

UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE ESTUDOS SOCIAIS
DEPARTAMENTO DE ECONOMIA E ANÁLISE

ANÁLISE DE DETERMINANTES SOCIOECONÔMICOS DA
CRIMINALIDADE NO BRASIL

VITÓRIA MARIA DE SOUZA MAFRA

MANAUS
2023

VITORIA MARIA DE SOUZA MAFRA

ANÁLISE DE DETERMINANTES SOCIOECONÔMICOS DA
CRIMINALIDADE NO BRASIL

Monografia apresentada à Faculdade de Estudos Sociais da Universidade Federal do Amazonas como requisito obrigatório à obtenção do título de Bacharel em Ciências Econômicas.

Orientador: Prof. Me. Geasi Morais

MANAUS
2023

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

M187a Mafra, Vitória Maria de Souza
Análise de determinantes socioeconômicos da criminalidade no
Brasil / Vitória Maria de Souza Mafra . 2023
39 f.: il. color; 31 cm.

Orientador: Geasi Morais
TCC de Graduação (Ciências Econômicas) - Universidade
Federal do Amazonas.

1. Criminalidade. 2. Homicídios. 3. Variáveis socioeconômicas. 4.
Brasil. I. Morais, Geasi. II. Universidade Federal do Amazonas III.
Título

RESUMO

Entende-se que o fenômeno da criminalidade pode gerar problemas intangíveis para a economia e a sociedade, à vista disto, este trabalho buscou investigar o impacto que as variáveis socioeconômicas exercem no indicador de criminalidade no Brasil para os censos de 1991, 2000 e 2010 sob a perspectiva da economia do crime. Dessa forma, foi utilizada a matriz de correlação e para a escolha modelo entre Pooled, Efeitos Fixos ou Efeitos Aleatórios foram realizados os testes de Chow, Breush-Pagan e Hausman. Foram coletados dados dos 27 estados do Brasil. Através da estimação do modelo econométrico de dados em painel com Efeitos Fixos, os resultados indicam que há uma relação significativa entre a taxa de analfabetismo, a expectativa de vida e o Índice de Gini e a taxa de homicídios, essas variáveis socioeconômicas afetam diretamente a criminalidade nos estados brasileiros.

Palavras-chave: criminalidade, homicídios, variáveis socioeconômicas.

ABSTRACT

Considering that the phenomenon of criminality can lead to intangible problems for the economy and society, this work sought to investigate the impact that socioeconomic variables exert on the indicator of crime in Brazil for the 1991, 2000 and 2010 censuses from the perspective of the crime economy. Thus, the use of the correlation matrix and for the model choice between pooled, fixed effects or random effects were performed the tests of Chow, Breush-Pagan and Hausman. Data from the 27 states of Brazil were collected. Through the estimation of the panel data economy with fixed effects, the results indicate that there is a significant relationship between the illiteracy rate, the life expectancy and the gini index and the homicide rate, these socioeconomic variables directly affect crime in the Brazilian states.

KEYWORDS: Criminality, Homicides, Socioeconomic Variables.

LISTA DE TABELAS E QUADROS

Quadro 1 - Quadro metodológico, segundo a organização das variáveis, definição e fonte dos dados.	28
Tabela 1 - Tabela de Estatística Descritiva	30
Tabela 2 - Matriz de Correlação	31
Tabela 3 - Resultados do Modelo de dados em Painel com Efeitos Fixos	33

LISTA DE ABREVIATURAS E SIGLAS

ALF - Taxa de Analfabetismo de 15 anos ou mais

BP - Breusch-Pagan

BPG - Breusch-Pagan-Godfrey

DATASUS - Departamento de Informática do SUS

EA - Efeitos Aleatórios

EF - Efeitos Fixos

EXP - Expectativa de vida

FBSP - Fórum Brasileiro de Segurança Pública

GINI - Índice de Gini relativo

GRETSL - *Gnu Regression, Econometrics and Time-series Library*

HOM - Taxa de Homicídios por 100 mil habitantes

IPEA - Instituto de Pesquisa Econômica Aplicada

LM - Multiplicador de Lagrange

MQG - Mínimos Quadrados Generalizados

MQO - Mínimos Quadrados Ordinários

POP - População relativa

RIPSA - Rede Interagencial de Informações para a Saúde

SQE - Soma dos Quadrados Explicados

SQR - Soma dos Quadrados dos Resíduos

UNESCO - Organização das Nações Unidas para a Educação, a Ciência e a Cultura

SUMÁRIO

1.	INTRODUÇÃO.....	9
2.	REVISÃO DA LITERATURA.....	12
2.1	Premissas da Teoria Econômica do Crime.....	12
2.2	Criminalidade relacionada aos fatores socioeconômicos.....	14
3.	METODOLOGIA.....	19
3.1.	Matriz de correlação.....	19
3.2.	Dados em painel.....	20
3.2.1.	Modelo Pooled.....	21
3.2.2.	Modelo de Dados em Painel com Efeitos Fixos.....	22
3.2.3.	Modelo de Dados em Painel com Efeitos aleatórios.....	22
3.3.	Teste para escolha do modelo.....	23
3.3.1.	Teste de Chow.....	23
3.3.2.	Teste de Breusch-Pagan.....	24
3.3.3.	Teste de Hausman.....	25
3.4.	Fonte e Tratamento dos dados.....	25
4.	ANÁLISE E INTERPRETAÇÃO DOS RESULTADOS.....	30
4.1	Análise das Estatísticas Descritivas.....	30
4.2	Análise da Matriz de Correlação.....	31
4.3	Análise dos Testes.....	32
4.4	Análise dos Dados em Painel.....	33
5.	CONSIDERAÇÕES FINAIS.....	36
	REFERÊNCIAS.....	38

1. INTRODUÇÃO

A criminalidade está presente em todas as sociedades e civilizações, em todos os lugares ao redor do mundo, é por conta desse aspecto, em particular, que se constata uma forte influência dos fatores econômicos nos fenômenos da criminalidade. O crime é um fenômeno social complexo, pois, gera custos significativamente altos para toda a sociedade, reduz o estoque de capital humano, reduz a qualidade de vida, reduz a atividade turística e, ainda, pode resultar na perda de atratividade de novos investimentos produtivos.

Nas últimas duas décadas, a violência e o crime aumentaram significativamente em todo o mundo. Alguns estudiosos tratam este cenário como um fenômeno global que está se tornando cada vez mais notório em diferentes partes do mundo. Nos últimos anos, guerras interestatais, guerras civis, ataques terroristas e crimes urbanos levaram a um aumento no número de mortes por causas externas.

Desde a década de 1980, a criminalidade vem crescendo no Brasil, tornando-se um dos mais violentos do mundo. De acordo com as Informações sobre Mortalidade do Ministério da Saúde, DATASUS, em 2017, o Brasil teve uma taxa de 30,7 homicídios dolosos por 100.000 habitantes, foi o ano com número recorde de homicídios.

Diante deste cenário, entende-se que a criminalidade é um dos maiores problemas enfrentados pelo país. A violência é um problema social, econômico e político, e está diretamente ligado às condições econômicas, afetando assim o desenvolvimento potencial das nações (FAJNZYLBER E ARAÚJO JR 2001).

A criminalidade também pode gerar custos a um país para garantir uma maior segurança para a melhoria do bem estar social. Com isso, a participação ativa das esferas governamentais é necessária para orientação das políticas de segurança pública. As estimativas indicam que a diminuição de estoque de capital humano resultante de homicídios é de quase 2,25 bilhões de dólares para o detrimento do Brasil para 2001, os custos do crime para a sociedade podem ser consideravelmente altos.

No Brasil, a violência é mais visível na forma de crimes urbanos, ou seja, crimes contra o patrimônio, como roubo, furto e sequestro, e crimes contra a pessoa, como o homicídio.

Nos últimos anos, houve a redução da taxa de homicídios no Brasil, contudo, isso não ocorre de forma homogênea para todas as unidades federativas, inclusive, de acordo com o Anuário Brasileiro de Segurança Pública, publicado em 2022 pelo Fórum Brasileiro de Segurança Pública (FBSP), em 2021, houve uma de redução de 6,5% da taxa de homicídios

em comparação ao ano anterior, porém, houve um aumento significativo no Amapá com 53,8%, na Bahia com 44,9% e no Amazonas com 39,1%. É por isso que é importante compreender o comportamento da criminalidade para uma análise dos estados brasileiros, pois os estados apresentam resultados desiguais no que tange a criminalidade.

Ainda em seu anuário, a FBSP demonstra que, apesar de o Brasil apresentar uma tendência de queda no número de mortes violentas, o país ainda segue no ranking mundial de homicídios, isso significa que o país ainda ocupa uma posição alarmante no *ranking* dos países mais violentos do mundo. Apenas os países Jamaica, Honduras, África do Sul, México, Santa Lucia, Belize e Colômbia têm taxas respectivamente maiores do que a brasileira, é possível notar que, todos os países, exceto a África do Sul, incluído o Brasil, estão na América Latina e no Caribe, o que significa que a região ainda é uma das mais perigosas do planeta.

Após sua redemocratização, o Estado brasileiro passou por drásticas mudanças econômicas e mudanças sociais que podem ter impactado a decisão de um indivíduo de cometer um crime ou não. Assim sendo, a disponibilidade de dados possibilitou aos pesquisadores dedicar alguns esforços explorando os principais determinantes do crime no Brasil. Além do mais, não é novidade que o crime e a violência tornaram-se questões relacionadas à saúde pública, bem-estar social, economia, justiça e governança.

