

Replicação: A redução da maioria penal diminui a violência? Evidências de um estudo comparado

Artur Damião*

16 de dezembro de 2025

Sumário

1	Introdução	2
2	Hipóteses	2
3	Interpretação dos coeficientes da regressão	3
4	Checagem dos modelos	4
4.1	Linearidade dos Resíduos	4
4.1.1	É possível ajustar a linearidade do modelo?	5
4.2	Colinearidade (VIF) no Modelo	7
4.3	Homocedasticidade	8
4.4	Normalidade dos resíduos	8
5	Analisando o erro padrão robusto	9
6	Avanços em relação ao trabalho original	11
	Referências	13

*Programa de Pós-Graduação em Sociologia da Universidade de São Paulo. N° USP 10701251. arturcardoso@usp.br.

1 Introdução

O artigo de Lins, Figueiredo Filho, e Silva (2016) tem como objetivo investigar a possível associação entre a diminuição da maioridade penal e os níveis de violência em perspectiva comparada. Para tanto, os autores utilizaram de análise espacial, correlação de Pearson e de modelos de regressão linear de mínimos quadrados ordinários (MQO).

Em termos de avanço do artigo original, propomos a interpretação das variáveis de controle, checagem dos pressupostos do modelo, a análise do erro padrão robusto e a estimação por modelos de Regressão Robusta com M-estimador de Huber.

Após a leitura dos dados diretamente pelo [GitHub](#), realizamos uma transformação nas variáveis, substituindo os valores 999 por NA nos dados de Cipriani.

2 Hipóteses

O objetivo dos autores é testar a hipótese de que quanto menor a maioridade penal, menor o nível de violência. A Equação 1 apresenta o modelo estimado:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \epsilon \quad (1)$$

Onde,

Y : Taxa de Homicídio por 100 mil habitantes (`homi_rate_unodc`)

X_1 : Idade de Maioridade Penal ou Responsabilidade Criminal (`acm_hazel`, `acm_gv`, `acr_hazel` ou `acr_cipriani`)

X_2 : Índice de Desenvolvimento Humano (`idh`)

X_3 : Índice Gini (`ginmar_solt`)

X_4 : Desemprego de longo termo (`desemprego_longo`)

Os modelos estão representados na Tabela 1 abaixo:

Tabela 1: Resultados das Regressões Lineares

	Variável dependente:			
	Taxa de Homicídio (100 mil hab.)			
	(1)	(2)	(3)	(4)
ACM Hazel	-0.365* (0.197)			
ACM GV		-0.224 (1.310)		
ACR Hazel			-0.271* (0.139)	
ACR Cipriani				-0.048 (0.093)
IDH	-13.778** (5.689)	-28.974*** (7.718)	-28.575*** (7.243)	-29.454*** (7.648)
Desemprego Longo	-0.025 (0.028)	-0.103*** (0.033)	-0.086** (0.032)	-0.102*** (0.033)
Índice de Gini	0.025 (0.053)	-0.012 (0.074)	-0.010 (0.069)	-0.009 (0.073)
Constant	19.743*** (6.325)	33.800 (24.764)	32.400*** (7.445)	30.529*** (7.854)
Observations	29	37	37	37
R ²	0.364	0.382	0.447	0.387
Adjusted R ²	0.258	0.305	0.378	0.310
Residual Std. Error	1.563 (df = 24)	2.501 (df = 32)	2.365 (df = 32)	2.492 (df = 32)
F Statistic	3.439** (df = 4; 24)	4.947*** (df = 4; 32)	6.476*** (df = 4; 32)	5.042*** (df = 4; 32)

Note:

*p<0.1; **p<0.05; ***p<0.01

3 Interpretação dos coeficientes da regressão

Mantendo todas as variáveis constantes iguais à zero (teoricamente), os modelos (valor de α) indicariam uma taxa de 19.74 (Modelo 1), 33.80 (Modelo 2), 32.40 (Modelo 3) e 30.52 (Modelo 4) homicídios para cada grupo de 100 mil habitantes.

