponta a maior concentração de vítimas e envolvidos em eventos de violência letal entre grupos populacionais mais jovens.

⁷ Conforme evidenciam os dados da Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua) disponibilizados pelo IBGE, a taxa de informalidade da região Nordeste permaneceu superior à média nacional nos anos disponíveis entre 2016 e 2023.

⁸ Segundo o Decreto nº 11.016/2022, o CadÚnico identifica e caracteriza famílias de baixa renda, reunindo informações sobre domicílio, composição familiar, escolaridade, trabalho e renda, além de servir como instrumento para seleção e inclusão em programas sociais.

Apesar da busca por contemplar as principais dimensões apontadas pela literatura, algumas limitações relacionadas à disponibilidade de dados restringiram a inclusão de determinados indicadores no modelo. Variáveis orçamentárias, como o gasto municipal *per capita* em segurança pública, não apresentaram padronização e cobertura suficientes para todos os municípios e anos da amostra. A desigualdade de renda e os níveis de urbanização, por sua vez, figuram entre os fatores mais frequentemente associados à criminalidade na literatura empírica, mas sua incorporação neste estudo foi limitada pela indisponibilidade de séries anuais municipais compatíveis com o recorte temporal analisado. Adicionalmente, embora o avanço e a interiorização de facções criminosas sejam fatores relevantes para compreender a dinâmica recente da violência, a ausência de informações sistematizadas em nível municipal impossibilitou sua inclusão direta no modelo. Essas limitações devem ser consideradas na interpretação dos resultados obtidos, uma vez que o modelo busca analisar os fatores associados aos homicídios a partir das informações disponíveis para o recorte espacial e temporal definido.

A partir das dimensões teóricas representadas pelas variáveis *proxy* incorporadas ao modelo, o Quadro II sintetiza os sinais esperados dos coeficientes estimados. Essas expectativas se baseiam nos postulados teóricos da economia do crime, bem como em abordagens relacionadas à anomia e ao capital humano.

Quadro II: Relação esperada entre as dimensões analisadas e a taxa de homicídios

Dimensão	Sinal esperado	Justificativa	Referência base
Renda	+ / -	Efeito ambíguo: pode ampliar o retorno esperado das atividades ilícitas, mas também refletir melhores condições econômicas e oportunidades legais	Becker (1968); Merton (1938); Ehrlich (1973)
Emprego formal	-	Aumenta o custo de oportunidade do crime pela inserção no mercado formal.	Becker (1968); Ehrlich (1973); Hirschi (1969)
Proteção social	+ / -	Efeito ambíguo: pode refletir maior vulnerabilidade socioeconômica, mas também maior alcance de políticas sociais e mecanismos de proteção institucional.	Zhang (1997); Shaw e McKay (1942)

Educação	+	Indica fragilidades na trajetória escolar, associadas à redução do acúmulo de capital humano (reprovação) e à interrupção do vínculo escolar (abandono), podendo limitar as expectativas de renda futura e ampliar a exposição a contextos de ilegalidade.	Becker (1968); Ehrlich (1973); Hirschi (1969); Block e Heineke (1975)
Composição demográfica	+	A maior proporção de jovens pode estar associada a menor estabilidade educacional e laboral, o que pode reduzir o custo de oportunidade das atividades ilícitas.	Becker (1968); Hirschi (1969)

Fonte: elaboração própria, com base no referencial teórico discutido na seção 2.1 e 2.2.

3.3. Especificação econométrica

A partir da estrutura unidade-período, modelos com dados em painel permitem analisar a relação entre uma variável dependente e um conjunto de variáveis explicativas observáveis em diferentes unidades e ao longo do tempo. Seguindo a notação proposta por Baltagi (2005), uma formulação geral pode ser expressa por:

$$y_{it} = \beta X_{it} + u_{it} \quad (5)$$

Nessa formulação, para cada unidade i no período t , y_{it} representa a variável dependente, X_{it} corresponde ao conjunto de variáveis explicativas observáveis incluídas no modelo, e β representa o vetor de parâmetros a serem estimados. O ponto central dessa especificação reside no termo u_{it} , que representa o componente de erro do modelo e reúne fatores não observados diretamente pelas variáveis explicativas incluídas na estimação. Como destaca Baltagi (2005), esse termo pode ser decomposto em dois elementos: um componente não observado específico da unidade e um distúrbio remanescente, representados, respectivamente, por μ_i e v_{it} . Algebricamente:

$$u_{it} = \mu_i + v_{it} \quad (6)$$

Em outras palavras, o termo μ_i capta a heterogeneidade não observada associada a cada unidade, ou seja, as características próprias da unidade i que não são mensuradas diretamente, mas que podem influenciar a variável dependente. Já o termo v_{it} corresponde ao

distúrbio idiossincrático, associado aos fatores não observados que variam entre as unidades e ao longo do tempo.

De acordo com Gujarati e Porter (2011), ao incorporar μ_i como um componente específico e invariante no tempo para cada unidade, essa estrutura passa a ser interpretada como um modelo de efeitos fixos unidirecionais. Diferentemente de um modelo *pooled* tradicional⁹, o modelo de efeitos fixos permite controlar a heterogeneidade não observada por meio da atribuição de interceptos específicos às unidades analisadas, capturando características próprias de cada unidade que permanecem constantes ao longo do tempo¹⁰. A especificação pode ainda incorporar efeitos fixos de tempo, ampliando o controle sobre componentes não observados associados à dimensão temporal. Dessa forma, Baltagi (2005) apresenta a estrutura bidirecional do componente de erro, na qual são considerados, simultaneamente, componentes não observados associados à unidade e ao período considerado:

$$u_{it} = \mu_i + \lambda_t + v_{it} \quad (7)$$

Nesse caso, λ_t representa o efeito específico do período t , comum a todas as unidades analisadas; μ_i representa o efeito invariante no tempo associado a cada unidade; e v_{it} permanece como distúrbio idiossincrático. Substituindo a decomposição do erro na formulação geral, obtém-se uma especificação para dados em painel com efeitos associados à unidade i e ao período t , representada por:

$$y_{it} = X_{it}\beta + \mu_i + \lambda_t + v_{it} \quad (8)$$

Embora a unidade observada neste estudo seja o município, a aplicação de efeitos fixos municipais não se mostra a alternativa mais adequada, uma vez que grande parte das variáveis explicativas apresenta baixa variação temporal dentro dos municípios. Conforme aponta Wooldridge (2010), a aplicação de efeitos fixos municipais nesses casos poderia absorver grande parte da variação dos dados durante a estimação, resultando em parâmetros imprecisos e erros-padrão¹¹ elevados. Diante dessa limitação, o modelo empírico adapta a

⁹ O modelo *pooled* trata todas as unidades como uma única amostra, assumindo um único intercepto e desconsiderando a heterogeneidade não observada específica de cada unidade. Dessa forma, esses efeitos permanecem incorporados ao termo de erro e, quando correlacionados com as variáveis explicativas, podem gerar estimativas viesadas e inconsistentes (GUJARATI; PORTER, 2011).

¹⁰ Os interceptos associados aos efeitos fixos não são apresentados diretamente nas tabelas de resultados, pois são absorvidos durante o processo de estimação do modelo (BALTAGI, 2005).