Partindo-se das questões supracitadas, surge o seguinte problema: Os fatores socioeconômicos exercem influencia nos níveis de criminalidade do país? Neste contexto, o objetivo geral apoia-se em investigar o impacto que as variáveis socioeconômicas exercem no indicador de criminalidade no Brasil para os censos de 1991, 2000 e 2010 sob a perspectiva da economia do crime. Mais especificamente, propõe-se: a) identificar os fatores socioeconômicos que influenciam diretamente nas taxas de criminalidade no Brasil. A vista disso, a premissa norteadora deste trabalho é voltado para uma análise de dados secundários em que se possa verificar se os fatores socioeconômicos desempenham uma influência direta no comportamento da criminalidade no país, apresentando algumas variáveis socioeconômicas e o impacto que elas exercem na taxa e homicídios que é usada como o indicador da criminalidade.

A presente pesquisa efetuou-se em razão do alto nível da criminalidade no país, que por sua vez, afeta o bem-estar de toda a população e pode gerar custos intangíveis à sociedade. Assim, este estudo empenha-se em contribuir para promover discussões sobre os fatores socioeconômicos e seus efeitos relacionados à criminalidade. Vale ressaltar que não há muitos estudos econométricos perante o viés da economia do crime voltado para uma análise

mais aprofundada no que tange o Brasil, portanto, esta pesquisa é importante para superar esta lacuna, servindo como um incentivo para outros estudos com uma nova perspectiva para esta área. É um estudo importante também pelo fato de se fazer uma análise de caráter econométrica sob a óptica da Economia do Crime, pois, a partir do uso da econometria pode-se entender a relação entre os indicadores socioeconômicos e a criminalidade através da aplicação de modelos de regressão, sendo possível obter resultados concretos na pesquisa.

Como modo metodológico, para se chegar aos objetivos propostos, foram coletados dados de 27 estados brasileiros para os anos de 1991, 2000 e 2010. O uso do modelo econométrico de dados em painel, utiliza para análise as variáveis socioeconômicas, como a população relativa, a taxa de analfabetismo, a expectativa de vida e o índice de Gini. Para o indicador de criminalidade, considera-se a variável sendo a taxa de homicídios, sendo assim, verifica-se como esta variável impacta as outras variáveis mencionadas anteriormente.

Quanto a estrutura, o trabalho está dividido em cinco seções. Além desta introdução tem-se, na seguinte ordem: a revisão da literatura, a metodologia, a análise e interpretação dos os resultados e as considerações finais.

2. REVISÃO DA LITERATURA

2.1 Premissas da Teoria Econômica do Crime

Diversos fatores econômicos e sociais podem influenciar o poder de decisão do indivíduo que tende a cometer atos ilícitos, dentre esses estimuladores, é possível citar a pobreza, a fome e a miséria, educação e analfabetismo, a política, o desemprego, guerras, urbanização e densidade demográfica e muitos outros que afetam todo o bem estar de uma sociedade. Com base nisso, Gary Becker se preocupou em analisar o poder de decisão que leva um indivíduo a delinquência, em seu artigo “*Crime and Punishment: An Economic Approach*” do economista Gary Becker, explica que os indivíduos que cometem crimes agem de maneira racional, pois, ponderam os benefícios e os custos ao tomar essa decisão, ou seja, quando analisam que cometer essa ação ilícita trará um resultado vantajoso, um resultado em que o benefício é maior que o custo, esses indivíduos estão agindo de forma racional, do ponto de vista econômico (BECKER, 1968). É o *trade off*¹ em cometer um crime ou não, ou seja, o indivíduo pondera o que pode ganhar e perder ao praticar um crime, leva-se em consideração que os custos em questão é, a probabilidade de ser pego, o tempo para planejar e executar o crime e o custo moral em ser pego. Em tese, o estudo de Becker constatou que uma das motivações para um indivíduo cometer um crime é porque o ganho financeiro poderá ser maior que o risco em praticar este ato ilegal, após a publicação de seu artigo no ano de 1968, a teoria econômica do crime teve um grande destaque, este artigo foi um marco no estudo sobre as causas da criminalidade.

Posteriormente, baseado nas ideias de Becker, Ehrlich (1973) leva-nos a compreender a relação entre o crime e a distribuição do tempo dos criminosos entre setores econômicos legais e ilegais, um estudo voltado para os Estados Unidos para três décadas, 1940, 1950 e 1960. O autor inclui em seu artigo as variáveis de rendimentos das atividades legais e ilegais, e o desemprego, os resultados demonstram que, a elasticidade do crime em relação ao valor esperado da punição do infrator são, respectivamente, -0,5 e -0,3, confirma então as ideias iniciais de Becker, além disso, os anos de estudo de uma população adulta e os crimes contra propriedade apresentam uma relação positiva, quando as taxas de desemprego e a

¹ Decisão que consiste na escolha de uma opção em detrimento de outra. Refere-se à análise daquilo que se perde em comparação com o que se ganha ao fazer determinada escolha.

desigualdade de renda caem, a criminalidade reduz, apesar da renda não apresentar um efeito significativo.

Em contrapartida, Block e Heinecke (1975), estudaram uma nova abordagem voltada para aspectos éticos e psicológicos que interfere na decisão de praticar uma atividade legal ou ilegal, diante dessa ideia, os autores consideram que as descobertas de Becker e Ehrlich sobre as oportunidades de lucro no mercado legal são válidas somente se as atividades legais e ilegais tiverem valor monetário equivalente e forem independentes da quantidade de riqueza. Eles concluem que os indivíduos avaliam suas escolhas em termos de tempo necessário para gerar um determinado nível de riqueza.

Cohen e Felson (1979), em suas pesquisas, foram um dos primeiros a tratar a criminalidade do ponto de vista da vitimização e não do infrator. Os autores se concentraram em pesquisar o comportamento das vítimas e de que forma a vida cotidiana das vítimas afetavam a probabilidade para o crime ocorrer. Os resultados da pesquisa sugerem que a criminalidade é entendida como um fator exógeno, pois, quando há uma melhoria dos indicadores sociais, as taxas de criminalidade aumentam, para eles, o crime é causado pela riqueza, não pela pobreza, uma melhor condição de vida do criminoso gera uma melhor condição para se cometer um crime.

Freeman (1982) buscou analisar a criminalidade relacionada ao mercado de trabalho, em termos de pesquisa de séries temporais, ele não encontrou consenso sobre o tópico com problemas de colinearidade, no que se refere as análises longitudinais com dados agregados regionalmente, mostram-se mais consistentes, representando uma relação positiva entre crime e desemprego.

Entorf e Spengler (2000) realizaram um estudo em painel para a Alemanha sobre a taxa de criminalidade relacionada ao PIB per capita de cada estado em relação ao geral nacional, a taxa de desemprego e a desigualdade de renda, para a demografia foi considerada o percentual de estrangeiros em relação à população, percentual de jovens entre 15 e 24 anos da população geral e a densidade demográfica e a taxa de esclarecimento do crime. Os resultados mostraram uma correlação negativa significativa da variável taxa de esclarecimento de crimes sendo considerada para a pesquisa a variável de dissuasão.

Diante disso, a economia do crime é uma área que estuda a criminalidade ligada a fatores econômicos, parte da ideia de que a tomada de decisão de se cometer um crime é algo racional que, por sua vez, interfere as esferas econômicas, tais como, as variáveis socioeconômicas. Este campo de estudos, que atualmente é objeto de estudo de economistas, é imprescindível para um país como o Brasil.

2.2 Criminalidade relacionada aos fatores socioeconômicos

Para Kliksberg (2002 apud SILVA 2018) o aumento da criminalidade na América Latina ocorre em razão do desemprego juvenil, educação e a articulação familiar, onde as crianças e os jovens não têm expectativas para um futuro melhor e que os leva a praticar atos ilícitos, pois não há uma estrutura familiar, não há uma base educacional, assim como, o difícil ingresso ao mercado de trabalho.

O papel da família é também é importante para a socialização do indivíduo, pois é no núcleo familiar que o crescimento da criança encontrará as condições necessárias para o seu desenvolvimento e aprendizagem social e valores de vida na sociedade e sua cultura. A falta de articulação familiar pode dificultar o processo de formação da criança que não encontrará condições para seu desenvolvimento, sem uma projeção para um futuro melhor, a tendências a entrar para meios ilícitos é maior.

É claro que há inúmeros fatores ligados ao aumento da criminalidade, entretanto, a falta de uma qualidade de vida está inteiramente ligada a este cenário. Em vista disso, Santiago (2016) acredita que quando há condições desfavoráveis de moradia, incidência de analfabetismo e desemprego, há uma grande chance para que a criminalidade se instale. A criminalidade gera custos a toda sociedade, diante disso, a sensação de insegurança da sociedade faz com que as esferas governamentais invistam em segurança pública para combater ou controlar a criminalidade e a violência, e assim, proporcionar o bem-estar social.

