PUNGGAWA
LAW REVIEW
Punggawa Law Review
Volume 2, Number 1, 2027
E-ISSN: 3125-0580
Open Access: punggawajournal.com/plr

# When Homicide Statistics Contradict Themselves: A Reproducible Simulation of Within–Between Reversals in Comparative Criminology

**Zul Khaidir Kadir**[1]
[1] Fakultas Hukum, Universitas Muslim Indonesia, Indonesia
*Correspondence:* zulkhaidir.kadir@umi.ac.id

Article Info

Article history:
Received Sept 28th, 2026
Revised Nov 6th, 2026
Accepted Nov 14th, 2026

Keyword:
Comparative criminology; Homicide; Inferential validity; Panel simulation; Within–between effects.

ABSTRACT

International homicide comparisons routinely place countries with different income levels on the same graph. The association across those countries is sometimes discussed as though it described what happens when income changes inside a country. The two comparisons, however, draw on different information. We examine their distinction through a reproducible simulation, alongside published regional homicide estimates from the United Nations Office on Drugs and Crime. The experiment contains 120 synthetic units observed over 14 periods. A negative between-unit coefficient and a positive within-unit coefficient are imposed in the generating process, then estimated under sampling noise and alternative analytic choices. The reference estimates are −0.606 and +0.516, with a within-minus-between contrast of +1.122 and a unit-bootstrap 95% interval of +1.041 to +1.203. Further simulations examine sampling precision, incomplete coverage and specification changes. The UNODC estimates establish that homicide rates vary across regions. They do not identify an income–homicide relationship. Opposing coefficients in the generated panel demonstrate an inferential possibility, not a reversal observed among actual countries. The resulting contribution is a replicable test of whether the statistical comparison reported in comparative homicide research matches the proposition used to interpret it.

© 2026 The Authors. Published by Punggawa Legacy Center. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

## INTRODUCTION

A country with high average income and low homicide may appear to support the proposition that prosperity protects against lethal violence. Its position relative to poorer countries, however, says little about what happened to homicide as that same country grew wealthier. The familiar comparison draws on long-standing differences between social settings, while the claim about economic change concerns movement through time. These distinctions have practical consequences for criminological theory. Messner (1982) questioned simple associations between poverty, inequality and urban homicide. Pridemore (2011) later showed how the handling of poverty alters interpretations of the cross-national inequality–homicide relationship. Income remains a plausible correlate of homicide. The problem begins when a relationship estimated among countries is made to explain changes inside them without further evidence.

GDP per capita compresses distinct social processes into one national average. Economic growth can coincide with institutional reform, uneven distributional gains or changes in population composition, yet the GDP series does not tell us which of these occurred. Messner (1989) associated societal homicide with economic discrimination, while Fajnzylber, Lederman and Loayza (2002a) examined income inequality and violent crime without treating prosperity as a complete social explanation. Their broader analysis further established that different crime outcomes call for different explanatory models (Fajnzylber, Lederman and Loayza 2002b). A GDP coefficient may still be

informative. On its own, however, it cannot identify the institutional or social pathway behind an association with killing. Population composition adds another complication. Kadir (2026a) found that age-conditioned homicide rankings were sensitive to adjustment assumptions, without establishing an effect of ageing on violence.

Mundlak (1978) supplied a way to distinguish differences between units from departures around their own means. Longitudinal analysis now uses this separation to clarify what an estimate represents. Bell and Jones (2015) emphasised the substantive value of reporting within and between associations, while Schunck (2013) described how they can be obtained through hybrid specifications. Bell, Fairbrother and Jones (2019) also cautioned against reducing model choice to a routine contest between fixed and random effects. None of these approaches makes a within coefficient causal by definition. Their immediate contribution is more specific. They keep the researcher from assigning one comparison the interpretation of another. Stating the two terms separately also prevents an apparently stable pooled coefficient from hiding an entirely different temporal association.

International homicide figures have another weakness before regression even begins. Police statistics and mortality records may apply different classification rules, and the coverage of either system varies. Harrendorf (2018) examined the resulting problems for comparative criminal justice data. Rogers and Pridemore (2023) found that homicide measurement choices matter for estimated cross-national associations. More recently, Buil-Gil, Pina-Sánchez and Aebi (2026) quantified consequences of national counting rules. Homicide is comparatively visible because it involves a death, but visibility is not equivalence. Kadir (2026g) found disagreement between UNODC and mortality-based homicide measures despite their strong cross-national correlation. A precise coefficient can still rest on outcomes measured differently across settings.

The contrast is visible in published homicide estimates, even before economic data are introduced. UNODC reports a global rate of 5.8 victims per 100,000 people in 2010 and 5.0 in 2023. Its 2023 estimates range from 19.3 in Latin America and the Caribbean to 0.6 in Eastern and South-Eastern Asia (UNODC 2025). Those regional values establish a difference in reported homicide burden, not a difference attributable to national income. Comparative studies have identified links with policy, political legitimacy and social conditions. Neumayer (2003) examined homicide in a country panel, LaFree and Tseloni (2006) investigated long-run trajectories, and Nivette and Eisner (2013) studied political legitimacy. The range of explanations makes it especially hazardous to let a single GDP association stand for an economic mechanism. Eisner (2003) also documents the historical scale of changes in violent crime, beyond short panel windows.