Os coeficientes negativos da Redução da Maioridade Penal (ACM) ou Responsabilização Criminal (ACR) indicam que quanto maior a idade penal, menor a taxa de homicídios. Dessa forma, rejeitamos a H_1 que menor maioridade = menor violência. Os modelos 2 (`acm_gv`) e 4 (`acr_cipriani`) não apresentam significância estatística, ao passo que os modelos 1 (`acm_hazel`) e 3 (`acr_hazel`) apresentam significância a um nível de 10% ($p < 0.1$). É um nível de significância fraco. A partir desses dados, não é possível afirmar que a redução da maioridade penal ou responsabilidade criminal diminui as taxas de homicídio.

O Índice de Desenvolvimento Humano, ao contrário, apresenta grande significância estatística ($p < 0.05$ e $p < 0.01$), sendo a variável que mais explica a redução na taxa de homicídios. No primeiro modelo, a cada um ponto em que o IDH aumenta, reduz-se 13,788 pontos na taxa de homicídio; no segundo modelo, a cada aumento de um ponto no IDH, a taxa de homicídio reduz 28,974 ponto - redução semelhante para os modelos três (28,575 pontos) e quatro (29,454 pontos).

O Desemprego Longo também apresenta sinal negativo nos coeficientes, sugerindo que um maior grau de desemprego a longo prazo está associado a uma menor taxa de homicídios. Esse resultado, contraintuitivo, pode ser causado pelas características da nossa amostra (nosso n varia entre 29 ou 37) ou colinearidade com outras variáveis (especificamente o IDH).

Por fim, o Índice de Gini não apresentou nenhum valor significativo, sugerindo que a desigualdade não explica a variação da taxa de homicídio. Provavelmente isso se dá pelo n pequeno. Ao interpretamos o R^2 , vemos que nosso modelo explica, em média, 0.395% da variação nas taxas de homicídio. Ou seja, mais da metade da violência é causada por fatores que não estão em nossa equação.

4 Checagem dos modelos

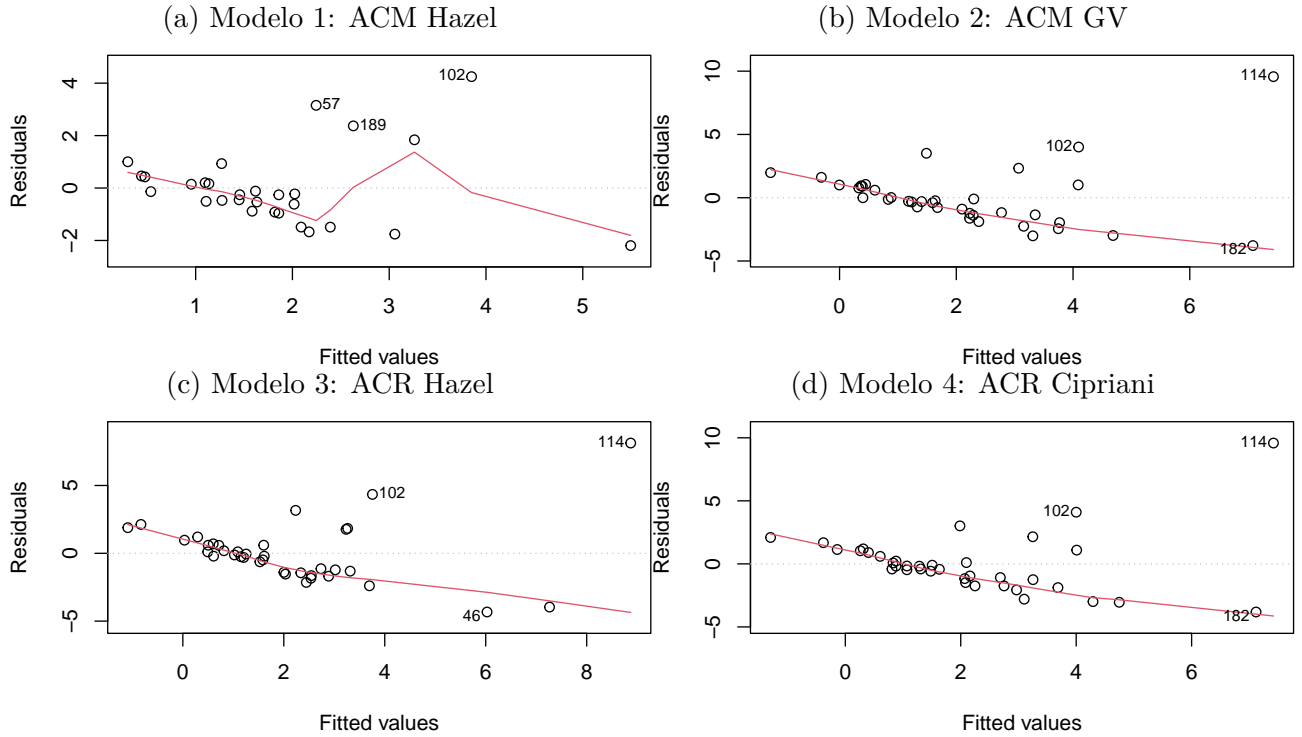
Nesta seção, avaliamos se os modelos atendem aos pressupostos da regressão linear:

1. Linearidade
2. Colinearidade (VIF)
3. Homocedasticidade
4. Normalidade dos resíduos

4.1 Linearidade dos Resíduos

O pressuposto da linearidade avalia se a relação entre os preditores e a variável resposta é, em média, linear. Supõe-se que $\mathbb{E}[\epsilon|X] = 0$ (a média condicional dos erros é zero, logo, o preditor e o erro não estão correlacionados).

Figura 1: Checagem da Linearidade (Resíduos X Valores Previstos)



Conforme a Figura 1, era esperado que a linha vermelha tivesse um comportamento reto horizontalmente, ou seja, com a média condicional dos erros iguais a zero. Todos os casos sugerem que não há relação linear entre a variável resposta e os preditores. Ademais, a presença de valores *outliers* é preocupante.

4.1.1 É possível ajustar a linearidade do modelo?

A não-linearidade prejudica a confiança na interpretação dos estimadores (β) do modelo. A não-linearidade, no nosso caso, pode se dar devido ao tamanho da amostra e a presença de *outliers*, constatada anteriormente. Para tentar solucionar este problema, aplicaremos o logaritmo natural na variável dependente¹. A Tabela 2 sintetiza o resultado:

¹Aplicando o log na taxa de homicídio, assumimos que as variáveis explicativas tem um efeito multiplicativo e não aditivo sobre o crime.

Tabela 2: Regressões com Log Natural

	Variável dependente:			
	Log Natural da Taxa de Homicídio (100 mil hab.)			
	(1)	(2)	(3)	(4)
ACM Hazel	-0.135 (0.080)			
ACM GV		-0.448 (0.359)		
ACR Hazel			-0.062 (0.040)	
ACR Cipriani				-0.035 (0.026)
IDH	-6.247** (2.312)	-8.337*** (2.116)	-8.550*** (2.072)	-8.902*** (2.097)
Desemprego Longo	-0.009 (0.011)	-0.019** (0.009)	-0.015 (0.009)	-0.018* (0.009)
Índice de Gini	0.007 (0.022)	-0.004 (0.020)	0.002 (0.020)	0.002 (0.020)
Constant	7.991*** (2.570)	16.050** (6.790)	8.617*** (2.130)	8.557*** (2.153)
Observations	29	37	37	37
R ²	0.366	0.384	0.400	0.389
Adjusted R ²	0.261	0.307	0.325	0.312
Residual Std. Error	0.635 (df = 24)	0.686 (df = 32)	0.677 (df = 32)	0.683 (df = 32)
F Statistic	3.467** (df = 4; 24)	4.989*** (df = 4; 32)	5.338*** (df = 4; 32)	5.090*** (df = 4; 32)

Note:

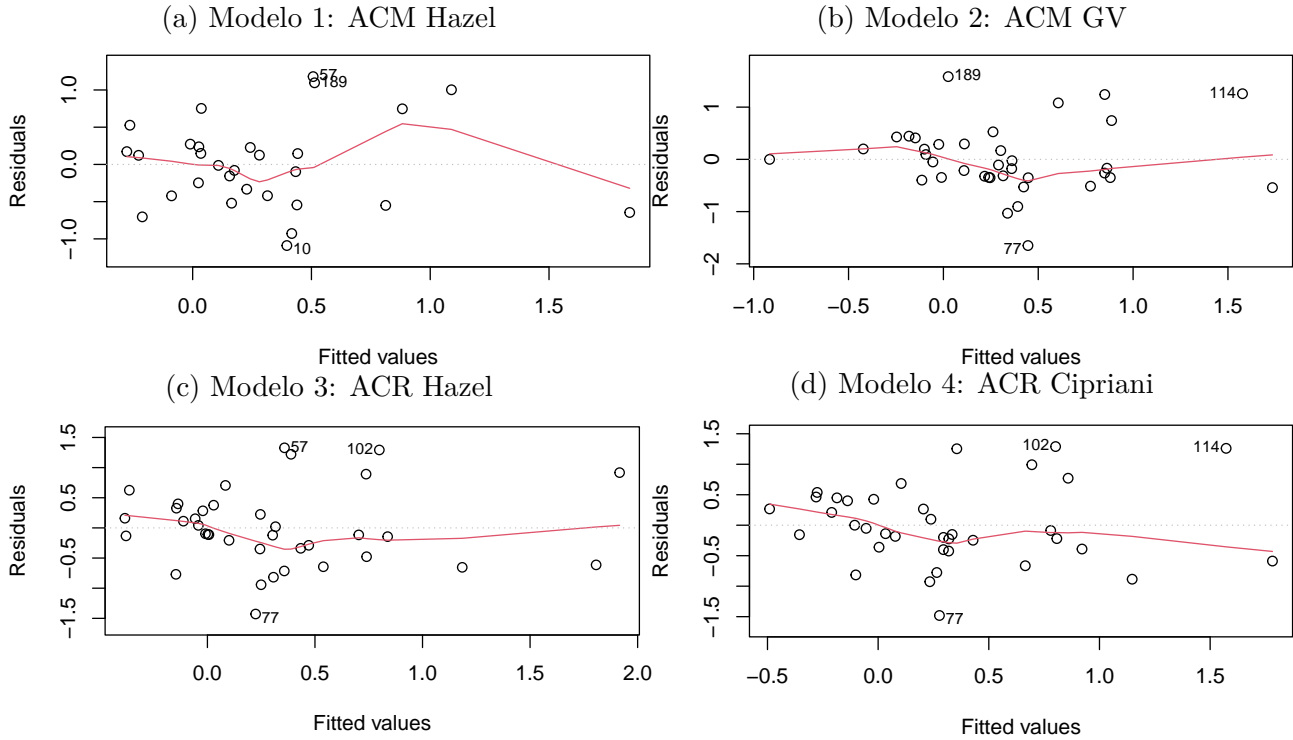
*p<0.1; **p<0.05; ***p<0.01

Em termos de interpretação do modelo ajustado, os dados de ACM Hazel, ACM GV, ACR Hazel e ACR Cipriani não apresentam significância estatística. Ou seja, rejeitamos a H_1 . Interpretando as variáveis de controle², constatamos que o Índice de Gini também não influencia de maneira significativa a taxa de homicídios. O Índice de Desenvolvimento Humano continua

²Geralmente, as variáveis de controle são utilizadas não para serem interpretadas. Entretanto, realizamos este exercício para fins pedagógicos.

significativo: a cada um ponto (um centésimo, já que o IDH varia entre 0 e 1) em que o IDH aumenta, reduz-se a taxa de homicídio em 6.2% (Modelo 1), 8.3% (Modelo 2), 8.5% (Modelo 3) e 8.9% (Modelo 4).

