RESEARCH PAPER

Armed Societies, Unequal Outcomes: A Global Statistical Investigation of Civilian Firearm Ownership, Gun Violence, and Socioeconomic Risk

Akshay Byagathvalli

November 2026

DATA	96 countries · SAS civilian firearms 2017 · UNODC firearm homicide 2013-2022 · WDI · WGI · UCDP
METHODS	12 regression specifications · 14 ML models, nested CV · ablations · SHAP/ALE · Moran's I · 34-spec robustness matrix
HEADLINE	Raw ρ = 0.01 · adjusted +9.5% per 10 guns/100 (fragile) · best MAE 3.53

All estimates are ecological, observational associations between countries. Firearms per 100 residents is an estimated stock, not the share of people who own a gun. No causal effects are identified.

Contents

Abstract 3

1. Introduction 3

2. Related Literature 4

3. Data and Measurement 5

4. Methodology 7

5. Descriptive Results 9

6. Econometric Results 23

7. Machine Learning Results 29

8. Geographic and Country-Level Analysis 36

9. Robustness Analysis 44

10. Discussion 46

11. Limitations 47

12. Conclusion 48

References 49

Appendix 51

Abstract

This study asks how strongly civilian firearm prevalence is associated with firearm homicide across countries, and how much of the cross-national variation is instead explained by socioeconomic, institutional and regional conditions. Small Arms Survey estimates of civilian firearms per 100 residents (reference year 2017) are linked to UNODC firearm homicide rates averaged over 2013-2022 for 96 countries (88 with complete covariates), supplemented with World Bank development indicators, Worldwide Governance Indicators and UCDP conflict data. Across countries the raw association is essentially zero (Spearman $\rho = 0.01$, 95% CI -0.19 to 0.21), because Latin America and the Caribbean combine moderate firearm stocks with the world's highest firearm homicide rates. After adjusting for socioeconomic conditions, regional fixed effects and governance, an additional 10 firearms per 100 residents is associated with 9.5% higher (1 + firearm homicide rate) (95% CI 1.2% to 18.5%, $p = 0.024$); a negative binomial count model gives a rate ratio of 1.32 (1.13-1.53). On the untransformed rate scale the adjusted coefficient is +0.28 per 100,000 and not distinguishable from zero. The adjusted association is positive in 97% of 34 robustness specifications but its confidence interval excludes zero in only 62%; it weakens when the United States is removed, when covariates are multiply imputed, or when continental rather than World Bank regions are used. Adding region effects and the governance index raises explained variance in log rates from 0.05 to 0.78. In nested cross-validation the best model (Random forest) reaches a mean absolute error of 3.53 per 100,000 against 3.67 for linear regression and 4.42 for a median baseline; adding firearm prevalence to all other predictors changes random-forest MAE by -0.5%. Firearm prevalence is therefore a weak, fragile conditional correlate of firearm homicide between countries, while regional context carries most of the explanatory and predictive information. All results are ecological associations and do not identify causal effects.

Keywords: firearm homicide; civilian firearms; cross-national analysis; governance; ecological regression; machine learning; spatial autocorrelation

1. Introduction

Firearm homicide rates differ by more than three orders of magnitude between countries. Some countries with very large civilian arsenals record low rates of gun killing, while some countries with comparatively modest civilian stocks record rates above 30 per 100,000. This variation has long motivated cross-national comparisons of firearm availability and violence (Killias, 1993; Hemenway & Miller, 2000; Hepburn & Hemenway, 2004), and equally long-standing criticism of what such comparisons can show (Kleck, 2015).

International comparisons are difficult for at least four reasons. First, the exposure is poorly measured: no country counts its civilian firearms directly, and the most widely used estimates (Karp, 2018) combine surveys, registration data, expert judgement and, for many countries, extrapolation from similar countries. Second, the outcome is unevenly reported: mechanism-specific homicide data are routinely published by European and Latin American governments but rarely by countries in Africa and South Asia. Third, firearm prevalence is entangled with income, inequality, urbanisation, state capacity, organised crime and armed conflict, all of which plausibly affect homicide (Fajnzylber, Lederman, & Loayza, 2002; Neumayer, 2003; Nivette, 2011). Fourth, country-level associations cannot be read as statements about individual risk (Robinson, 1950).

This paper therefore aims at a precise descriptive and predictive answer rather than a causal one. The central research question is:

To what extent is civilian firearm ownership associated with firearm-related homicide rates across countries, and how much additional variation can be explained by socioeconomic, demographic, geographic, and institutional characteristics?

The contribution is a transparent, reproducible pipeline that (i) builds a carefully aligned country-level dataset with explicit reference years, (ii) estimates twelve complementary statistical specifications, (iii) measures incremental

predictive value with fourteen learning algorithms on identical cross-validation folds, (iv) tests for spatial dependence, and (v) reports how the headline association moves across more than thirty robustness specifications.

2. Related Literature

Firearm availability and violence. Early cross-national work found positive correlations between household gun ownership and homicide in small samples of developed countries (Killias, 1993). Hemenway and Miller (2000) reported that, across 26 high-income countries, firearm availability was associated with higher homicide rates, a pattern reviewed by Hepburn and Hemenway (2004) and revisited with mortality data by Grinshteyn and Hemenway (2016). Bangalore and Messerli (2013) reported a correlation between guns per capita and firearm deaths across 27 developed countries. Kleck (2015) argued that most such studies fail to address reverse causality (people may arm themselves in response to violence) and rely on weak proxies for ownership. A recurring feature of this literature is that positive associations are clearest in samples of high-income countries and much weaker when middle-income countries are included.

International homicide variation. UNODC (2019, 2023) documents that the Americas account for a disproportionate share of global homicide, much of it by firearm and linked to organised crime and gangs. Long-run historical series show that Western European homicide fell by more than an order of magnitude between the late Middle Ages and the twentieth century (Eisner, 2003), a reminder that levels of interpersonal violence change with institutions and social control, not only with weapons.

Inequality, institutions and crime. Fajnzylber et al. (2002) found that income inequality predicts higher homicide and robbery rates in a cross-national panel. Neumayer (2003) emphasised the role of good governance and economic policy, and Nivette's (2011) meta-analysis identified inequality and Latin American location as among the most consistent cross-national predictors of homicide. The Worldwide Governance Indicators (Kaufmann, Kraay, & Mastruzzi, 2011) provide a standard summary of rule of law, government effectiveness, political stability and corruption control.

Cross-national ecological research and prediction. Country-level regression is subject to ecological fallacy (Robinson, 1950), omitted-variable bias and small samples. Policy-focused reviews (Santaella-Tenorio, Cerdá, Villaveces, & Galea, 2016) therefore weight within-country legislative changes more heavily than cross-sectional comparisons. Predictive modelling of crime rates with flexible learners is less common in this literature; with around one hundred observations, the main methodological risk is optimistic error estimates from tuning on the evaluation data (Varma & Simon, 2006), which motivates nested cross-validation here.

3. Data and Measurement

3.1 Sources

The primary dataset is the Kaggle collection *Global Gun Violence and Civilian Firearms*, which contains four files: Small Arms Survey (SAS) civilian firearm estimates for 230 locations (reference year 2017); a country profile; a 187-country UNODC/World Bank panel for 2000-2024 (3,120 country-years); and a long-run file of total homicide rates for Western European countries (1250-2023). Every column was inspected programmatically; the data dictionary and an HTML quality report are generated in `outputs/`. Three external sources were integrated because definitions, identifiers and reference years could be aligned: the World Bank Worldwide Governance Indicators (four estimates, 2013-2022 means), the UCDP Battle-Related Deaths dataset v25.1 (conflict exposure), and Natural Earth 1:110m boundaries (maps and contiguity). The source manifest records URLs, retrieval dates and SHA-256 checksums.

3.2 Outcome definition

The primary outcome is the **firearm homicide rate**: intentional homicides committed with a firearm per 100,000 population, as reported to UNODC, averaged over the available years in 2013-2022. The window brackets the 2017 firearm estimate; the number of observed years is retained for every country (median 8; 14 countries have fewer than three). Reported rates agree with counts divided by population to a median absolute difference of 0.0003 per 100,000; 6 country-years report a firearm rate above the total homicide rate and are excluded from the derived non-firearm rate only. Total homicide, the firearm share of homicides and the non-firearm homicide rate are analysed as clearly labelled secondary outcomes. The dataset contains **no firearm suicide, accidental, undetermined, total firearm mortality, non-fatal injury or mass-shooting measures**, so none of these is analysed and no outcome is substituted for another.

3.3 Exposure definition

The exposure is the SAS estimate of **civilian-held firearms per 100 residents** in 2017. This is an estimate of the *stock* of firearms, not the share of people or households that own a gun; a value of 120.5 (United States) means more guns than residents, not universal ownership. Estimates are produced by three methods: 23 locations use survey and expert estimates, 158 expert estimates and 49 analogous comparison with similar countries. The method is retained and used in robustness checks. Firearm prevalence is measured once; any panel model therefore cannot use within-country variation in ownership.

3.4 Country harmonisation

All sources were mapped to ISO 3166-1 alpha-3 codes; the audit table records 919 source rows and 32 manual decisions. SAS reports the United Kingdom as three units (England & Wales, Scotland, Northern Ireland), which were aggregated by summing firearms and population (4.9 per 100). Puntland, Somaliland and Northern Cyprus are not UNODC reporting units and were dropped; Natural Earth codes of -99 (France, Norway, Kosovo) were corrected; UCDP conflict locations were matched by name. Regions follow the World Bank classification recorded in the data (the Middle East group includes Afghanistan and Pakistan); income groups are World Bank groups.

3.5 Analytical samples

Table 1. Analytical samples. Primary = S1; complete-covariate = S2 (used for adjusted models).

Sample	Countries	Regions	Median firearms/100	Median firearm homicide
Full available (union of SAS 2017 and UNODC panel units)	226	7	5.4	0.38
Primary: firearms per 100 + firearm homicide (2013-2022)	96	7	10.0	0.38
Complete core covariates (GDP, urban, unemployment, male 15-64, rule of law)	88	7	10.1	0.34
Core covariates + Gini	75	7	10.1	0.37
Region-comparable (regions with >= 5 primary countries)	91	5	10.0	0.37
Income-comparable (income groups with >= 5)	95	7	10.0	0.37
Secondary: firearms per 100 + total homicide	162	7	7.5	0.38
Panel: country-years with firearm homicide (2005-2023)	99	7		0.49

Firearm homicide data exist for only 96 of the 230 SAS locations. Coverage is concentrated in Europe and Central Asia (38 countries) and Latin America and the Caribbean (31), with only 5 Sub-Saharan African and 2 South Asian countries and a single low-income country (Figure 1). The findings therefore describe countries that report mechanism-specific homicide, not the world. No country was removed because of an extreme rate.

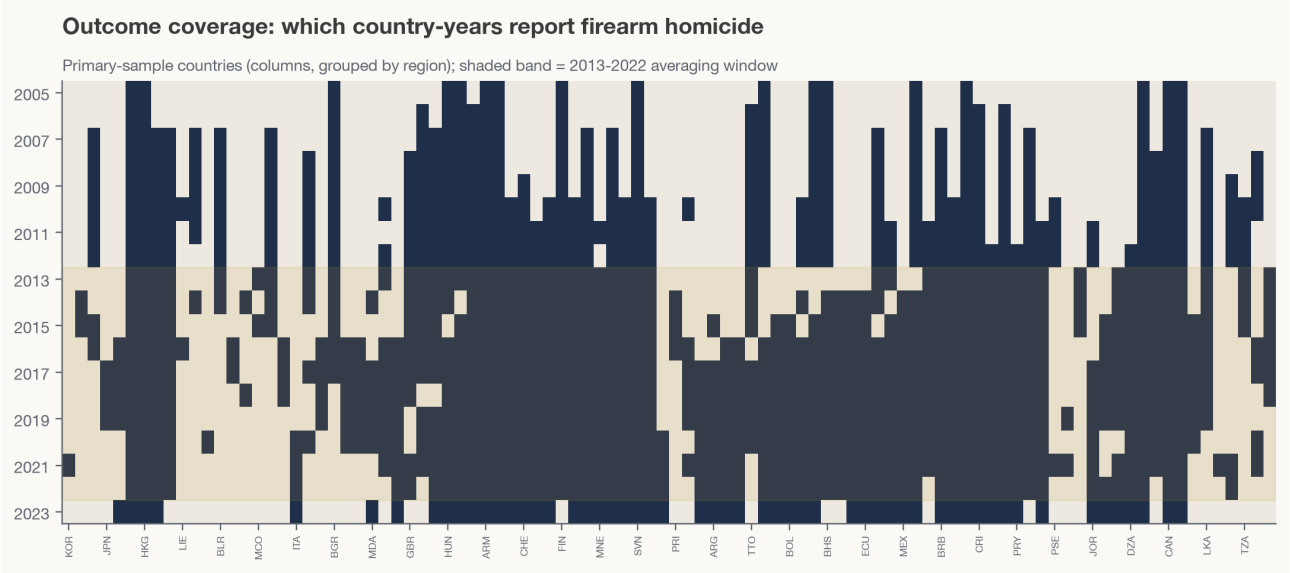


Figure 1. Which primary-sample countries report firearm homicide in each year, 2005-2023; the shaded band is the 2013-2022 averaging window.

Table 2. Descriptive statistics, primary sample (window means; pairwise non-missing).

Variable	n	mean	sd	min	median	max
Firearms per 100 residents	96	13.38	15.00	0.20	10.00	120.50
Firearm homicide rate	96	4.55	8.82	0.00	0.38	38.36
Total homicide rate	94	8.20	11.61	0.26	2.49	50.01
Firearm share of homicides (%)	92	31.77	25.76	0.00	26.76	89.89
GDP per capita (US$)	96	24,800	32,192	834.31	10,728	177,689
Gini index	76	36.88	7.21	24.79	35.70	54.60
Unemployment (%)	88	8.29	5.49	0.65	6.77	25.55
Urban population (%)	96	66.75	21.02	14.54	69.53	100.00
Rule of law (WGI)	96	0.34	0.90	-1.28	0.37	2.04

4. Methodology

4.1 Hypotheses

H1 tests whether firearms per 100 residents has a non-zero conditional association with firearm homicide; H2 whether socioeconomic characteristics add explanatory or predictive information; H3 whether the association differs across regions; H4 whether a non-linear form fits better; H5 whether ownership adds out-of-sample predictive information once other features are known; and H6 whether model residuals are spatially dependent. Directions were not assumed.

4.2 Regression models

Ownership enters per 10 firearms per 100 residents. The baseline linear association (Model 1) and its log-outcome analogue (Model 2) are

$$y_i = \beta_0 + \beta_1 \mathrm{Firearms}_i/10 + \varepsilon_i, \qquad \log(1+y_i) = \beta_0 + \beta_1 \mathrm{Firearms}_i/10 + \varepsilon_i \tag{1}$$

where y_i is the firearm homicide rate. Because $\log(1+y)$ is not a log of y, $100(e^{\beta_1}-1)$ is interpreted as the percentage change in $(1+y)$, which approximates a percentage change in y only when rates are well above one. Models 3-5 add covariates sequentially:

$$\log(1+y_i) = \beta_0 + \beta_1 \mathrm{Firearms}_i/10 + \mathbf{x}_i'\boldsymbol{\gamma} + \alpha_{r(i)} + \delta\,\mathrm{Gov}_i + \varepsilon_i \tag{2}$$

with socioeconomic controls $\mathbf{x}_i$ (log GDP per capita, urban share, unemployment, males aged 15-64), World Bank region effects α_r (Model 4) and a governance index (Model 5), the first principal component of the four WGI estimates (89% of their variance). Gini is available for fewer countries and enters a separate specification. Inference uses HC3 standard errors (MacKinnon & White, 1985). Model 6 interacts ownership with region groups, high income and above-median inequality; Model 7 uses restricted cubic splines, a GAM (Wood, 2017) and a hinge search with bootstrap re-selection of the threshold; Model 8 is quantile regression at τ = 0.25, 0.5, 0.75, 0.9 (Koenker & Bassett, 1978); Model 9 weights by population; Model 10 has regional random intercepts (and slopes); Model 11 is a Poisson/negative binomial model for summed annual counts with a log person-years offset (Cameron & Trivedi, 2013); and Model 12 uses the country-year panel.

$$\log E[C_i] = \log(\mathrm{PY}_i) + \beta_1 \mathrm{Firearms}_i/10 + \mathbf{x}_i'\boldsymbol{\gamma} + \alpha_{r(i)} + \delta\,\mathrm{Gov}_i \tag{3}$$

Because ownership is measured only in 2017, the panel fixed-effects model ($\log(1+y_{it}) = \mu_i + \tau_t + \mathbf{z}_{it}'\boldsymbol{\theta} + u_{it}$) cannot identify an ownership coefficient; ownership is instead entered in a correlated random-effects (Mundlak, 1978) specification with country means of the time-varying covariates and standard errors clustered by country.

4.3 Machine learning

Fourteen learners are compared: median baseline, linear regression, ridge, lasso, elastic net, Huber, random forest (Breiman, 2001), extra trees, gradient boosting, XGBoost (Chen & Guestrin, 2016), LightGBM (Ke et al., 2017), CatBoost (Prokhorenkova et al., 2018), support vector regression and a penalised-spline GAM. Features are firearm prevalence, eight socioeconomic and demographic indicators, four WGI estimates, income-group rank and region; country names and every outcome-derived variable are excluded. Missing covariates are median-imputed with missingness indicators, and scaling and encoding are fitted inside each training fold. Models are trained on $\log(1+y)$ and back-transformed. Generalisation is estimated by nested cross-validation: 5 repeats of 5-fold outer CV (25 test folds) with a 3-fold randomised hyperparameter search inside each training fold. Two stress tests complement this: leave-one-region-out validation without region indicators, and a single stratified 80/20 hold-out used once as a final test.

Feature-group ablations A-H (ownership only; socioeconomic only; institutional only; ownership + socioeconomic; socioeconomic + institutional; all; all except ownership; all except region) are evaluated with ridge, random forest and gradient boosting on the same 50 folds, so differences are paired. Interpretation uses held-out permutation importance, SHAP values (Lundberg & Lee, 2017), partial dependence and ICE curves, and accumulated local effects, which are less distorted by correlated predictors (Apley & Zhu, 2020).

4.4 Spatial methods and robustness

Global Moran's I (Moran, 1950) and local Moran's I (Anselin, 1995) are computed with 999 permutations under four weight definitions: 4 and 6 nearest neighbours by great-circle distance, inverse distance within 2,500 km, and polygon contiguity. Lagrange multiplier tests for spatial lag and error dependence follow Anselin (1988), and a spatial lag model is estimated by spatial two-stage least squares (Kelejian & Prucha, 1998). The robustness matrix re-estimates the adjusted ownership coefficient under 34 specifications covering weighting, outcome transformations, influential and conflict-affected countries, firearm-estimate quality, region and income classifications, covariate sets, multiple imputation with Rubin's (1987) rules, robust and quantile estimators, functional form, clustered inference, spatial dependence and alternative temporal windows.

5. Descriptive Results

5.1 Firearm prevalence

Across the 225 locations with an estimate, the median is 5.4 civilian firearms per 100 residents, the mean 9.8 and the skewness 3.9 (Figure 2). The United States (120.5) holds 46% of the estimated global civilian stock, and ten countries hold 76%. Among countries with outcome data, the next highest values are Montenegro (39.1), Serbia, Canada and Uruguay; the lowest are in East Asia, where Japan and the Republic of Korea report fewer than one firearm per 100 residents (Figure 3).

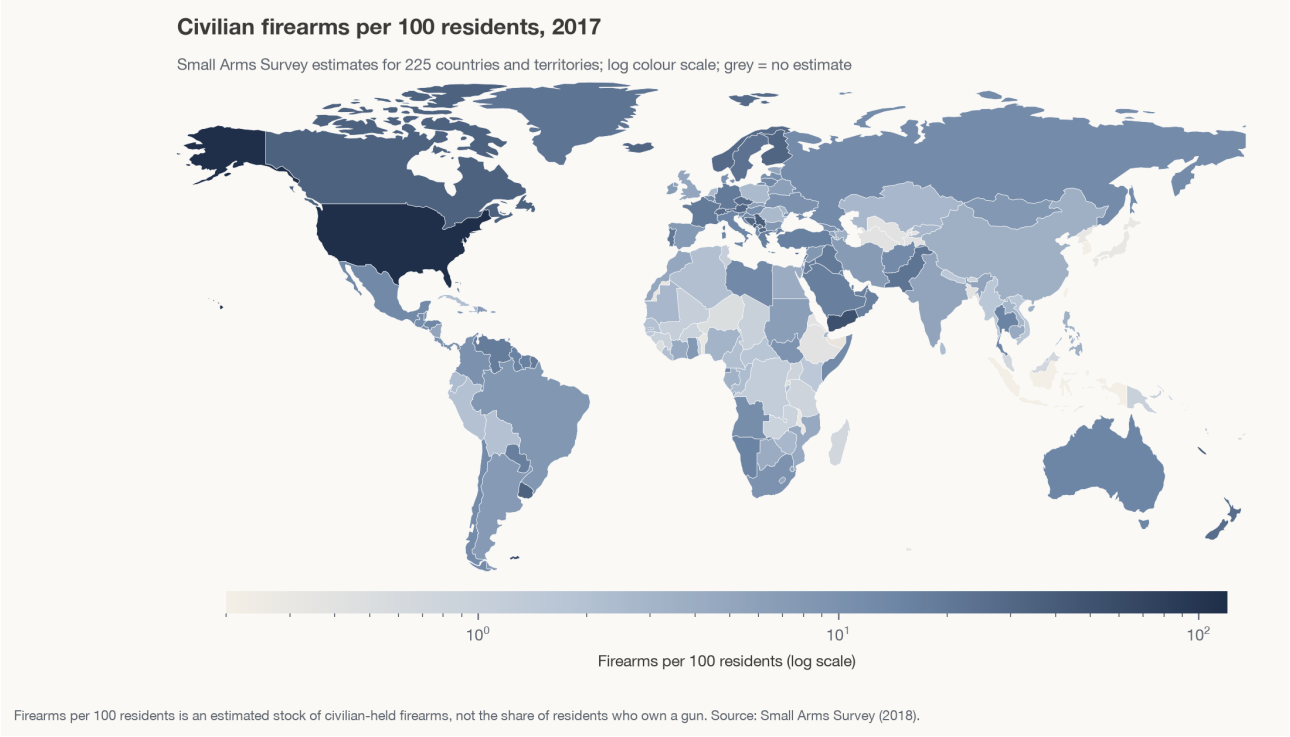


Figure 3. Civilian firearms per 100 residents, 2017 (Small Arms Survey). Log colour scale; grey = no estimate.

5.2 Firearm homicide

The primary-sample median firearm homicide rate is 0.38 per 100,000 and the mean 4.55; the population-weighted mean is 3.39. All twenty highest rates are in Latin America and the Caribbean, led by El Salvador (38.4), Jamaica (38.3) and Honduras (37.0); 6 small or East Asian jurisdictions report zero. Counts and rates rank countries differently: Mexico (186,594 firearm homicides over the observed window years) and the United States (117,429) dominate counts, whereas small Caribbean states dominate rates (Figure 4, Figure 5).

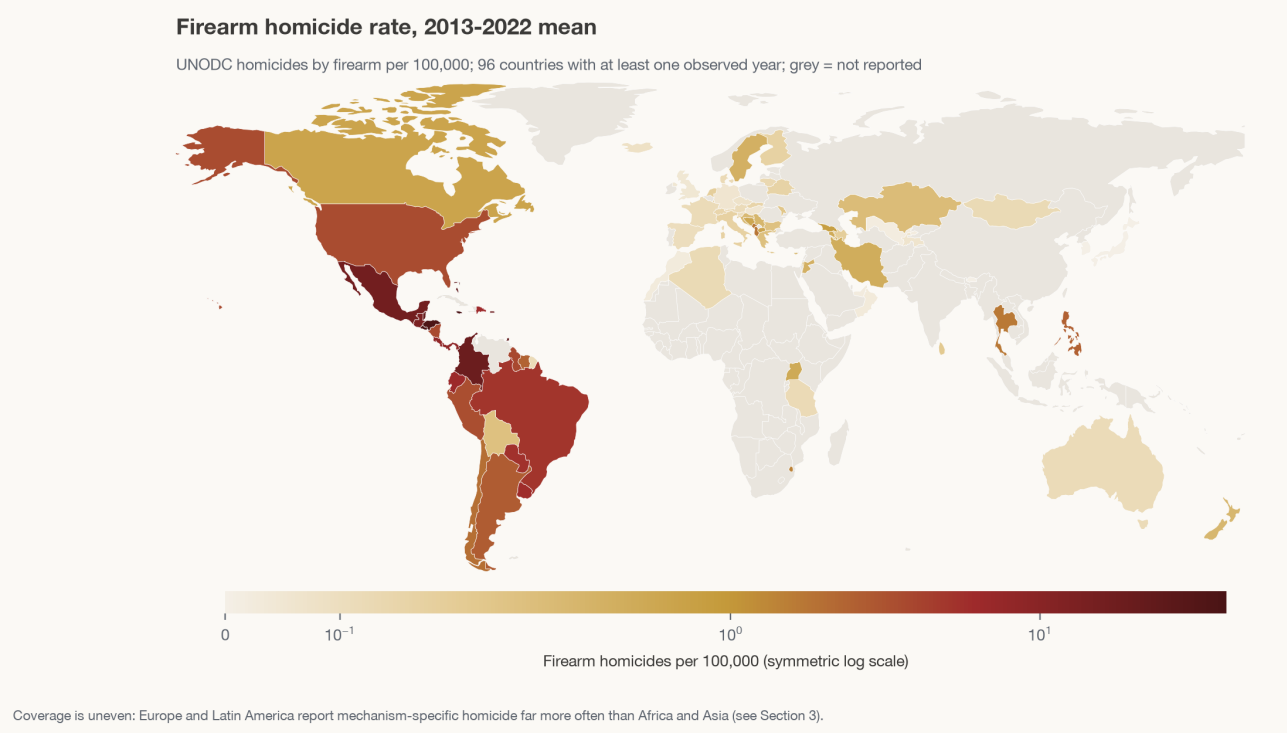


Figure 4. Firearm homicide rate per 100,000, 2013-2022 mean (UNODC). Symmetric-log colour scale; grey = not reported.