Pinto et al. (2018) afirma que o investimento em segurança pública é a política mais utilizada para reduzir a crescente criminalidade no Brasil, contudo, as medidas para combater a criminalidade não geram resultado imediato, é preciso que as instituições públicas tenham uma integração entre os órgãos, essa integração não seria benefício apenas para as instituições de segurança pública como também para todas as demais instituições públicas que tem por fim melhorar a sua eficiência e o bem estar da sociedade. Os autores chegaram a essa conclusão ao analisar a taxa de homicídios por cem mil habitantes para explicar a criminalidade e variáveis explicativas que representam as condições socioeconômicas, demográficas e do sistema de segurança pública e defesa social e assistência social, estimando os determinantes do nível de crimes de homicídios nos Estados brasileiros.

Junior et al. (2017) investigam o impacto dos investimentos públicos, principalmente referente a segurança pública e assistência social, bem como os fatores socioeconômicos, sobre a criminalidade nas 27 unidades federativas do Brasil para os anos de 2004 a 2012. A

partir do modelo econométrico de dados em painel com efeitos aleatórios, foi observado que a variável desigualdade de renda, medida pelo coeficiente de Gini, é a que mais contribuiu para o aumento da criminalidade, ou seja, em seu resultado constatou ser uma variável estatisticamente significativa para o modelo utilizado, é compreendido que os indivíduos agem de forma racional, comparando os ganhos e perdas possíveis, e estas variáveis servem de incentivos na opção do indivíduo pelo crime. No que se refere a densidade demográfica, foram estatisticamente significativos e positivos, isto significa que esta afeta positivamente a criminalidade, estados onde existe um número maior de habitantes por km², maior será o nível de crimes.

Em cenários de altos índices de analfabetismo, altas taxas de desemprego, maiores densidades demográficas e maior desigualdade social e concentrações de riquezas, são grandes fatores de estímulos das ocorrências criminais, diminuindo o custo de oportunidade do indivíduo para entrar em uma atividade ilícita.

Silva (2018), por sua vez, analisou a criminalidade em Alagoas por meio de variáveis socioeconômicas, sendo a taxa de homicídios por cem mil habitantes como determinante da criminalidade e variáveis socioeconômicas como as variáveis explicativas, concluiu que as políticas voltadas à educação e qualificação profissional junto com a urbanização e desenvolvimento local podem contribuir na luta contra a criminalidade, com isso, a população terá mais chances de entrar no mercado de trabalho, além de contribuir para uma melhor qualidade de vida, esta é uma forma de desincentivo para adentrar no mercado ilícito.

Já Cruz e Araújo (2012) evidenciam em sua pesquisa que o impacto que as variáveis socioeconômicas exercem sobre a criminalidade pode ser significativo, acredita-se que políticas sociais de combate à pobreza e um sistema preventivo e repressivo de segurança pública eficaz trazem resultados satisfatório à sociedade e ao desenvolvimento de uma região. Esta pesquisa teve enfoque para a cidade de Palmas, então, o autor chegou a esse resultado mediante a uma análise econométrica da criminalidade diante da violência urbana da região, fez-se modelos de regressão linear múltipla, inserindo os determinantes da criminalidade como variável dependente e os indicadores socioeconômicos como variáveis independentes.

Em contrapartida, Mariano (2010) reitera que a criminalidade afeta o desenvolvimento econômico, pois os investimentos que deveriam ser utilizados para o bem-estar da sociedade são destinados ao combate da criminalidade. Em seu estudo empírico, buscou verificar o impacto de variáveis socioeconômicas no nível de crimes contra o patrimônio nos municípios de São Paulo, ele considerou as seguintes variáveis socioeconômicas: PIB per capita, emprego formal, densidade demográfica, escolaridade, concentração de renda e o número de policiais.

Contudo, a conclusão do resultado econométrico de Mínimos Quadrados Ordinários foi que apenas o PIB per capita e o número de policiais não foram estatisticamente significativos para explicar o modelo.

Jales (2020), por sua vez, ao analisar os determinantes socioeconômicos da taxa de homicídios no Rio Grande do Norte (RN), diante de uma análise econométrica, constatou que o Índice de Gini PIB per capita, taxa de analfabetismo e taxa de desemprego, apresentaram relação positiva com a taxa de homicídios, além de serem variáveis determinantes para o aumento da criminalidade no estado. Outra questão a ser observada é a relação estatisticamente significativa do valor do benefício do programa Bolsa Família e a taxa de homicídios, isto é, quando o Valor dos benefícios do programa Bolsa Família aumenta, a taxa de homicídios também cresce, reafirma a ideia que é necessária que Estado e Municípios se mobilizem para desenvolverem políticas eficazes para combater o crime.

Santos (2008), ao analisar os aspectos socioeconômicos dos indivíduos, aborda o quanto os indicadores de desigualdade de renda podem influenciar na vitimização das pessoas. A partir da estimação dos modelos de regressão negativo binomial e zeros inflados, os resultados indicam que morar em área urbana torna a pessoa mais vulnerável a qualquer tipo de crime de roubo ou furto e quanto maior a renda concentrada na área, maior a desigualdade e, portanto, maior a probabilidade de uma pessoa ser vitimada mais de uma vez. A autora descreve também que, os anos de estudo estão positivamente associados à renda de uma pessoa, o que também afeta a frequência de vitimização, porque quanto mais você estudar, mais chance terá de encontrar um bom emprego e aumentar sua renda, essas pessoas se tornam alvos potenciais para algum tipo de crime.

Com relação ao estudo de Oliveira (2008), ao realizar uma análise espacial da criminalidade no Estado do Rio Grande do Sul a partir do modelo econométrico espacial para os determinantes da criminalidade, é considerado que a desigualdade de renda e das aglomerações urbanas são fatores que aumentam a criminalidade em cidades. Inclusive, em seu estudo destaca-se o papel da família e da escola na explicação da criminalidade. Conclui-se que problemas na estrutura familiar e a ineficiência do ensino no Estado afetam positivamente a criminalidade. O resultado positivo da educação indica que o papel da escola na redução da criminalidade não foi o esperado, com isso, sugere que isso ocorre pelo fato de que a escola pode não estar cumprindo com seu papel de inserir o indivíduo no mercado de trabalho e de passar valores morais aos mesmos.

Santos (2019) realizou um estudo para verificar a relação existente entre o nível de renda e o aumento da criminalidade em Minas Gerais nos anos de 2000 e 2010 mediante a

análise da Matriz de Dados Espaciais, pela Análise Exploratória de Dados Espaciais (AEDE) e também com a estimação de modelos econométricos espaciais. A partir da sua análise, é possível compreender que o nível de renda se relaciona de forma positiva ao nível de criminalidade dos municípios quando estão associados a outra variável de controle, da mesma forma que o analfabetismo se demonstra como estatisticamente significativo para índices de violência nos municípios mineiros no período estudado.

Mendonça (2001) obteve um modelo econométrico da criminalidade para os estados brasileiros no período 1985 a 1995, a partir dos resultados, acredita-se que, especificamente para o caso brasileiro, quanto se leva em consideração questões ligadas a heterogeneidade existente, o índice de Gini é um fator que acentua a criminalidade.

Cabe ressaltar que poucos trabalhos examinaram a relação entre desigualdade de renda e crime, pois, os dados criminais são frequentemente calculados usando metodologias diferentes ao longo do tempo e do espaço, ao ponto de evidenciar a heterogeneidade e erro de medição, especialmente para os casos de países menos desenvolvidos, leva pesquisadores a usar dados sobre tudo de registros de óbitos para homicídios para estudar crimes.

Há limitações quanto o uso da teoria econômica para análise de crimes violentos como homicídios, porém, ainda são os que mais incomodam a sociedade e pesquisadores. Seillier (2010) analisa a relação entre quatro diferentes taxas de criminalidade e algumas variáveis socioeconômicas e demográficas em 27 unidades federativas brasileiras no período 2001-2005, a autora considera os crimes separados entre crimes patrimoniais violentos e não violentos, crimes letais e crimes envolvendo drogas. A partir disso, constata-se que para as taxas de crimes letais a única variável que se mostrou estatisticamente significativa tanto no estimador de Efeitos Fixos quanto nas regressões com estimador SUR foi a esperança de vida ao nascer, em que a taxa de crimes letais e a esperança de vida nos estados apresentam uma relação negativa.

Segundo Glaeser e Sarcedote (1999), regiões mais urbanizadas sofrem mais com a violência, áreas mais densamente povoadas tendem a apresentar maior interação entre os indivíduos o que acaba por reduzir os custos de execução e planejamento das atividades criminosas, possibilita uma maior troca de informações entre os grupos de criminosos e os potenciais criminosos. Ainda, conforme Cano e Santos (2000), a probabilidade de fracasso do crime é baixa por conta de um maior anonimato dos indivíduos nas áreas urbanas, diminuindo a capacidade de controle social e aumenta a impunidade.

É perceptível que, diante de determinados estudos da criminalidade relacionada aos determinantes socioeconômicos, as variáveis socioeconômicas são significativas para o

modelo, em contrapartida, em outros estudos, as variáveis socioeconômicas não se mostram significativas para explicar o fenômeno da criminalidade.

3. METODOLOGIA

Neste tópico, é apresentada a metodologia aplicada para a elaboração deste estudo e as definições de métodos e modelos econométricos, sendo explanada respectivamente, a análise de correlação, modelos de Dados em Painel, testes para a escolha do modelo e, por fim, a fonte e o tratamento de dados utilizados.