Figura 2: Checagem da Linearidade (Resíduos X Valores Previstos) com Log Natural



Agora, com o modelo ajustado, é possível perceber que, diferente dos gráficos 1, a Figura 2 apresenta os valores muito mais próximo da reta. Ou seja, a média condicional dos erros está mais próxima de zero ao transformar a variável dependente em seu logaritmo natural.

Outra modificação possível seria a transformação quadrática da variável IDH. Neste caso, a Equação 2 representaria nosso modelo:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 IDH^2 + \epsilon \quad (2)$$

Dessa forma, consideraríamos as relações não lineares entre as variáveis, ajustando uma curva parabólica ao invés de uma curva reta.

4.2 Colinearidade (VIF) no Modelo

A Tabela 3 sintetiza os valores de Variation Inflation Factor, indicando que as variáveis são independentes entre si. Ou seja, elas contribuem de maneira distinta para o modelo. Os valores encontram-se significativamente abaixo do limite crítico de 5. Logo, desconsideramos a colinearidade suspeitada na Seção 3.

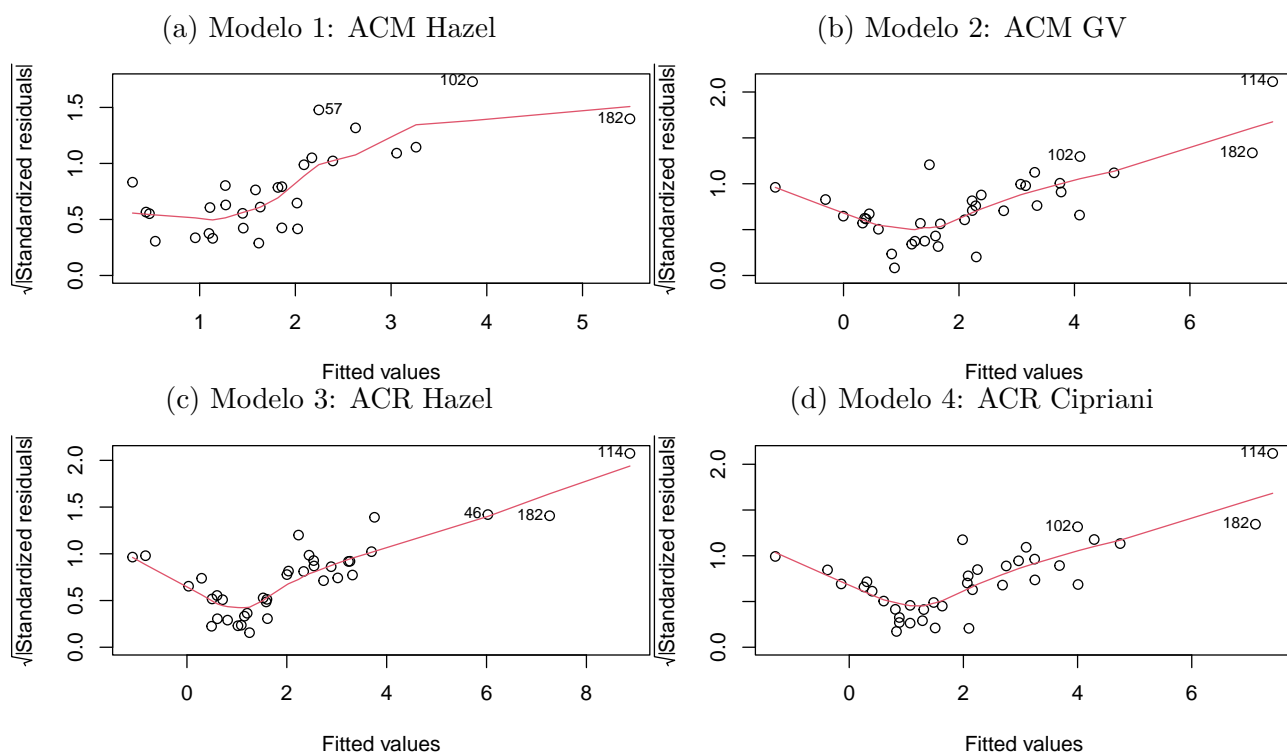
Tabela 3: Tabela de Multicolinearidade (VIF)