Table 3. Regional summary, primary sample (Spearman ρ shown only where n >= 4).

Region	n	Median firearms/100	Median firearm homicide	Pop.-weighted firearm homicide	Spearman ρ
Latin America & Caribbean	31	8.80	7.45	9.82	0.10
North America	3	34.70	3.52	3.22	
Sub-Saharan Africa	5	4.80	0.52	0.32	-0.05
Middle East & North Africa	8	12.00	0.24	0.27	0.19
Europe & Central Asia	38	14.25	0.17	0.14	0.14
East Asia & Pacific	9	3.60	0.12	0.98	0.66
South Asia	2	1.60	0.11	0.21	

Regional differences in homicide are far larger than regional differences in firearm prevalence (Figure 6, Figure 7). The Kruskal-Wallis test across regions gives $H = 60.9$ ($p < 0.001$).

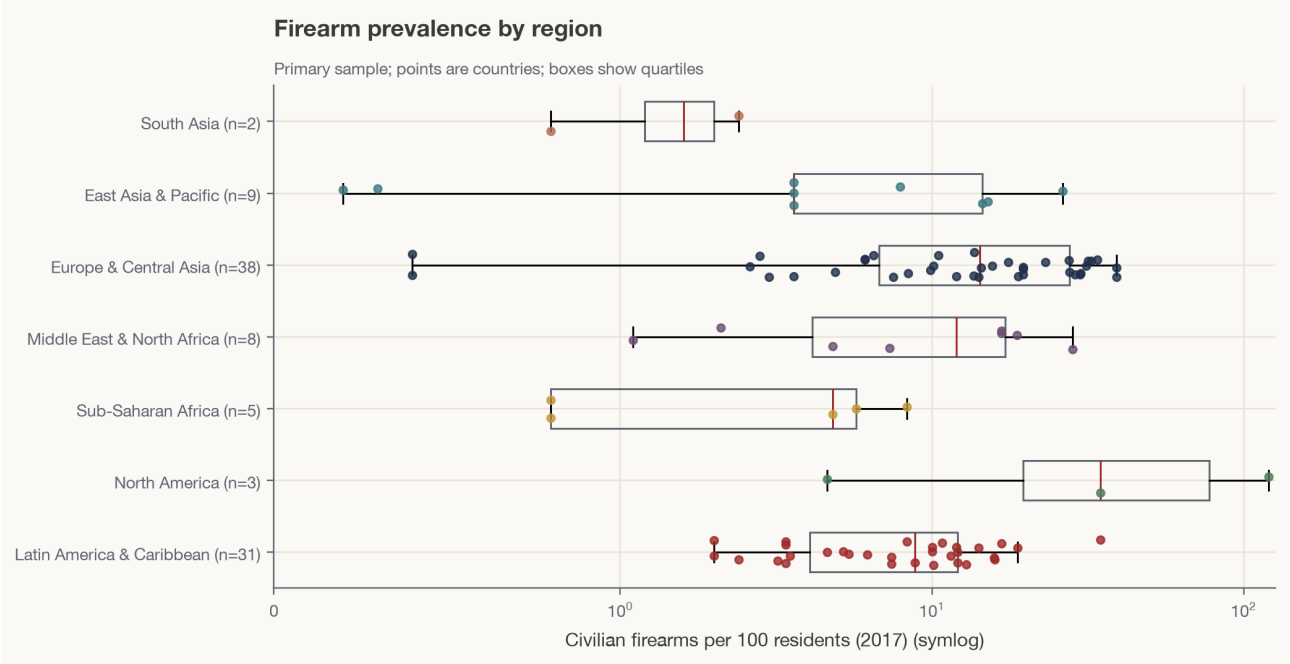


Figure 6. Firearm prevalence by World Bank region (primary sample).

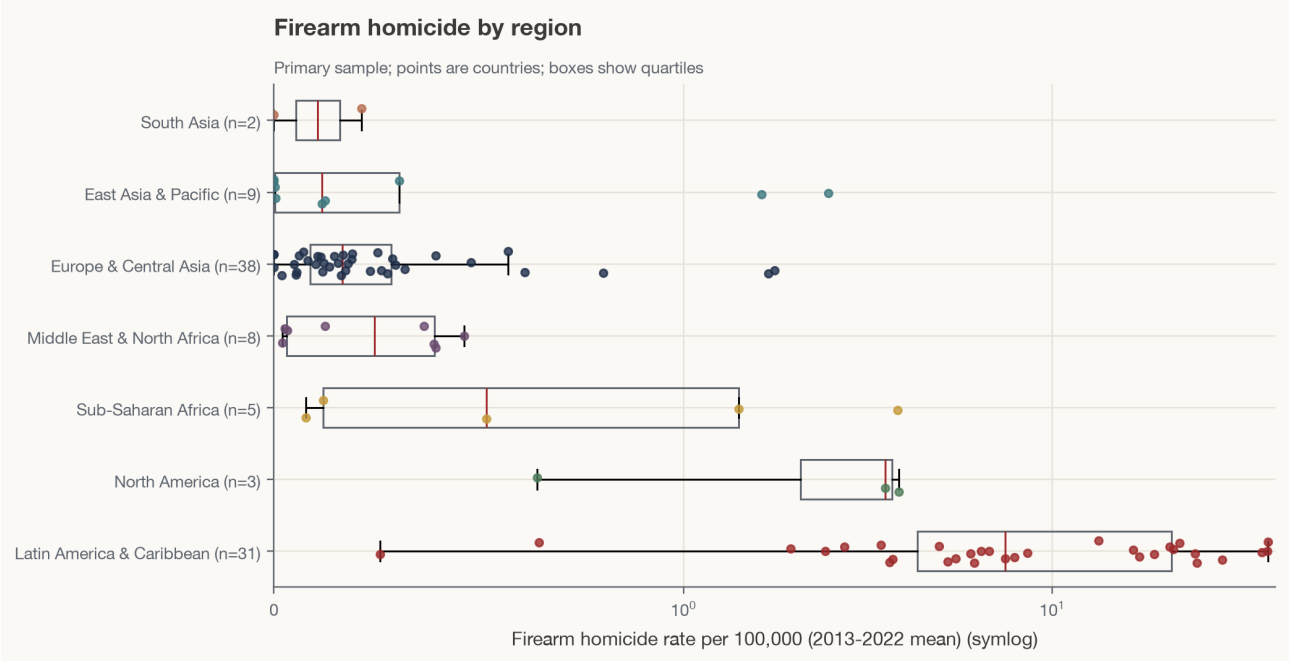


Figure 7. Firearm homicide by World Bank region (primary sample, symmetric-log axis).

5.3 Ownership and homicide

The unadjusted cross-national association is flat (Figure 8): Pearson r = -0.09 (p = 0.378), Spearman ρ = 0.01 (p = 0.921) and Kendall τ = 0.01. Permutation and bootstrap inference agree (Table 4). On log scales the cloud remains flat (r = 0.05; Figure 9), and the ownership quartiles have similar medians (Kruskal-Wallis p = 0.222; Figure 10).

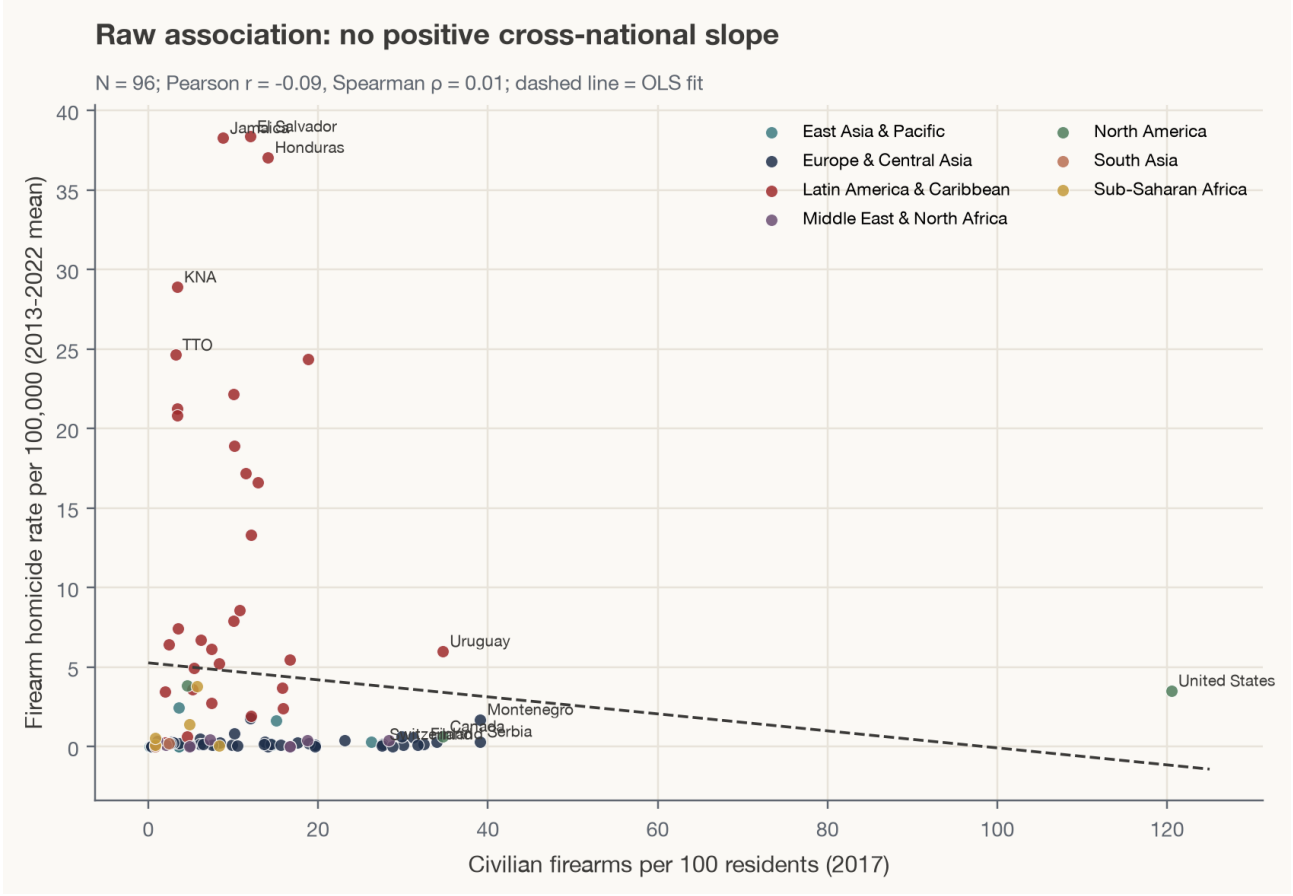


Figure 8. Firearm prevalence and firearm homicide, primary sample. Dashed line: OLS fit.

Table 4. Formal hypothesis tests with Holm-adjusted p-values (bootstrap 95% CIs in brackets).

Test	n	Stat.	Value	p	Holm p	Effect
Pearson correlation (rate)	96	r	-0.091	0.378	1.000	r = -0.09 [-0.21, -0.01]
Spearman correlation	96	rho	0.010	0.921	1.000	rho = 0.01 [-0.18, 0.21]
Kendall tau	96	tau	0.014	0.839	1.000	tau = 0.01
Welch t-test: top-quartile vs other ownership (log1p rate)	96	t	-2.815	0.007	0.065	Cohen's d = -0.59 [-1.06, -0.15]
Mann-Whitney U: top-quartile vs other ownership (rate)	96	U	699.000	0.164	1.000	rank-biserial r = -0.19
Kruskal-Wallis: firearm homicide across regions	96	H	60.947	<0.001	<0.001	epsilon^2 (approx.) = 0.62
Kruskal-Wallis: firearm homicide across ownership quartiles	96	H	4.389	0.222	1.000	epsilon^2 (approx.) = 0.02
Kruskal-Wallis: ownership across regions	96	H	14.783	0.011	0.102	
Partial correlation (ownership, log1p rate; controlling socioeconomic + governance)	88	r	0.063	0.562	1.000	partial r = 0.06
Nested F: socioeconomic block added to ownership (log1p)	88	F	1.111	0.357	1.000	delta R2 = 0.05
Nested F: region FE + governance added to M3 (log1p)	88	F	35.091	<0.001	<0.001	delta R2 = 0.73
Joint Wald: socioeconomic block in Model 5 (HC3)	88	chi2	5.998	0.199	1.000	
Joint Wald: region effects in Model 5 (HC3)	88	chi2	159.225	<0.001	<0.001	

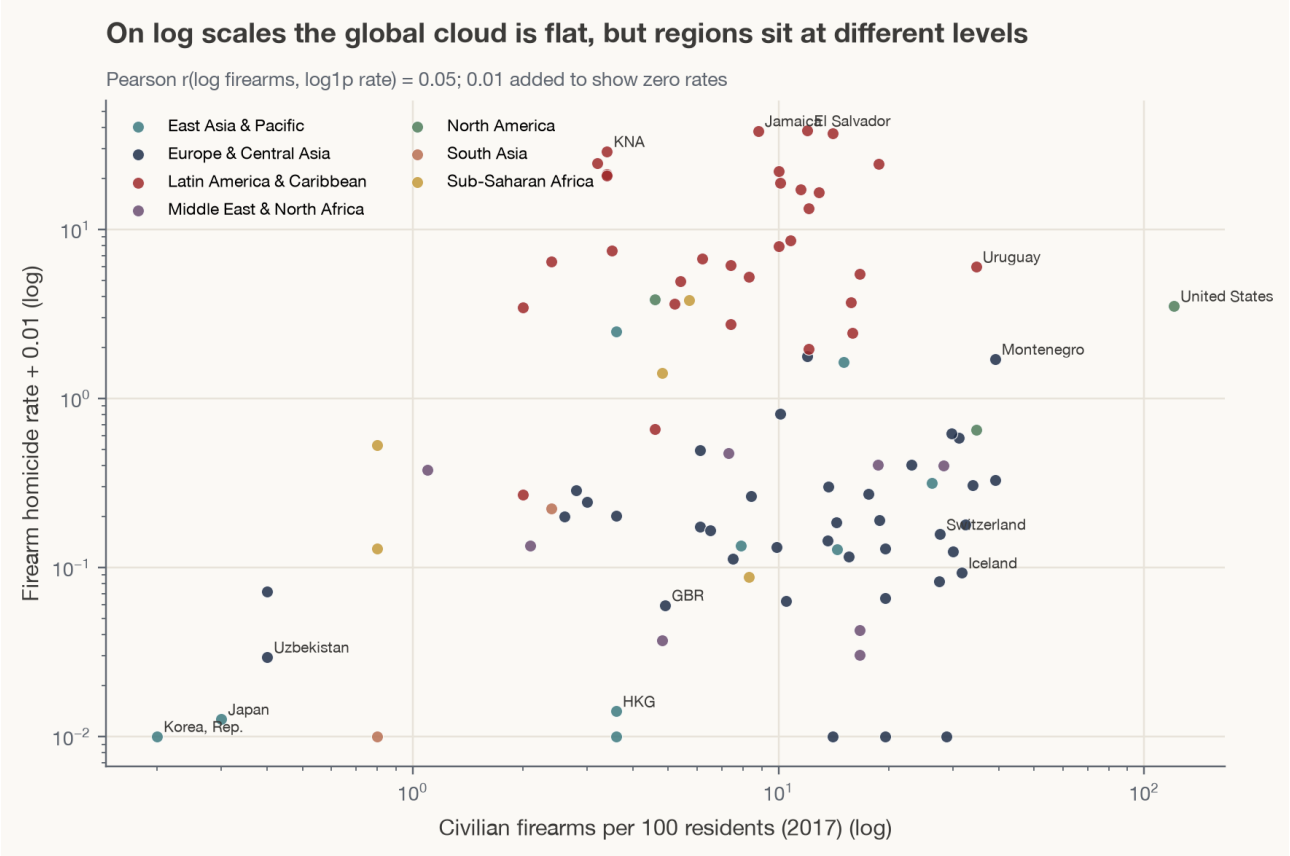


Figure 9. Log-log view; 0.01 added to show zero rates. Colours denote regions.

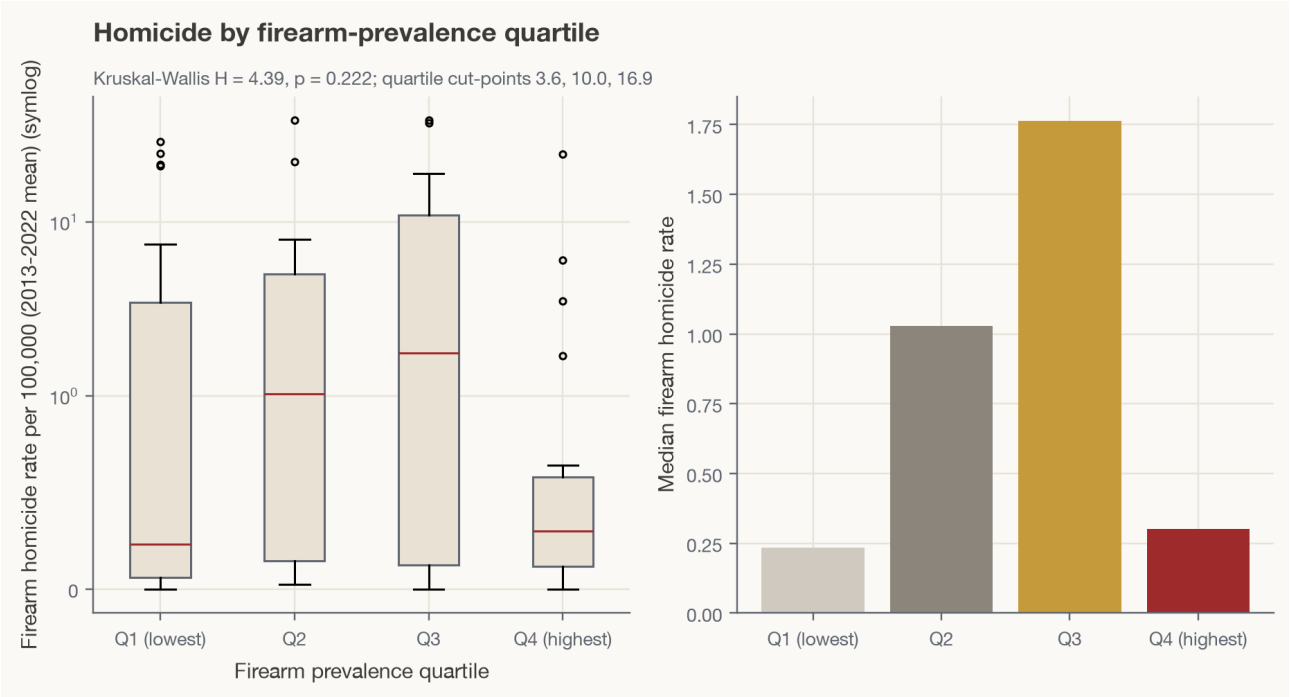


Figure 10. Firearm homicide by firearm-prevalence quartile.

The flat global pattern hides structure. Excluding Latin America and the Caribbean, the Pearson correlation is 0.31 ($p = 0.011$, $n = 65$); among high-income countries outside that region it is 0.52 ($p < 0.001$, $n = 37$; Spearman 0.34), consistent with the high-income literature. Within regions, rank correlations are positive but small and imprecise: 0.14 in Europe and Central Asia ($n = 38$) and 0.10 in Latin America and the Caribbean ($n = 31$) (Figure 11, Figure 12). Population weighting does not change the conclusion (Figure 13).



Figure 11. Within-region scatterplots with OLS fits on log firearm prevalence.

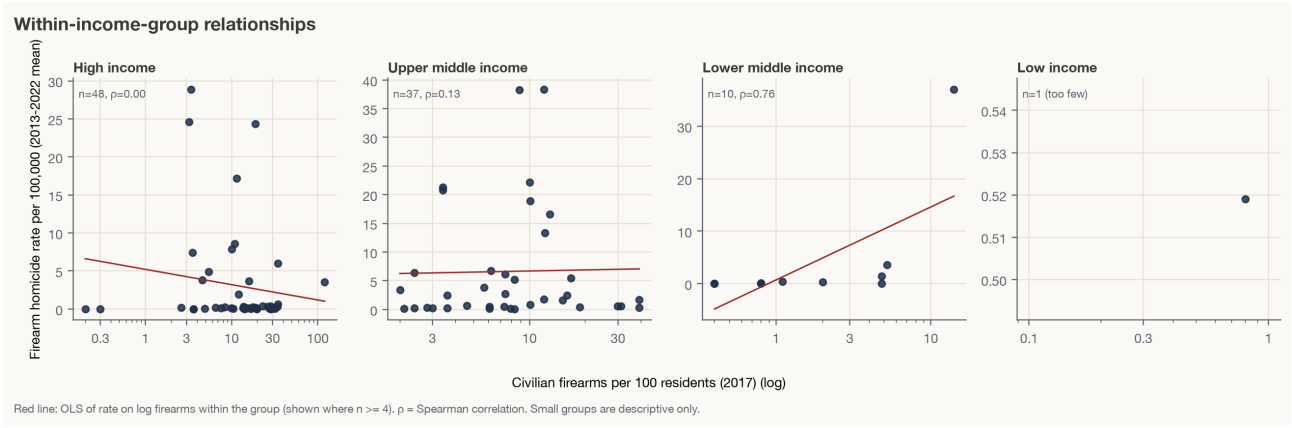


Figure 12. Within-income-group scatterplots.

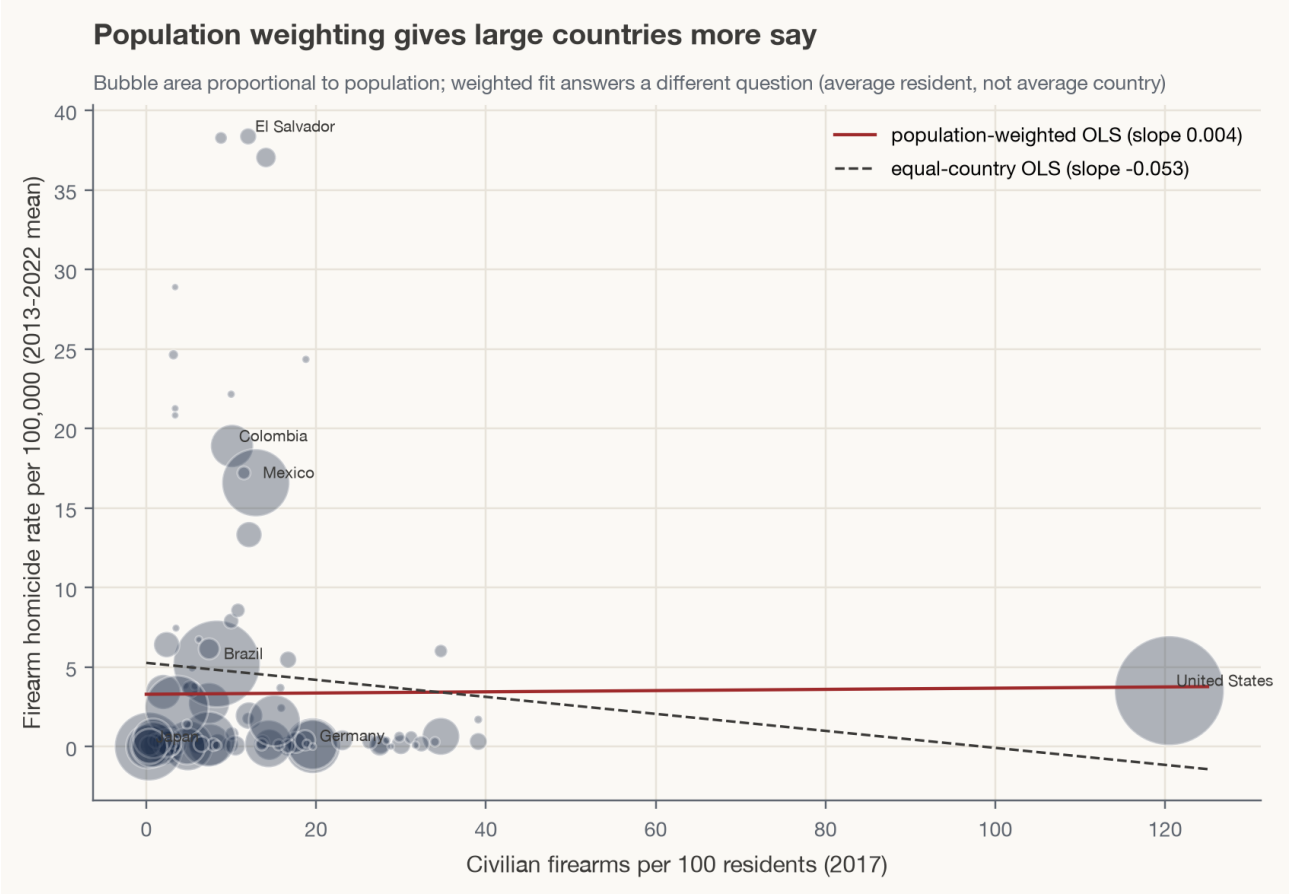


Figure 13. Population-weighted versus equal-country fits; bubble area proportional to population.