¹¹ O erro padrão indica o grau de incerteza associado a um coeficiente estimado. Valores mais elevados sugerem menor precisão da estimativa, enquanto valores menores indicam maior precisão (WOOLDRIDGE, 2012).

estrutura da Equação (8) ao incorporar efeitos fixos¹² de UF, permitindo controlar diferenças institucionais e estruturais compartilhadas pelos municípios de um mesmo estado. Essa escolha também se justifica pelas características institucionais da segurança pública brasileira, uma vez que, conforme estabelece o art. 144 da Constituição Federal, as Polícias Cíveis e Militares, responsáveis pela apuração de infrações penais e pela preservação da ordem pública, estão vinculadas aos governos estaduais (BRASIL, 1988).

Adicionalmente, o modelo também incorpora efeitos fixos de ano, permitindo controlar fatores temporais comuns a todos os municípios da amostra, como mudanças institucionais, conjunturas econômicas e outros choques agregados ocorridos ao longo do período analisado. Dessa forma, a representação algébrica do modelo estimado pode ser expressa como:

$$Hom_{it} = \beta_1 PIB_{it} + \beta_3 Cad_{it} + \beta_2 Emp_{it} + \beta_4 Reprv_{it} + \beta_5 Aband_{it} + \beta_6 Jovens_{it} + \alpha_{UF(i)} + \lambda_t + v_{it} \quad (9)$$

Em que Hom_{it} corresponde à taxa de homicídios por 100 mil habitantes do município i no ano t . As variáveis explicativas são representadas pelo log do PIB real *per capita* (PIB_{it}), pela taxa de cobertura do CadÚnico (Cad_{it}), pela taxa de emprego formal (Emp_{it}), pelas taxas de reprovação ($Reprv_{it}$) e abandono no ensino médio ($Aband_{it}$), e pela proporção de jovens de 15 a 24 anos ($Jovens_{it}$). O termo $\alpha_{UF(i)}$ representa os efeitos fixos associados às UF, enquanto λ_t representa os efeitos fixos de ano. Por fim, v_{it} corresponde ao termo de erro idiossincrático.

3.4. Procedimentos de diagnóstico e inferência

Além da especificação dos modelos, foram adotados procedimentos voltados à avaliação da adequação da inferência estatística. Esses procedimentos concentram-se em dois aspectos principais: a avaliação das relações entre as variáveis explicativas e a análise do comportamento do termo de erro.

No âmbito das relações entre as variáveis explicativas, destaca-se a possibilidade de multicolinearidade, que se refere à existência de associação linear elevada entre dois ou mais regressores. Quando essa associação é elevada, os coeficientes podem ser estimados com menor precisão, uma vez que os erros-padrão tendem a ser maiores, dificultando a

¹² Embora testes formais de especificação possam auxiliar na escolha entre modelos de efeitos fixos e efeitos aleatórios, a adoção dos efeitos fixos considera a possibilidade de correlação entre características não observadas das unidades e as variáveis explicativas. Conforme Baltagi (2005), essa abordagem permite controlar a heterogeneidade persistente entre as unidades mesmo quando esses efeitos não observados estão correlacionados com os regressores.

identificação da significância estatística individual dos regressores (GUJARATI; PORTER, 2011; WOOLDRIDGE, 2012). Por esse motivo, possíveis indícios de multicolinearidade foram avaliados por meio da matriz de correlação de Pearson entre as variáveis explicativas. Nessa matriz, são apresentados os coeficientes de correlação entre os pares de variáveis, que variam entre -1 e +1 e permitem identificar associações lineares elevadas, especialmente quando iguais ou superiores a 0,8 em valor absoluto (GUJARATI; PORTER, 2011).

Quanto ao comportamento do termo de erro, a análise considerou a possibilidade de heterocedasticidade e autocorrelação serial. Segundo Gujarati e Porter (2011), a hipótese clássica de homocedasticidade pressupõe que os erros apresentem variância¹³ constante entre as observações, o que pode ser restritivo em painéis nos quais as unidades observadas apresentam diferenças significativas de tamanho e estrutura. Desse modo, quando a variância dos erros não é constante, tem-se a presença de heterocedasticidade. No contexto deste estudo, essa possibilidade é relevante porque os municípios analisados, ainda que majoritariamente de pequeno porte, apresentam diferenças relevantes de escala populacional, estrutura socioeconômica e dinâmica territorial, o que pode contribuir para a violação da hipótese de homocedasticidade.

A autocorrelação serial, por sua vez, está associada à dimensão temporal do painel e refere-se à possibilidade de correlação dos termos de erro ao longo do tempo. De acordo com Gujarati e Porter (2011), o modelo clássico de regressão linear pressupõe que os erros sejam independentes entre si. Esse pressuposto, no entanto, torna-se particularmente sensível em dados em painel, uma vez que as mesmas unidades são acompanhadas ao longo de diferentes períodos. Nessas circunstâncias, choques ou fatores não observados podem persistir ao longo do tempo, fazendo com que os erros de uma unidade guardem relação com os erros dessa mesma unidade em períodos seguintes (BALTAGI, 2005).

De acordo com Wooldridge (2012), a presença de heterocedasticidade e autocorrelação serial não implica, por si só, viés nos coeficientes estimados, mas pode comprometer a precisão da inferência estatística, uma vez que os erros-padrão convencionais podem ser calculados de forma inadequada. Para lidar com essa limitação, Cameron e Miller (2015) recomendam o uso de erros-padrão robustos clusterizados, procedimento que ajusta a estimação dos erros-padrão sem alterar os coeficientes estimados, permitindo considerar a

¹³ A variância é uma medida de dispersão calculada a partir do quadrado dos desvios de cada observação em relação à média da distribuição (GUJARATI; PORTER, 2011).

possibilidade de correlação entre os erros pertencentes a uma mesma unidade, ao mesmo tempo em que os erros de unidades distintas são assumidos como independentes.

Embora o modelo incorpore efeitos fixos de UF, a clusterização nesse nível de agregação não se mostra adequada devido ao número reduzido de estados na região analisada. Dessa forma, considerando que a unidade acompanhada ao longo do tempo corresponde ao município, as estimações foram conduzidas com erros-padrão robustos clusterizados por município¹⁴, permitindo considerar a heterocedasticidade entre as unidades e a correlação serial dos erros dentro de cada município.

3.5. Estratégia de estimação e composição dos modelos

A partir da especificação econométrica apresentada na Equação (9), a estratégia empírica foi organizada em duas etapas. Inicialmente, estimou-se um modelo geral (Modelo 1) utilizando todos os municípios da amostra. Esse modelo serve como referência comparativa para as estimações que incorporam variáveis relacionadas ao porte populacional, definido a partir da população municipal observada em 2013, de modo a manter os grupos fixos ao longo do período analisado. A composição dos modelos estimados é sintetizada na Tabela I.

Tabela I: Composição dos modelos estimados

Modelo	Especificação	Municípios (n)	Observações (N)
(1)	Geral	1.790	19.669
(2)	<i>Dummies</i> de porte populacional	1.790	19.669
(3a)	Inferior a 50 mil hab.	1610	17.689
(3b)	50 mil a 100 mil hab.	120	1.320
(3c)	100 mil a 300 mil hab.	42	462
(3d)	Igual ou superior a 300 mil hab.	18	198

Fonte: Elaboração própria, com base em dados do SIM/DataSUS, IBGE, Ipeadata, MDS/SAGICAD, RAIS, MS e INEP, 2013–2023.