3.1. Matriz de correlação

A definição de correlação, para Larson e Farber (2015, p. 438), consiste em uma “relação entre duas variáveis”, em outras palavras, a correlação explica a intensidade em que as variáveis se relacionam, por conta disso, para sua análise é importante examinar e identificar se duas ou mais variáveis são independentes ou variam juntas. Cabe ressaltar que há três tipos de correlação, a correlação simples está relacionada ao grau entre duas variáveis, a correlação múltipla é o grau de relação entre a variável dependente e duas ou mais variáveis independentes, e a correlação parcial que se refere à regressão múltipla, porém, neste caso, em específico, acontece quando é identificada a relação pura entre duas variáveis, depois que é eliminada estatisticamente a influencia de outras variáveis independentes. É possível interpretar a correlação a partir do diagrama de dispersão, contudo, é uma análise mais visual, então, outra possibilidade é calcular o coeficiente de correlação, pois, este cálculo dispõe de uma compreensão mais detalhada para medir a força de uma correlação entre variáveis. De acordo com Gujarati e Porter (2011, p.98), o coeficiente de correlação pode ser calculado tanto por:

$$r = \pm\sqrt{r^2} \quad (1)$$

Outra forma para o coeficiente de correlação amostral, conforme sua de definição é:

$$r = \frac{\sum x_i y_i}{\sqrt{(\sum x_i^2)(\sum y_i^2)}} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2][n \sum y_i^2 - (\sum y_i)^2]}} \quad (1.2)$$

O coeficiente de determinação r^2 (duas variáveis) ou R^2 (regressão múltipla), se refere à qualidade do ajustamento da linha de regressão mediante a um conjunto de dados, espera-se que os resíduos \hat{u}_i em torno da linha de regressão sejam os menores possíveis.

Através da seguinte relação $0 \leq r^2 \leq 1$, pode-se concluir que, se a correlação for igual a 1, há uma correlação positiva perfeita, se a correlação for igual a zero, não há qualquer relação entre regressando e regressor. Portanto, a correlação mais próxima de 1 sugere que há uma boa qualidade de ajustamento da reta de regressão.

3.2. Dados em painel

O método econométrico de dados em painel corresponde a uma série de tempo para cada membro do corte transversal do conjunto de dados, em outras palavras, é a união dos dados de séries temporais e os dados de corte, o seu diferencial é que os dados em painel abrangem uma dimensão temporal e outra espacial, uma vez que, é analisada a mesma unidade de corte transversal ao longo do tempo.

Com relação às observações, há dois tipos, painel balanceado e painel desbalanceado, o primeiro se refere à mesma quantidade de observações para todas as unidades de análise, já o segundo tipo é usado quando a quantidade de observações não é o mesmo para as unidades de análise, é desbalanceado por conta da falta de dados por observação ao longo do tempo, com isso, deve-se ter cautela quanto ao painel desbalanceado, pois pode gerar viés na análise e, ainda, diminuir a eficiência do estimador.

Com relação à dimensão do painel, Beck (2001) explica que há duas especificações, o painel curto, onde o número de períodos no tempo é menor que número de observações na amostra ($T < N$) e o painel longo no qual o número de períodos no tempo é maior que número de observações na amostra ($T > N$).

A regressão geral para dados em painel é reescrita da seguinte forma:

$$y_{it} = \alpha_i + \beta_1 x_{1,it} + \dots + \beta_k x_{k,it} + e_{it} \quad t = 1, \dots, T. \quad (2)$$

Sendo que i são as unidades observacionais e t os períodos de tempo. O elemento, y_{it} , corresponde a variável dependente, é a variável de resposta da unidade i , no período t . O termo, $x_{k,it}$, refere-se a k -ésima variável explicativa para a unidade i , no período t . O elemento, e_{it} , é o termo de erro e α_i sendo o intercepto.

De acordo com Hsiao (1986), a regressão de dados em painel dispõe de algumas vantagens, dentre elas, a possibilidade de ocorrer heterogeneidade das unidades *crosssection* considerando variáveis individuais específicas, ou seja, controla a heterogeneidade nos indivíduos, outra questão é que com a combinação de dados ao longo do tempo e observações

de corte transversal, há uma maior amostra, então, proporciona estimadores mais eficientes, dessa forma, também gera maior informação, maior variabilidade e menor colinearidade entre variáveis, propicia maior controle dos efeitos das variáveis não observadas. Além do mais, os dados em painel são importantes para investigar a dinâmica da mudança, isto é, como as variáveis, ou as relações entre elas, mudam com o tempo.

Para a estimação do modelo de regressão seguindo o método de dados em painel, há três possibilidades clássicas, sendo elas, respectivamente, modelo MQO (Mínimos Quadrados Ordinários) para dados empilhados (*pooled*), Efeitos Fixos (EF) e Efeitos Aleatórios (EA).

3.2.1. Modelo Pooled

O Modelo Pooled despreza a natureza de corte transversal e de séries temporais dos dados, não levam em consideração as dimensões de tempo e espaço, as observações são empilhadas obtendo uma regressão, isto significa que o intercepto e seus coeficientes angulares são constantes ao longo do tempo e no espaço. Cabe ressaltar que, as observações são tratadas como não correlacionadas para os indivíduos, incluindo erros homocedásticos para os indivíduos, ocasiona problemas para ocultar a heterogeneidade que possa existir entre as variáveis, inclusive, pode ser que o termo de erro possa estar correlacionado com alguns dos regressores incluídos no modelo, os coeficientes estimados na equação podem ser tendenciosos e inconsistentes.

Mesquita, Fernandes e Figueiredo Filho (2020), descrevem que o modelo é a união de todas as observações em uma mesma regressão, independente de diferenças na unidade i , as variáveis independentes abrangem todas as informações necessárias. Os autores consideram a seguinte expressão:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \quad (3)$$

Os componentes i e t , demonstram que as observações são indexadas por caso e tempo, Y é a variável dependente, α é o intercepto, β é a variação observada na variável dependente quando a variável independente, X , aumenta em uma unidade, e ε é a natureza estocástica do modelo. Entretanto, a equação (3), revela heterogeneidades individuais não modeladas, há apenas um intercepto para toda a população e o restante não observável é representado como ruído no fator de erro. Vale enfatizar que a independência das observações

está em risco, pois elas se repetem no tempo, a repetição da observação do mesmo sujeito pode resultar a heterogeneidade.

3.2.2. Modelo de Dados em Painel com Efeitos Fixos

O modelo de Efeitos Fixos considera que os coeficientes angulares são constantes e o intercepto varia entre os indivíduos, é a combinação das observações em que cada unidade de corte transversal tenha seu próprio intercepto, sem variar com o tempo. Concebe a heterogeneidade entre indivíduos, permite que cada um tenha seu próprio intercepto.

Seguindo a ideia da equação (3), quando se decompõe o tempo de erro ε_{it} , verifica-se a seguinte relação:

$$\varepsilon_{it} = \mu_i + v_{it} \quad (3.1)$$

Em que μ_i é o efeito fixo (heterogeneidade não-observada) e o v_{it} é efeito variável, este termo varia para as unidades e o tempo. Logo, de forma resumida, pode-se considerar a seguinte equação para demonstrar os efeitos fixos (WOOLDRIDGE, 2010):

$$y_{it} = \alpha_i + \beta_1 x_{1,it} + \dots + \beta_k x_{k,it} + \mu_i \quad t = 1, \dots, T \quad (4)$$

Neste caso, é possível notar que a variável α_i não apresenta o subscrito t, pois, não muda ao longo do tempo, é fixa ao longo do tempo.

3.2.3. Modelo de Dados em Painel com Efeitos aleatórios

No modelo de Efeitos Aleatórios pressupõe diferentes termos de intercepto para cada observação, também, esses interceptos são constantes ao longo do tempo, mas, diferentemente do modelo de Efeitos Fixos, na estimativa modelo de Efeitos Aleatórios, os interceptos de cada unidade de corte transversal são assumidos como surgindo a partir de uma intersecção comum. Vale destacar que é um estimador de MQG (Mínimos Quadrados Generalizados) que considera a correlação entre os erros de cada unidade.

Ainda, em conformidade com Wooldridge (2010), quando se trata de um modelo de modelo de Efeitos Aleatórios, pode-se admitir que o efeito não-observado α_i da equação (4) é não correlacionado com cada variável explicativa. Portanto, ao ponderar o termo de erro

composto tem-se $v_{it} = \alpha_i + \mu_i$. Quando o intercepto α_i é correlacionado com as variáveis explicativas em qualquer período de tempo, usa-se o modelo de Efeitos Fixos.

3.3. Teste para escolha do modelo

Para Torres-Reyna (2010), a análise de dados em painel prescinde da avaliação do modelo mais adequado. Para isto, foram realizados os testes de Chow, de Breusch-Pagan e o Teste de Hausman. O Teste de Chow é utilizado para decidir entre o modelo Pooled ou o modelo de Efeitos Fixos, pois, depende do resultado do teste de hipótese, ou seja, se houver a rejeição significativa da hipótese nula, de que todos os coeficientes são iguais à zero, então, admite-se a utilização do modelo de Efeitos Fixos, por outro lado, se for admitida a hipótese nula, o modelo mais adequado é o modelo Pooled. O Teste de Breusch-Pagan esboça a escolha entre o modelo Pooled ou o modelo de Efeitos Aleatórios, logo, a rejeição da hipótese nula indica que o modelo com Efeitos Aleatórios é mais significativo, em contrapartida, se for considerado a hipótese nula, o modelo mais adequado é o modelo Pooled. O Teste de Hausman é aplicado para a escolha entre o modelo de Efeitos Fixos ou o modelo de Efeitos Aleatórios, isto significa que, se for admitido a rejeição da hipótese nula, indica que o modelo de Efeitos Fixos será mais adequado, no entanto, se for verificado a hipótese nula, utiliza-se o modelo de Efeitos Aleatórios.