The secondary outcomes sharpen the picture (Figure 14). Firearm prevalence is unrelated to total homicide (ρ = -0.04, n = 162) but positively related to the **share** of homicides committed with firearms (ρ = 0.27, p = 0.008; log-scale r = 0.47) and negatively related to the non-firearm homicide rate (ρ = -0.30, p = 0.003). Descriptively, where firearms are more common a larger fraction of killings involve them, without a corresponding difference in overall homicide. This pattern is compatible with weapon substitution but also with reporting and classification differences; it is not evidence of either.

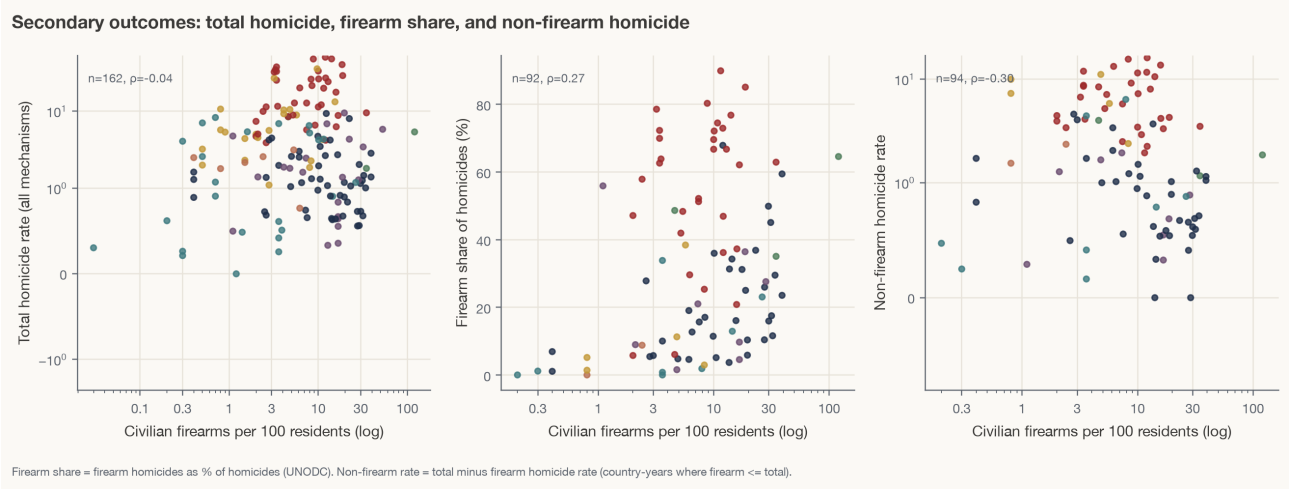


Figure 14. Secondary outcomes against firearm prevalence: total homicide, firearm share, and non-firearm homicide.

5.4 Socioeconomic and institutional correlates

Firearm homicide is more strongly related to inequality (ρ = 0.65, n = 76) and governance (rule of law ρ = -0.37) than to income (ρ = -0.30) or urbanisation (ρ = -0.20) (Figure 15, Figure 16, Figure 17, Figure 18). The four WGI indicators are nearly collinear (Figure 19), which is why they are combined into one index in the main models.

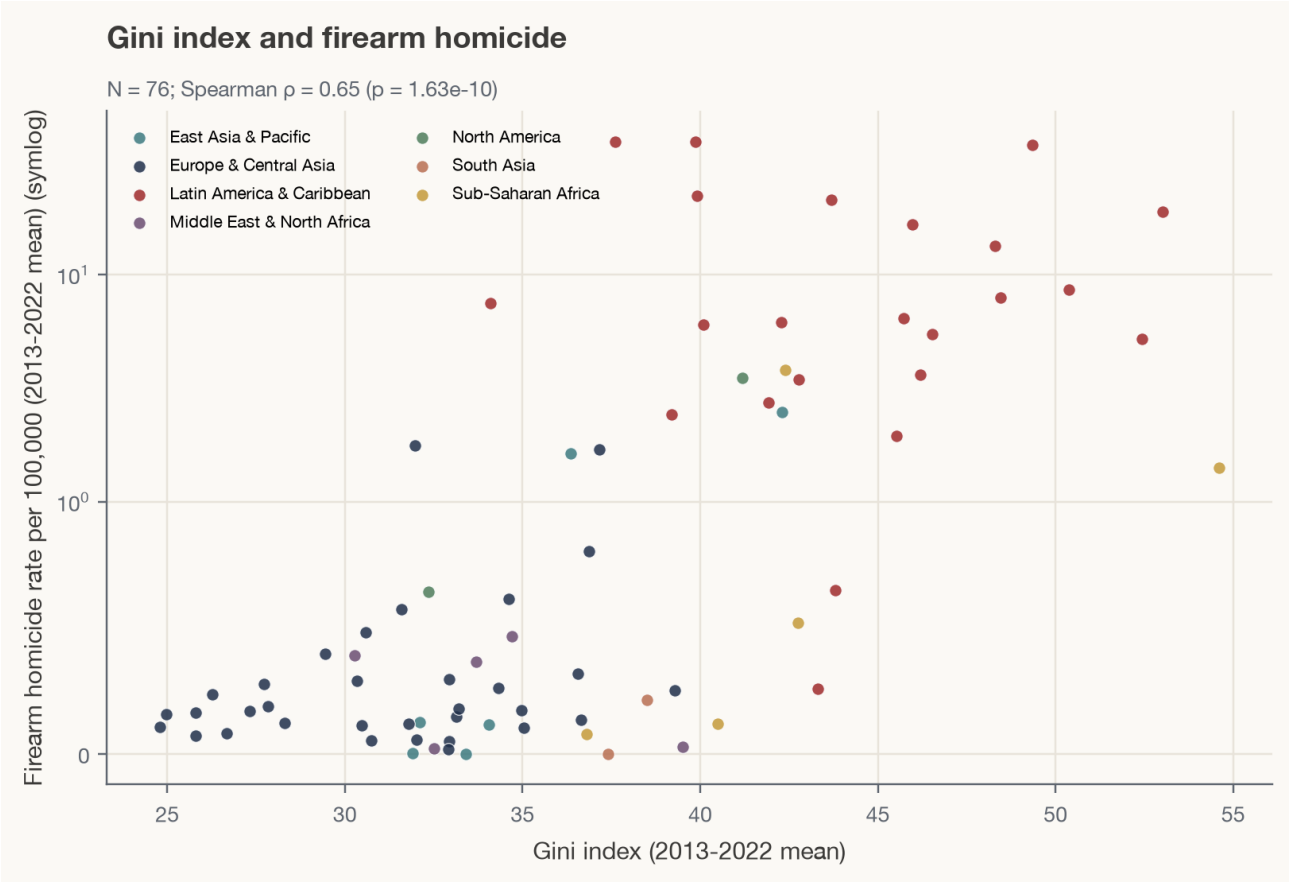


Figure 15. Income inequality (Gini) and firearm homicide.

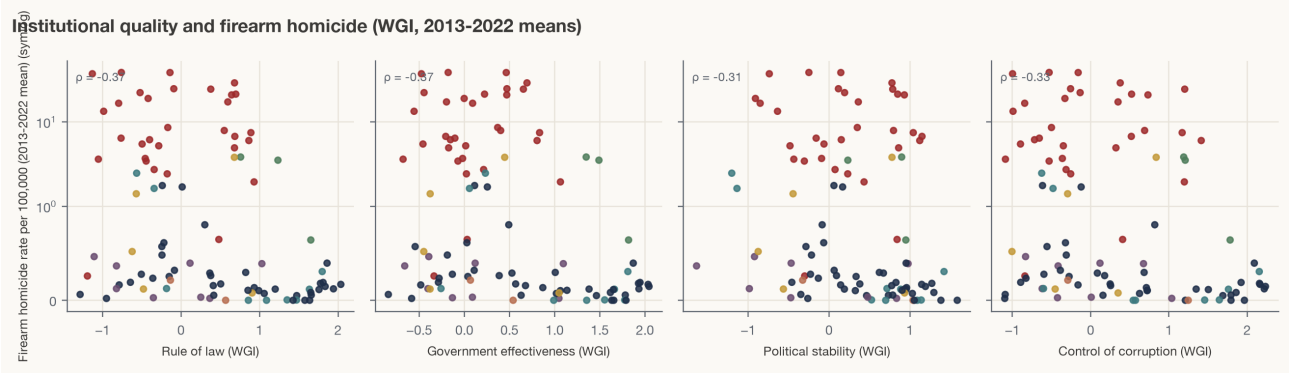


Figure 16. Worldwide Governance Indicators and firearm homicide.

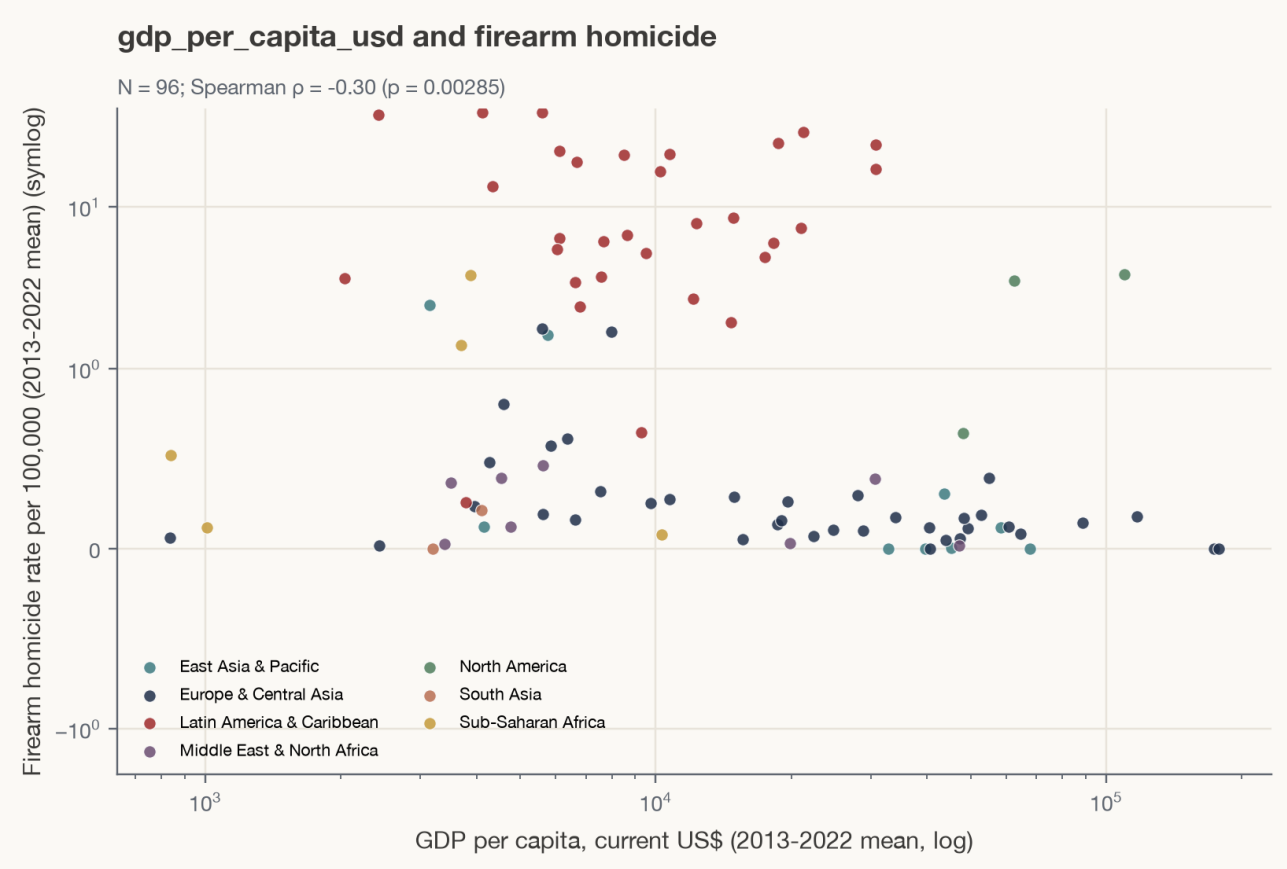


Figure 17. GDP per capita and firearm homicide.

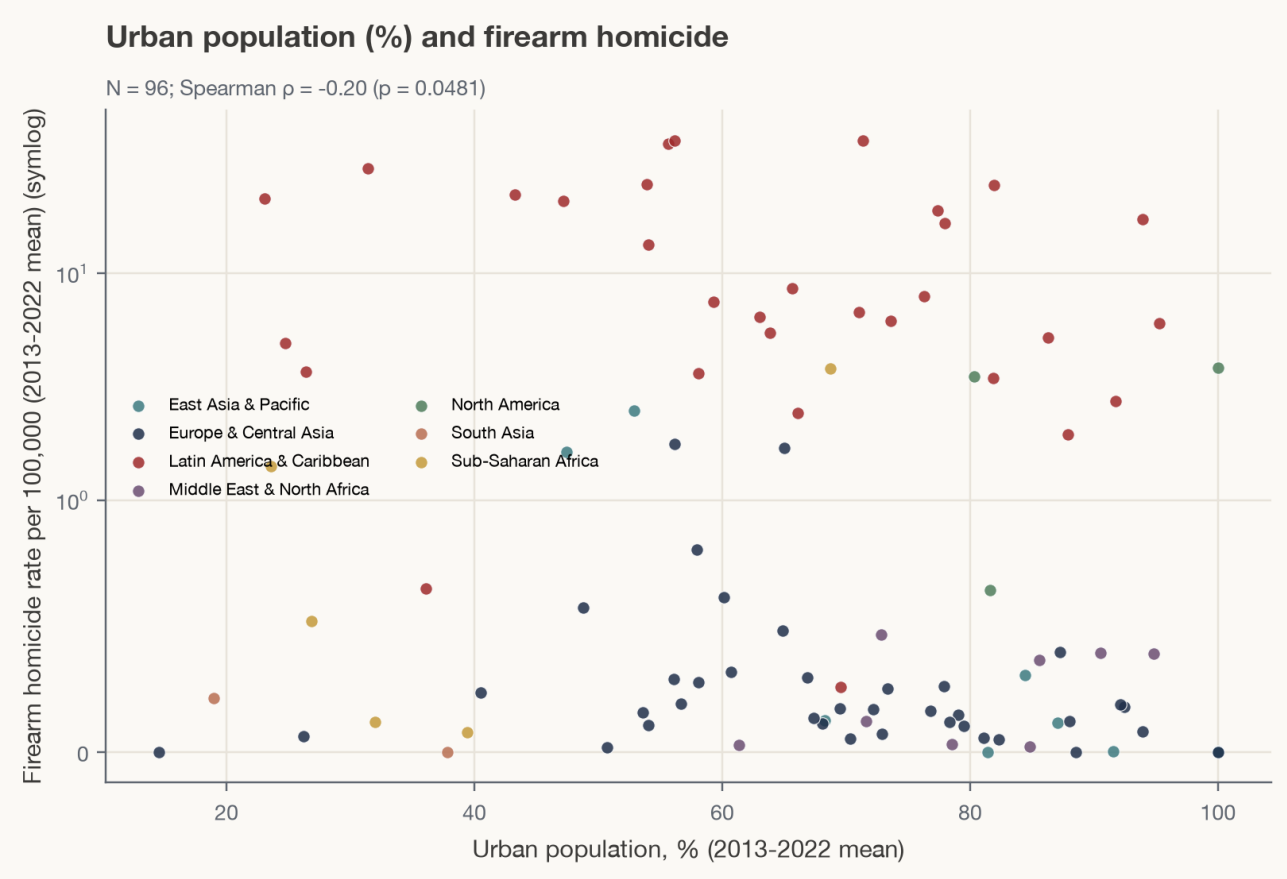


Figure 18. Urbanisation and firearm homicide.



Figure 19. Spearman correlation matrix of outcomes and predictors (primary sample).

5.5 Country profiles

Splitting at the medians (10.0 firearms per 100 and 0.38 firearm homicides per 100,000) gives four profiles of similar size (Figure 20). *High ownership, low homicide* includes Andorra, United Arab Emirates, Australia, Austria, Switzerland, Cyprus, Czechia, Germany; *low ownership, high homicide* includes Argentina, Armenia, Antigua and Barbuda, Bermuda, Brazil, Barbados, Cabo Verde, Dominica. Among the top quartile of firearm prevalence, firearm homicide ranges from zero to the high thirties per 100,000 (Figure 21).

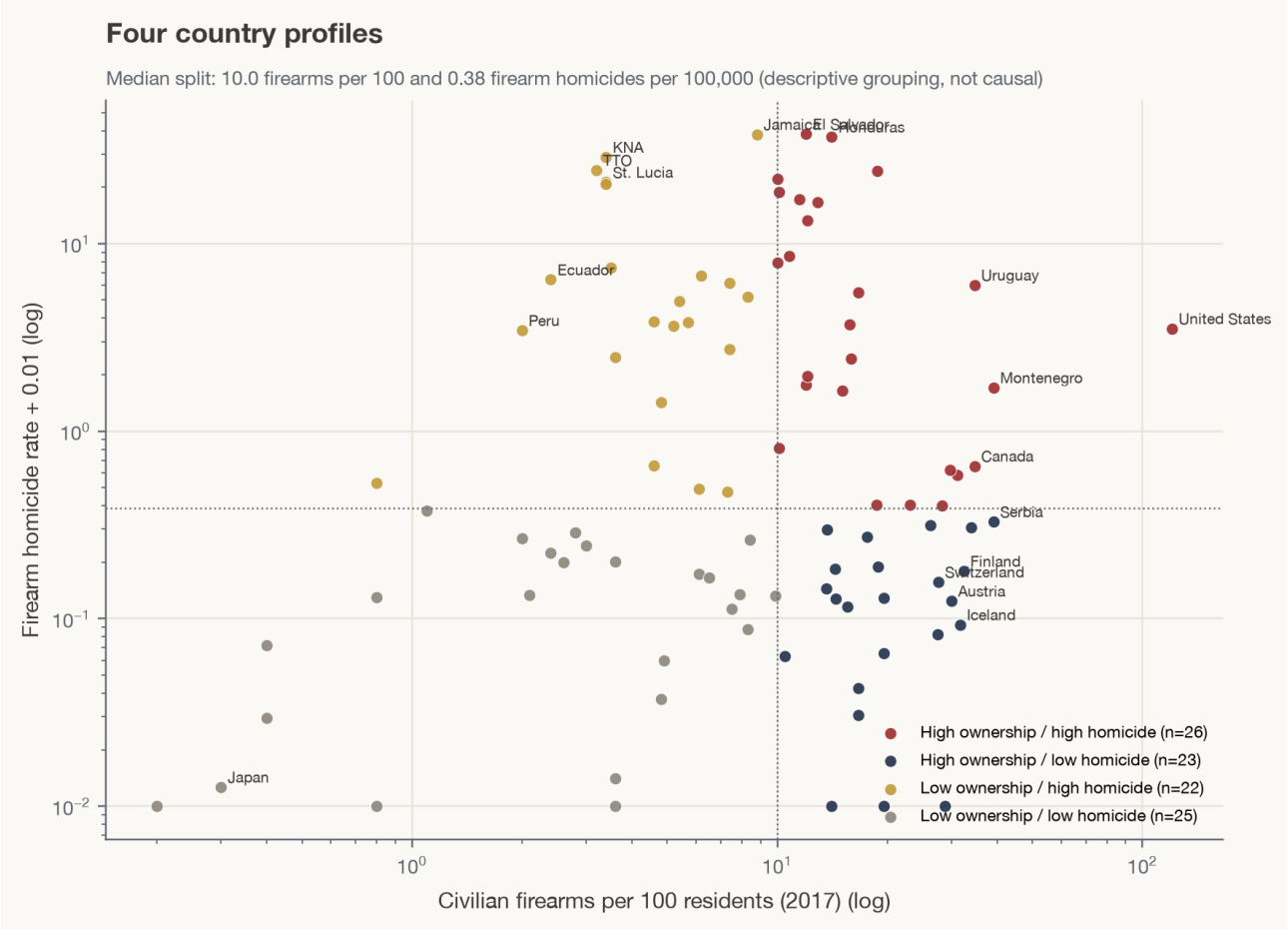


Figure 20. Four descriptive country profiles defined by median splits (not causal categories).

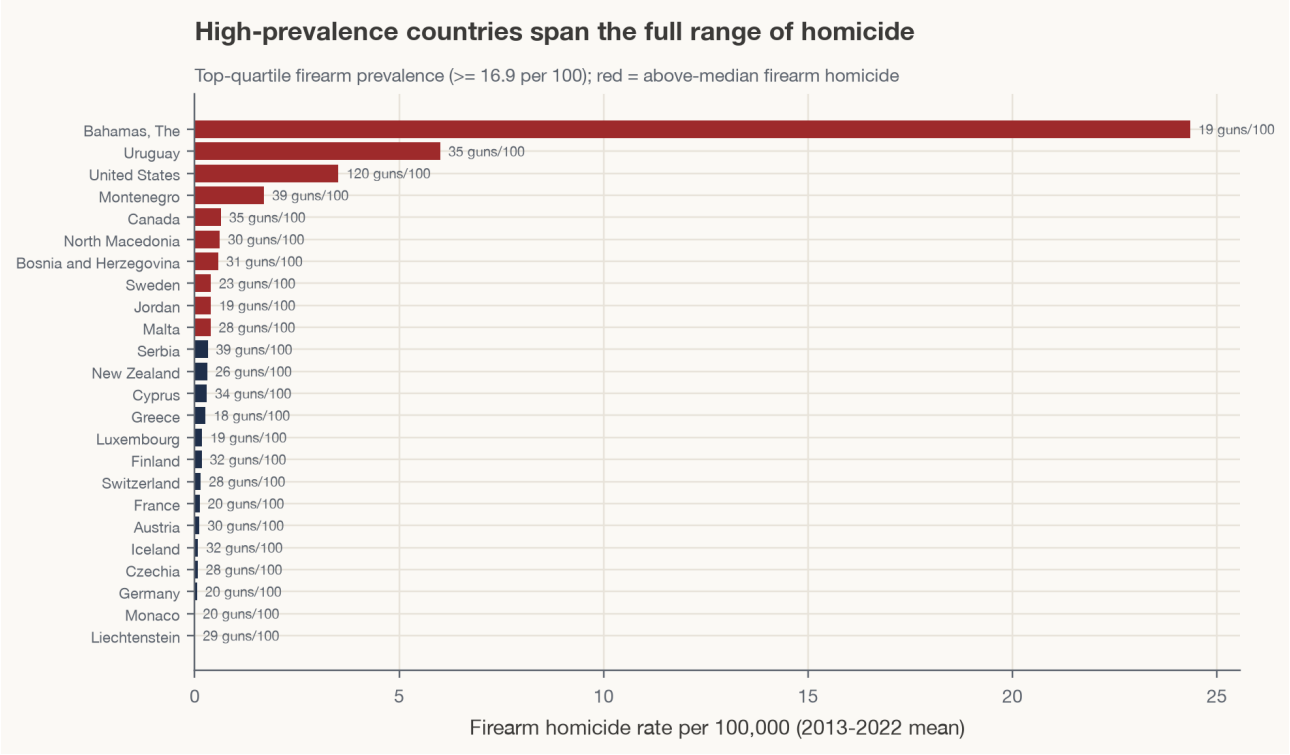


Figure 21. Firearm homicide among the top quartile of firearm prevalence.

Where a genuine panel exists, firearm homicide trends are regionally distinct; the Latin American median is persistently an order of magnitude above the European median (Figure 22). Ownership is not shown as a trend because it is measured once. As historical context, total homicide in Western Europe fell by one to two orders of magnitude since the late Middle Ages (Figure 23).

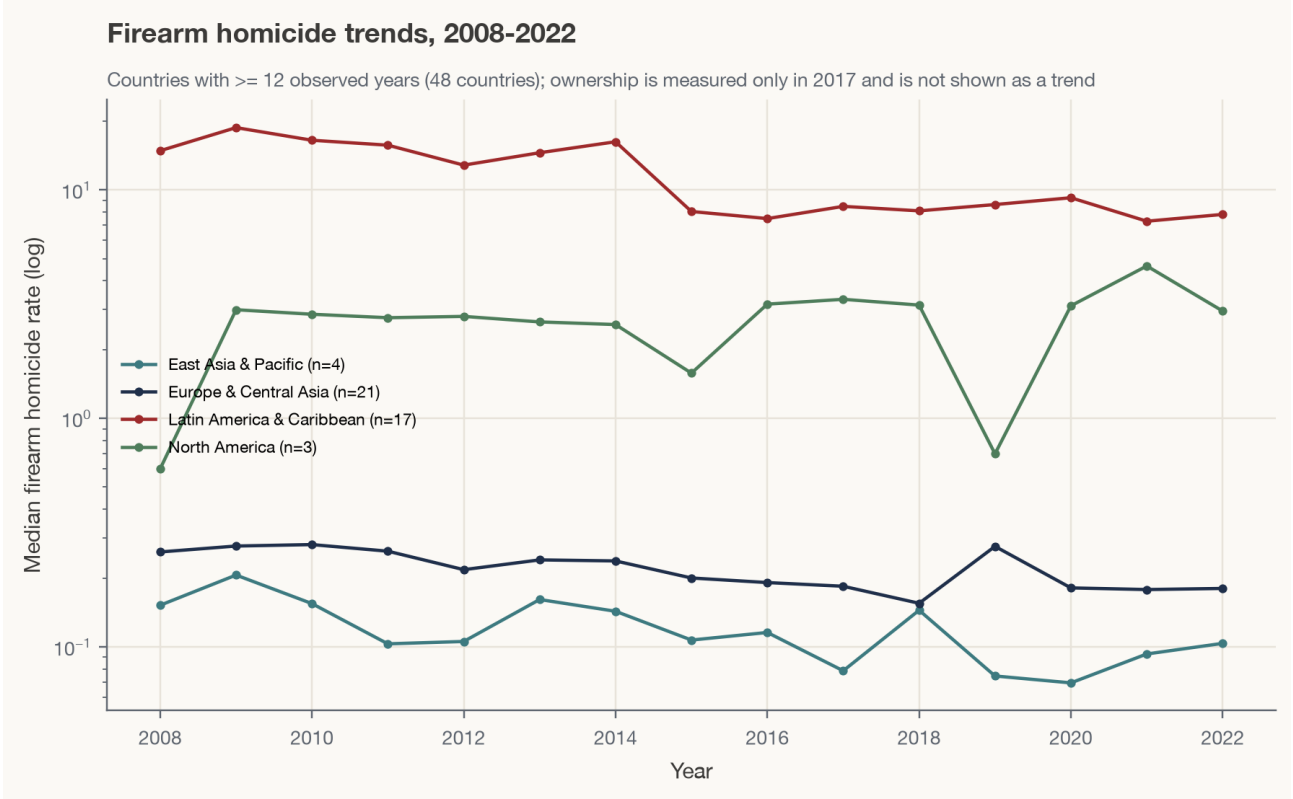


Figure 22. Median firearm homicide rate by region, 2008-2022, countries with at least 12 observed years.