Na primeira etapa, o Modelo 1 é comparado ao Modelo 2, que incorpora variáveis binárias¹⁵ (*dummies*) de porte populacional. Nesse modelo, a categoria de referência corresponde aos municípios com população inferior a 100 mil habitantes. As variáveis *dummy* identificam dois grupos adicionais: municípios de porte médio, com população igual ou

¹⁴ As estimações foram realizadas no *software* R, por meio do pacote *fixest*, com erros-padrão robustos clusterizados no nível municipal, especificados pelo argumento *vcov = ~ id_municipio*.

¹⁵ Variáveis binárias, ou *dummies*, assumem valores 0 e 1 para indicar, respectivamente, a ausência ou presença de determinado atributo ou categoria (GUJARATI; PORTER, 2011).

superior a 100 mil e inferior a 300 mil habitantes, e municípios de grande porte, com população igual ou superior a 300 mil habitantes. Essa comparação possui dois objetivos principais. O primeiro consiste em verificar se a inclusão do porte populacional altera os sinais, as magnitudes e os níveis de significância dos regressores do Modelo 1. O segundo consiste em verificar se municípios de diferentes portes apresentam taxas de homicídios distintas em relação à categoria de referência, mesmo após o controle pelas demais variáveis explicativas e pelos efeitos fixos de UF e ano.

Na segunda etapa, o Modelo 1 é comparado às estimações dos Modelos 3a, 3b, 3c e 3d, realizadas separadamente para cada uma das faixas populacionais apresentadas na Tabela I. Diferentemente da estratégia baseada em variáveis *dummy*, essa abordagem permite que os coeficientes associados às variáveis explicativas variem entre os grupos de municípios analisados.

4. RESULTADOS E ANÁLISE

Esta seção apresenta os resultados da pesquisa e sua análise, iniciando pela descrição das variáveis selecionadas para o estudo e pela matriz de correlação entre os regressores. Em seguida, são discutidos os resultados das estimações econométricas, considerando o modelo geral, o modelo com inclusão de variáveis binárias e os modelos estimados por faixas populacionais.

4.1. Análise descritiva das variáveis

A Tabela II sintetiza as estatísticas descritivas das variáveis utilizadas nos modelos econométricos, considerando a amostra efetivamente empregada nas estimações, composta por 19.669 observações município-ano no período de 2013 a 2023. Além das medidas de tendência central e dispersão, a análise considera os quartis da distribuição¹⁶, o que permite observar a concentração da parte central dos dados e indícios da presença de valores extremos.

Tabela II: Estatística descritiva das variáveis analisadas, 2013 a 2023

Estatística	Mínimo	Q1	Mediana	Média	Q3	Máximo	Desvio-Padrão
Homicídios	0	5,32	19,16	24,95	36,85	307,56	25,37
Log PIB pc	6,37	9,31	9,49	9,6	9,75	13,44	0,47
Emp. Formal	0,05	5,14	6,79	8,34	9,68	81,26	5,66
CadÚnico	24,95	66,69	74,4	74,19	81,99	306,06	13,01
Reprov. EM	0	1,5	5,5	7,37	11,1	61,5	7,29
Aband. EM	0	2,1	6,2	7,47	11,1	53,5	6,54
Jovens 15-24	10,94	16,18	17,5	17,51	18,85	23,99	1,86

Fonte: Elaboração própria, com base em dados do SIM/DataSUS, IBGE, Ipeadata, MDS/SAGICAD, RAIS, MS e INEP, 2013–2023.

A taxa média de homicídios observada na amostra, no período analisado, foi de 24,95 por 100 mil habitantes, enquanto a mediana foi de 19,16. A análise dos quartis mostra que a parte central da distribuição se concentra em patamares relativamente inferiores, com o primeiro quartil em 5,32 e o terceiro quartil em 36,85 homicídios por 100 mil habitantes. No

¹⁶ Os quartis são medidas de posição obtidas a partir da ordenação dos dados e permitem avaliar como as observações se distribuem ao longo da série. O primeiro quartil (Q1) representa o ponto que separa os 25% menores valores da distribuição, enquanto o terceiro quartil (Q3) representa o ponto a partir do qual se encontram os 25% maiores valores (BUSSAB; MORETTIN, 2017).

entanto, o valor máximo de 307,56 está substancialmente acima do terceiro quartil, sugerindo a presença de valores extremos na parte superior da distribuição. Esse comportamento, combinado à média superior à mediana, sugere uma distribuição assimétrica à direita, uma vez que valores elevados tendem a deslocar a média para cima em relação à mediana¹⁷. Além disso, o desvio-padrão¹⁸ de 25,37, de magnitude semelhante à média, sinaliza forte dispersão e heterogeneidade das taxas de violência letal entre as observações município-ano analisadas.

O logaritmo natural do PIB real *per capita*, utilizado como indicador das condições econômicas municipais, apresentou média de 9,60, mediana de 9,49 e desvio-padrão de 0,47. A proximidade entre média e mediana, associada à concentração da metade central das observações entre 9,31 e 9,75, sugere uma menor dispersão relativa quando comparada às demais variáveis analisadas. Esse resultado reflete, em parte, o efeito da transformação logarítmica na atenuação das diferenças de escala entre os municípios. Ainda assim, a amplitude observada entre o valor mínimo, de 6,37, e o valor máximo, de 13,44, indica que persistem diferenças econômicas relevantes no território analisado, mesmo após a transformação logarítmica.

A taxa de emprego formal apresentou média de 8,34% e mediana de 6,79%, indicando baixa proporção de vínculos formais em relação à população municipal. Os quartis mostram que a metade central das observações está situada entre 5,14% e 9,68%, o que reforça a predominância de baixos níveis de emprego formal na amostra. No entanto, o valor máximo de 81,26% revela a existência de municípios com níveis de formalização substancialmente superiores ao padrão observado na maior parte da amostra. Além disso, o desvio-padrão de 5,66, relativamente elevado em relação à média, indica que o emprego formal não se distribui uniformemente no território nordestino, sugerindo que a inserção formal no mercado de trabalho ocorre de maneira desigual entre os municípios, possivelmente refletindo diferenças na dinâmica econômica local.

A taxa de cobertura do CadÚnico, calculada como a proporção da população cadastrada em relação à população municipal estimada, apresentou média expressiva de 74,19% e mediana de 74,40%. A proximidade entre essas medidas indica que a elevada presença relativa de pessoas cadastradas no CadÚnico não é um fenômeno restrito a casos

¹⁷ Segundo Bussab e Morettin (2017), a assimetria pode ser observada pela distância dos quartis em relação à mediana: maior afastamento à direita indica assimetria à direita, enquanto maior afastamento à esquerda indica assimetria à esquerda.

¹⁸ O desvio-padrão corresponde à raiz quadrada positiva da variância e expressa a dispersão dos valores na mesma unidade de medida da variável analisada. Quanto maior o desvio-padrão, maior a dispersão dos dados em torno do valor médio (GUJARATI; PORTER, 2011).

isolados, mas uma característica recorrente na amostra analisada. Essa interpretação também é sustentada pelos quartis, uma vez que a metade central das observações está situada entre 66,69% e 81,99%, sinalizando níveis elevados de cobertura cadastral em parcela expressiva das observações município-ano. O desvio-padrão de 13,01 indica dispersão moderada em torno da média, inferior à observada em outras variáveis deste estudo. O valor máximo de 306,06%, por sua vez, deve ser interpretado com cautela, uma vez que provavelmente reflete descompasso entre os registros administrativos do CadÚnico e as estimativas populacionais oficiais, especialmente em municípios de pequeno porte ou por diferenças de atualização entre as bases.