Many longitudinal studies estimate within-country relationships, but fewer present them beside between-country estimates obtained from the same units and years. The gap worth testing is therefore precise. Does an apparently stable income gradient among countries survive when the analysis uses changes inside those countries instead? A difference need not mean either calculation was incorrect. It may mean the two calculations addressed different propositions. Imai and Kim (2019) discuss the assumptions required for causal interpretation of unit fixed effects, while Hamaker, Kuiper and Grasman (2015) show how longitudinal models can misrepresent distinct levels of variation. These concerns warrant explicit comparison rather than a general accusation of faulty regression. A descriptive comparison by Kadir (2026c) also found divergent short-period homicide trajectories among countries with similar starting rates, without attributing those trajectories to a common cause.

An observed reversal in global homicide cannot be established without matched national income and homicide records. The regional UNODC figures available for this exercise contain no country-year GDP observations. We therefore do not estimate a real-world income effect. Instead, a controlled simulation allows the two components to be assigned known values before the estimators are applied. The exercise can establish when the resulting slopes disagree and whether common modelling choices conceal that difference. It cannot tell us how often this occurs among countries.

The question tested here is whether one generated panel can produce a negative association among unit averages and a positive association within units, while its pooled estimate resembles the former. Repeated experiments vary uncertainty, sample composition and the time window to expose how that conclusion depends on the design. The contribution lies in matching each inferential statement to the comparison that supports it. Published UNODC estimates supply genuine geographic context. Simulated observations supply a separate demonstration of coefficient behaviour. Keeping those forms

of evidence distinct gives the ensuing discussion a narrower, more defensible claim about comparative criminology.

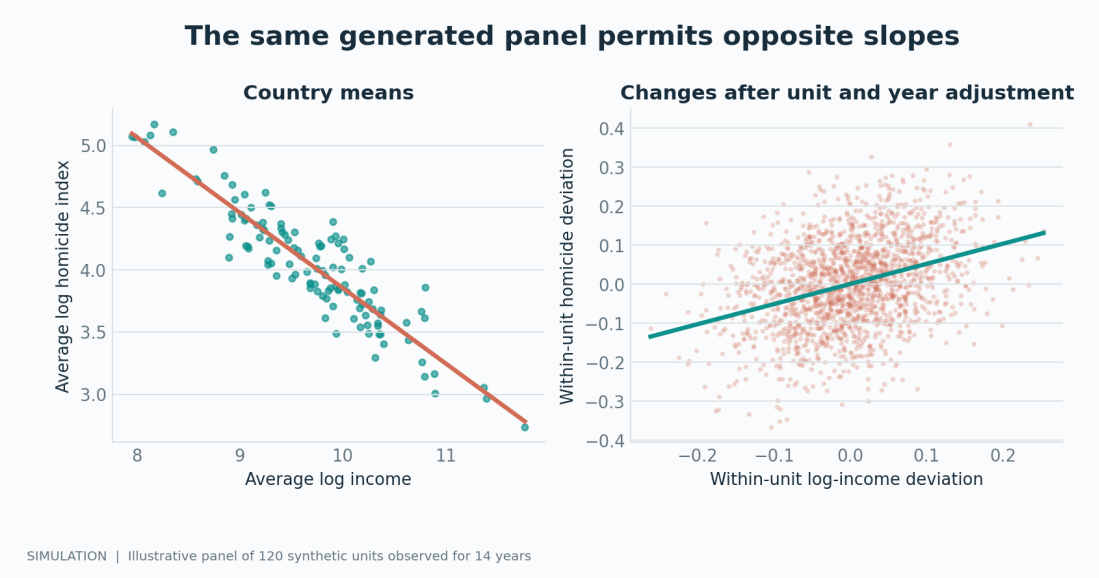
## RESEARCH METHOD

Published regional homicide rates and a synthetic panel serve different purposes in the design. The observed benchmarks come from Figure 7 of a UNODC research brief published in November 2025 and cover 2010, 2020 and 2023. They support descriptive regional comparisons only. The statistical experiment generates 120 country-like units over 14 indexed periods, displayed as 2010–2023 for convenience rather than as historical observations. There are 1,680 synthetic unit-period records. Log-income means vary among units, and each income trajectory contains a unit-specific trend and annual shocks. The transformed homicide-like outcome has an imposed between-income parameter of −0.65 and a within-income parameter of +0.45, with persistent and transitory disturbances. The simulation also includes a common time component. Numeric random seeds are saved for replication.