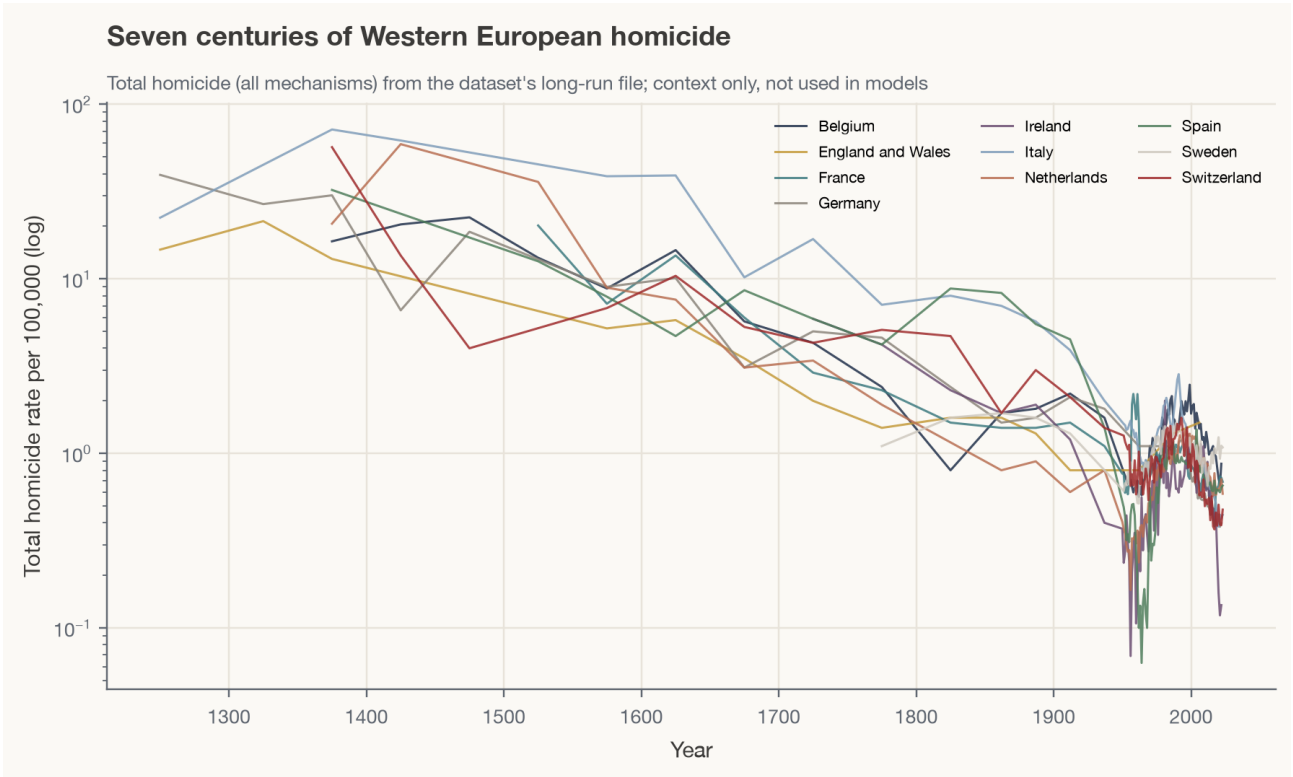


Figure 23. Total homicide rates in Western Europe, 1250-2023 (long-run file; context only).

6. Econometric Results

6.1 Unadjusted and adjusted associations

Table 5. Ownership coefficient across Models 1-5 (HC3 confidence intervals). M3g adds the Gini index on its smaller sample.

Spec	n	Coef. per +10	95% CI	p	R^2	% change in (1+y)
M1 (level, S1)	96	-0.535	[-1.682, +0.613]	0.361	0.008	
M1 (level, S2)	88	-0.395	[-1.302, +0.513]	0.394	0.005	
M3 (level, S2)	88	-0.030	[-0.914, +0.854]	0.947	0.029	
M3g (level, SG)	75	+0.127	[-1.092, +1.346]	0.838	0.208	
M4 (level, S2)	88	+0.288	[-0.478, +1.053]	0.461	0.496	
M5 (level, S2)	88	+0.283	[-0.511, +1.078]	0.485	0.497	
M2 (log1p, S1)	96	-0.041	[-0.275, +0.193]	0.731	0.003	-4.0%
M2 (log1p, S2)	88	-0.012	[-0.197, +0.172]	0.897	0.000	-1.2%
M3 (log1p, S2)	88	+0.050	[-0.118, +0.219]	0.559	0.052	+5.2%
M3g (log1p, SG)	75	+0.053	[-0.072, +0.178]	0.408	0.454	+5.4%
M4 (log1p, S2)	88	+0.091	[+0.009, +0.172]	0.029	0.778	+9.5%
M5 (log1p, S2)	88	+0.091	[+0.012, +0.170]	0.024	0.778	+9.5%

Without adjustment, an additional 10 firearms per 100 residents is associated with a -0.53 difference in the firearm homicide rate (95% CI -1.68 to +0.61; p = 0.361). Socioeconomic controls move the log-scale coefficient from -0.041 to +0.050; the decisive change comes from region effects, which move it to +0.091 (p = 0.029). In Model 5 the coefficient is **0.091 (95% CI 0.012 to 0.170; p = 0.024)**, i.e. 9.5% higher (1 + rate) per 10 firearms per 100 (CI 1.2% to 18.5%). For a country at the sample median rate (0.38), this corresponds to roughly 0.13 additional firearm homicides per 100,000 per 10 firearms per 100. In levels the adjusted coefficient is +0.28 per 100,000 (CI -0.51 to +1.08; p = 0.485) (Figure 24).

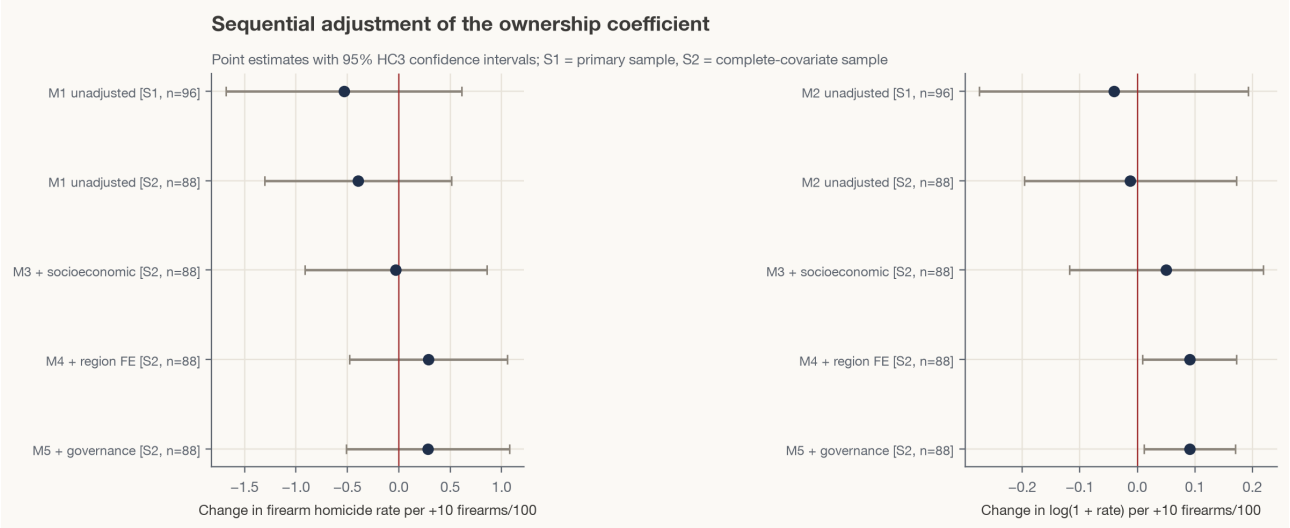


Figure 24. Sequential adjustment of the ownership coefficient on the rate (left) and log(1 + rate) (right) scales.

The adjustment pattern has a clear interpretation: Latin America and the Caribbean has both moderate firearm prevalence and very high homicide, so pooling regions masks a positive within-region gradient. Region effects are jointly decisive (Wald $\chi^2(6) = 159$, $p < 0.001$); the socioeconomic block is not ($\chi^2(4) = 6.0$, $p = 0.199$). Within Model 5, governance has a coefficient of -0.042 per SD (p = 0.780) and log GDP +0.019 (p = 0.914) (Figure 25, Figure 26). Adding Gini on its smaller sample (n = 75) gives a strong inequality association (+0.095 per Gini point, $p < 0.001$) but no region effects in that specification. The maximum VIF in Model 5 is 7.7 (ownership 2.1).

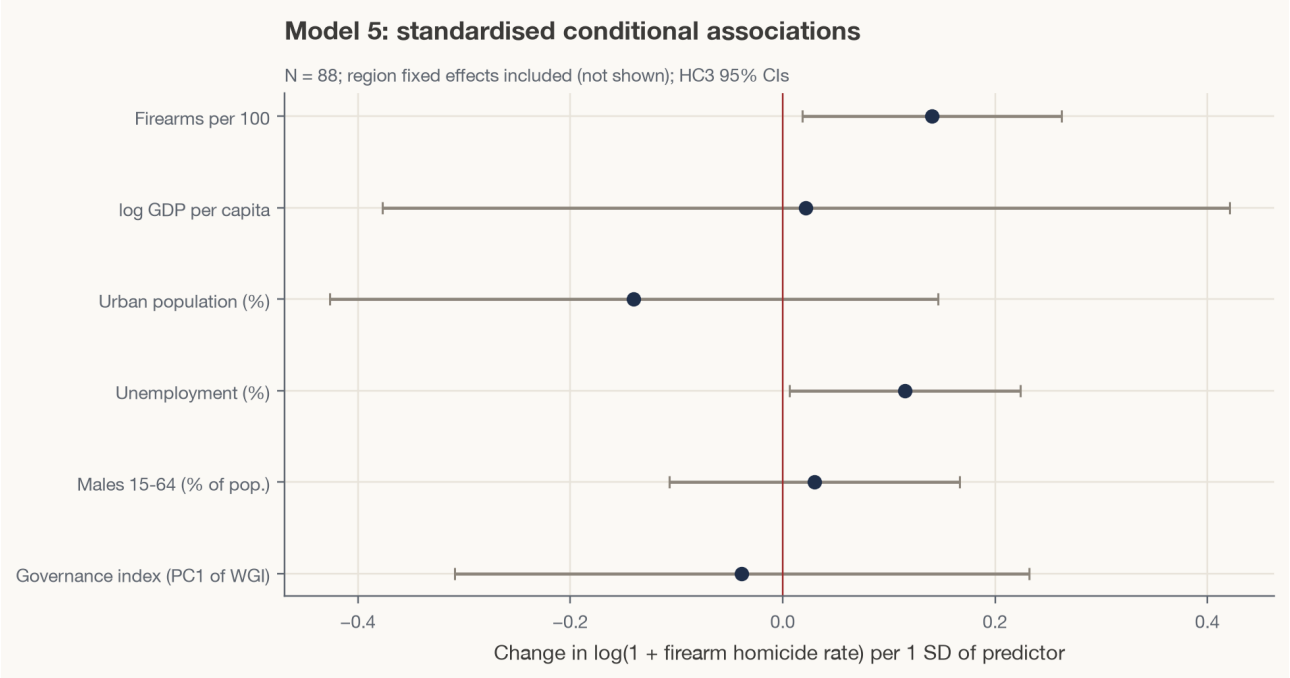


Figure 25. Model 5 standardised coefficients (change in log(1 + rate) per 1 SD), HC3 95% CIs.

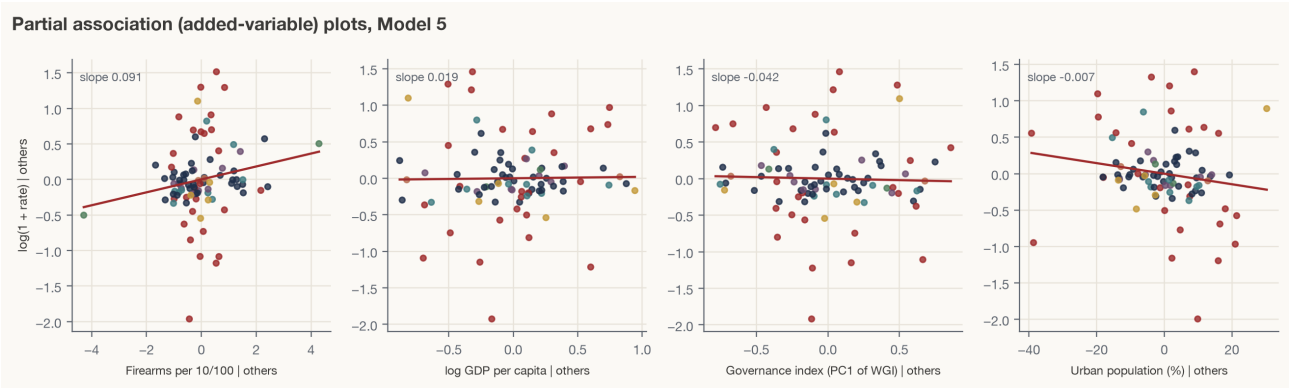


Figure 26. Added-variable plots from Model 5.

Diagnostics favour the log outcome (Figure 27): residual skewness is -0.09 versus 1.67 in levels, and R^2 is 0.78 versus 0.50. Breusch-Pagan tests still indicate heteroskedasticity ($p = 0.004$), justifying HC3 errors; the RESET test does not reject the functional form ($p = 0.186$).



Figure 27. Residual-versus-fitted and Q-Q diagnostics for Model 5 on the level and log1p scales.

6.2 Regional heterogeneity

Interaction models give implied slopes (per +10 firearms, log scale) of +0.039 in Europe and Central Asia, +0.146 in Latin America and the Caribbean and +0.112 in other regions; the joint interaction test does not reject equal slopes (p = 0.331; p = 0.560 without the United States, whose leverage dominates the 'other' group). The income interaction (p = 0.669) and inequality interaction (p = 0.459) are also small (Figure 28). Bivariate regional slopes under two classifications are shown in Figure 29; the equal-slope test across World Bank regions with at least five countries has p = 0.995 (classical F p = 0.972). Regions differ strongly in levels, not detectably in slopes; with 5-38 countries per region, the power to detect slope differences is low.

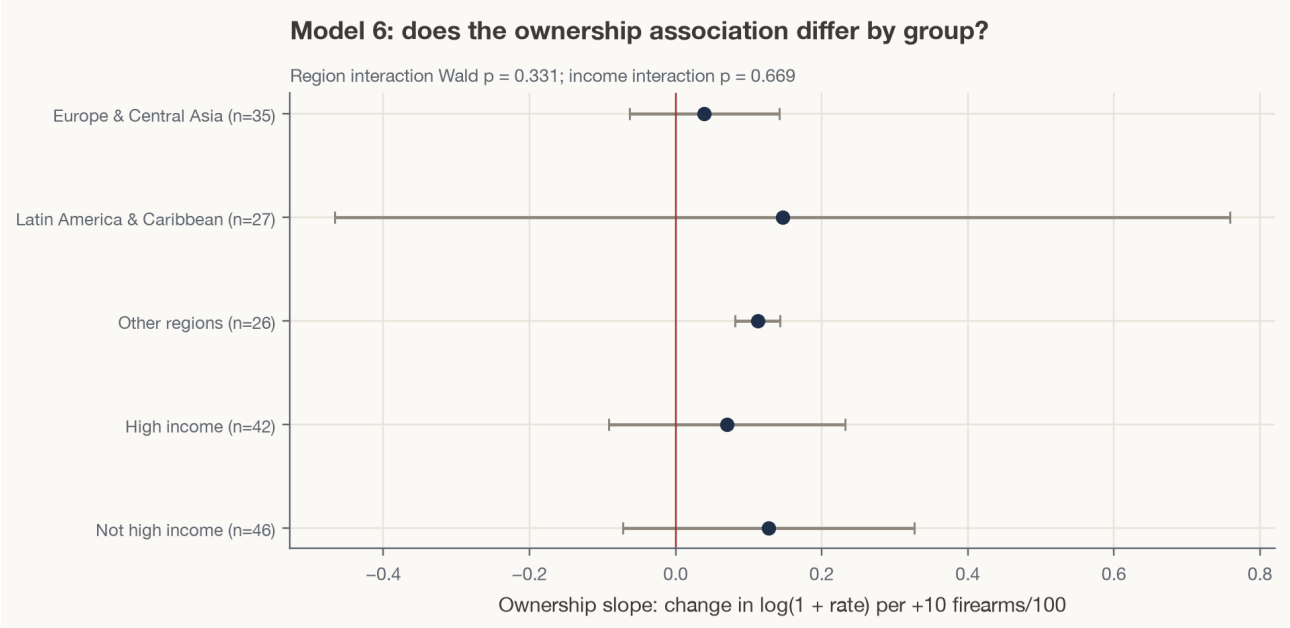


Figure 28. Implied ownership slopes from Model 6 interaction specifications.

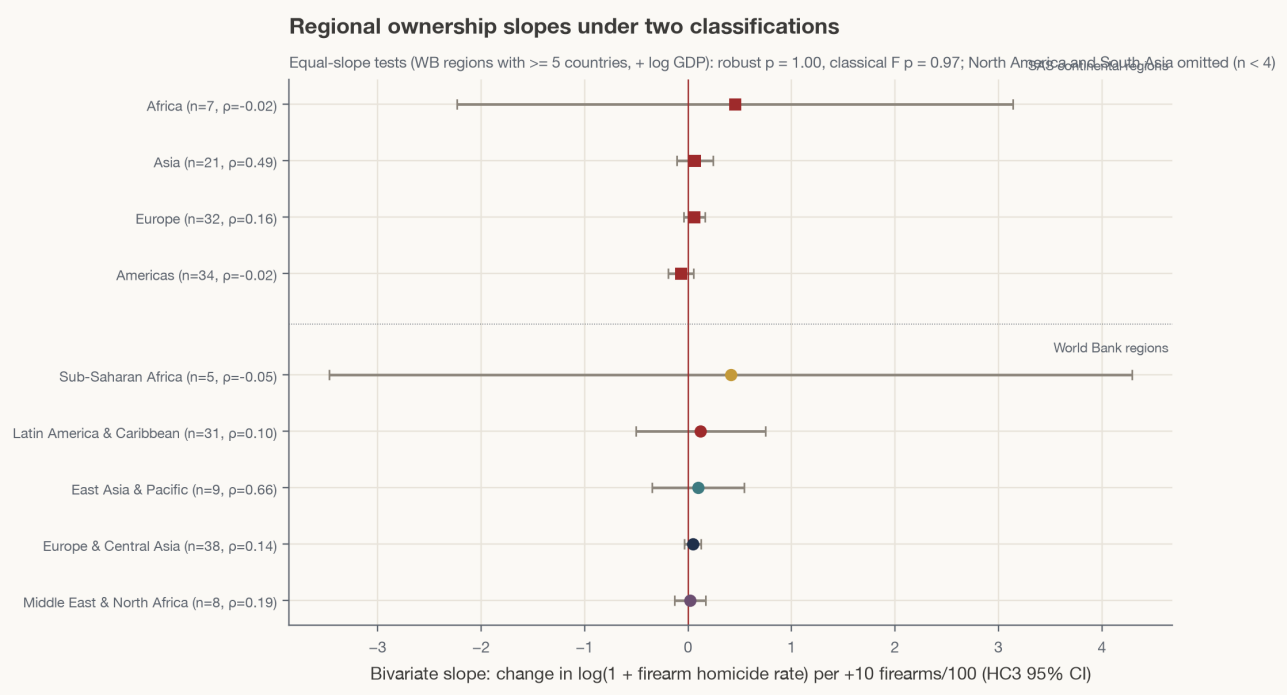


Figure 29. Bivariate ownership slopes by World Bank and Small Arms Survey region, HC3 95% CIs.

A multilevel model with regional random intercepts gives a similar ownership coefficient (0.090, 95% CI -0.013 to 0.193) and an intra-region correlation of 0.67: about two-thirds of residual variance in log rates lies between regions. The largest regional intercept is Latin America and the Caribbean (+1.65). A random-slope variance is estimable (0.0007) but weakly identified with seven groups.

6.3 Nonlinearity, quantiles, weighting and counts

Non-linear forms do not fit materially better (Figure 30). A four-df spline improves on the linear term by LR = 4.24 (p = 0.120); the GAM's AIC (162.0) is no better than its linear counterpart (160.9); and log firearm prevalence fits about as well as the linear term (AIC 158.5 vs 159.7). The best hinge is at 3.4 firearms per 100, but bootstrap re-selection places it anywhere between 2.6 and 22.3; there is no stable tipping point. Above 40 firearms per 100 the curve rests on one country.

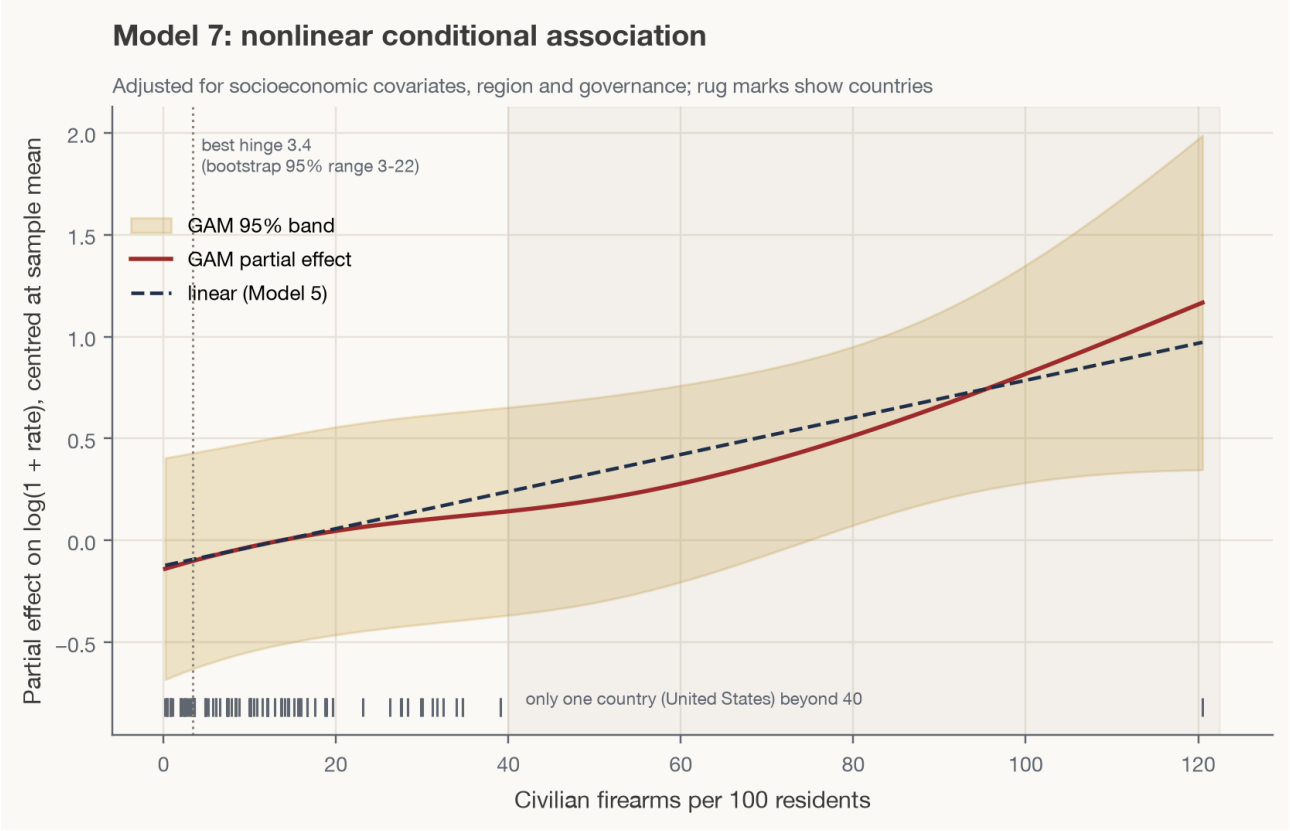


Figure 30. GAM partial effect of firearm prevalence with 95% band, linear Model 5 fit, and best hinge (bootstrap range shown).