As taxas de reprovação e abandono no ensino médio apresentaram comportamentos estatísticos semelhantes, com médias de 7,37% e 7,47% e medianas de 5,50% e 6,20%, respectivamente. A diferença entre média e mediana em ambas as variáveis sugere a presença de assimetria positiva, indicando que observações com valores mais elevados exercem influência relevante sobre a distribuição. Essa interpretação é consistente com os quartis observados, uma vez que a metade central das observações está situada entre 1,60% e 11,10% na reprovação e entre 2,10% e 11,10% no abandono. Ao mesmo tempo, os valores máximos de 61,50% e 53,50%, respectivamente, revelam a existência de observações substancialmente acima do padrão predominante na amostra. Além disso, os desvios-padrão de 7,29 e 6,54, respectivamente, apresentam magnitude semelhante às médias observadas para essas variáveis, indicando dispersão relevante entre as observações analisadas.

A proporção de jovens entre 15 e 24 anos destaca-se como a variável de maior estabilidade estatística da amostra. A média de 17,51% e a mediana de 17,50% são praticamente idênticas, indicando ausência de assimetria relevante na distribuição. A concentração da metade central das observações entre 16,18% e 18,85% reforça essa leitura, revelando um intervalo relativamente estreito quando comparado às demais variáveis do estudo. Além disso, o desvio-padrão de 1,86 evidencia baixa dispersão, sugerindo que a estrutura demográfica juvenil apresenta pouca variabilidade entre as observações município-ano da amostra.

4.2. Matriz de correlação e diagnóstico de multicolinearidade

Após a análise descritiva das variáveis, foi elaborada uma matriz de correlação de Pearson entre os regressores do modelo, com o objetivo de identificar associações lineares

elevadas que pudessem indicar problemas de multicolinearidade. Os coeficientes de correlação são apresentados na Tabela III.

Tabela III: Matriz de correlação entre os regressores

	Log PIB pc	Emp. formal	CadÚnico	Reprov. EM	Aband. EM	Jovens 15-24
Log PIB pc	1,000					
Emp. formal	0,685	1,000				
CadÚnico	-0,233	-0,270	1,000			
Reprov. EM	0,075	0,051	-0,071	1,000		
Aband. EM	0,009	-0,014	0,078	0,185	1,000	
Jovens 15-24	-0,044	-0,093	0,038	0,170	0,183	1,000

Fonte: Elaboração própria, com base em dados do SIM/DataSUS, IBGE, Ipeadata, MDS/SAGICAD, RAIS, MS e INEP, 2013–2023.

De modo geral, os resultados não indicam correlações elevadas entre os regressores, o que reduz a preocupação com problemas severos de multicolinearidade. A maior correlação observada ocorre entre o logaritmo natural do PIB real *per capita* e a taxa de emprego formal, com coeficiente de 0,685. Esse achado é esperado, pois municípios com maior atividade econômica tendem a apresentar maior proporção de vínculos formais em relação à população municipal. Ainda assim, o coeficiente não é alto o suficiente para indicar que as duas variáveis representem a mesma dimensão, o que justifica sua permanência conjunta no modelo.

A taxa de cobertura do CadÚnico apresenta correlação negativa com o logaritmo natural do PIB real *per capita* (-0,233) e com a taxa de emprego formal (-0,270). Esse resultado é coerente com a interpretação da variável como indicador do alcance cadastral da população de baixa renda, uma vez que municípios com maior dinamismo econômico e maior proporção de emprego formal tendem a apresentar menor participação relativa da população cadastrada. Apesar dessa associação, os coeficientes observados permanecem relativamente baixos, indicando que a variável adiciona informações complementares às fornecidas pelos indicadores econômicos. Assim, sua permanência no modelo contribui para incorporar aspectos relacionados à proteção social sem gerar preocupação relevante de multicolinearidade.

As variáveis educacionais apresentam baixa correlação tanto entre si quanto em relação aos demais regressores. A correlação entre as taxas de reprovação e abandono no

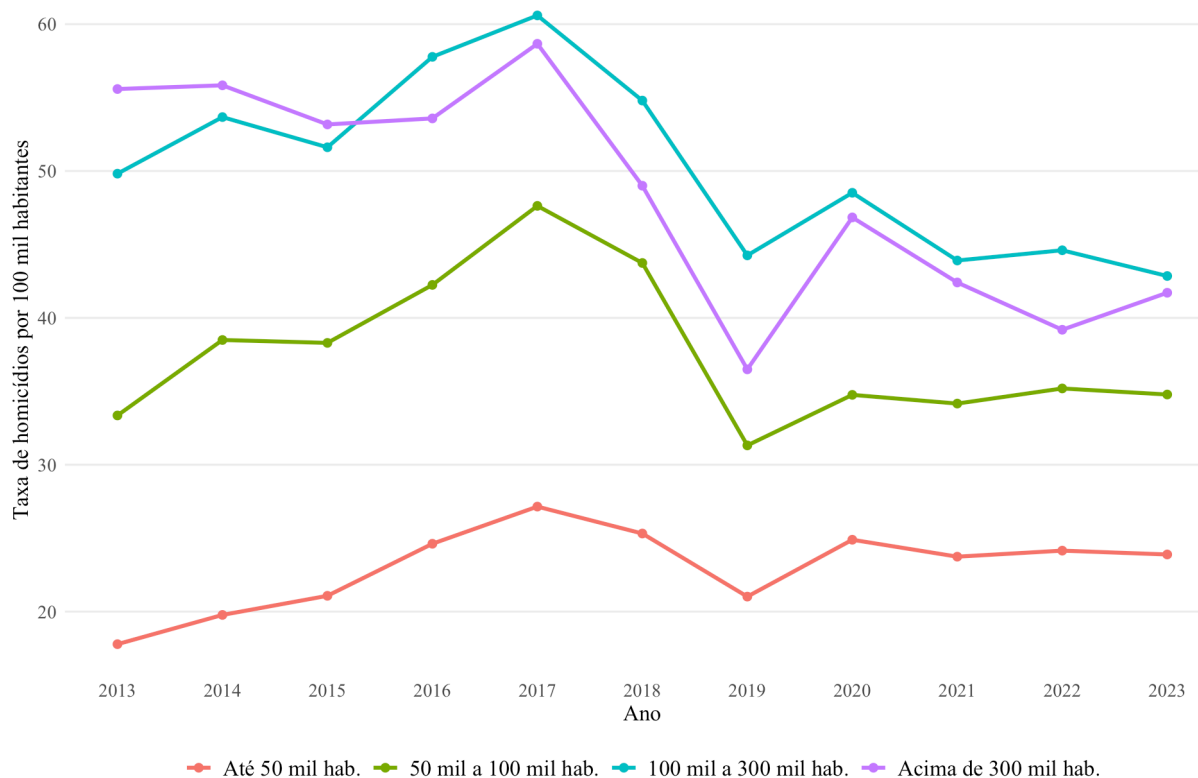
ensino médio foi de 0,185, indicando que, embora ambas estejam relacionadas à trajetória escolar, captam aspectos distintos das condições educacionais dos municípios. De forma semelhante, a proporção de jovens de 15 a 24 anos apresenta correlações reduzidas com os demais regressores, indicando que a composição etária dos municípios constitui uma dimensão relativamente independente das condições econômicas, educacionais e sociais captadas pelas demais variáveis.