Variável	ACM Hazel (1)	ACM GV (2)	ACR Hazel (3)	ACR Cipriani (4)
Maioridade/Responsabilidade	1.18	1.07	1.09	1.01
IDH	1.22	1.10	1.09	1.09
Desemprego Longo	1.45	1.16	1.25	1.16
Índice de Gini	1.17	1.13	1.08	1.08

4.3 Homocedasticidade

A homocedasticidade pressupõe que os resíduos variem de forma constante ao longo dos valores previstos. Os gráficos da Figura 3 apresentam que os erros não variam igualmente ao longo da reta. Logo, há a presença de heterocedasticidade.

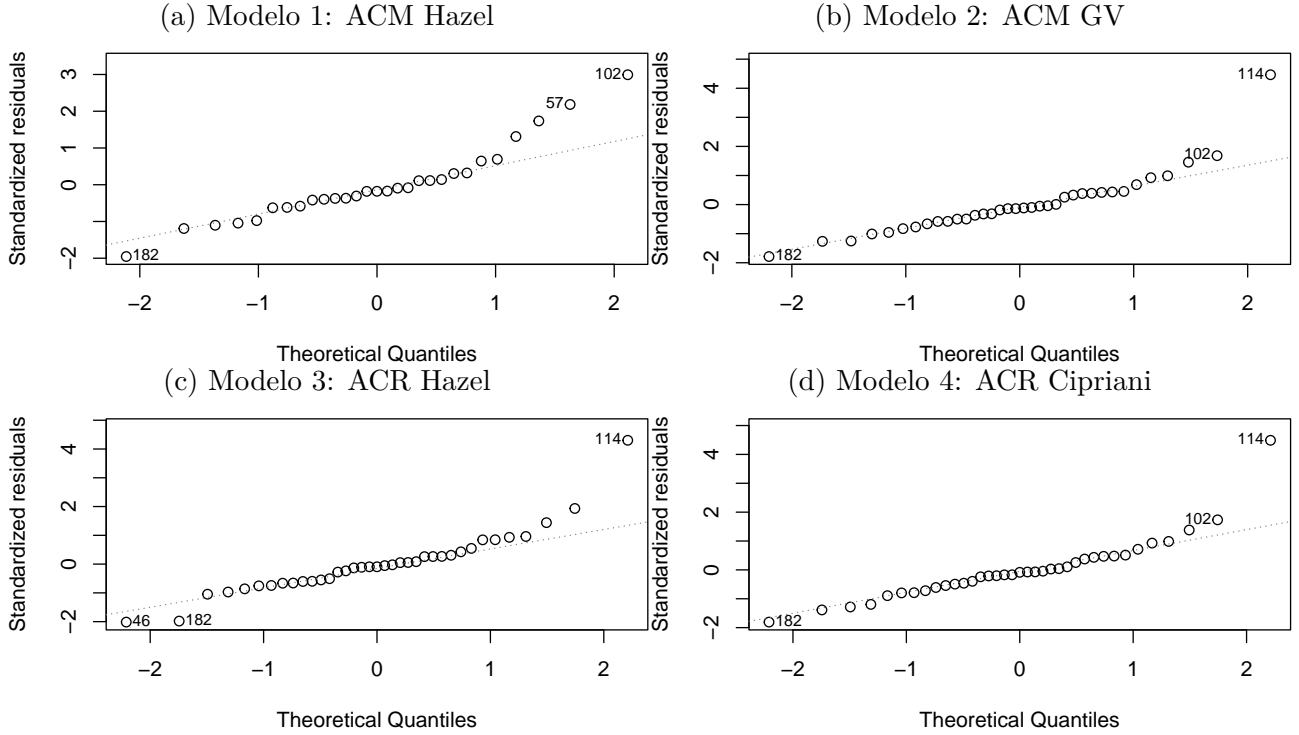
Figura 3: Checagem da Homocedasticidade



4.4 Normalidade dos resíduos

O teste de normalidade dos resíduos, apresentado na Figura 4, verifica se os resíduos do modelo seguem uma distribuição normal. Nosso gráfico abaixo indica que há uma assimetria nas pontas, sugerindo que nosso modelo erra mais para cima do que para baixo.