Quantile regressions (socioeconomic and governance controls) give coefficients of +0.065 at $\tau = 0.25$, +0.126 at $\tau = 0.5$, +0.076 at $\tau = 0.75$, +0.019 at $\tau = 0.9$; none is individually distinguishable from zero (Figure 31). Population weighting changes the estimand to the average resident and makes the estimate depend heavily on a few populous countries: the weighted adjusted coefficient is 0.109 (CI -0.226 to 0.445), rising to 0.229 without the United States. The Poisson model is massively overdispersed (Pearson dispersion 2960); the negative binomial rate ratio per 10 firearms per 100 is 1.32 (95% CI 1.13-1.53; $p < 0.001$) adjusted, versus 0.89 (0.79-1.01) unadjusted.

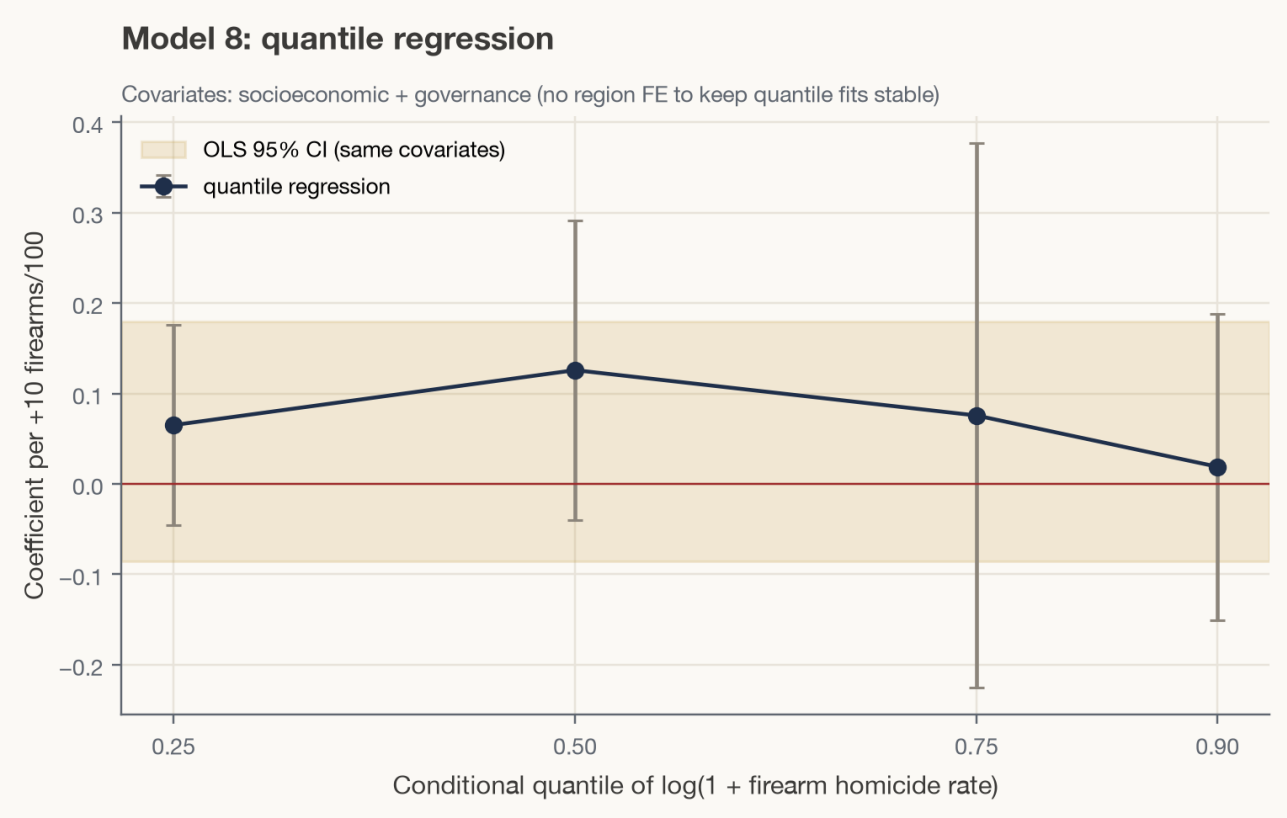


Figure 31. Quantile regression coefficients on firearm prevalence with 95% CIs; band = OLS CI with the same covariates.

In the country-year panel (982 observations, 91 countries, 2005-2023), country and year fixed effects absorb ownership. Among time-varying covariates, within-country improvements in rule of law are associated with lower firearm homicide (-0.361, CI -0.652 to -0.070). The correlated random-effects estimate for ownership, which relies on between-country variation, is +0.037 (CI -0.064 to +0.138; p = 0.470) without region effects.

7. Machine Learning Results

Table 6. Nested cross-validated performance (25 outer folds; rate scale unless noted). Validation scores, not final-test scores.

Model	MAE	MAE 95% CI (OOF bootstrap)	RMSE	R^2 (rate)	MedAE	R^2 (log1p)
Random forest	3.53	[2.25, 4.87]	7.15	-0.051	0.46	0.550
XGBoost	3.59	[2.23, 4.97]	7.46	-0.123	0.56	0.525
Lasso	3.59	[2.25, 5.09]	7.65	-0.076	0.62	0.532
Extra Trees	3.59	[2.23, 4.94]	7.24	0.017	0.48	0.552
Elastic Net	3.61	[2.19, 5.06]	7.68	-0.055	0.68	0.530
GAM (pygam)	3.66	[2.33, 5.14]	7.48	0.042	0.73	0.501
Linear regression	3.67	[2.24, 5.22]	7.78	-0.192	0.68	0.535
Gradient boosting	3.68	[2.36, 5.03]	7.48	-0.031	0.61	0.525
Huber	3.69	[2.23, 4.99]	7.70	-0.153	0.62	0.534
CatBoost	3.69	[2.30, 5.09]	7.68	-0.151	0.62	0.513
LightGBM	3.74	[2.33, 5.06]	7.40	-0.049	0.66	0.522
Ridge	3.78	[2.29, 5.27]	8.03	-0.167	0.78	0.495
Support vector regression	3.85	[2.34, 5.26]	7.89	-0.127	0.89	0.356
Dummy (median)	4.42	[2.80, 6.19]	9.11	-0.259	0.58	-0.374

The best model is **Random forest** (MAE 3.53, R^2 on the log scale 0.55), followed closely by several others (Figure 32). It beats the median baseline (4.42) clearly, but its advantage over linear regression (3.67) is small: the fold-paired MAE difference is -0.15 (2.5-97.5% fold range -1.40 to +1.13), and it is better in 56% of folds. Fold-to-fold variation is larger than most between-model differences (Figure 33). Rate-scale R^2 is modest because a handful of Caribbean and Central American countries dominate squared error; on the log scale the leading models explain roughly 55% of variance. On the single stratified hold-out (20 countries, used once) Random forest obtains MAE 2.55 (linear regression 8.36, ridge 8.23), consistent with but noisier than the cross-validated estimates.

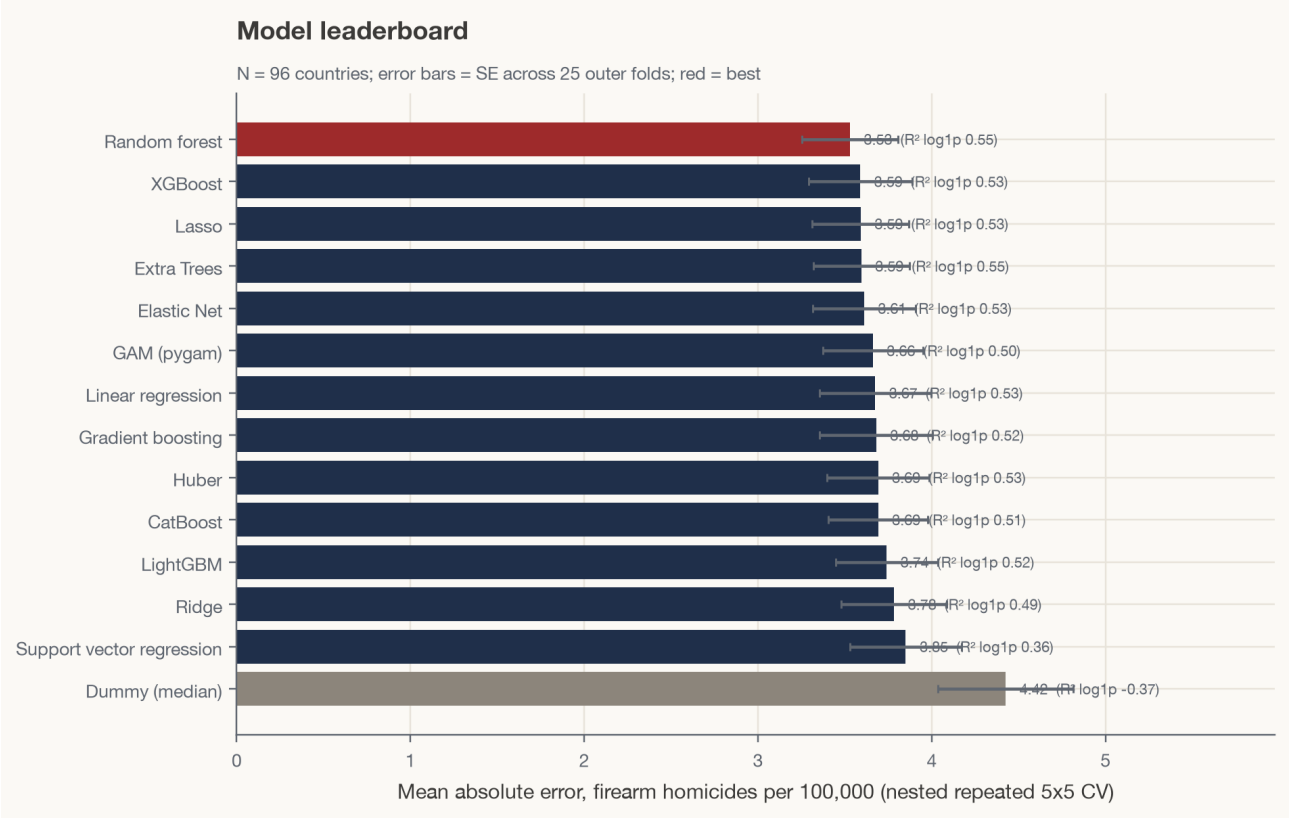


Figure 32. Nested cross-validated MAE for fourteen models; error bars are standard errors across outer folds.

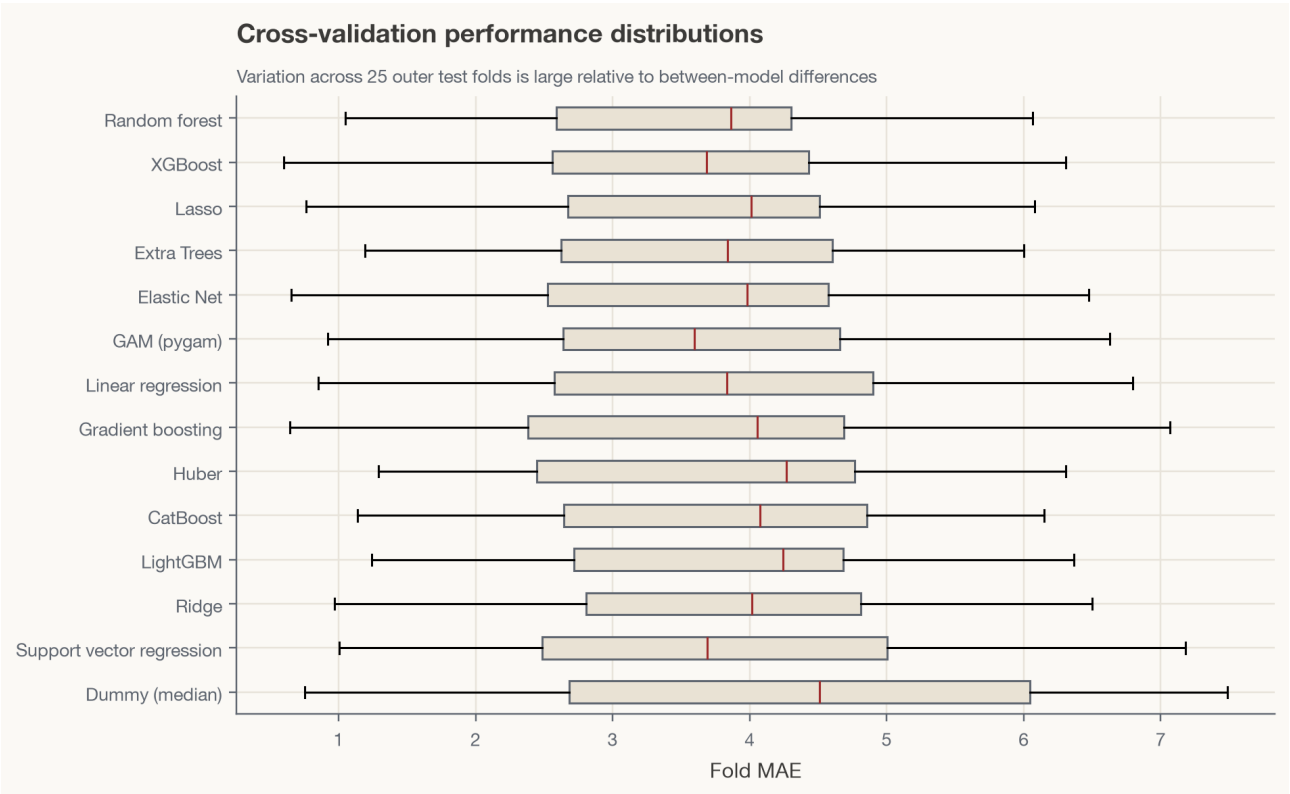


Figure 33. Distribution of fold-level MAE across the outer test folds.

Generalising to an unseen region is much harder (Figure 34). With region indicators removed and each region predicted from the others, pooled MAE is 5.31 for the random forest and 5.78 for ridge, compared with 3.53 and 3.78 under random folds. Latin America and the Caribbean, predicted from the other regions, is systematically under-predicted, which is the same regional level shift seen in the regressions.

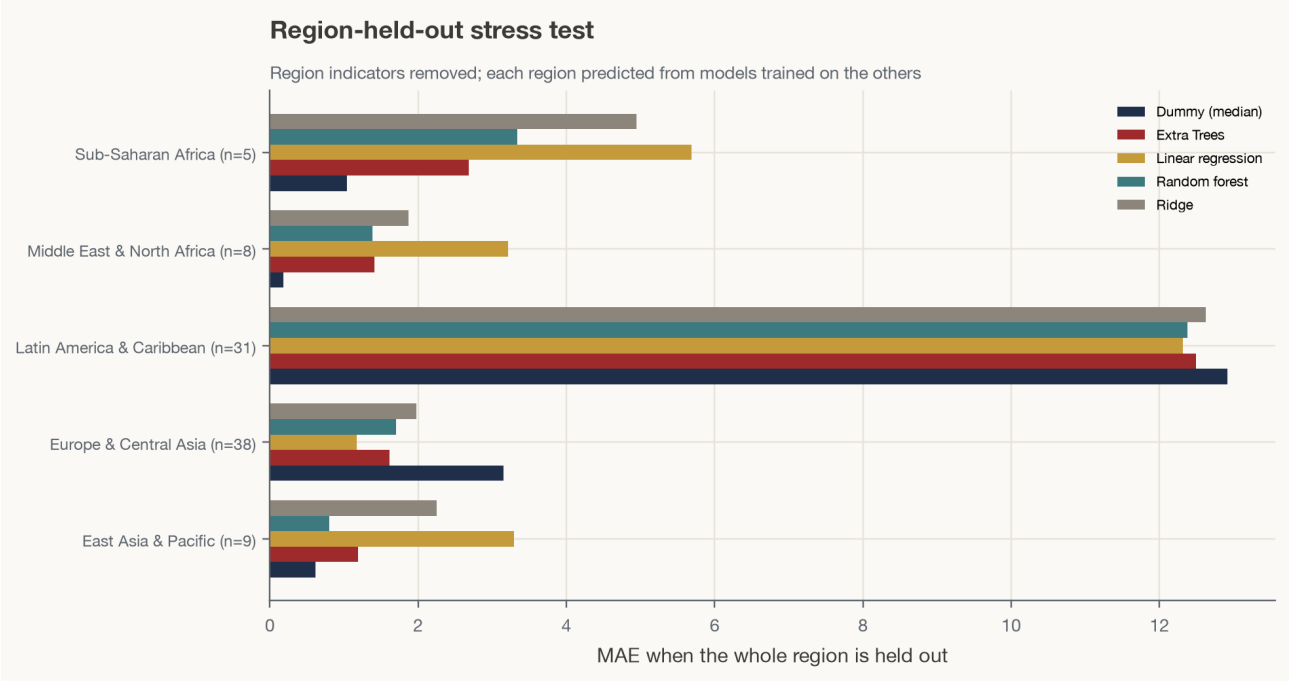


Figure 34. Leave-one-region-out MAE by held-out region.

7.1 Feature ablations

Table 7. Feature-group ablations on identical folds (50 folds).

Feature group	MAE: Gradient boosting	MAE: Random forest	MAE: Ridge	R^2 log1p: Gradient boosting	R^2 log1p: Random forest	R^2 log1p: Ridge
A: Firearm ownership only	4.71	4.49	4.71	-0.225	-0.067	-0.129
B: Socioeconomic only	4.64	4.16	4.31	0.065	0.196	0.132
C: Institutional only	5.18	4.55	4.62	-0.244	0.002	-0.077
D: Ownership + socioeconomic	4.49	4.02	4.33	0.134	0.254	0.115
E: Socioeconomic + institutional	4.39	4.08	4.23	0.147	0.261	0.16
F: All verified predictors	3.75	3.47	4.01	0.559	0.614	0.463
G: All except firearm ownership	3.73	3.49	3.98	0.56	0.606	0.478
H: All except regional identity	4.3	3.97	4.15	0.207	0.306	0.199

Firearm prevalence on its own (A) predicts no better than a constant. Socioeconomic indicators alone (B) and institutional indicators alone (C) do considerably better, and region carries large additional information (H vs F). The central ablation adds firearm prevalence to all other predictors (G → F): MAE changes from 3.49 to 3.47 for the random forest (-0.5%; fold-paired 95% range -0.18 to +0.26; improved in 60% of folds), from 3.98 to 4.01 for ridge (+0.8%) and from 3.73 to 3.75 for gradient boosting (+0.7%). On the log scale the corresponding R^2 changes are +0.008, -0.015 and -0.000. By contrast, adding region to all other predictors changes random-forest MAE by -12.6% (Figure 35). A small incremental predictive contribution is not evidence of no causal effect: ownership is correlated with region, income and governance, and a predictor can matter causally yet be redundant for prediction.

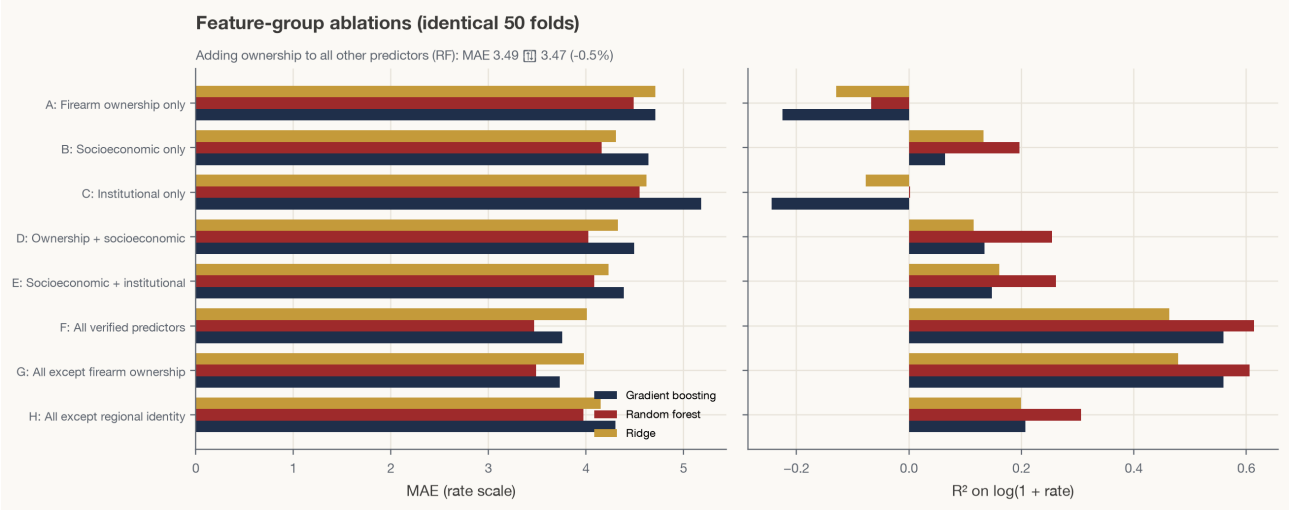


Figure 35. Feature-group ablations: MAE and log-scale R^2 for ridge, random forest and gradient boosting.

7.2 Interpretability

Held-out permutation importance ranks firearm prevalence 9th for the random forest and 7th for ridge, behind region and, for the random forest, inequality and schooling; individual governance indicators add little once region is known because they are highly correlated with it and with one another (Figure 36). In SHAP terms (Figure 37), firearm prevalence ranks 5th and accounts for 2.3% of total mean absolute attribution, versus 75.5% for region indicators and 4.5% for the four governance indicators. The SHAP dependence, partial dependence, ICE and ALE curves (Figure 38, Figure 39, Figure 40) show a shallow, mostly monotone increase in predicted log rate with firearm prevalence; the random-forest ALE range for ownership is 0.11 log points, compared with 0.03 for rule of law and 0.05 for log GDP. The strongest SHAP interaction partner of ownership is Region: Latin America & Caribbean. These attributions describe the fitted models, not causal mechanisms, and correlated predictors share credit in ways that are partly arbitrary.

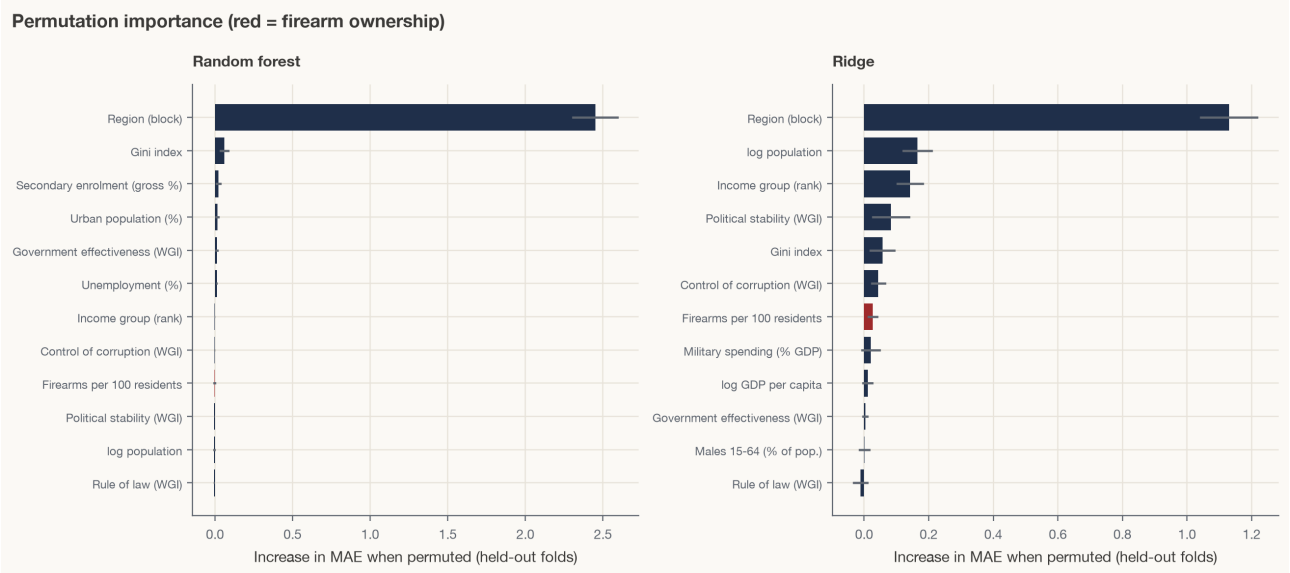


Figure 36. Held-out permutation importance (random forest and ridge).