Por conseguinte, a rejeição e não rejeição da hipótese nula implica na utilização de outro modelo que está sendo testado para a hipótese alternativa.

3.3.1. Teste de Chow

O Teste de Chow é usado para se verificar a ocorrência de quebras estruturais, pode determinar se uma função de regressão múltipla se diferencia entre dois grupos e pode ser aplicado em dois períodos de tempo diferentes (WOOLDRIDGE, 2010). Para Gujarati e Porter (2011), essa mudança estrutural corresponde aos valores dos parâmetros do modelo que não se mantêm iguais durante todo o período de tempo, podem ocorrer por diferenças no intercepto ou no coeficiente angular, ou em ambos, esta mudança se verifica por meio de uma regressão de modelos de séries temporais. Em vista disso, o teste pressupõe que os termos de erro da regressão dos períodos A e B (equações 5 e 6) se distribuem normalmente com a mesma variância homocedástica: $u_{1t} \sim N(0, \sigma^2)$ e $u_{2t} \sim N(0, \sigma^2)$. Além disto, os termos de erro se distribuem de forma independente.

$$\text{Período A:} \quad Y_t = \lambda_1 + \lambda_2 X_1 + u_{1t} \quad (5)$$

$$\text{Período B:} \quad Y_t = \gamma_1 + \gamma_2 X_1 + u_{2t} \quad (6)$$

$$\text{Período A+B:} \quad Y_t = \alpha_1 + \alpha_2 X_1 + u_t \quad (7)$$

Quando é estimada a regressão (7), gera-se a Soma dos Quadrados dos Resíduos (SQR), atendendo aos graus de liberdade $n_1 + n_2 - k$ em que k são os números de parâmetros estimados. Quando é estimada a equação (5), gera-se a SQR_1 , com os graus liberdade sendo $n_1 - k$. Quando é estimada a equação (6), gera-se a SQR_2 , com os graus liberdade sendo $n_2 - k$. Ao somar SQR_1 e SQR_2 , obtêm-se SQR_{SR} que é a soma sem restrições dos quadrados dos resíduos, desta forma, os graus de liberdade seguem assim $n_1 + n_2 - 2k$. Conseqüentemente, para identificar a mudança durante o período analisado, têm-se:

$$F = \frac{(SQR_R - SQR_{SR})/k}{(SQR_{SR})/(n_1 + n_2 - 2k)} \sim F_{[k, (n_1 + n_2 - 2k)]} \quad (8)$$

Isto posto, se o F calculado for menor que o F tabelado ($F_{calc} < F_{tab}$) em determinado nível de significância ou valor p, não rejeita a hipótese nula de estabilidade dos parâmetros que se refere a ausência de mudança estrutural, então, pode-se utilizar a regressão (7). Contudo, se o F calculado for maior que o F tabelado ($F_{calc} > F_{tab}$), rejeita-se a hipótese nula de estabilidade dos parâmetros, assim, sugere que as regressões (5) e (6) são distintas.

3.3.2. Teste de Breusch-Pagan

Como Hoffman e Vieira (1987) explicam, o teste de Breusch-Pagan analisa a natureza da variância do termo de erro (σ_u^2), investiga a hipótese de homocedasticidade do termo de erro de uma regressão linear, admitem-se os resíduos sendo elevados ao quadrado.

Para explicar o Teste de Breusch-Pagan (BP) baseado no Multiplicador de Lagrange (LM), é importante antes abordar o teste de Breusch-Pagan-Godfrey (BPG). Para obtenção do teste, segundo Gujarati e Porter (2011), estima-se o a regressão (9) por Mínimos Quadrados Ordinários (MQO), então, é obtido os resíduos para as variáveis ($\hat{u}_1, \hat{u}_2, \dots, \hat{u}_n$), em seguida, é obtida a estimativa de máxima verossimilhança da variância do erro, depois que é construída variáveis p_i , será feito um novo modelo de regressão, a partir disto, é calculado a Soma dos Quadrados Explicados (SQE), pode-se concluir que, ao considerar os erros u_i com distribuição normal, há distribuição assintótica de qui-quadrado com $(m - 1)$ graus de

liberdade, baseado nisso, se SQE calculado for maior que o valor crítico X^2 , em determinado nível de significância, rejeita-se a hipótese de homocedasticidade.

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad (9)$$

Com relação ao teste LM de Breusch e Pagan, é possível analisar a hipótese de que não há efeitos aleatórios, nesse caso, admite os σ_u^2 da equação (10) sendo iguais à zero. Se a hipótese nula ($H_0 = \sigma_u^2 = 0$) em que LM tem distribuição qui-quadrado, com um grau de liberdade for aceita, significa que há uma variância constante e o modelo de Mínimos Quadrados Empilhados é preferível ao modelo de efeitos aleatórios.

$$var(w_{it}) = \sigma_\varepsilon^2 + \sigma_u^2 \quad (10)$$

3.3.3. Teste de Hausman

O Teste de Hausman é utilizado para decidir qual modelo é mais apropriado, trata-se da escolha dos modelos de Efeitos Fixos ou Efeitos Aleatórios. Conforme Greene (1993), consiste em um teste de especificação de erro para analisar se há correlação entre os termos idiossincráticos α_i e as variáveis explicativas, neste sentido é descritas as hipóteses a seguir:

- $H_0: \alpha_i \rightarrow$ São consistentes, ou seja, não há correlação com as variáveis explicativas.
- $H_1: \alpha_i \rightarrow$ Inconsistência, ou seja, há correlação entre as variáveis explicativas.

A aceitação da hipótese nula sugere que o modelo de Efeitos Aleatórios deve ser eficiente, deve ser utilizado. Em compensação, se houver rejeição da hipótese nula, quer dizer, aceitação da hipótese alternativa, o modelo de Efeitos Fixos deve ser utilizado.

3.4. Fonte e Tratamento dos dados

Esta pesquisa contém uma análise quantitativa, segundo Richardson (1999) trata-se de uma abordagem metodológica em que tudo pode ser quantificável, onde a quantificação é feita por meio de técnicas estatísticas, como percentual, média, desvio-padrão, até as técnicas mais complexas, como coeficiente de correlação, análise de regressão e assim por diante. Esta abordagem é utilizada para descrever e explicar populações, fenômenos e a verificação da existência de relação entre variáveis, logo, é possível mensurar as relações entre variáveis,

acerca da associação ou causa e efeito. Logo, conforme o autor, a pesquisa descritiva investiga a relação entre variáveis, procura explicar e interpretar um fenômeno.

Fundamentado nisso, os dados desta pesquisa foram elaborados em painel para análise da criminalidade nos 26 estados brasileiros e no Distrito Federal, para os anos censitários de 1991, 2000 e 2010.

Para tratamento dos dados, foi utilizado os dados em painel balanceado, pois, a quantidade de observações foi a mesma para todas as unidades de análise. Para o tamanho do painel, foi considerado o painel curto, em que, o número de observações no corte transversal foi maior que o número de períodos de tempo. Cabe destacar que, os dados foram empilhados em 81 observações dos 26 estados brasileiros e no Distrito Federal.

Os dados foram provenientes do Instituto de Pesquisa Econômica Aplicada (IPEA), banco de dados do IPEADATA, Rede Interagencial de Informações para a Saúde (RIPSA), todos disponíveis no Departamento de Informática do SUS (DATASUS).

Neste estudo os dados coletados foram analisados por meio do software *Gnu Regression, Econometrics and Time-series Library* (Gretl).

A variável dependente, como indicador da criminalidade, é a taxa de homicídios por 100 mil habitantes. Essa variável foi escolhida pelo fato de ser frequentemente utilizada para comparar os níveis de criminalidade entre unidades geográficas, e foi escolhida, também, por abranger o somatório do número de homicídio doloso, roubo seguido de morte (latrocínio), lesão corporal seguida de morte e mortes por intervenção policial, cuja frequência absoluta anual desses eventos é dividida pela população residente e multiplicados por 100 mil para se obter a taxa bruta de mortalidade (TBM) para Unidade da Federação, em cada ano observados, os dados foram extraído do IPEADATA.

As variáveis explicativas selecionadas foram as populações relativa, analfabetos de 15 anos ou mais, bem como a expectativa de vida e o Índice de Gini.

Para a escolha da variável população, foi considerada a variável população relativa que se refere à população urbana sobre a população residente total. Conforme os estudos de Zanutelli et al. (2011), essa escolha se deu pelo fato de que a expansão urbana, na perspectiva da urbanização brasileira, e o aumento da taxa de homicídios, constituem entre si entre si tendências temporais semelhantes. Os dados para essa variável foram auferidos pelo IPEADATA.