Em síntese, a matriz de correlação não evidencia associações lineares suficientemente elevadas para caracterizar problemas severos de multicolinearidade. Dessa forma, os resultados sustentam a permanência dos regressores na especificação econométrica adotada.

4.3. Heterogeneidade da violência letal por porte populacional

Com o objetivo de explorar possíveis diferenças nos níveis de violência letal entre municípios de distintos portes populacionais, o Gráfico I apresenta a evolução da taxa média de homicídios por 100 mil habitantes segundo o porte populacional dos municípios nordestinos entre 2013 e 2023.

Gráfico I: Evolução da taxa média de homicídios por porte populacional nos municípios do Nordeste brasileiro (2013–2023)



Fonte: Elaboração própria, com base em dados do SIM/DataSUS e IBGE

Nota: Os municípios foram agrupados em quatro faixas populacionais: até 50 mil habitantes, de 50 mil a 100 mil habitantes, de 100 mil a 300 mil habitantes e acima de 300 mil habitantes.

Observa-se que os municípios com até 50 mil habitantes registraram as menores taxas médias de homicídios durante todo o período analisado. Em contrapartida, os grupos de maior porte registraram as taxas mais elevadas, com uma mudança importante a partir de 2016: municípios entre 100 mil e 300 mil habitantes superaram os municípios com mais de 300 mil habitantes, mantendo essa posição nos anos seguintes. Embora todos os grupos apresentem trajetória temporal semelhante, marcada por crescimento até 2017 e posterior redução, as diferenças entre os estratos permanecem expressivas ao longo de toda a série histórica. Diante dessas diferenças, a análise econométrica a seguir busca verificar se essas disparidades permanecem após o controle pelas demais características socioeconômicas, demográficas e institucionais incorporadas ao modelo.

4.3.1. *Modelo geral e modelo com dummies de porte populacional*

A Tabela IV apresenta os resultados do Modelo 1 e do Modelo 2, ambos estimados com efeitos fixos de UF e ano e erros-padrão robustos clusterizados por município. O Modelo 1 corresponde à especificação geral, enquanto o Modelo 2 incorpora variáveis *dummy* de porte populacional, tendo como categoria de referência os municípios com população inferior a 100 mil habitantes.

Tabela IV: Estimções com variáveis *dummy* de porte populacional

Variável dependente: taxa de homicídios por 100 mil habitantes		
	(1)	(2)
Renda	5,522*** (1,189)	5,312*** (1,151)
Emprego Formal	0,586*** (0,093)	0,524*** (0,093)
Proteção social	-0,188*** (0,038)	-0,156*** (0,038)
Reprovação EM	0,280*** (0,045)	0,278*** (0,045)
Abandono EM	0,225*** (0,042)	0,228*** (0,042)
Composição demográfica	1,083*** (0,278)	1,102*** (0,276)

Cidade média (<i>Dummy</i>)		12,259*** (2,167)
Cidade grande (<i>Dummy</i>)		2,066 (2,604)
Observações	19.669	19.669
R ²	0,188	0,193
Within R ²	0,087	0,093

Fonte: Elaboração própria, com base em dados do SIM/DataSUS, IBGE, Ipeadata, MDS/SAGICAD, RAIS, MS e INEP, 2013–2023.

Nota: *** Significante a 1%. ** Significante a 5%. * Significante a 10%.

Nota: Os valores em negrito representam os coeficientes estimados; os valores entre parênteses correspondem aos erros-padrão robustos clusterizados por município.

Nota: A categoria de referência das *dummies* corresponde aos municípios com menos de 100 mil habitantes em 2013.

A comparação entre os modelos indica que a inclusão das variáveis binárias de porte populacional não altera os sinais nem os níveis de significância dos regressores incluídos nos modelos. As magnitudes também permanecem bastante próximas entre as duas especificações, sugerindo que a inclusão do porte populacional não altera substancialmente os resultados do Modelo 1. Dessa forma, a análise dos coeficientes toma como referência o Modelo 2, por incorporar a dimensão do porte populacional e apresentar a especificação considerada mais completa.

No bloco econômico, representado pelas *proxies* de renda e inserção no mercado de trabalho, tanto o logaritmo natural do PIB real *per capita* quanto a taxa de emprego formal apresentam associação positiva e estatisticamente significativa a 1% com a taxa de homicídios. No caso do PIB, como a variável está em logaritmo, o coeficiente do Modelo 2, de 5,312, indica que um aumento de 1% no PIB real *per capita* está associado a um acréscimo aproximado de 0,053 ponto na taxa de homicídios por 100 mil habitantes, mantidas constantes as demais variáveis do modelo. Já a taxa de emprego formal, com coeficiente de 0,524, indica que o aumento de 1 ponto percentual na proporção de vínculos formais em relação à população municipal está associado a um acréscimo médio de 0,524 ponto na taxa de homicídios por 100 mil habitantes. A associação positiva observada para ambos os indicadores econômicos sugere que municípios economicamente mais dinâmicos tendem a registrar maiores taxas de homicídios. Esse resultado não corrobora, de forma direta, a expectativa teórica tradicional de que melhores condições econômicas estejam associadas à

redução da criminalidade. Contudo, essa associação exige cautela, pois pode refletir características estruturais de municípios economicamente mais dinâmicos, incluindo aspectos relacionados à desigualdade socioeconômica que não puderam ser incorporados à especificação empírica devido à limitação de dados, e não necessariamente um efeito direto dos indicadores econômicos sobre a violência.

Em sentido oposto aos indicadores econômicos, a *proxy* de proteção social, representada pela taxa de cobertura do CadÚnico, apresenta associação negativa com a taxa de homicídios nas duas especificações. No Modelo (2), o coeficiente de -0,156, estatisticamente significativo a 1%, indica que o aumento de 1 ponto percentual na cobertura do CadÚnico está associado a uma redução média de 0,156 ponto na taxa de homicídios por 100 mil habitantes, mantidas constantes as demais variáveis. O sinal negativo observado sugere que municípios com maior cobertura cadastral da população de baixa renda tendem a apresentar menores taxas de homicídios. Embora a associação observada não permita estabelecer relações causais, o resultado é compatível com interpretações que relacionam a ampliação do alcance de mecanismos de proteção social à redução da vulnerabilidade social.

Nos indicadores educacionais, observa-se uma associação positiva entre fragilidades no fluxo escolar e maiores taxas de violência letal. Tanto a reprovação quanto o abandono no ensino médio apresentam coeficientes positivos e estatisticamente significativos a 1%, com valores de 0,278 e 0,228 no Modelo 2, respectivamente. Em termos marginais, isso significa que aumentos de 1 ponto percentual nesses indicadores estão associados a acréscimos médios de 0,278 ponto e 0,228 ponto na taxa de homicídios por 100 mil habitantes, mantidas constantes as demais variáveis. A associação positiva observada para os indicadores de reprovação e abandono escolar é compatível com abordagens que associam trajetórias escolares fragilizadas à menor integração social, conforme discutido pela teoria do controle social de Hirschi (1969), e à ampliação das oportunidades econômicas e do custo de oportunidade associado ao envolvimento em atividades ilícitas.

A proporção de jovens de 15 a 24 anos, utilizada como *proxy* da composição demográfica, apresenta coeficiente positivo e estatisticamente significativo a 1%, alcançando coeficiente de 1,102 no Modelo 2. Esse resultado sugere que municípios com maior participação relativa desse grupo etário tendem a registrar taxas mais elevadas de homicídios, resultado compatível com a literatura que aponta maior exposição da população jovem às dinâmicas da violência letal. A magnitude do coeficiente indica que o aumento de 1 ponto

percentual da população jovem está associado a acréscimo médio de 1,102 pontos na taxa de homicídios por 100 mil habitantes.