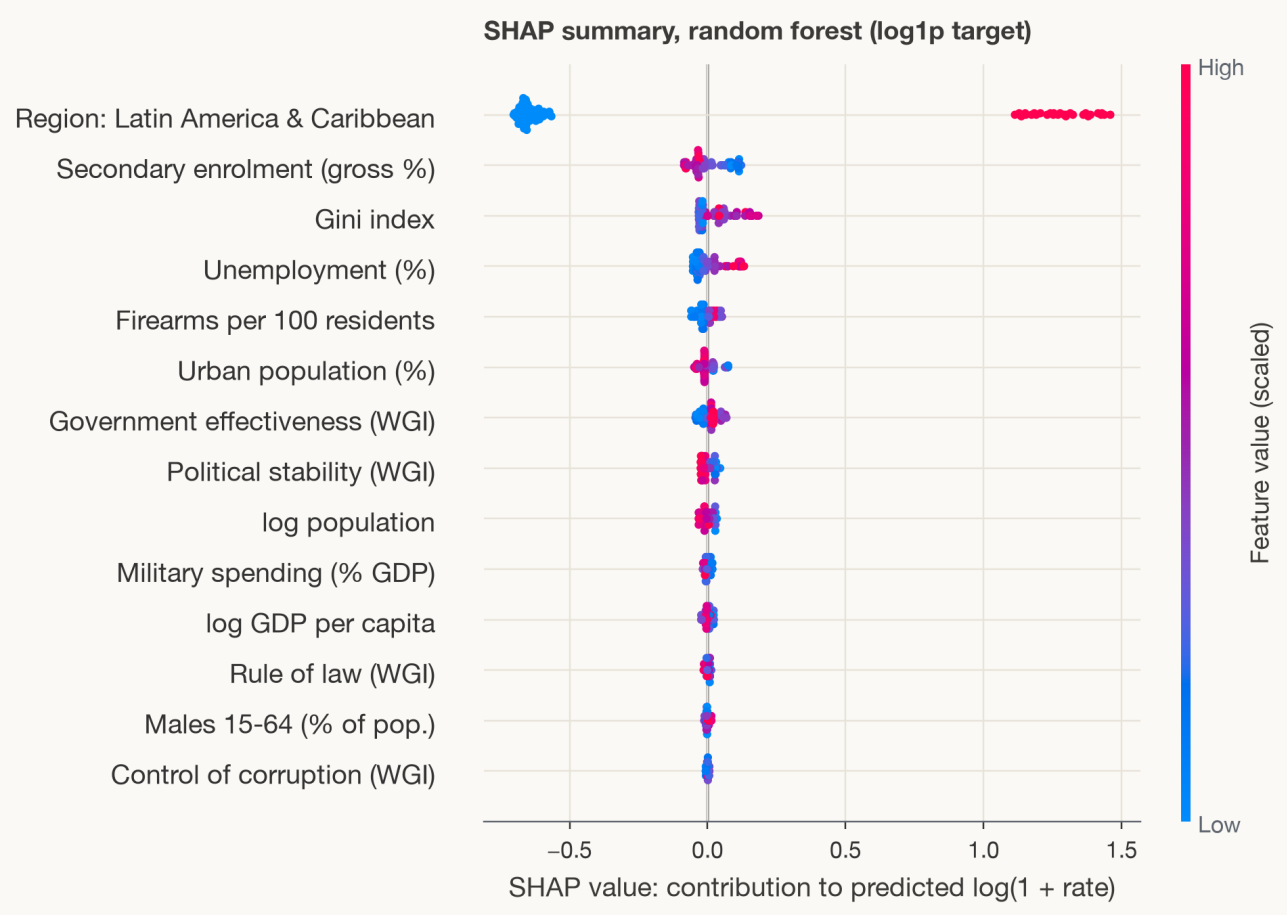


Figure 37. SHAP summary for the random forest (log1p target).

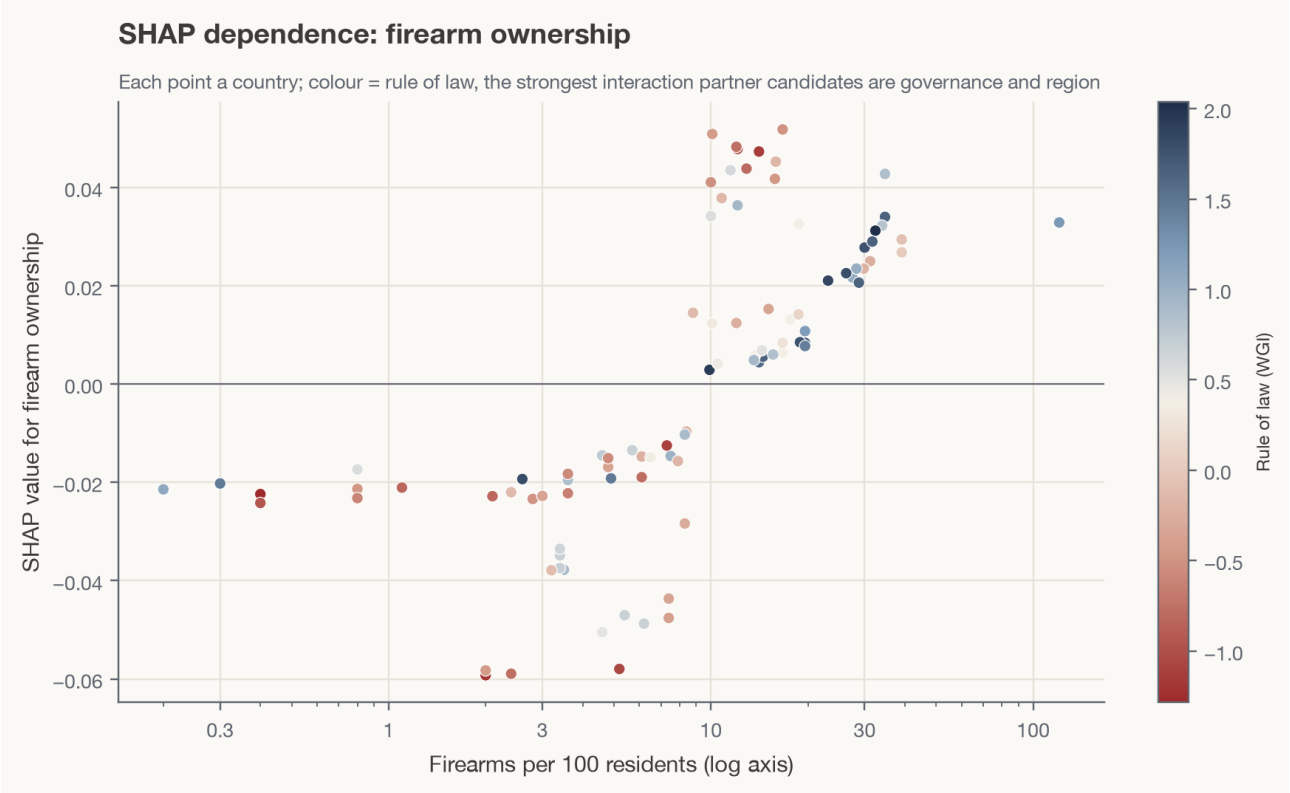


Figure 38. SHAP dependence for firearm prevalence, coloured by rule of law.

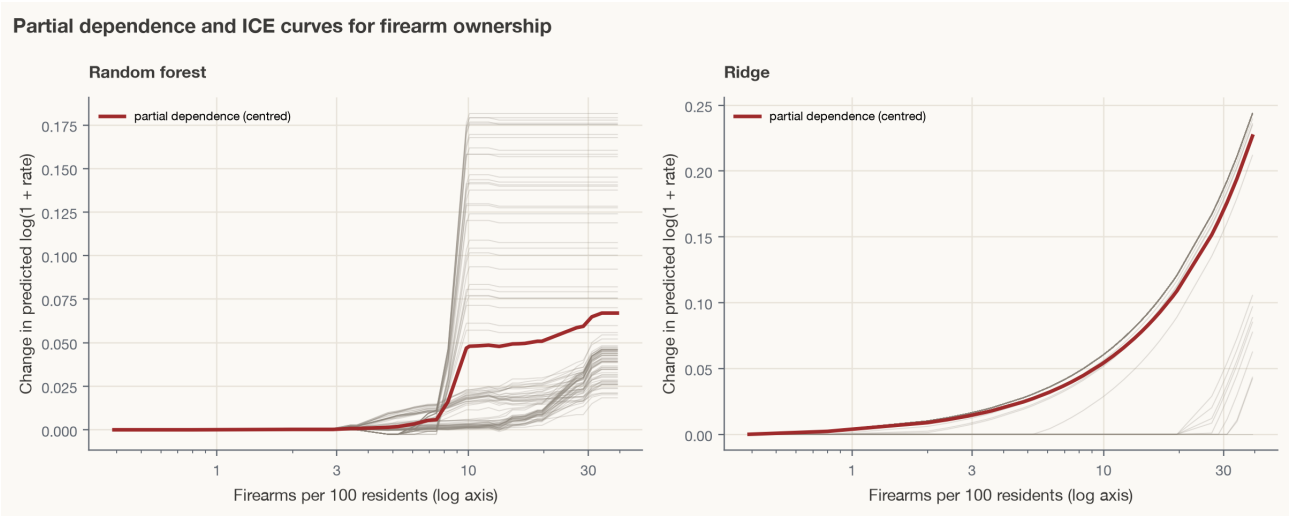


Figure 39. Partial dependence (red) and individual conditional expectation curves for firearm prevalence.

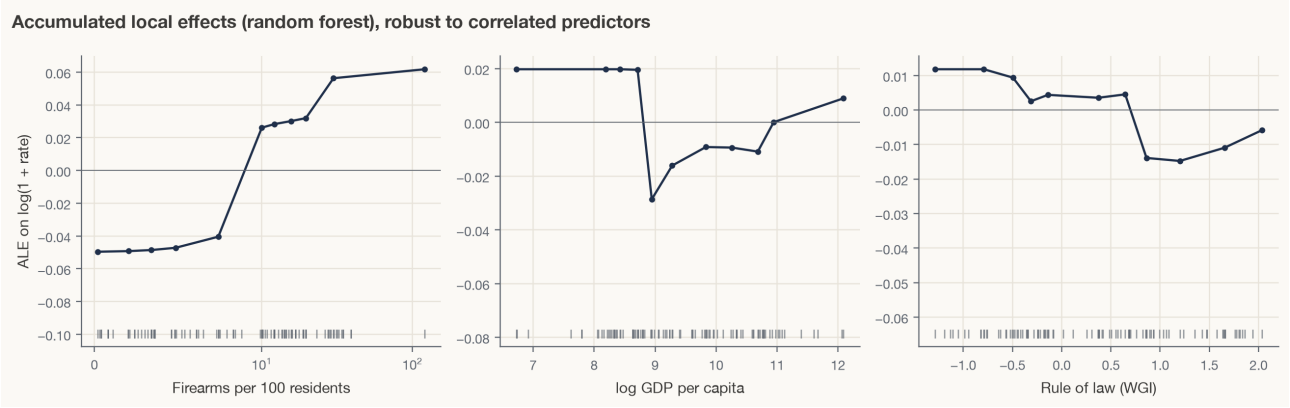


Figure 40. Accumulated local effects for firearm prevalence, log GDP and rule of law.

Prediction errors are concentrated where rates are high (Figure 41): Latin America and the Caribbean has by far the largest mean absolute error, and the top quintile of observed rates contributes most of the total error.

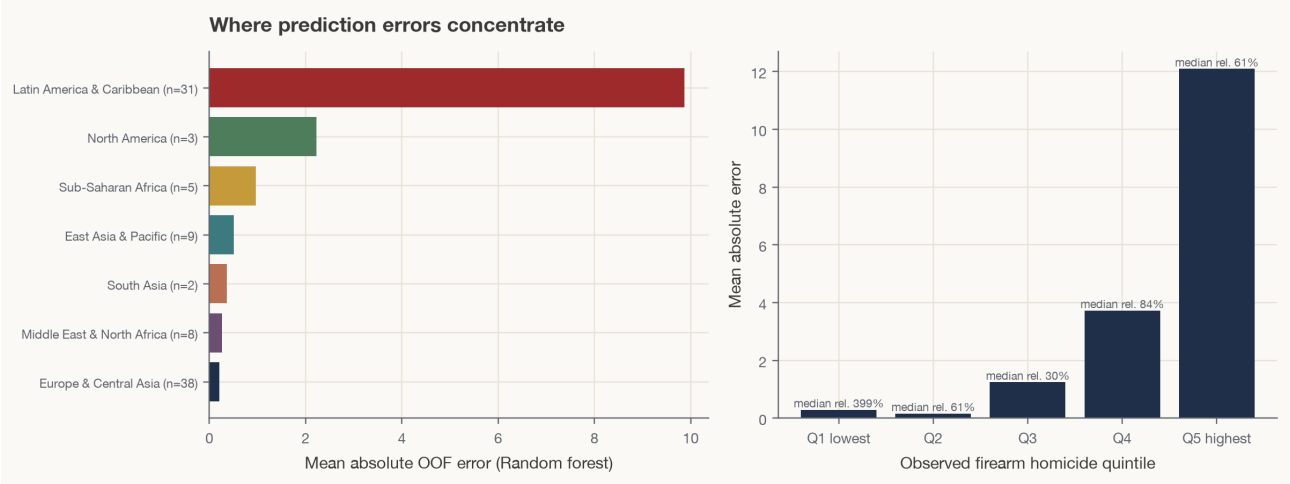


Figure 41. Out-of-fold absolute error of Random forest by region and by quintile of observed rate.

8. Geographic and Country-Level Analysis

8.1 Spatial dependence

Firearm homicide is strongly spatially clustered (Moran's I = 0.85 with 4 nearest neighbours, permutation p = 0.001), much more than firearm prevalence (I = 0.19). Residuals from a model without region effects inherit almost all of that clustering (I = 0.79); Model 5 residuals retain a smaller but significant amount (I = 0.26, p = 0.001; Figure 42). The result holds with 6 neighbours (I = 0.20), inverse-distance weights (I = 0.32) and contiguity (I = 0.42; 61 countries with at least one in-sample land neighbour). Lagrange multiplier tests on Model 5 reject both no-lag (LM = 16.1) and no-error dependence (LM = 14.7). A spatial lag model estimated by S2SLS gives ρ = 0.69 and an ownership coefficient of 0.089 (CI 0.029 to 0.148), close to Model 5. H6 is therefore supported. Spatial clustering does not imply spillovers; it may reflect shared unmeasured conditions such as drug-trafficking routes.

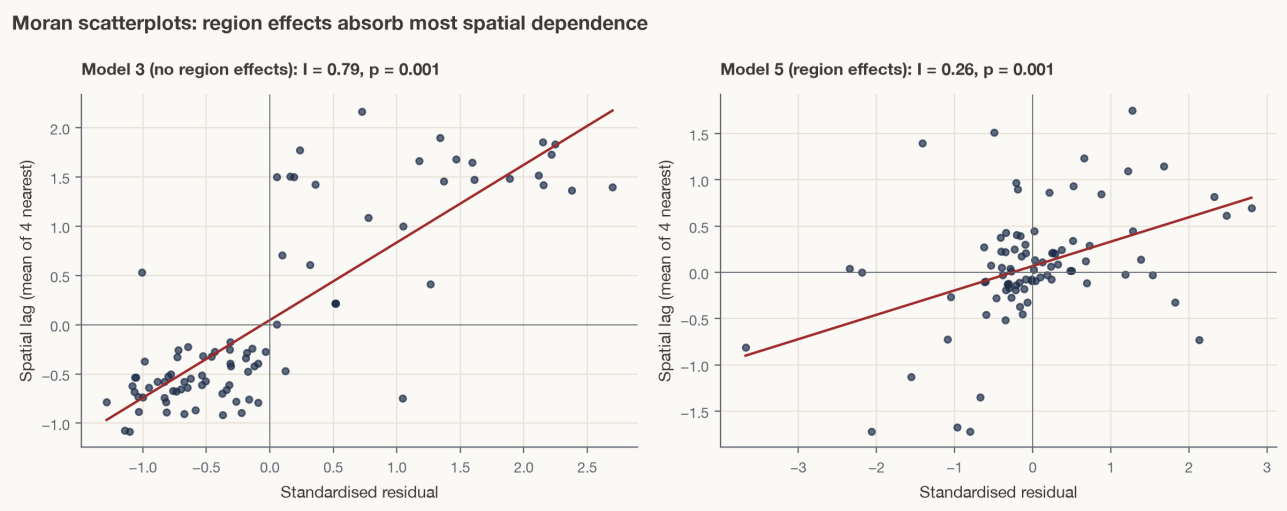


Figure 42. Moran scatterplots of residuals with and without region effects (kNN-4 weights).

Table 8. Global Moran's I under alternative spatial weights.

Weights	Variable	n	I	z	Permutation p
knn4	log(1 + firearm homicide)	87	0.849	13.7	0.001
knn4	Firearms per 100	87	0.187	3.6	0.003
knn4	Model 3 residuals	87	0.788	11.5	0.001
knn4	Model 5 residuals	87	0.264	4.0	0.001
invdist2500	log(1 + firearm homicide)	85	0.856	12.8	0.001
invdist2500	Firearms per 100	85	0.265	4.6	0.005
invdist2500	Model 3 residuals	85	0.796	12.0	0.001
invdist2500	Model 5 residuals	85	0.316	5.0	0.001
contiguity	log(1 + firearm homicide)	61	0.829	7.8	0.001
contiguity	Firearms per 100	61	0.306	3.7	0.004
contiguity	Model 3 residuals	61	0.750	7.0	0.001
contiguity	Model 5 residuals	61	0.425	4.0	0.002

Before region adjustment, local Moran statistics identify a large high-high cluster of residuals spanning North America, Central America and the northern Caribbean basin (22 countries). After adjustment, 11 countries remain in significant local clusters: high-high Bahamas, The, Belize, Guatemala, Mexico; low-low Argentina, Brazil, Chile, Paraguay, Uruguay. The model therefore still under-predicts Mexico and northern Central America and over-predicts much of South America relative to a single Latin American intercept (Figure 43, Figure 44).

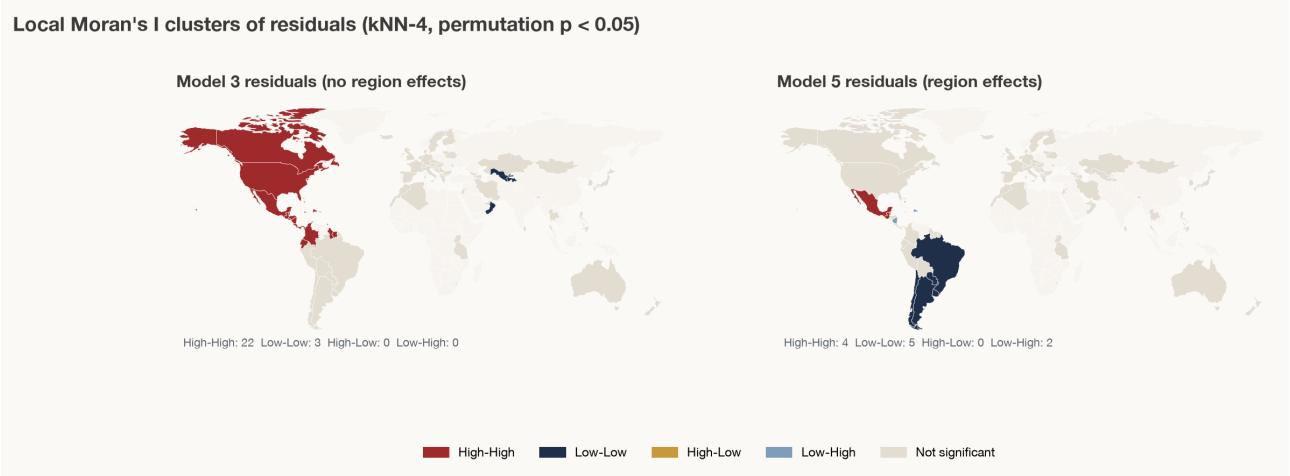


Figure 43. Local Moran's I cluster maps of residuals (kNN-4, permutation p < 0.05).

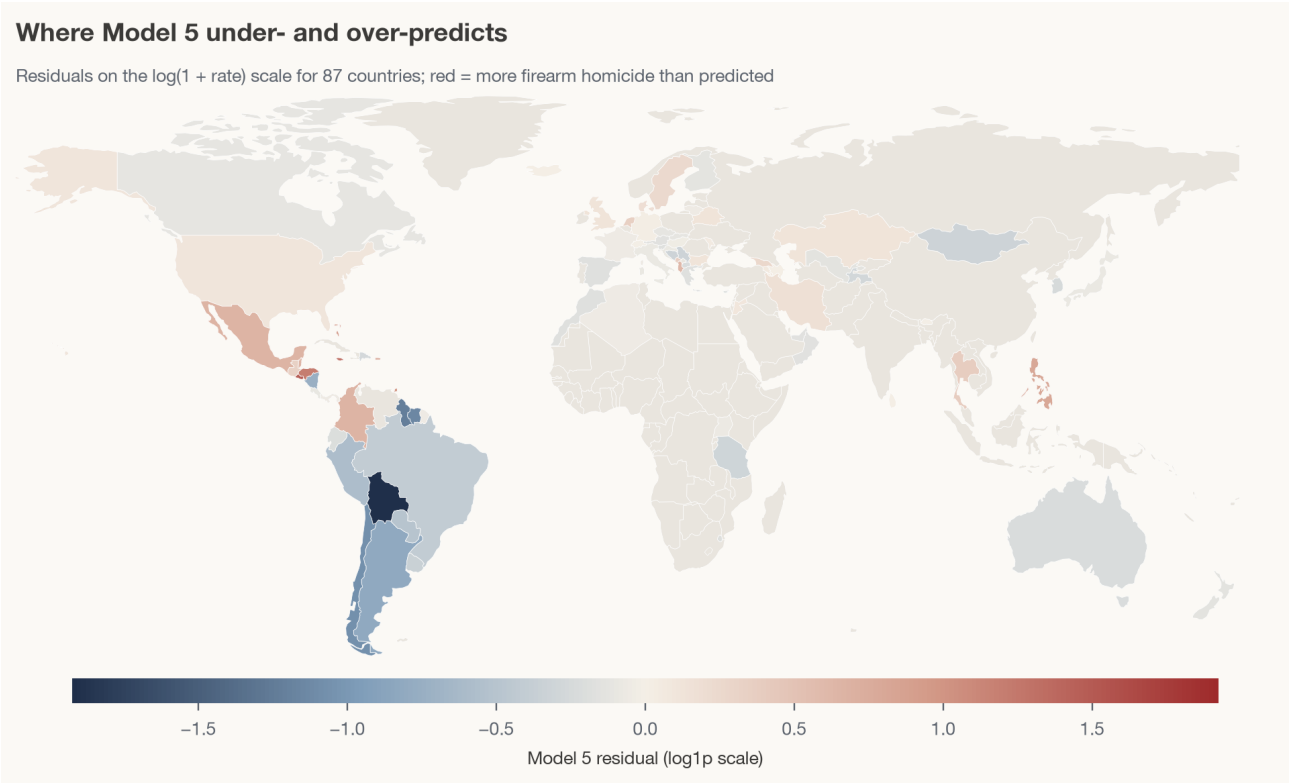


Figure 44. Model 5 residuals by country (log1p scale).

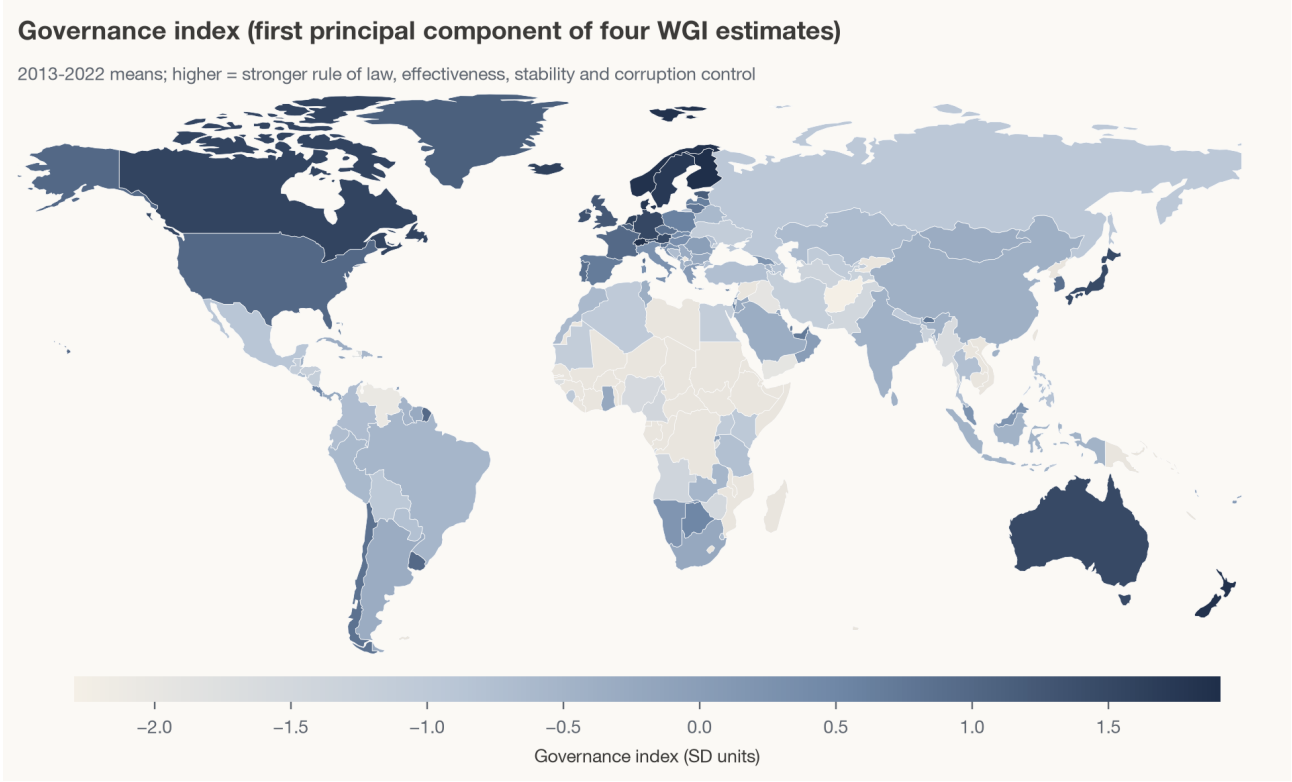


Figure 45. Governance index (first principal component of four WGI estimates), 2013-2022 means.