Foi escolhida a variável analfabetos de 15 anos ou mais, pois, Pinto et al. (2018) sugere em seu estudo, uma relação inversa do nível de escolaridade da população com a

criminalidade, demonstra que a taxa de crimes de homicídios sofre um efeito redutor quando aumentamos a escolaridade das pessoas.

A escolha da expectativa de vida, de acordo com a literatura, se deu pelo fato de que em determinado ano a taxa de homicídios por 100 mil habitantes no Brasil aumentou e a prevalência dos homicídios observada nesse ano foi responsável pela perda de expectativa de vida ao nascer (CERQUEIRA e SOARES, 2011). A expectativa de vida pode ter uma relação direta com a taxa de homicídios.

Para a escolha da variável Índice de Gini, foi utilizado o Coeficiente de Gini que serve para medir o grau de concentração de renda de determinado grupo populacional. Neste estudo este indicador foi selecionado do banco de dados da RIPASA para cada UF.

Para analisar a criminalidade nos estados brasileiros, estimou-se um modelo econométrico, por meio de uma análise de regressão múltipla, conforme Gujarati e Porter (2011, p.205), consistem em um modelo de regressão “em que a variável dependente, ou regressando, Y, depende de duas ou mais variáveis explicativas, ou regressoras”. Desta forma esta análise considerou como variável dependente a taxa bruta de homicídios, como indicador de criminalidade, ao passo que os indicadores socioeconômicos (população relativa - % de população residente, taxa de analfabetismo de população de 15 anos ou mais de idade, expectativa de vida ao nascer, Índice de Gini) como variáveis independentes ou explicativas. Portanto, o número de homicídio é uma função das demais variáveis.

Desta maneira será estimado o seguinte modelo de regressão:

$$\ln HOM(Y_{it}) = \beta_0 + \beta_1 POP_{it} + \beta_2 ALF_{it} + \beta_3 EXP_{it} + \beta_4 GINI_{it} + u_{it}$$

onde,

- $\ln HOM(Y_{it})$ = variável dependente logaritmizada que representa a taxa de homicídios por 100 mil habitantes no estado i , no ano t .
- β_0 = constante;
- $\beta_1, \beta_2, \beta_3$ e β_4 = parâmetros a serem estimados, são os coeficientes angulares;
- POP = variável independente que representa a população relativa;
- ALF = variável independente que representa pessoas analfabetas de 15 anos ou mais;
- EXP = variável independente que representa a expectativa de vida ao nascer;
- $GINI$ = variável independente que representa o Índice de Gini;

- u_{it} = representa o termo erro, com i sendo a i -ésima observação, no tempo t .

Vale destacar que a variável dependente escolhida para estimação do modelo está sendo apresentada na sua forma logaritmizada, obtendo um entendimento do termo de elasticidade para uma melhor interpretação dos resultados obtidos.

Quadro 1 - Quadro metodológico, segundo a organização das variáveis, definição e fonte dos dados.

Variável Dependente	Sigla	Definição	Fonte
Taxa de Homicídios por 100 mil habitantes	HOM	Óbito por causa externa ou não-natural, indiferente do tempo entre o evento lesivo e a morte propriamente, é categorizado como consequente de lesão provocada por violência (acidentes, homicídios, suicídios ou morte suspeita).	IPEADATA, 2022
Variáveis Independentes	Sigla	Definição	Fonte
População relativa	POP	As características gerais da população compreendem informações sobre tamanho e estrutura, por idade e sexo, cor ou raça, distribuição da população, densidade e urbanização. Inclui-se em população os grupos populacionais específicos, os quais compreendem as estatísticas agrupadas por segmentos específicos da população a partir dos grupos geracionais, étnico raciais ou regionais.	IPEADATA, 2017
Taxa de Analfabetismo de 15 anos ou mais	ALF	Percentual de pessoas de 15 ou mais anos de idade que não sabem ler nem escrever um bilhete simples.	IPEADATA, 2015
Expectativa de vida	EXP	Expectativa de anos de vida de uma pessoa nascida no ano de referência supondo que as taxas de mortalidade por idade estimadas para anos anteriores se mantivessem constantes nos anos posteriores.	IPEADATA, 2013

Índice de Gini	GINI	Mede o grau de desigualdade na distribuição da renda domiciliar per capita entre os indivíduos. Seu valor pode variar teoricamente desde 0, quando não há desigualdade (as rendas de todos os indivíduos têm o mesmo valor), até 1(ou 100), quando a desigualdade é máxima (apenas um indivíduo detém toda a renda da sociedade e a renda de todos os outros indivíduos é nula).	IPEADATA, 2016
----------------	------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------

Fonte: Elaboração própria.

4. ANÁLISE E INTERPRETAÇÃO DOS RESULTADOS

4.1 Análise das Estatísticas Descritivas

A tabela 1 apresenta as estatísticas descritivas das variáveis utilizadas para a pesquisa. Como mencionado anteriormente, a pesquisa foi realizada com as 27 unidades federativas do Brasil para os anos censitários de 1991, 2000 e 2010, que, por sua vez, foram empilhados em 81 observações para cada variável. De acordo com Reis (2002), a Estatística Descritiva é um método que serve para organizar, resumir e descrever os aspectos importantes das variáveis selecionadas ou comparar tais características entre dois ou mais conjuntos. Dentre as medidas de tendência central, foi considerada a média. Para as medidas de dispersão, foi considerado o desvio padrão, serve para quantificar a dispersão de um conjunto de dados, ele também é expresso na mesma unidade utilizada na média. O valor mínimo e máximo das variáveis são os valores extremos de um conjunto de dados, se refere ao menor e o maior valor.

Tabela 1 - Tabela de Estatística Descritiva

Variável	Média	Desvio Padrão	Mínimo	Máximo
HOM	25,15	12,93	4,29	64,57
POP	1,97	3,79	0,01	19,65
ALF	26,45	78,52	3,59	718,19
EXP	68,30	8,49	0,60	76,93
GINI	61,81	3,62	49,42	68,68

Fonte: Elaboração própria com base nos dados coletados do software Gretl.

Pode-se observar que para os anos 1991, 2000 e 2010, os 27 estados brasileiros apresentaram uma média da taxa de 25,15 homicídios, um valor consideravelmente baixo se considerar variações de 1 a 100, o desvio padrão de 12,93, indica que os valores amostrais são poucos distribuídos em torno da média, com base nisto, Alagoas foi o estado que apresentou maior número de homicídios no ano de 2010, sendo representado pelo valor máximo de 64,57 das 81 observações para os anos censitários analisados.

Em compensação, a menor média é da população relativa com 1,97, valor baixo. O valor máximo de 19,65 da amostra indica que o estado de São Paulo tem a maior população relativa entre os estados analisados, conforme o censo de 1991. Vale ressaltar ainda, que as

variáveis POP, EXP e GINI, apresentam desvios padrão pequenos, de respectivamente, 3,79, 8,49 e 3,62, revela que, os dados estão próximos da média ou do valor esperado, ou seja, os dados são mais homogêneos.

A média da variável ALF é de 26,45, sendo assim, é baixo o percentual de pessoas analfabetas de 15 anos ou mais para os anos e para as observações analisadas, apresenta o maior desvio padrão de 78,58, significa que os valores amostrais estão bem distribuídos em torno da média, outra questão importante é que o maior valor da amostra é 718,19, mostra que, dentre os estados do Brasil, o estado de Tocantins apresentou maior número de pessoas analfabetas de 15 anos ou mais no ano de 2010.

A maior média é da variável EXP, com 68,30, em contrapartida, seu valor mínimo de 0,60, indica que o estado de Tocantins obteve a menor expectativa de vida em comparação aos outros estados analisados, novamente no ano de 2010.

A variável GINI tem a média de 61,81, indica que Índice de Gini relativo se classifica os estados como desiguais, pois, quanto mais perto de 100 (índice relativo), maior é a desigualdade de renda entre a população. O valor máximo de 68,68 na análise descritiva indica que o estado de Alagoas apresentou uma maior desigualdade de renda entre os estados brasileiros de acordo com o censo de 2000. O valor mais baixo da amostra, com 49,42, sugere que, entre os estados analisados, o estado de Santa Catarina é mais igualitário referente à riqueza da população.

4.2 Análise da Matriz de Correlação

A Tabela 2 apresenta a matriz de correlação das variáveis independentes, mostra como as mesmas correlacionam entre si, mede a intensidade de relação entre as variáveis. As variáveis quando correlacionadas perfeitamente com elas mesmas o resultado é igual a 1.

Tabela 2 - Matriz de Correlação

	POP	ALF	EXP	GINI
POP	1			
ALF	-0,0741	1		
EXP	0,1390	-0,9376	1	
GINI	-0,2706	0,0523	-0,2460	1

Fonte: Elaboração própria com base nos dados coletados do software Gretl.

A tabela 2 demonstra que na segunda linha, o analfabetismo (ALF) e a população relativa (POP) têm uma correlação negativa muito fraca de $-0,0741$, o resultado indica que quanto maior é o analfabetismo, menor tende a ser a população relativa, é uma relação inversamente proporcional.

Na terceira linha, a expectativa de vida ao nascer (EXP) e a população relativa (POP) têm uma correlação positiva muito fraca de $0,1390$, pois, quando a expectativa de vida aumenta, a população relativa irá aumentar, porém, com um grau fraco. Na mesma linha, a variável EXP e a ALF têm uma correlação negativa muito forte de $-0,9376$, então, quando a expectativa de vida aumenta, o analfabetismo diminui.