Os resultados associados ao porte populacional revelam diferenças importantes entre os grupos de municípios analisados. Mesmo após o controle pelas demais variáveis do modelo e pelos efeitos fixos de UF e ano, os municípios de porte médio apresentam uma associação positiva e estatisticamente significativa com a taxa de homicídios em relação aos municípios da categoria de referência. O coeficiente estimado para esse grupo é de 12,259, estatisticamente significativo a 1%, o que corresponde a uma diferença média de aproximadamente 12,26 pontos na taxa de homicídios por 100 mil habitantes em relação à categoria de referência. Por outro lado, embora o coeficiente dos municípios de grande porte seja positivo, de 2,066, ele não apresenta significância estatística. Dessa forma, não é possível afirmar que municípios de grande porte apresentem taxas de homicídios diferentes das observadas nos municípios de menor porte. Esse resultado deve ser interpretado com cautela, uma vez que o grupo de municípios de grande porte é substancialmente menor que os demais, o que pode contribuir para menor precisão das estimativas associadas a essa categoria. Em conjunto, as diferenças observadas entre os estratos populacionais reforçam a importância de considerar o porte populacional na análise da violência letal, especialmente pela associação diferenciada observada entre os municípios de porte médio e as taxas de homicídios.

Em síntese, os resultados do Modelo 2 indicam que a violência letal está associada a fragilidades educacionais, à maior proporção de jovens de 15 a 24 anos na população e a um maior dinamismo econômico, representado pelo PIB *per capita* e pela taxa de emprego formal. Em contrapartida, maiores níveis de proteção social, representados pela cobertura do CadÚnico, apresentaram associação com menores taxas de homicídios. Um aspecto importante das estimações refere-se à incorporação das variáveis *dummy* de porte populacional que preservou, em grande medida, as associações identificadas no Modelo 1. Além disso, essa especificação evidenciou que os municípios de porte médio apresentam taxas de homicídios significativamente superiores às observadas nos municípios com menos de 100 mil habitantes, evidência que converge com os resultados apresentados pelo Atlas da Violência 2026. De forma distinta, não foram encontradas diferenças estatisticamente significativas entre os municípios de grande porte e a categoria de referência. Em conjunto, esses achados reforçam a interpretação de que a violência letal não se concentra apenas nas

grandes metrópoles, podendo assumir maior intensidade em municípios de porte intermediário.

4.3.2. *Modelo geral e modelos por faixas populacionais*

A Tabela V apresenta os resultados das estimações por faixas populacionais, tendo o Modelo 1 como referência geral. Essa comparação permite observar se os fatores associados à taxa de homicídios mantém o mesmo comportamento quando a amostra é segmentada segundo o porte dos municípios.

Tabela V: Estimações por estrato de faixas populacionais

Variável dependente: taxa de homicídios por 100 mil habitantes					
	(1)	(3a)	(3b)	(3c)	(3d)
Renda	5,522*** (1,189)	4,542*** (1,181)	6,378 (4,191)	15,492*** (4,509)	12,722 (9,413)
Emprego formal	0,586*** (0,093)	0,494*** (0,104)	0,552 (0,364)	0,299 (0,234)	0,049 (0,210)
Proteção social	-0,188*** (0,038)	-0,114*** (0,039)	-0,200 (0,152)	-0,366* (0,203)	-0,134 (0,386)
Reprovação EM	0,280*** (0,045)	0,279*** (0,046)	0,045 (0,174)	-0,083 (0,294)	-0,989 (0,773)
Abandono EM	0,225*** (0,042)	0,229*** (0,043)	0,336* (0,192)	0,796** (0,315)	0,892 (0,913)
Composição demográfica	1,083*** (0,278)	1,095*** (0,285)	0,073 (1,231)	1,003 (2,371)	2,585 (1,578)
Observações	19.669	17.689	1.320	462	198
R ²	0,188	0,149	0,237	0,449	0,462
Within R ²	0,087	0,047	0,120	0,250	0,092

Fonte: Elaboração própria, com base em dados do SIM/DataSUS, IBGE, Ipeadata, MDS/SAGICAD, RAIS, MS e INEP, 2013–2023.

Nota: *** Significante a 1%. ** Significante a 5%. * Significante a 10%.

Nota: Os valores em negrito representam os coeficientes estimados; os valores entre parênteses correspondem aos erros-padrão robustos clusterizados por município.

Nota: Faixas definidas com base na população municipal de 2013: (3a) menos de 50 mil habitantes; (3b) 50 mil a menos de 100 mil habitantes; (3c) 100 mil a menos de 300 mil habitantes; (3d) 300 mil habitantes ou mais.

Considerando a predominância de municípios de pequeno porte na estrutura municipal nordestina, o Modelo 3a, referente aos municípios com população inferior a 50 mil habitantes, concentra a maior parte das observações e, por isso, tende a se aproximar mais do padrão observado no Modelo 1. Essa proximidade se confirma pelos sinais, magnitudes e níveis de significância observados nos coeficientes estimados: no Modelo 3a, o logaritmo natural do PIB real *per capita*, a taxa de emprego formal, a cobertura do CadÚnico, a reprovação, o abandono no ensino médio e a proporção de jovens de 15 a 24 anos permanecem estatisticamente significativos, com sinais e magnitudes semelhantes aos do Modelo 1. Com isso, verifica-se que a associação média observada no modelo geral é fortemente influenciada pelos municípios pequenos, dada sua elevada participação no conjunto de observações.

Nas faixas populacionais seguintes, os coeficientes passam a apresentar maior instabilidade, principalmente em termos de magnitude e significância estatística. Esse comportamento se manifesta, inicialmente, no Modelo 3b, referente aos municípios com população igual ou superior a 50 mil e inferior a 100 mil habitantes. Embora a maior parte dos coeficientes preserve os sinais observados no Modelo 1, a maioria deles deixa de apresentar significância estatística. A exceção é a taxa de abandono no ensino médio, que permanece positiva e estatisticamente significativa a 10%, indicando que, mesmo nesse estrato, a fragilidade na permanência escolar continua associada à violência letal. Dessa forma, os resultados do Modelo 3b indicam que as associações observadas no modelo geral se tornam menos evidentes nesse estrato populacional, embora o abandono escolar ainda apareça como dimensão relevante.

No Modelo 3c, referente aos municípios com população igual ou superior a 100 mil e inferior a 300 mil habitantes, a heterogeneidade torna-se mais evidente. Diferentemente do Modelo 3b, esse estrato volta a apresentar coeficientes estatisticamente significativos em dimensões específicas. O principal destaque está no logaritmo natural do PIB real *per capita*, que permanece significativo a 1% e apresenta coeficiente de 15,492, valor substancialmente superior ao observado no Modelo 1 (5,522), sugerindo que a associação entre dinamismo econômico e violência letal é particularmente intensa nesse grupo de municípios. A fragilidade educacional também permanece relevante, uma vez que a taxa de abandono no ensino médio apresenta coeficiente positivo de 0,796 e significância estatística a 5%. Além disso, a cobertura do CadÚnico mantém coeficiente negativo e estatisticamente significativo a 10%, sugerindo que o alcance cadastral da população de baixa renda continua associado a

menores taxas de homicídios, ainda que com menor nível de significância estatística. Assim, o Modelo 3c sugere que, nos municípios com população entre 100 mil e 300 mil habitantes, fatores econômicos, educacionais e de proteção social apresentam associações mais evidentes com a violência letal.