Figura 4: Checagem da Normalidade dos Resíduos (Q-Q Plot)



Isso se dá, provavelmente, pelo caráter da amostra. É mais provável que poucos países concentrem taxas de homicídios mais altas, enquanto a maior parte dos países da amostra de Lins, Figueiredo Filho, e Silva (2016) tem taxas baixas ou médias. Como o n amostral é relativamente pequeno ($n < 40$, para cada país), os poucos casos *outliers* distorcem a normalidade dos resíduos.

5 Analisando o erro padrão robusto

Para corrigir a heterocedasticidade presente nos modelos, utilizaremos os erros padrão robustos. Especificamente, utilizaremos o *HC3*, dado o tamanho de nossa amostra. O *HC3* é apropriado para amostras pequenas porque ele penaliza os *outliers*, evitando falsos positivos³. A fórmula para o estimador *HC3* é:

$$\frac{\hat{u}_i^2}{(1 - h_{ii})^2}$$

Leia-se: o quadrado do resíduo da observação i , dividido pelo quadrado do complemento da alavancagem (*leverage*) dessa mesma observação.⁴

³Parte dessa informação foi dada pelo professor Manoel Galdino. De forma complementar, perguntei ao Gemini por quê usar o *HC3* ao invés do *HC1* e comparei o resultado dos dois testes nos modelos. No *HC1*, algumas variáveis (como Índice de Desenvolvimento Humano), continuaram estatisticamente significativas. Com o *HC3*, nenhuma variável manteve sua significância.

⁴Pedi ajuda para o Gemini sobre como ler essa fórmula. Basicamente: o *HC3* pega o erro que o modelo cometeu (\hat{u}^2) e o “infla” (aumenta) baseado no quanto aquela observação é influente (h_{ii}). Quanto mais próximo o h estiver de 1, menor será o denominador e maior será o resultado total (a penalidade).

A Tabela 4 apresenta o erro robusto para o Modelo 1:

Tabela 4: Estimativas do Modelo ACM Hazel com Erros Robustos (HC3)

Variável	Estimativa	Erro Padrão	t	p-valor
Constante	19.743	12.328	1.602	0.122
Maioridade (Hazel)	-0.365	0.248	-1.470	0.155
IDH	-13.778	10.016	-1.376	0.182
Desemprego	-0.025	0.021	-1.183	0.248
Gini	0.025	0.061	0.416	0.681

A Tabela 5 apresenta o erro robusto⁵ para o Modelo 2:

Tabela 5: Estimativas do Modelo ACM GV com Erros Robustos (HC1)

Variável	Estimativa	Erro Padrão	t	p-valor
Constante	33.800	13.617	2.482	0.018
Maioridade (GV)	-0.224	0.254	-0.882	0.384
IDH	-28.974	14.079	-2.058	0.048
Desemprego	-0.103	0.059	-1.740	0.091
Gini	-0.012	0.052	-0.227	0.822

A Tabela 6 apresenta o erro robusto para o Modelo 3:

Tabela 6: Estimativas do Modelo ACR Hazel com Erros Robustos (HC3)

Variável	Estimativa	Erro Padrão	t	p-valor
Constante	32.400	20.248	1.600	0.119
Responsabilidade (Hazel)	-0.271	0.269	-1.007	0.321
IDH	-28.575	17.108	-1.670	0.105
Desemprego	-0.086	0.060	-1.443	0.159
Gini	-0.010	0.061	-0.159	0.875

A Tabela 7 apresenta o erro robusto para o Modelo 4:

⁵Para este caso, utilizamos o *HC1*, dado que o *HC3* é incalculável visto que a observação 86, o valor h_{ii} é igual a 1 e o resultado é NaN.