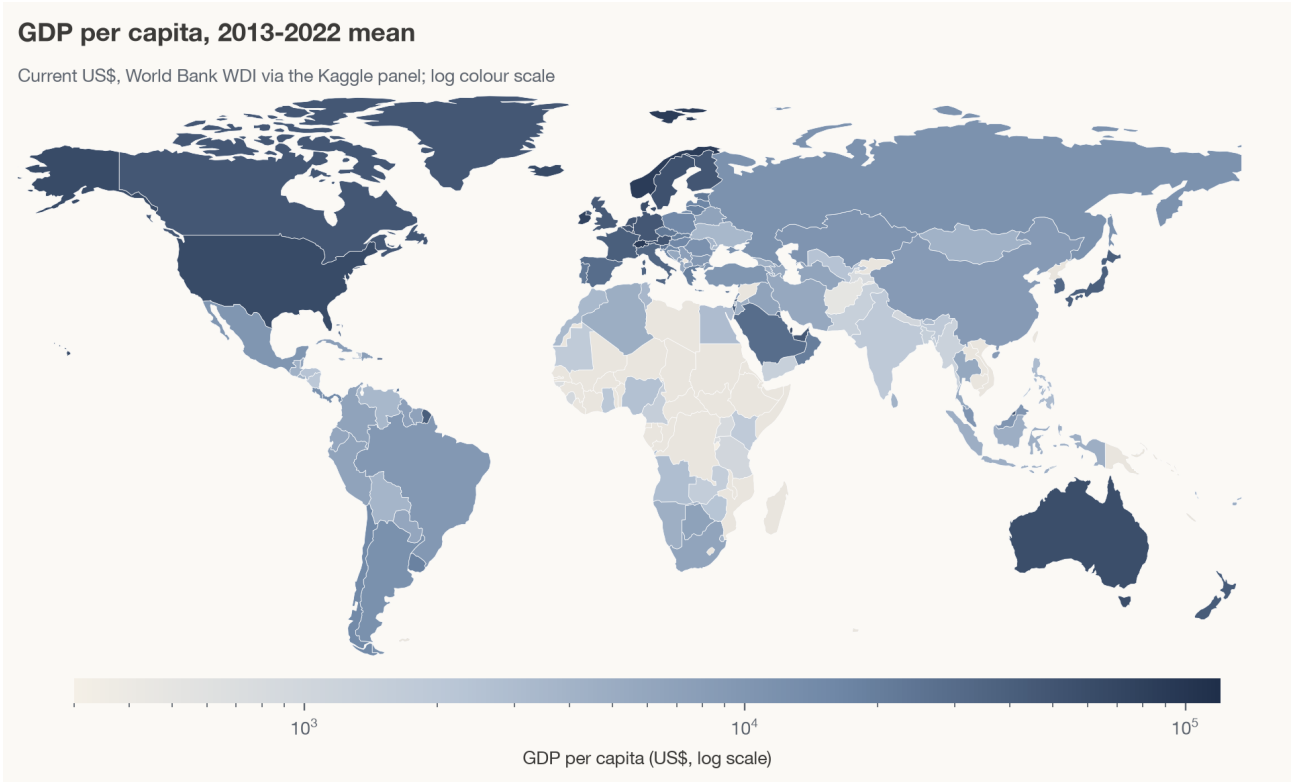


Figure 46. GDP per capita, 2013-2022 mean (log colour scale).

8.2 Countries above and below expectation

Table 9. Largest positive and negative Model 5 residuals (back-transformed predictions; prediction intervals from the log model).

Country	Firearms/100	Observed	Predicted	95% prediction interval	Stud. resid.	Leverage	Cook's D
El Salvador	12.0	38.36	8.10	[1.87, 27.85]	+2.83	0.07	0.041
Jamaica	8.8	38.27	9.71	[2.37, 32.99]	+2.49	0.07	0.034
Cabo Verde	5.7	3.80	0.58	[-0.56, 4.68]	+2.48	0.32	0.205
Honduras	14.1	37.04	10.27	[2.56, 34.68]	+2.31	0.07	0.028
Trinidad and Tobago	3.2	24.63	8.87	[2.04, 30.99]	+1.83	0.11	0.031
Bahamas, The	18.8	24.35	9.56	[2.33, 32.48]	+1.64	0.07	0.015
Nicaragua	5.2	3.61	8.65	[2.03, 29.73]	-1.38	0.08	0.012
Argentina	7.4	2.73	7.42	[1.63, 25.93]	-1.53	0.09	0.017
Chile	12.1	1.95	7.68	[1.69, 26.95]	-2.07	0.10	0.035
Suriname	15.9	2.42	9.75	[2.43, 32.65]	-2.14	0.05	0.016
Guyana	15.8	3.69	15.00	[3.77, 52.62]	-2.49	0.17	0.094
Bolivia	2.0	0.26	7.62	[1.72, 26.27]	-3.87	0.07	0.068

Countries with more firearm homicide than Model 5 predicts include El Salvador, Jamaica, Cabo Verde, Honduras, Trinidad and Tobago; countries with less include Bolivia, Guyana, Suriname, Chile, Argentina (Figure 47, Figure 48). Only 8 countries have |studentised residual| > 2 and 6 fall outside their 95% prediction intervals, so few are statistically exceptional once uncertainty is considered. Residuals, leverage and influence are distinct (Figure 49): the United States has very high leverage (0.70) but a small residual, while Cabo Verde, Guyana, Bolivia have Cook's distance above 4/n. Removing those influential countries leaves the ownership coefficient at 0.094 (CI 0.032 to 0.155); removing the United States moves it to 0.075 (CI -0.032 to 0.181).

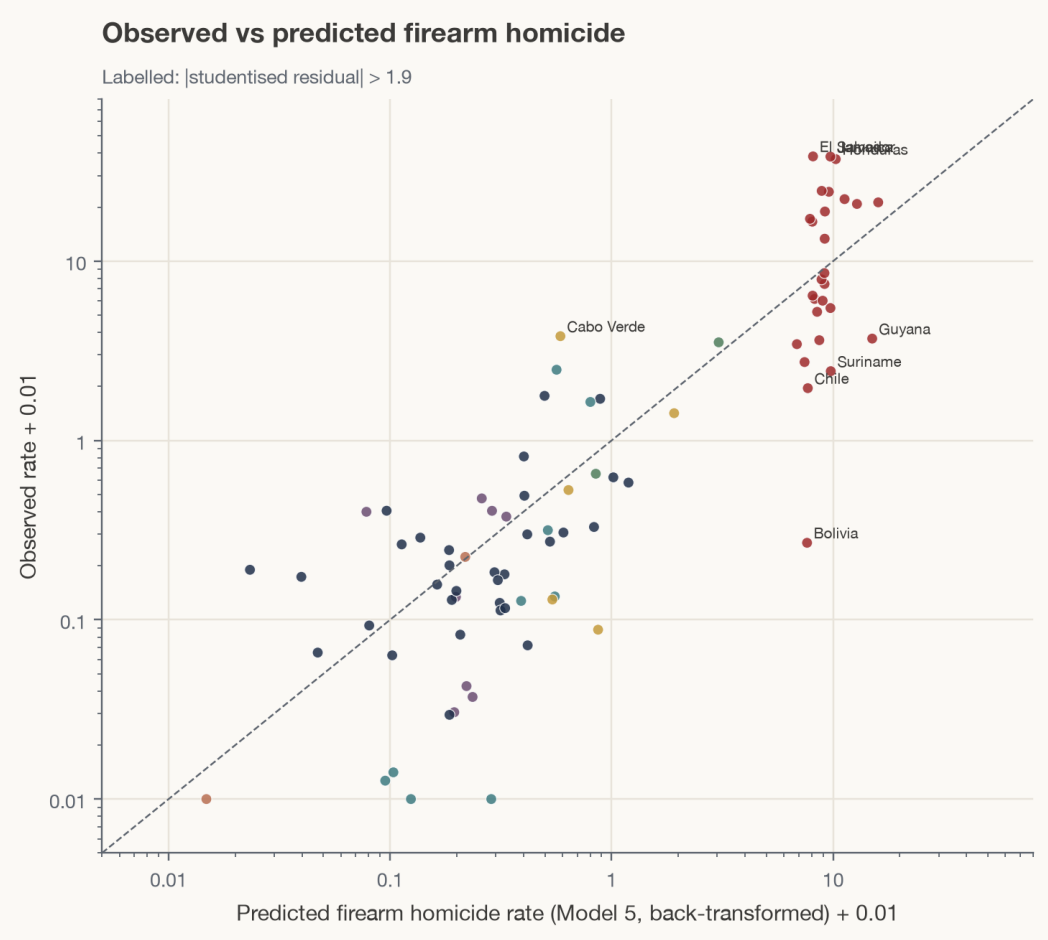


Figure 47. Observed versus Model 5 predicted firearm homicide (log axes).

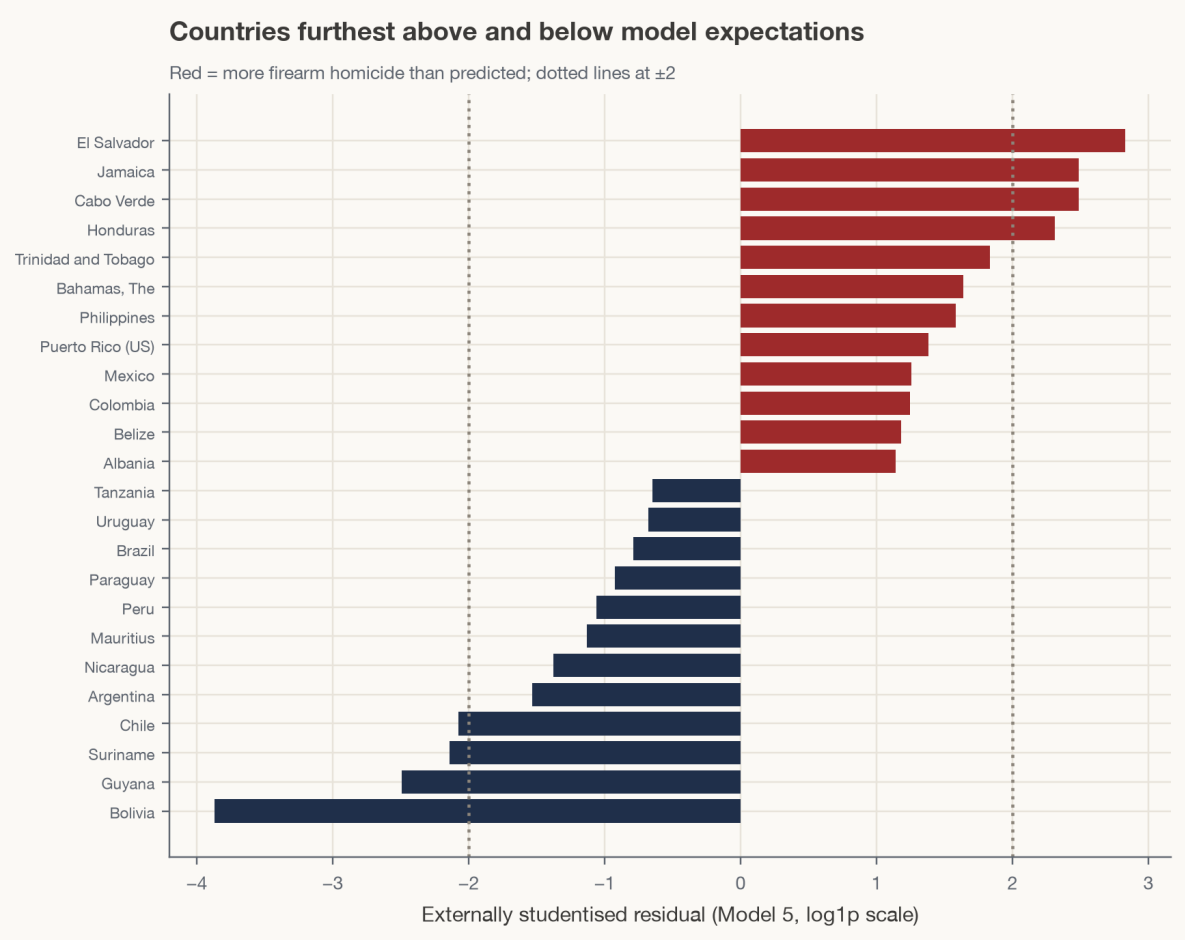


Figure 48. Countries ranked by externally studentised residual.

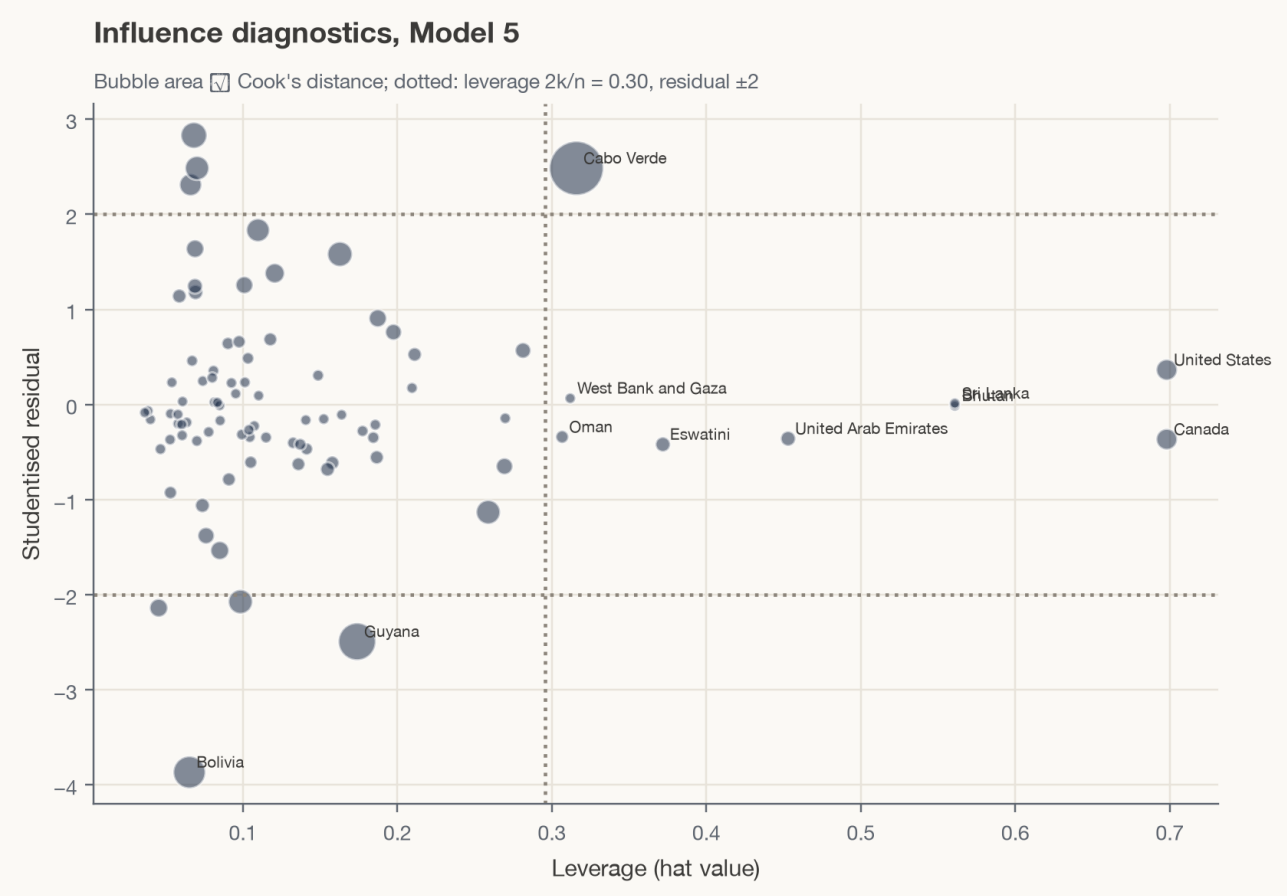


Figure 49. Leverage, studentised residuals and Cook's distance.

Country-level SHAP explanations (Figure 50) show why: the random forest's largest contributions for El Salvador are Region: Latin America & Caribbean, Secondary enrolment (gross %), Gini index, and it assigns the United States' very high firearm prevalence a contribution of +0.03 log points.

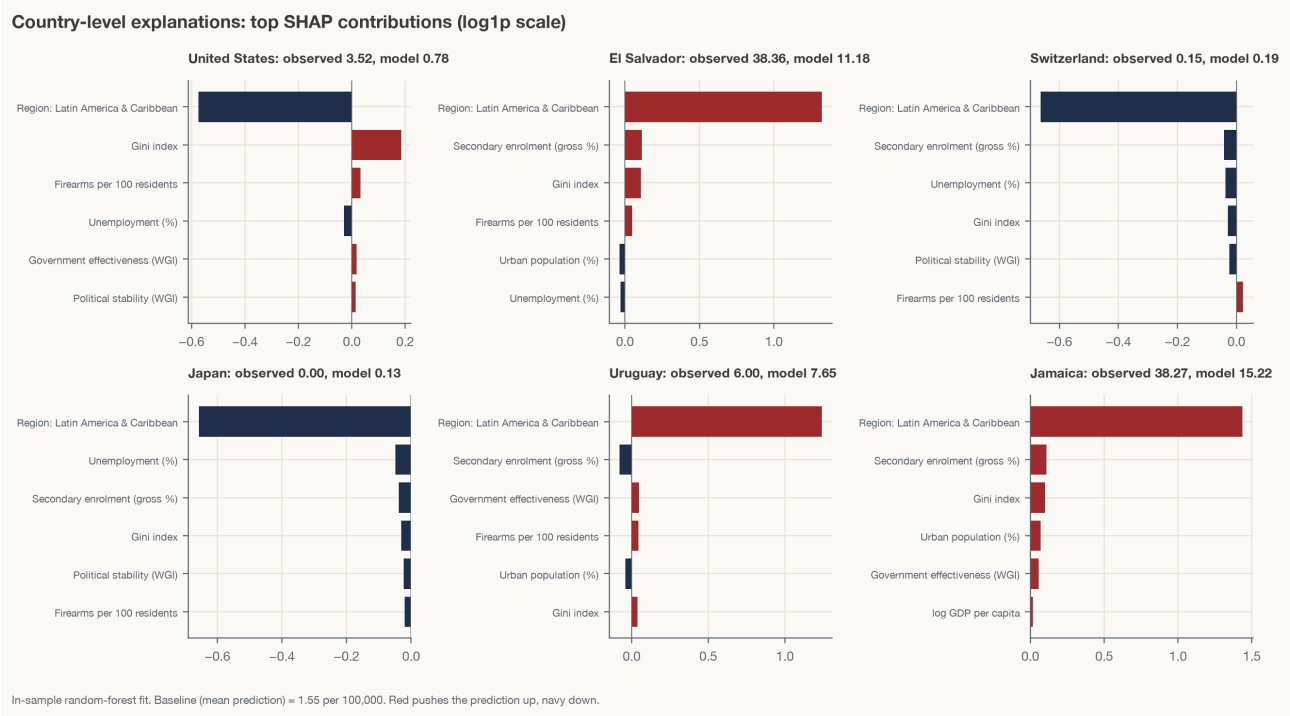


Figure 50. Top SHAP contributions for six countries (random forest, in-sample).

Contextual notes. The model cannot identify why a country deviates from expectation. For the largest positive residuals in Central America and the Caribbean, UNODC (2019, 2023) documents high levels of gang- and organised-crime-related lethal violence and trafficking, which are not measured here. Large negative residuals in South America (for example Bolivia, Guyana, Suriname and Chile) may partly reflect classification of homicide mechanisms: where a large share of homicides has an unknown mechanism, firearm homicide is undercounted. These notes are hypotheses for further work, not findings of this study.

9. Robustness Analysis

Table 10. Robustness matrix for the adjusted ownership coefficient (log1p scale per +10 firearms/100 unless marked 'different scale').

Group	Specification	n	Coef.	95% CI	p	R^2	Interpretation
Baseline	Baseline Model 5 (equal country weights)	88	+0.091	[+0.012, +0.170]	0.024	0.78	positive, CI excludes 0
Weighting	Population weights	88	+0.109	[-0.226, +0.445]	0.524	0.87	positive, CI includes 0
Weighting	Population weights, excluding United States	87	+0.229	[+0.032, +0.426]	0.022	0.86	positive, CI excludes 0
Outcome transformation	Outcome: asinh(rate)	88	+0.120	[+0.018, +0.222]	0.022	0.80	positive, CI excludes 0
Outcome transformation	Outcome: log(rate + 0.1)	88	+0.221	[+0.091, +0.351]	<0.001	0.81	positive, CI excludes 0
Outcome transformation	Outcome: log1p of pooled rate (sum counts / person-years)	88	+0.090	[+0.009, +0.171]	0.030	0.78	positive, CI excludes 0
Outcome transformation	Outcome: rate in levels (units: per 100,000)	88	+0.283	[-0.511, +1.078]	0.485	0.50	different scale
Outcome transformation	Negative binomial count model (log rate ratio)	88	+0.274	[+0.125, +0.424]	<0.001		positive, CI excludes 0
Sample	Exclude influential countries (Cook's D > 4/n)	85	+0.094	[+0.032, +0.155]	0.003	0.85	positive, CI excludes 0
Sample	Exclude United States	87	+0.075	[-0.032, +0.181]	0.168	0.78	positive, CI includes 0
Sample	Exclude conflict-affected (UCDP >= 1,000 battle deaths in any year)	86	+0.089	[+0.008, +0.171]	0.032	0.78	positive, CI excludes 0
Sample	Exclude small states (population < 1 million)	74	+0.091	[+0.023, +0.159]	0.009	0.79	positive, CI excludes 0
Sample	Outcome observed in >= 3 window years	76	+0.081	[-0.009, +0.172]	0.077	0.77	positive, CI includes 0
Exposure measurement	Exclude 'analogous comparison' firearm estimates	73	+0.083	[-0.004, +0.170]	0.062	0.79	positive, CI includes 0
Exposure measurement	Only survey-based firearm estimates	16	+0.022	[-0.034, +0.078]	0.434	0.96	positive, CI includes 0
Classification	Region: Small Arms Survey continental regions	88	-0.032	[-0.124, +0.060]	0.495	0.74	negative, CI includes 0
Classification	Region: three groups (Europe & CA, LatAm & Caribbean, other)	88	+0.099	[+0.037, +0.161]	0.002	0.77	positive, CI excludes 0
Classification	Income-group FE instead of region FE	88	+0.041	[-0.137, +0.219]	0.649	0.15	positive, CI includes 0
Classification	Region FE + income-group FE	88	+0.079	[-0.024, +0.182]	0.133	0.79	positive, CI includes 0
Covariates	No region FE (Model 3 + governance)	88	+0.047	[-0.086, +0.179]	0.489	0.10	positive, CI includes 0
Covariates	Governance: rule of law instead of PC index	88	+0.091	[+0.009, +0.173]	0.029	0.78	positive, CI excludes 0
Covariates	Governance: all four WGI separately	88	+0.090	[+0.016, +0.164]	0.017	0.79	positive, CI excludes 0
Covariates	Add schooling and military spending	75	+0.116	[+0.042, +0.190]	0.002	0.76	positive, CI excludes 0
Missing data	Complete case incl. Gini	75	+0.103	[+0.031, +0.174]	0.005	0.77	positive, CI excludes 0
Missing data	Multiple imputation of covariates (m = 20, incl. Gini), Rubin's rules	96	+0.047	[-0.058, +0.152]	0.381		positive, CI includes 0
Estimator	Robust regression (Huber M-estimator)	88	+0.080	[+0.011, +0.149]	0.023		positive, CI excludes 0
Estimator	Median regression	88	+0.068	[+0.008, +0.129]	0.027		positive, CI excludes 0
Functional form	Nonlinear: log firearms per 100 (per log unit)	88	+0.121	[+0.014, +0.228]	0.026	0.78	different scale
Functional form	Nonlinear: quadratic in firearms (linear term)	88	+0.078	[-0.658, +0.813]	0.836	0.78	different scale
Inference	Clustered SE by region (7 clusters)	88	+0.091	[+0.031, +0.151]	0.003	0.78	positive, CI excludes 0
Spatial	Spatial lag model (S2SLS, kNN-4)	87	+0.089	[+0.029, +0.148]	0.004		positive, CI excludes 0
Temporal alignment	Window 2015-2019 for outcome and covariates	79	+0.088	[+0.016, +0.161]	0.017	0.76	positive, CI excludes 0
Temporal alignment	Window 2010-2022 for outcome and covariates	90	+0.099	[+0.026, +0.171]	0.008	0.78	positive, CI excludes 0
Temporal alignment	Window 2017-2017 for outcome and covariates	72	+0.074	[-0.017, +0.165]	0.112	0.73	positive, CI includes 0

Across 34 specifications, 97% of estimates are positive, 62% have a 95% interval above zero, and none is significantly negative (Figure 51, Figure 52). Comparable estimates range from -3.2% to +31.6% per 10 firearms per 100 with a median coefficient of 0.090. The association is robust to outcome transformation (asinh, log(rate + 0.1), pooled rates, negative binomial), to excluding influential, conflict-affected or small countries, to robust and

median regression, to alternative governance measures, to the spatial lag model, and to the 2015-2019 and 2010-2022 windows. It is **not** robust to: excluding the United States (+0.075, p = 0.168); population weighting (p = 0.524); multiple imputation of covariates on the full primary sample (+0.047, p = 0.381); replacing World Bank regions by Small Arms Survey continental regions (-0.032), which merges North and South America; income-group rather than region effects; or dropping firearm estimates based on analogous comparison (p = 0.062). The adjusted association therefore depends on how regional context is controlled and on a few observations, which is the defining sense in which it is fragile.