No que se refere aos municípios com população igual ou superior a 300 mil habitantes, o Modelo 3d não apresenta coeficientes estatisticamente significativos. Embora alguns sinais permaneçam positivos, como no caso do PIB real *per capita*, do abandono escolar e da proporção de jovens, os erros-padrão mais elevados e o número reduzido de observações recomendam cautela na interpretação. Dessa forma, a ausência de significância estatística nesse grupo não deve ser interpretada como ausência de associação, mas analisada com cautela diante da menor precisão das estimativas obtidas para esse estrato populacional.

De modo geral, as estimações por faixa populacional evidenciam que os padrões identificados no Modelo 1 não se distribuem de forma homogênea entre os municípios nordestinos. Enquanto os municípios com até 50 mil habitantes reproduzem, em grande medida, o comportamento observado para o conjunto da amostra, os estratos de maior porte apresentam alterações relevantes na magnitude e na significância dos coeficientes. Destaca-se, sobretudo, o grupo de municípios com população entre 100 mil e 300 mil habitantes, no qual os indicadores econômicos, educacionais e de proteção social exibem relações mais intensas com a violência letal. Em conjunto, esses resultados mostram que o porte populacional constitui um elemento relevante para compreender a heterogeneidade da violência letal, uma vez que os efeitos estimados variam em magnitude e significância entre os diferentes estratos populacionais.

5. CONSIDERAÇÕES FINAIS

Em conjunto, os resultados permitem responder aos objetivos propostos neste trabalho, identificando fatores associados à violência letal nos municípios nordestinos entre 2013 e 2023. As estimações evidenciaram associações entre os homicídios e indicadores econômicos, educacionais, demográficos e de proteção social, sugerindo que a dinâmica da violência resulta da interação entre diferentes dimensões socioeconômicas. No entanto, destaca-se que a associação positiva observada para os indicadores econômicos diverge da expectativa mais tradicional da teoria econômica do crime, sugerindo que o dinamismo econômico, isoladamente, não é suficiente para explicar a redução da violência.

Contudo, a principal contribuição do estudo não reside apenas na identificação dessas associações, mas na constatação de que elas não se manifestam de forma homogênea entre os municípios nordestinos. A análise descritiva indicou diferenças relevantes nas taxas médias de homicídios entre os estratos populacionais, resultado posteriormente corroborado pelas estimações econométricas. A incorporação do porte populacional à análise revelou que municípios com população entre 100 mil e 300 mil habitantes apresentaram taxas médias de homicídios mais elevadas em comparação aos municípios com menos de 100 mil habitantes, mesmo após o controle pelas demais variáveis do modelo. Além disso, as estimações por faixas populacionais indicaram que a magnitude e a significância estatística dos coeficientes estimados variam conforme o porte dos municípios, sendo mais evidentes justamente nos municípios com população entre 100 mil e 300 mil habitantes. As evidências obtidas demonstram que análises agregadas podem ocultar diferenças relevantes na dinâmica da violência letal, uma vez que as associações identificadas para o conjunto da região não se reproduzem necessariamente em todos os estratos populacionais. Dessa forma, a violência letal no Nordeste brasileiro revela-se como um fenômeno marcado por significativa heterogeneidade territorial, cujos fatores associados variam conforme as características dos municípios analisados. Em consequência, estratégias de enfrentamento baseadas em diagnósticos agregados podem não capturar adequadamente as especificidades locais que influenciam a dinâmica dos homicídios.

Apesar das contribuições apresentadas, algumas limitações devem ser reconhecidas. A indisponibilidade de séries anuais municipais compatíveis com o período analisado restringiu a incorporação de variáveis frequentemente utilizadas na literatura, como indicadores de desigualdade de renda e urbanização. Além disso, a ausência de informações padronizadas

sobre gastos municipais em segurança pública e de indicadores sistematizados acerca da atuação de facções criminosas em nível municipal impediu a inclusão direta desses fatores nos modelos estimados.

Por fim, a heterogeneidade observada entre municípios de distintos portes populacionais indica que a dinâmica da violência letal no Nordeste brasileiro ainda apresenta aspectos que merecem investigação mais aprofundada. Para isso, pesquisas futuras podem incorporar fatores não contemplados neste estudo, como medidas de desigualdade socioeconômica, níveis de urbanização, indicadores de capacidade institucional dos governos locais e informações sobre a atuação de organizações criminosas. A incorporação desses fatores pode auxiliar na compreensão das diferenças observadas entre os municípios nordestinos, contribuindo para o avanço das evidências sobre a heterogeneidade que caracteriza a violência letal na região.

6. REFERÊNCIAS

BALTAGI, Badi H. **Econometric analysis of panel data**. 3. ed. Chichester: John Wiley & Sons, 2005.

BATISTA, Thiago Ramos. **Violência e produtividade policial na região Nordeste**. 2021. 29 f. Trabalho de Conclusão de Curso (Graduação em Ciências Econômicas) – Faculdade de Economia, Administração, Atuária e Contabilidade, Universidade Federal do Ceará, Fortaleza, 2021. Disponível em: https://repositorio.ufc.br/bitstream/riufc/64457/1/2021_tcc_trbatista.pdf. Acesso em: 21 jun. 2026.

BECKER, Gary S. Crime and punishment: an economic approach. **Journal of Political Economy**, v. 76, n. 2, p. 169-217, 1968.

BLOCK, Michael K.; HEINEKE, John M. A labor theoretic analysis of criminal choice. **American Economic Review**, v. 65, n. 3, p. 314-325, 1975.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, [1988]. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 9 jul. 2026.

BRASIL. **Decreto nº 11.016, de 29 de março de 2022**. Regulamenta o Cadastro Único para Programas Sociais do Governo Federal, instituído pelo art. 6º-F da Lei nº 8.742, de 7 de dezembro de 1993. Brasília, DF: Presidência da República, 2022. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2022/decreto/d11016.htm. Acesso em: 21 jun. 2026.

BRASIL. **Decreto-Lei nº 3.914, de 9 de dezembro de 1941**. Lei de introdução do Código Penal (Decreto-Lei nº 2.848, de 7 de dezembro de 1940) e da Lei das Contravenções Penais (Decreto-Lei nº 3.688, de 3 de outubro de 1941). Brasília, DF: Presidência da República, 1941. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto-lei/del3914.htm. Acesso em: 21 jun. 2026.

BRASIL. Ministério da Economia. **Relação Anual de Informações Sociais (RAIS): Microdados Vínculos**. Brasília, DF: Ministério da Economia, [s. d.]. Disponível em: <https://basedosdados.org/dataset/3e7c4d58-96ba-448e-b053-d385a829ef00?table=dabe5ea8-3bb5-4a3e-9d5a-3c7003cd4a60>. Acesso em: 21 jun. 2026.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS. **Mortalidade - desde 1996 pela CID-10**. Brasília, DF: Ministério da Saúde, [s. d.]. Disponível em: <https://datasus.saude.gov.br/mortalidade-desde-1996-pela-cid-10>. Acesso em: 21 jun. 2026.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS. **População Brasileira: município**. Brasília, DF: Ministério da Saúde, [s. d.]. Disponível em: <https://basedosdados.org/dataset/1e2b9a88-9dc7-4f0e-a3a5-e8d2a13869bf?table=1a8d9636-c11d-443b-ae83-1b00576f0b70>. Acesso em: 21 jun. 2026.