Tabela 7: Estimativas do Modelo ACR Cipriani com Erros Robustos (HC3)

Variável	Estimativa	Erro Padrão	t	p-valor
Constante	30.529	18.490	1.651	0.108
Responsabilidade (Cipriani)	-0.048	0.102	-0.470	0.642
IDH	-29.454	17.640	-1.670	0.105
Desemprego	-0.102	0.074	-1.379	0.177
Gini	-0.009	0.059	-0.146	0.885

Analisando as tabelas geradas, após a constatação da violação dos pressupostos de homocedasticidade, aplicamos o estimador *HC3* para investigar os erros padrão do modelo. As variáveis explicativas do modelo 1, 3 e 4 perderam significância estatística. O IDH, que antes era uma variável que contribuía para explicar o fenômeno da queda das taxas de homicídio, agora não é mais significativo.

Uma exceção analítica foi necessária para o Modelo 2 (ACM GV), no qual a presença de uma observação com alavancagem perfeita inviabilizou matematicamente o cálculo do estimador *HC3*. Ao adotar-se o estimador *HC1*, que aplica uma correção menos conservadora, observou-se que o IDH preservou sua significância estatística ($p = 0,048$) com um coeficiente de -28,974, enquanto a variável de desemprego de longo prazo apresentou significância marginal ($p = 0,091$). Esse resultado sugere que o desenvolvimento humano permanece como o preditor mais resiliente da taxa de homicídios, resistindo a correções moderadas de heterocedasticidade, embora sua significância seja suprimida sob os critérios mais estritos do *HC3*.

6 Avanços em relação ao trabalho original

Para avançar na interpretação dos resultados, estimamos os modelos de Regressão Robusta com M-estimador de Huber. Nesse modelo de regressão, é possível mitigar a influência dos valores extremos sem recorrer à exclusão amostral. A Tabela 8 apresenta o resultado:

Tabela 8: Regressão Robusta (estimador-M de Huber)

	Variável dependente:			
	Taxa de Homicídio (100 mil hab.)			
	(1)	(2)	(3)	(4)
ACM Hazel	-0.135 (0.096)			
ACM GV		-0.284 (0.390)		
ACR Hazel			-0.051 (0.045)	
ACR Cipriani				-0.015 (0.029)
IDH	-9.036*** (2.761)	-10.482*** (2.295)	-10.889*** (2.367)	-10.894*** (2.356)
Desemprego Longo	-0.013 (0.014)	-0.023** (0.010)	-0.022** (0.011)	-0.023** (0.010)
Índice de Gini	0.010 (0.026)	-0.003 (0.022)	0.0003 (0.023)	0.001 (0.022)
Constant	11.456*** (3.069)	16.068** (7.364)	11.783*** (2.433)	11.321*** (2.420)
Observations	29	37	37	37
Residual Std. Error	0.533 (df = 24)	0.594 (df = 32)	0.643 (df = 32)	0.635 (df = 32)

Note:

*p<0.1; **p<0.05; ***p<0.01

Sob a estimação robusta, o IDH volta a apresentar significância estatística ($p < 0.01$) em todas as quatro especificações, apresentando coeficientes elevados (variando entre -9,036 e -10,894). Adicionalmente, a Regressão Robusta identificou uma associação negativa e significativa para o desemprego de longa duração nos Modelos 2, 3 e 4.

Entretanto, o resultado é consistente para as variáveis independentes principais. Em todas as estratégias de estimação que utilizamos até aqui - OLS, erros padrão robusto (HC) e Regressão Robusta (RLM) - as medidas de alteração na maioridade ou responsabilidade penal não apresentam significância estatística em nenhum dos cenários estimados. Concluímos, então

que, controlando-se pelas condições socioeconômicas estruturais, as variações nos marcos legais de imputabilidade penal não exercem impacto perceptível sobre a taxa de homicídios no contexto analisado.

Referências

Lins, Rodrigo, Dalson Figueiredo Filho, e Lucas Silva. 2016. «A Redução Da Maioridade Penal Diminui a Violência? Evidências de Um Estudo Comparado». *Opinião Pública* 22 (1): 118–39. <https://doi.org/10.1590/1807-0191201622118>.