Finalmente, na quarta linha, o índice de Gini relativo (GINI) e a população relativa (POP) têm uma fraca correlação negativa de $-0,2706$, logo, quanto maior for o GINI, ou seja, a desigualdade de renda entre a população, menor será a população relativa. O GINI e o ALF têm uma correlação positiva muito fraca de $0,0523$, isto é, quando se verifica um aumento da desigualdade de renda e da concentração de riqueza, maior é a taxa de analfabetismo. O GINI e a EXP têm uma fraca correlação negativa de $-0,2460$, pois, quando a desigualdade de renda diminui, a expectativa de vida aumenta.

4.3 Análise dos Testes

A partir do tratamento dos dados, foi analisado qual modelo se adequa melhor para explicar o comportamento da taxa de homicídios. Para o estudo, referente aos dados em painel, foi considerado os modelos Pooled, Efeitos Fixos e Efeitos Aleatórios. Para saber qual destes modelos é o mais indicado, foi examinado os testes de Chow, Breusch-Pagan e Hausman.

O resultado do teste Chow indica que a hipótese nula (hipótese nula é de que há igualdade nos interceptos e nas inclinações para todas as unidades da federação, ou seja o modelo de dados agrupados (pooled) é melhor) é rejeitada ao nível de 1% de significância, assim o modelo de Efeitos Fixos se ajusta melhor aos dados do que o modelo Pooled.

Quanto aos resultados do teste de Breusch-Pagan tem-se que a hipótese nula (o modelo de dados agrupados (pooled) é preferível ao modelo de efeitos aleatórios) também é rejeitada ao nível de 1% de significância, assim o modelo de Efeitos Aleatórios também se ajusta melhor que o Pooled aos dados da pesquisa.

Por fim o teste de Hausman para escolha entre o modelo de Efeitos Fixos e Efeitos Aleatórios indica que o modelo de Efeitos Fixos é superior, considerando um nível de

significância de 5%, uma vez que o p-valor foi de 0,01456. Ou seja, a hipótese nula de que o efeito aleatório é superior é rejeitada. Desse modo, usaremos o modelo de Efeitos Fixos para estimar os resultados do modelo de dados em painel da pesquisa.

4.4 Análise dos Dados em Painel

Na Tabela 3 estão os resultados obtidos através do modelo de dados em painel com efeitos fixos. Como a variável dependente (taxa de homicídios - HOM) está em forma logarítmica e as variáveis independentes estão em forma linear, o modelo apresenta a forma funcional de um modelo semilogarítmico Log-Lin. Com exceção da variável POP, é possível notar que todas as outras variáveis independentes foram estatisticamente significativas para o modelo. Por isso, pode-se levar em conta que, a taxa de analfabetismo (ALF) e a expectativa de vida são estatisticamente significativas ao nível de 99% de confiança, já a variável do Índice de Gini relativo (GINI) é estatisticamente significativo ao nível de 90% de confiança, um indicativo de que estas variáveis tendem a influenciar a taxa de homicídio. Com relação a qualidade de ajustamento do modelo, o coeficiente de determinação de 0,7530, presume que, cerca de 75,3% da variação da taxa de homicídios podem ser explicadas pelas variáveis independentes incluídas no modelo, quanto mais próximo de 1 ou 100%, melhor a reta se ajusta ao modelo de regressão, logo, o coeficiente de determinação de 75,3% indica que o ajustamento do modelo é relativamente alto.

Tabela 3 - Resultados do Modelo de dados em Painel com Efeitos Fixos

	Coefficiente	Erro Padrão	Razão-t	P-valor	
Constante	0,901056	2,64276	0,3410	0,7346	
POP	0,416851	0,345306	1,207	0,2330	
ALF	0,00604361	0,00215910	2,799	0,0073	***
EXP	0,0587194	0,0203921	2,880	0,0058	***
GINI	-0,0456602	0,0244301	-1,869	0,0675	*

Fonte: Elaboração própria com base nos dados coletados do software Gretl.

Conforme a Tabela 3, a variável POP, que representa a população relativa das 27 unidades federativas do Brasil, apresentou sinal positivo, ou seja, apesar de não ser

significativa para o modelo, ainda pode-se pressupor que é teoricamente condizente. Pois, o resultado demonstra que, um aumento de 1 unidade da população relativa, a taxa de homicídios tende a aumentar, em média, 41,68%, mantendo as demais variáveis constantes. Essa questão pode ser explicada pelo fato do crescimento populacional não planejado impactar diretamente a densidade demográfica populacional o que pode gerar em uma população deslocada em ambientes que tendem a influenciar o indivíduo a cometer atos ilícitos. Contudo, o resultado com o sinal positivo é o mesmo que foi obtido no estudo de Glaeser e Sarcedote (1999) e de Seillier (2010) ao analisar para o caso brasileiro.

Quanto a variável ALF, indica que o aumento de 1 ponto percentual da taxa da população analfabética de 15 anos ou mais, tende a aumentar, em média, 0,6% a taxa de homicídios, mantendo as demais variáveis constantes. Mostra-se condizente a literatura, pois Santos (2019), ao analisar a taxa de analfabetismo da população de 15 anos ou mais, obteve o mesmo resultado onde o analfabetismo também afeta os índices de violência nos municípios mineiros no período estudado de forma positiva, é estatisticamente significativo para o modelo utilizado pelo estudo. Dessa forma, quando há um cenário de piores condições de estudo a tendência é o aumento da criminalidade em determinado lugar. De acordo com a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), uma educação inadequada pode reforçar as divisões sociais, a intolerância e os preconceitos que resultam a conflitos violentos.

A variável EXP indica que, quando a expectativa de vida aumenta em 1 ano, a taxa de homicídios aumenta em média de 5,87%, mantendo as demais variáveis constantes. Embora tenha apresentado ser estatisticamente significativa para o modelo, nesse caso, mostra-se contrária ao que diz a literatura, com um sinal positivo, pois, o aumento da esperança de vida ao nascer indica melhoria das condições de vida e saúde da população, se há uma melhoria do bem-estar populacional, a tendência seria de uma diminuição da taxa de homicídio. Seillier (2010) em seu estudo voltado para 27 unidades federativas brasileiras no período 2001-2005, seguindo o modelo econométrico de dados em painel, concluiu que há uma relação inversa entre taxa de crimes letais e esperança de vida, logo, os estados que apresentam taxas de homicídios menores também apresentam maiores anos de vida, em média, esperados para um recém-nascido. Portanto o resultado obtido neste estudo não demonstra ser condizente com o resultado obtido pela autora ao analisar a mesma variável.

Com relação a variável GINI, quando há um aumento de 1 unidade do índice de Gini relativo, a taxa de homicídios diminui em média 4,56%, mantendo as demais variáveis constantes. O resultado sugere que quanto maior é o grau de desigualdade na distribuição da

renda domiciliar per capita entre os indivíduos, menor tende a ser a taxa de homicídios. Contudo, não demonstra ser condizente com outros estudos, o sinal negativo para essa variável não é esperado, ou seja, pressupõe-se teoricamente que essa variável aponta uma relação positiva, pois, de acordo com a análise de Mendonça (2001), ao verificar a influência que a desigualdade social sobre o fenômeno da criminalidade para o modelo de dados de painel para os estados brasileiros no período 1985-95, observou que a desigualdade social representada pelo índice de Gini exerce impacto positivo sobre a criminalidade no caso brasileiro, verifica-se que os estados com maior desigualdade econômica possuem maiores taxas de criminalidade.

Por fim, de acordo com os dados inferidos na Tabela 3, é estabelecido o seguinte modelo de regressão:

$$\ln\widehat{HOM}(Y_{it}) = 0,9010 + 0,4168POP_{it} + 0,0060ALF_{it} + 0,0587EXP_{it} - 0,0456GINI_{it}$$

5. CONSIDERAÇÕES FINAIS

Este estudo buscou apresentar os principais pontos sobre os determinantes da criminalidade relacionados a fatores socioeconômicos nos estados do Brasil.

A literatura do presente trabalho abordou premissas importantes da Teoria Econômica do Crime com os primeiros que autores que analisaram a criminalidade sob o viés da Ciência Econômica. Também foi apresentada literaturas de estudos mais recentes da criminalidade especificamente relacionada aos fatores socioeconômicos. Por meio da literatura, é possível observar que certas variáveis socioeconômicas mostram-se significativas para determinado modelo, em contrapartida, em outros estudos, alguns aspectos socioeconômicos não se mostram significativos para explicar o fenômeno da criminalidade.

Esta pesquisa teve como objetivo geral, por meio de uma análise de dados em painel, investigar o impacto que as variáveis socioeconômicas exercem no indicador de criminalidade no Brasil para os censos de 1991, 2000 e 2010 sob a perspectiva da economia do crime. Com base nos resultados encontrados, pode-se indicar que o objetivo proposto foi alcançado.