BRASIL. Ministério do Desenvolvimento e Assistência Social, Família e Combate à Fome. Secretaria de Avaliação, Gestão da Informação e Cadastro Único. **Pessoas cadastradas no Cadastro Único: número total**. Brasília, DF: Sagicad, [s. d.]. Disponível em: <https://aplicacoes.cidadania.gov.br/vis/data3/data-explorer.php>. Acesso em: 21 jun. 2026.

BUSSAB, Wilton de O.; MORETTIN, Pedro A. **Estatística básica**. 9. ed. São Paulo: Saraiva, 2017.

CAMERON, A. Colin; MILLER, Douglas L. A practitioner's guide to cluster-robust inference. **The Journal of Human Resources**, v. 50, n. 2, p. 317-372, 2015.

CERQUEIRA, Daniel; BUENO, Samira (coord.). **Atlas da violência**. São Paulo: IPEA; FBSP, 2026. 170 p. Disponível em: <https://repositorio.ipea.gov.br/entities/publication/9ab87dfc-33eb-4ac2-8a54-f4f1543dabbc>. Acesso em: 21 jun. 2026.

CERQUEIRA, Daniel; MOURA, Rodrigo Leandro de. **O efeito das oportunidades no mercado de trabalho sobre as taxas de homicídios no Brasil**. In: ENCONTRO NACIONAL DE ECONOMIA, 43., 2015, Florianópolis. Anais [...]. Florianópolis: ANPEC, 2015. Disponível em: https://www.anpec.org.br/encontro/2015/submissao/files_I/i12-0ce869e09e6385120c0146e239bb5bf8.pdf. Acesso em: 21 jun. 2026.

EHRlich, Isaac. Participation in illegitimate activities: a theoretical and empirical investigation. **Journal of Political Economy**, v. 81, n. 3, p. 521-565, 1973.

FLEISHER, Belton M. The effect of unemployment on juvenile delinquency. **Journal of Political Economy**, v. 71, n. 6, p. 543-555, 1963.

FLEISHER, Belton M. The effects of income on delinquency. **American Economic Review**, v. 56, n. 1/2, p. 118-137, 1966.

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

HIRSCHI, Travis. **Causes of Delinquency**. Berkeley, CA: University of California Press, 1969.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **População Brasileira: município**. Rio de Janeiro, RJ: IBGE, [s. d.]. Disponível em: <https://basedosdados.org/dataset/d30222ad-7a5c-4778-a1ec-f0785371d1ca?table=0c279444-165b-41da-92cd-50fd7e66baa1>. Acesso em: 21 jun. 2026.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Tabela 4708 - Taxa de informalidade das pessoas de 14 anos ou mais de idade ocupadas na semana de referência**. Rio de Janeiro, RJ: IBGE, [s. d.]. Disponível em: <https://sidra.ibge.gov.br/Tabela/4708>. Acesso em: 21 jun. 2026.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Tabela 5938 - Produto interno bruto a preços correntes, impostos, líquidos de subsídios, sobre produtos a preços correntes e valor adicionado bruto a preços correntes total e por atividade econômica, e respectivas participações - Referência 2010**. Rio de Janeiro, RJ: IBGE, [s. d.]. Disponível em: <https://sidra.ibge.gov.br/tabela/5938>. Acesso em: 21 jun. 2026.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Produto interno bruto (PIB) a preços de mercado - deflator implícito: variação anual (SCN10_DIPIBG10)**. Brasília, DF: Ipea, [s. d.]. Disponível em: <https://www.ipeadata.gov.br/ExibeSerie.aspx?serid=1184389724&module=M>. Acesso em: 21 jun. 2026.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Taxas de rendimento escolar**. Brasília, DF: Inep, 2025. Disponível em: <https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/indicadores-educacionais/taxas-de-rendimento-escolar>. Acesso em: 21 jun. 2026.

JORGE, Marco Antonio. O mistério da região Nordeste: o crescimento da taxa de homicídios no século XXI. **Revista Paranaense de Desenvolvimento**, Curitiba, v. 39, n. 134, p. 55-75, 2018.

LOUREIRO, André O. F.; CARVALHO JUNIOR, José Raimundo. **Uma análise econométrica do impacto dos gastos públicos sobre a criminalidade no Brasil**. Fortaleza: Laboratório de Estudos da Pobreza / CAEN / UFC, 2006. Ensaio sobre a Pobreza, n. 9.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003.

MERTON, Robert K. Social structure and anomie. **American Sociological Review**, v. 3, n. 5, p. 672-682, 1938.

NÓBREGA JÚNIOR, José Maria Pereira da. Violência homicida no Nordeste brasileiro: dinâmica dos números e possibilidades causais. **Dilemas: Revista de Estudos de Conflito e Controle Social**, Rio de Janeiro, v. 10, n. 3, p. 553-572, 2017.

NOGUEIRA, Lauro César Bezerra; VIEIRA, Carlos Rangel Pereira; VIEIRA, Bianca Alencar; MONTEIRO, Adriano David. Gastos públicos versus criminalidade: o caso da região Nordeste do Brasil. **Revista Brasileira de Segurança Pública**, v. 13, n. 2, p. 12-29, 2019.

OLIVEIRA, Cristiano Aguiar de. **Criminalidade e o tamanho das cidades brasileiras: um enfoque da economia do crime**. In: ENCONTRO NACIONAL DE ECONOMIA, 33., 2005, Natal. Anais [...]. Natal: ANPEC, 2005. Disponível em: <https://www.anpec.org.br/encontro2005/artigos/A05A152.pdf>. Acesso em: 21 jun. 2026.

PLASSA, Wander; PASCHOALINO, Pietro André Telatin; SANTOS, Moisés Pais dos. Determinantes socioeconômicos das taxas de homicídios no Nordeste brasileiro: uma análise espacial. **Planejamento e Políticas Públicas**, Brasília, n. 53, p. 479-507, 2019.

ROCHA, Luiz Eduardo Vasconcelos; LIMA, Eduardo Alvares de. **Evolução e dinâmica territorial da taxa de homicídios na região Nordeste no Brasil**. In: ENCONTRO NACIONAL DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS REGIONAIS E URBANOS, 21., 2023, Belém. Anais [...]. Belém: ABER, 2023. Disponível em: <https://brsa.org.br/wp-content/uploads/wpcf7-submissions/15690/Enaber-2023-ID.pdf>. Acesso em: 21 jun. 2026.

SHAW, Clifford R.; McKAY, Henry D. **Juvenile Delinquency and Urban Areas**. Chicago: University of Chicago Press, 1942.

UNITED NATIONS OFFICE ON DRUGS AND CRIME. **International classification of crime for statistical purposes (ICCS), version 1.0**. Vienna: United Nations, 2015.

Disponível em:

https://www.unodc.org/documents/data-and-analysis/statistics/crime/ICCS/ICCS_English_2016_web.pdf. Acesso em: 21 jun. 2026.

WOOLDRIDGE, Jeffrey M. **Introdução à econometria**: uma abordagem moderna. 4. ed. São Paulo: Cengage Learning, 2012.

ZHANG, Junsen. The effect of welfare programs on criminal behavior: a theoretical and empirical analysis. **Economic Inquiry**, Western Economic Association International, v. 35, n. 1, p. 120-137. 1997.