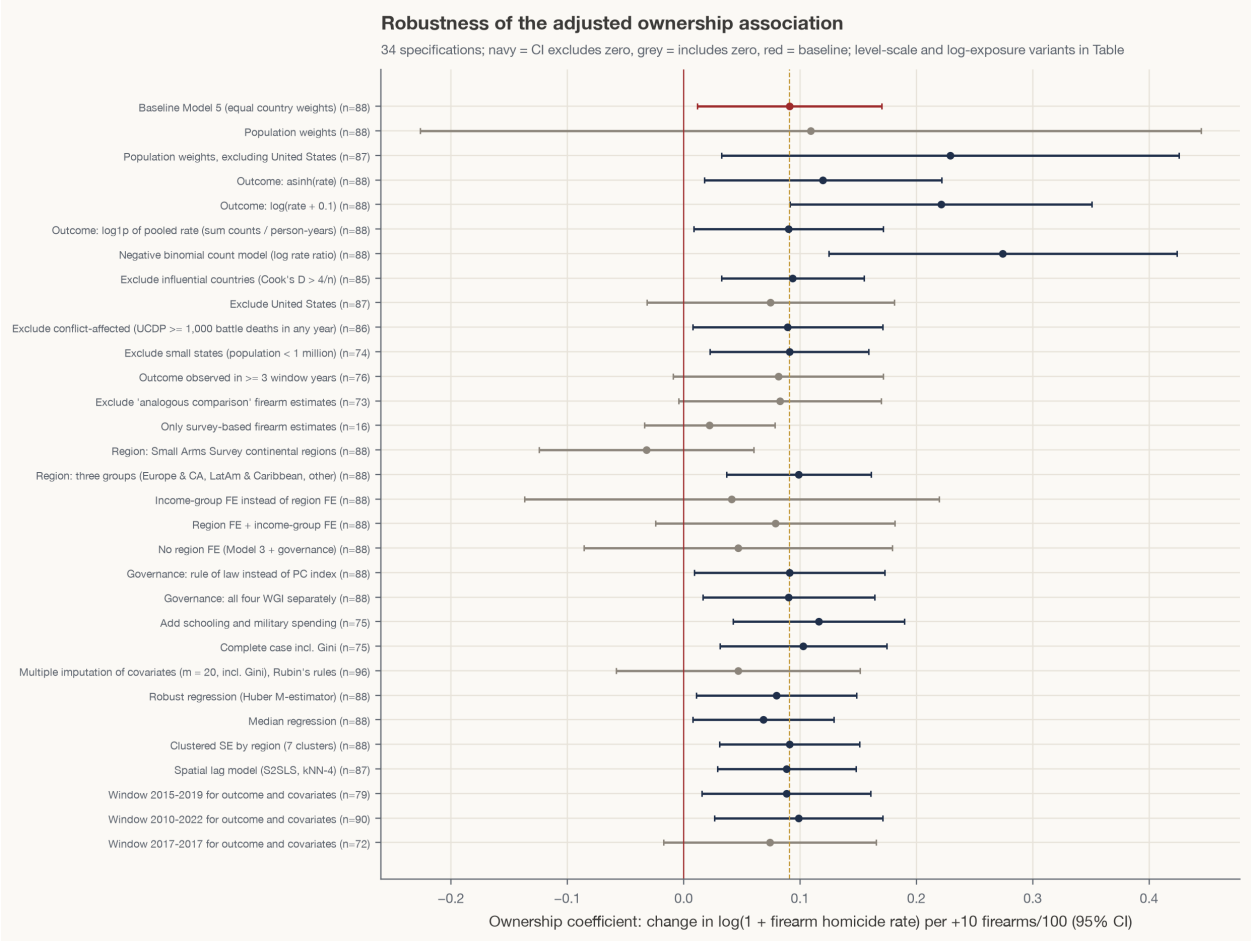


Figure 51. Adjusted ownership coefficient across robustness specifications (95% CIs).

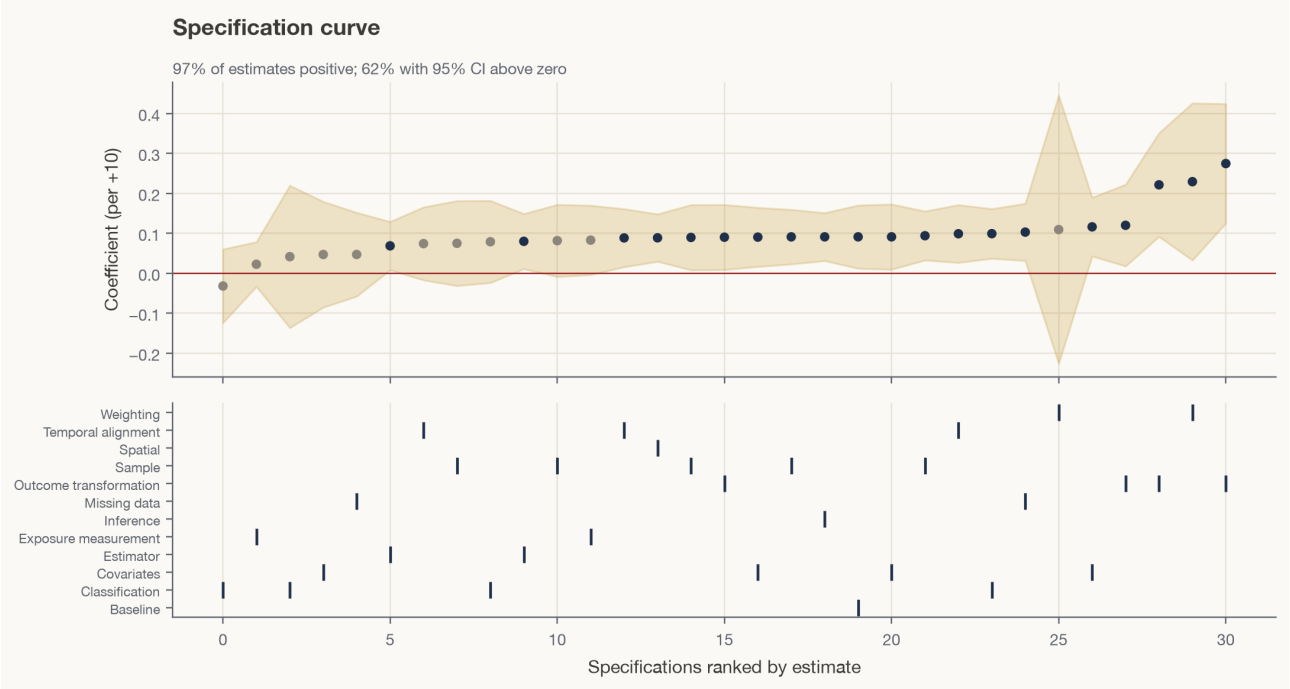


Figure 52. Specification curve: estimates ranked, with the specification group of each.

10. Discussion

How strongly is firearm ownership associated with firearm homicide? Unconditionally, not at all across the 96 countries studied: the correlation is indistinguishable from zero and its 95% interval (-0.19 to 0.21) excludes even moderate associations in either direction. This null global association is not evidence that firearms are irrelevant, because the global comparison mixes regions whose homicide levels differ for reasons unrelated to gun stocks.

How does adjustment change the relationship? Socioeconomic controls alone change little. Once regional differences in levels are removed, a positive within-region gradient emerges: about 10% higher (1 + rate) per 10 firearms per 100 on the log scale and a rate ratio of 1.32 in the count model. This is a modest association relative to the regional gaps; Latin America and the Caribbean's random intercept alone corresponds to a multiple of several times in (1 + rate). The adjusted association is consistent with the high-income literature, but in this sample it depends on the United States, on how regions are defined, and on the scale of measurement.

Which predictors improve prediction? Region carries most out-of-sample information, above all the Latin America and the Caribbean indicator; governance matters in the regressions but is largely redundant with region for prediction, and inequality is strongly associated with firearm homicide where measured. Adding firearm prevalence to everything else changes random-forest MAE by -0.5%, and it ranks 5th by SHAP attribution.

What explains regional heterogeneity? The data show heterogeneity in *levels*, concentrated in Latin America and the Caribbean, and spatial clustering of residuals even after region effects. Slopes do not differ detectably between regions, but the tests have little power. The level differences align with the regional patterns of organised-crime-related violence documented by UNODC rather than with differences in civilian firearm stocks.

How large is the association in practical terms? Moving from the sample's first to third quartile of firearm prevalence (3.6 to 16.9 per 100) corresponds, under Model 5, to a 13% difference in (1 + rate). The difference between the Latin American and the European random intercepts in the multilevel model is 2.10 log points, a factor of about 8.2 in (1 + rate). Even taken at face value, the interquartile ownership contrast is therefore an order of magnitude smaller than the regional contrast. For the United States, the adjusted coefficient implies that its extreme

firearm stock contributes materially to its position above other high-income countries, but the estimate is extrapolated from a single observation beyond 40 firearms per 100 and is the least certain part of the curve.

How do these results compare with earlier work? The pattern reconciles two strands of the literature. Studies restricted to high-income countries (Killias, 1993; Hemenway & Miller, 2000; Bangalore & Messerli, 2013) report positive associations, and the same is found here among high-income countries outside Latin America. Studies and critiques that emphasise confounding and heterogeneous samples (Kleck, 2015) are consistent with the null global correlation and the fragility of the adjusted estimate. The dominant role of Latin American location and the strong bivariate association with inequality agree with Fajnzylber et al. (2002) and Nivette (2011). Where this study adds to that literature is in quantifying, on identical cross-validation folds, how little predictive information firearm prevalence contributes once regional identity is known, and in showing that residual spatial clustering persists after regional adjustment.

Why might firearm prevalence relate to the share but not the level of homicide? If firearms mainly change the instrument used in violence that would occur anyway, they would raise the firearm share and lower non-firearm homicide without changing total homicide; if they also raise lethality or the frequency of attacks, total homicide should rise too. The secondary outcomes are closer to the first pattern, but differences in how countries classify the mechanism of a killing, and the small, selective sample, mean this is a descriptive regularity rather than a test of either mechanism.

What can and cannot be inferred about policy? These are cross-sectional, ecological associations between an estimated firearm stock in 2017 and averaged homicide rates. They cannot say what would happen to a country's homicide rate if its firearm stock changed, whether firearm stocks respond to violence, or how firearm laws affect outcomes. They are compatible with a modest causal contribution of firearm availability to firearm homicide within similar countries, and equally compatible with confounding by unmeasured regional or institutional factors. Policy evaluation requires within-country designs (Santaella-Tenorio et al., 2016).

11. Limitations

- **Ecological fallacy.** All units are countries. Nothing here describes the risk faced by individuals or households that own firearms (Robinson, 1950).
- **Cross-sectional design and reverse causality.** Ownership is measured once (2017). Firearm stocks may respond to violence, so the direction of any association is unidentified.
- **Omitted variables.** Organised crime, drug-trafficking routes, policing, firearm type (handguns versus long guns), legislation and illicit flows are not measured; region effects and governance proxy for some of these only crudely.
- **Measurement error in the exposure.** Most SAS estimates are expert judgements or analogous comparisons; only 23 use surveys. Classical error would bias coefficients toward zero; non-classical error could bias them in either direction.
- **Under-reporting and mechanism classification.** Firearm homicide depends on recording the mechanism; countries with many unknown-mechanism homicides undercount firearm homicide.
- **Selective outcome coverage.** Only 96 countries report firearm homicide; Africa and South Asia are nearly absent and one low-income country is included, so results do not generalise to them.
- **Different reference years.** The exposure refers to 2017; outcomes and covariates are averaged over 2013-2022, with country-specific observed years. Alternative windows give similar results.
- **Small samples and model uncertainty.** With 88 countries in adjusted models and seven regions, interaction and random-slope estimates are imprecise, and specification choices move the main estimate.

- **Missing covariates.** Gini is missing for many countries; complete-case and multiply imputed results differ, and imputation used only covariates, never outcome or exposure.
- **Aggregation and the averaging window.** Averaging annual rates over a decade smooths year-to-year volatility but also mixes periods before and after the 2017 exposure measurement; it cannot separate stable country differences from trends within the window.
- **Region definitions.** Regional effects are powerful but blunt. They absorb many unmeasured factors at once, and the main estimate changes when regions are defined differently, so region effects should not be read as explanations.
- **Limits of prediction.** Good predictive accuracy does not imply explanation; SHAP and partial dependence describe model behaviour, not causal effects.

12. Conclusion

Across 96 countries, civilian firearm prevalence in 2017 is essentially uncorrelated with firearm homicide in 2013-2022 (Spearman $\rho = 0.01$). After adjusting for socioeconomic conditions, World Bank region and governance in 88 countries, each additional 10 firearms per 100 residents is associated with 9.5% higher (1 + firearm homicide rate) (95% CI 1.2% to 18.5%), a negative binomial rate ratio of 1.32, and a level-scale difference of +0.28 per 100,000 that is not statistically distinguishable from zero. The strongest findings are:

1. **The global association is null; the adjusted association is small, positive and fragile.** It is positive in 97% of 34 specifications but significant in only 62%, and loses precision without the United States or with alternative regional controls.
2. **Regional context dominates.** Region and governance raise R^2 for log rates from 0.05 to 0.78; about 67% of residual variance lies between regions, and all twenty highest rates are in Latin America and the Caribbean.
3. **Firearm prevalence adds little predictive information beyond other country characteristics.** The best model (Random forest) attains cross-validated MAE 3.53 per 100,000; adding ownership to all other predictors changes random-forest MAE by -0.5%.
4. **Firearms shift the mechanism more clearly than the level of homicide.** Firearm prevalence correlates with the firearm share of homicides ($\rho = 0.27$) and negatively with non-firearm homicide ($\rho = -0.30$), but not with total homicide ($\rho = -0.04$).
5. **Firearm homicide is spatially clustered beyond what regions explain** (Moran's I of Model 5 residuals = 0.26), pointing to unmeasured, geographically shared conditions.

No causal identification is claimed. The defensible conclusion is that cross-national differences in firearm homicide are explained far more by where a country is and how it is governed than by how many civilian firearms it holds, while a modest positive conditional association with firearm prevalence cannot be excluded and is visible in several specifications.

References

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Kluwer Academic. https://doi.org/10.1007/978-94-015-7799-1

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis, 27*(2), 93-115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B, 82*(4), 1059-1086. https://doi.org/10.1111/rssb.12377

Bangalore, S., & Messerli, F. H. (2013). Gun ownership and firearm-related deaths. *The American Journal of Medicine, 126*(10), 873-876. https://doi.org/10.1016/j.amjmed.2013.04.012

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139013567

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). https://doi.org/10.1145/2939672.2939785

Eisner, M. (2003). Long-term historical trends in violent crime. *Crime and Justice, 30*, 83-142. https://doi.org/10.1086/652229

Fajnzylber, P., Lederman, D., & Loayza, N. (2002). Inequality and violent crime. *The Journal of Law and Economics, 45*(1), 1-39. https://doi.org/10.1086/338347

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed conflict 1946-2001: A new dataset. *Journal of Peace Research, 39*(5), 615-637. https://doi.org/10.1177/0022343302039005007

Grinshteyn, E., & Hemenway, D. (2016). Violent death rates: The US compared with other high-income OECD countries, 2010. *The American Journal of Medicine, 129*(3), 266-273. https://doi.org/10.1016/j.amjmed.2015.10.025

Hemenway, D., & Miller, M. (2000). Firearm availability and homicide rates across 26 high-income countries. *Journal of Trauma: Injury, Infection, and Critical Care, 49*(6), 985-988. https://doi.org/10.1097/00005373-200012000-00001

Hepburn, L. M., & Hemenway, D. (2004). Firearm availability and homicide: A review of the literature. *Aggression and Violent Behavior, 9*(4), 417-440. https://doi.org/10.1016/S1359-1789(03)00044-2

Karp, A. (2018). *Estimating global civilian-held firearms numbers* (Briefing Paper). Small Arms Survey. https://www.smallarmssurvey.org/resource/estimating-global-civilian-held-firearms-numbers

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The Worldwide Governance Indicators: Methodology and analytical issues. *Hague Journal on the Rule of Law, 3*(2), 220-246. https://doi.org/10.1017/S1876404511200046

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems 30* (pp. 3146-3154).

Kelejian, H. H., & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics, 17*(1), 99-121. https://doi.org/10.1023/A:1007707430416

Killias, M. (1993). International correlations between gun ownership and rates of homicide and suicide. *Canadian Medical Association Journal, 148*(10), 1721-1725.

Kleck, G. (2015). The impact of gun ownership rates on crime rates: A methodological review of the evidence. *Journal of Criminal Justice, 43*(1), 40-48. https://doi.org/10.1016/j.jcrimjus.2014.12.002

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*(1), 33-50. https://doi.org/10.2307/1913643

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30* (pp. 4765-4774).

MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics, 29*(3), 305-325. https://doi.org/10.1016/0304-4076(85)90158-7

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*(1/2), 17-23. https://doi.org/10.2307/2332142

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica, 46*(1), 69-85. https://doi.org/10.2307/1913646

Neumayer, E. (2003). Good policy can lower violent crime: Evidence from a cross-national panel of homicide rates, 1980-97. *Journal of Peace Research, 40*(6), 619-640. https://doi.org/10.1177/00223433030406001

Nivette, A. E. (2011). Cross-national predictors of crime: A meta-analysis. *Homicide Studies, 15*(2), 103-131. https://doi.org/10.1177/1088767911406397

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems 31* (pp. 6638-6648).

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*(3), 351-357. https://doi.org/10.2307/2087176

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley. https://doi.org/10.1002/9780470316696

Santaella-Tenorio, J., Cerdá, M., Villaveces, A., & Galea, S. (2016). What do we know about the association between firearm legislation and firearm-related injuries? *Epidemiologic Reviews, 38*(1), 140-157. https://doi.org/10.1093/epirev/mxv012

UNODC. (2019). *Global study on homicide 2019*. United Nations Office on Drugs and Crime. https://www.unodc.org/unodc/en/data-and-analysis/global-study-on-homicide.html

UNODC. (2023). *Global study on homicide 2023*. United Nations Office on Drugs and Crime. https://www.unodc.org/unodc/en/data-and-analysis/global-study-on-homicide.html

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics, 7*, 91. https://doi.org/10.1186/1471-2105-7-91

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279

Data sources: Small Arms Survey (2018) via Kaggle dataset `dhananjayk2609/global-gun-violence-and-civilian-firearms`; UNODC intentional homicide statistics and World Bank WDI via the same dataset; World Bank Worldwide Governance Indicators (API source 3, updated 2026-09-25); UCDP Battle-Related Deaths v25.1; Natural Earth 1:110m Admin 0. Retrieved 18 November 2026.

Appendix

A.1 Additional regression output

Table 11. Model 5 (log1p outcome) full coefficient table; reference region Europe & Central Asia.

Term	Coef.	HC3 SE	95% CI	p
Intercept	-0.0532	1.1752	[-2.357, +2.250]	0.964
Region: East Asia & Pacific	+0.3412	0.1713	[+0.005, +0.677]	0.046
Region: Latin America & Caribbean	+2.1433	0.1823	[+1.786, +2.501]	0.000
Region: Middle East & North Africa	-0.0011	0.2320	[-0.456, +0.454]	0.996
Region: North America	+0.3747	0.3547	[-0.320, +1.070]	0.291
Region: South Asia	-0.1742	0.2263	[-0.618, +0.269]	0.442
Region: Sub-Saharan Africa	+0.2828	0.4257	[-0.552, +1.117]	0.507
own10	+0.0909	0.0404	[+0.012, +0.170]	0.024
log_gdp	+0.0195	0.1807	[-0.335, +0.374]	0.914
urban_population_pct	-0.0073	0.0076	[-0.022, +0.008]	0.338
unemployment_pct	+0.0210	0.0101	[+0.001, +0.041]	0.037
male_15_64_share_pct	+0.0082	0.0191	[-0.029, +0.046]	0.668
governance_index	-0.0416	0.1491	[-0.334, +0.251]	0.780

A.2 Hypothesis decisions

Table 12. Summary of hypothesis tests H1-H6.

Hypothesis	Evidence	Decision
H1 ownership association	Unadjusted slope -0.53 per +10 (p = 0.36); Model 5 log1p 0.091 [0.012, 0.170], p = 0.024; level-scale p = 0.48; NB IRR 1.32	Reject H0 conditionally (log/count scales) but not unconditionally or on the level scale; association is fragile (robustness: 62% of specifications with CI above zero)
H2 socioeconomic adjustment	Socioeconomic block alone: F = 1.11, p = 0.36 (R2 0.000 to 0.052); region + governance: F = 35.1, p < 0.001 (R2 to 0.778)	Socioeconomic covariates alone add little explanatory power; regional context and governance add a great deal. Reject H0 for the broader country-context block
H3 regional heterogeneity	Equal-slope test across WB regions (n >= 5): p = 1.00; three-group interaction p = 0.33; income interaction p = 0.67	Fail to reject equal slopes; regions differ greatly in levels, not detectably in slopes (low power)
H4 nonlinearity	Spline vs linear LR p = 0.12; GAM AIC 162.0 vs linear GAM 160.9; hinge at 3.4 with bootstrap range 3-22	Fail to reject linearity; no stable threshold
H5 predictive contribution	Ridge: MAE 3.98 to 4.01 (+0.8%, fold 95% range -0.10 to +0.26); Random forest: MAE 3.49 to 3.47 (-0.5%, fold 95% range -0.18 to +0.26); Gradient boosting: MAE 3.73 to 3.75 (+0.7%, fold 95% range -0.30 to +0.41)	Fail to reject H0: adding ownership changes out-of-sample error by less than 3% and the fold-paired range includes zero for every learner
H6 spatial dependence	Moran's I of Model 5 residuals (kNN-4) = 0.26, p = 0.001; Model 3 (no region) = 0.79	Reject H0: residuals are spatially dependent; region effects absorb most but not all of it

A.3 Hyperparameters

Table 13. Hyperparameters selected most often in the inner searches.

Model	Most frequent tuned setting (outer folds)
Dummy (median)	untuned
Linear regression	untuned
Ridge	[('alpha', np.float64(1.743328822199989))]
Lasso	[('alpha', np.float64(0.035622478902624426))]
Elastic Net	[('alpha', np.float64(0.018329807108324356)), ('l1_ratio', 0.9)]
Huber	[('alpha', np.float64(1.232846739442066)), ('epsilon', 1.2)]
Random forest	[('max_depth', 3), ('max_features', 1.0), ('min_samples_leaf', 6)]
Extra Trees	[('max_depth', 3), ('max_features', 0.6), ('min_samples_leaf', 6)]
Gradient boosting	[('learning_rate', 0.02), ('max_depth', 1), ('n_estimators', 200), ('subsample', 0.7)]
XGBoost	[('learning_rate', 0.05), ('max_depth', 1), ('n_estimators', 200), ('reg_lambda', 20), ('subsample', 0.7)]
LightGBM	[('learning_rate', 0.02), ('min_child_samples', 10), ('n_estimators', 100), ('num_leaves', 8)]
CatBoost	[('depth', 2), ('l2_leaf_reg', 10), ('learning_rate', 0.03)]
Support vector regression	[('C', np.float64(28.48035868435799)), ('epsilon', 0.05), ('gamma', 0.01)]
GAM (pygam)	untuned

A.4 Feature definitions and provenance

Firearms per 100: Small Arms Survey 2017. Firearm homicide rate: UNODC, mean of available years 2013-2022. Socioeconomic covariates (GDP per capita, Gini, unemployment, urban share, males 15-64, secondary enrolment, military spending, population): World Bank WDI as provided in the Kaggle panel, 2013-2022 means. Governance: WGI estimates, 2013-2022 means; governance index = first principal component (loadings rule of law 0.52, gov effectiveness 0.50, political stability 0.47, control corruption 0.52). Conflict: UCDP BRD v25.1, country is the location of an incompatibility with at least 1,000 battle-related deaths in any window year. Excluded from all predictor sets as target-derived: fh_rate, log1p_fh, fh_count, fh_pooled_rate, fh_share_pct, hom_rate, log1p_hom, nonfirearm_rate, person_years, fh_years, fh_first_year, fh_last_year, hom_years.

A.5 Country coverage

Primary-sample countries (96): Albania, Algeria, Andorra, Antigua and Barbuda, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahamas, The, Barbados, Belarus, Belize, Bermuda, Bhutan, Bolivia, Bosnia and Herzegovina, Brazil, Bulgaria, Cabo Verde, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominica, Dominican Republic, Ecuador, El Salvador, Eswatini, Finland, France, Georgia, Germany, Greece, Grenada, Guatemala, Guyana, Honduras, Hong Kong SAR, China, Hungary, Iceland, Iran, Islamic Rep., Italy, Jamaica, Japan, Jordan, Kazakhstan, Korea, Rep., Liechtenstein, Lithuania, Luxembourg, Macao SAR, China, Malta, Mauritius, Mexico, Moldova, Monaco, Mongolia, Montenegro, Morocco, Netherlands, New Zealand, Nicaragua, North Macedonia, Oman, Panama, Paraguay, Peru, Philippines, Puerto Rico (US), Serbia, Slovak Republic, Slovenia, Spain, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Suriname, Sweden, Switzerland, Tajikistan, Tanzania, Thailand, Trinidad and Tobago, Uganda, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, West Bank and Gaza.

Reproducibility: `python run_pipeline.py` rebuilds every table, figure and this paper from the raw files; seed 20171100.