O principal resultado do modelo econométrico de dados em painel para efeitos fixos, mostrou que a taxa de analfabetismo, a expectativa de vida e o Índice de Gini têm um impacto significativo na taxa de homicídios. São fatores socioeconômicos que podem exercer influência nos níveis de criminalidade no país. Por outro lado, no que se refere à variável população relativa constatou-se que não é uma variável significativa para explicar o fenômeno da criminalidade, contudo, o resultado com sinal positivo demonstrou que quando a população relativa aumenta, a criminalidade tende a aumentar, cabe ressaltar que, no que tange a densidade demográfica, essa questão não é um problema, mas sim o bem estar populacional com maior motivo de preocupação para o aumento da criminalidade. Ainda, é possível notar que a qualidade de ajustamento do modelo é relativamente alta, significa que cerca de 75,3% da variação da taxa de homicídios podem ser explicadas pelas variáveis socioeconômicas incluídas no modelo.

Os resultados da pesquisa levam a contribuir e incentivar novos estudos acerca da criminalidade, pois, é um problema social que se agrava cada vez mais com o passar dos anos e que aflige não somente a vida dos brasileiros, mas toda a sociedade. Ademais os achados desse estudo são úteis, também, para promover discussões da criminalidade no Brasil sob um novo olhar para a Ciência Econômica.

Como limitações desse estudo, destacam-se as dificuldades do levantamento de dados de outras variáveis socioeconômicas para tornar o modelo econométrico mais adequado, outro aspecto limitador é que não há muitos estudos econométricos recentes sobre a economia do crime voltado para uma análise social mais aprofundada no que tange o Brasil.

Portanto, recomenda-se um maior aprofundamento sobre esse tema para identificar novos variáveis socioeconômicas que possam afetar a criminalidade. Além disso, trabalhos futuros podem abordar outros modelos econométricos e utilizar outros indicadores de criminalidade além da taxa de homicídios.

REFERÊNCIAS

- ARAÚJO JR., A. F.; FAJNZYLBBER, P. O que causa a criminalidade violenta no Brasil? Uma análise a partir do modelo econômico do crime: 1981 a 1996. Belo Horizonte: Universidade Federal de Minas Gerais, CEDEPLAR, 88p. Texto de Discussão, n. 162, 2001.
- BECK, Nathaniel. TIME-SERIES–CROSS-SECTION DATA: What Have We Learned in the Past Few Years?. **Annual Review of Political Science**, v.4, n.1, p. 271–293, jun. 2001.
- BECKER, G. Crime and Punishment: An Economic Approach. **Journal of Political Economy**, v.76, n.2, p. 169-217, mar/abr. 1968.
- BLOCK, M. K.; HEINECKE, J. M. A labor theoretic analysis of the criminal choice. **American Economic Review**, v. 65, n.3, p. 314-325, jun. 1975.
- BRASIL. Anuário Brasileiro de Segurança Pública. Fórum Brasileiro de Segurança Pública (FPSP) 2022. Disponível em: forumseguranca.org.br/wp-content/uploads/2022/06/anuario-2022.pdf?v=5. Acesso em: 17 jan.2023.
- Cano, I. & Santos, N. Violência letal, renda e desigualdade no Brasil. Fórum de Debate. Rio de Janeiro: IPEA/CESEC, 2000.
- CERQUEIRA, D. R. de C.; SOARES, R. R. **Custo de bem-estar da violência letal no Brasil e desigualdades regionais, educacionais e de gênero**. Brasília: IPEA, 2011.
- COHEN, L. E.; FELSON, M. Social change and crime rate trends: a routine activity approach. **American Sociological Review**, v. 44, n.4, p. 588-608, ago. 1979.
- CRUZ, Francisco Viana; ARAÚJO, Adriano Firmino Valdevino. Análise da Criminalidade em Palmas-TO: Uma abordagem Econométrica da Violência Urbana. **Informe Gepec**, v.16, n.2, p. 170-185, jul/dez. 2012.
- EHRlich, I. Participation in illegitimate activities: A theoretical and empirical investigation. **Journal of Political Economy**, University of Chicago Press, v. 81, n.3, p. 521-565, mai/jun. 1973.
- ENTORF, H.; SPENGLER, H. Socioeconomic and demographic factors of crime in Germany: evidence from panel data of the German states. **International Review of Law and Economics**, v. 20, n.1, p. 75-106, mar. 2000.
- FREEMAN, R. B. Crime and labor market. **National Bureau of Economic Research**. Cambridge, p. 1-32. nov. 1982.
- Glaeser, E.; Sacerdote, B. "Why is there more crime in cities?". **Journal Political Economy**, v.6, n.107, p. 225-258, 1999.
- GREENE, William H. **Econometric Analysis**. 2 ed. New York: MacMillan Publishing Company, 1993.

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

HOFFMANN, R. e VIEIRA, S. **Análise de regressão: uma introdução à econometria**. 2. ed. São Paulo: Hucitec. 1987.

HSIAO, CHENG. **Analysis of panel data**. 1. ed. Cambridge: Cambridge University Press, 1986.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA – **Ipeadata**. Dados macroeconômicos e regionais. Disponível em: www.ipeadata.gov.br. Acesso em: 30 out. 2022.

JALES, Leoncio do Vale. **Análise dos determinantes socioeconômicos da taxa de homicídios no Rio Grande do Norte**. 2020. Monografia (Graduação em Ciências Econômicas)- Universidade Federal do Rio Grande do Norte – UFRN, Natal, 2020.

JUNIOR, L. F. *et al.* La criminalidade no Brasil: avaliação do impacto dos investimentos públicos e dos fatores socioeconômicos. **Espacio Abierto**, Venezuela, v.26, n.2, p.219-243, abril/jun. 2017.

LARSON, R.; FARBER, B. **Estatística Aplicada**. 6. ed. São Paulo: Pearson Prentice Hall, 2015.

MARIANO, R. S. **Fatores Socioeconômicos da Criminalidade no Estado de São Paulo: Um Enfoque da Economia do Crime**. 2010. Dissertação (Mestrado em Economia) - Pontifícia Universidade Católica de São Paulo. São Paulo, 2010.

MENDONÇA, M. J. C. Um Modelo de Criminalidade para o Caso Brasileiro. In: Encontro Nacional de Economia, 2001, Salvador. **XXIX Encontro Nacional de Economia**, ANPEC, 2001.

MESQUITA, Rafael; FERNANDES, Antônio; FIGUEIREDO FILHO, Dalson B. Uma introdução à regressão com dados de painel. **Política Hoje**, Pernambuco, v.29, n.1, 2020.

OLIVEIRA, C. A. Análise Espacial da Criminalidade no Rio Grande do Sul. **Revista de Economia**, Curitiba, v. 34, p. 35-60, 2008.

PINTO, Antonio Marcos. et al. Uma análise dos determinantes da taxa de crimes de homicídios nos estados do brasil: uma aplicação em painel dinâmico. **Revista de Economia Regional, Urbana e do trabalho**, v. 7, n.2, p. 36-52, jun/dez. 2018.

REDE INTERAGENCIAL DE INFORMAÇÕES PARA A SAÚDE – **RIPSA**. Dados demográficos e socioeconômicos. Disponível em: fichas.ripsa.org.br/2012/. Acesso em: 30 out. 2022.

REIS, E.A., REIS I.A. **Análise Descritiva de Dados**. Relatório Técnico do Departamento de Estatística da UFMG, jun. 2002. Disponível em: www.est.ufmg.br. Acesso em: 12 out. 2022.

RICHARDSON, R. J. **Pesquisa social: métodos e técnicas**. 5. ed. São Paulo: Atlas, 1999.

SANTIAGO, A. C. R. de S. **Espacialização da criminalidade**: um estudo sobre a relação entre densidade demográfica e violência em Manaus. 2016. Dissertação (Mestrado em Geografia) – Universidade Federal do Amazonas, 2016.

SANTOS, Lucas Miranda Vilela. **A pobreza e a difusão da criminalidade nos municípios de Minas Gerais – MG**. 2019. Monografia (Graduação em Ciências Econômicas)- Universidade Federal de Ouro Preto, Mariana, 2019.

SANTOS, Sylvia Cristina Lavor dos. **Vitimização e desigualdade de renda no Brasil**. 2008. 82f. : Dissertação (mestrado) - Universidade Federal do Ceará, Programa de Pós Graduação em Economia, CAEN, Fortaleza, CE, 2008.

SILVA, A. J. **Economia do crime**: uma análise dos possíveis determinantes da criminalidade em Alagoas. 2018. Trabalho de Conclusão de Curso (Bacharelado em Curso de Ciências Econômicas) – Universidade Federal de Alagoas, Santana do Ipanema, 2018.

SEILLIER, Martha. **Criminalidade nos estados brasileiros no período 2001-2005**: uma análise econômica com dados em painel. 2010. Dissertação (Mestrado em Ciências Econômicas)-Universidade de Brasília, Brasília, 2010.

TORRES-REYNA, Oscar. **Getting started in fixed/random effects models using R**. Data & Statistical Services, Princeton University, 2010.

UNESCO. A crise oculta: conflitos armados e educação. Organização das Nações Unidas para Educação, a Ciência e Cultura (UNESCO), Paris, 2011. Disponível em: unesdoc.unesco.org/ark:/48223/pf0000191186_por. Acesso em: 10 nov.2022.

WOOLDRIDGE, J. M. **Introdução à econometria**: uma abordagem moderna. 4. ed. São Paulo: Cengage Learning, 2010.

ZANOTELLI, C. et al. **Atlas da criminalidade no Espírito Santo**. 1. ed. São Paulo: Annablume, 2011.