Between-unit association is estimated by regressing average transformed homicide-like outcomes on average log income for the 120 synthetic units. The within slope comes from a two-way fixed-effects regression that removes unit and period means. A pooled regression is retained as a diagnostic, not as a substitute for either coefficient. We calculate the within-minus-between contrast and obtain percentile intervals from 700 bootstrap resamples of entire units. To examine sampling behaviour, 350 new panels are generated and analysed using the same estimators. A further sweep assigns 11 distinct within parameters and repeats each specification 100 times. The substantive meaning of all estimated slopes is restricted to the declared generating process. None refers to observed national homicide.

The sensitivity analysis uses only features present in the generated data. Separate experiments change the noise level and the number of units, re-estimate the model over shorter time windows, remove influential units and select complete units with unequal probabilities. Four alternative within parameters test whether a similar pooled coefficient can arise from distinct underlying associations. Fifteen prespecified variants examine the reference-panel contrast across limited sample and income-series choices. Figures display the corresponding computed estimates, and all scripts retain their random seeds. The regional homicide figures are reproduced from the UNODC source, without filling missing annual observations or joining them to fictional income values. Consequently, the empirical benchmarks are descriptive, while the regression findings are statements about the simulation.

*Figure 1. Opposing associations identified from one synthetic panel*




Unit averages give the negative fitted relationship in the left panel. The right panel instead uses residuals after adjustment for unit and period means and gives a positive slope. Both panels contain information from the same generated dataset. Their divergence does not arise from changing the outcome or selecting an alternative source. It arises because the estimators use different variation. The

illustration cannot establish that any observed country experiences this reversal, since the relevant parameters were assigned in advance.

## RESULT AND DISCUSSION

### Published Homicide Variation and Its Evidentiary Limits

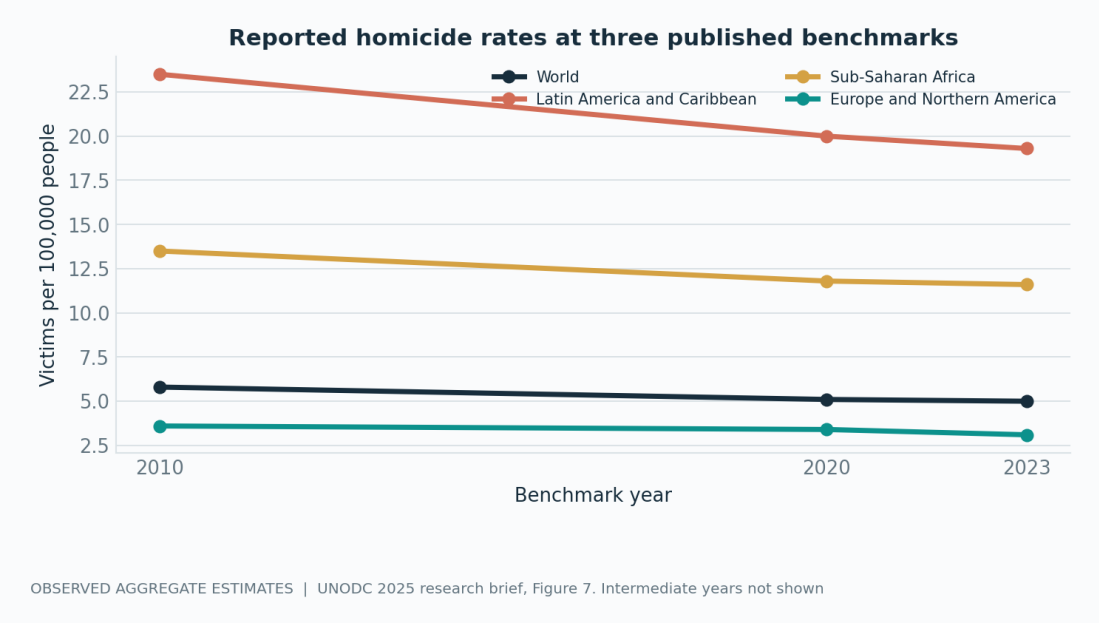
UNODC reports that the global intentional homicide rate declined from 5.8 per 100,000 people in 2010 to 5.0 in 2023. Its 2023 regional estimates differ sharply. Latin America and the Caribbean recorded 19.3, while Eastern and South-Eastern Asia recorded 0.6, both measured per 100,000 population. These observations describe published homicide levels, not outcomes produced by the simulation. The regional table includes neither national income nor matched country-year pairs. It therefore cannot establish a GDP gradient, much less determine whether between-country and within-country coefficients differ. UNODC also qualifies its regional estimates where data coverage is limited (UNODC 2025). Geographic variation is documented. An economic explanation remains untested. The values also leave open substantial differences among countries within each region, which the reported aggregation cannot resolve.

*Table 1. Published UNODC homicide rates by region and year*

| Region | 2010 | 2020 | 2023 |
|---|---|---|---|
| Latin America and Caribbean | 23.5 | 20.0 | 19.3 |
| Sub-Saharan Africa | 13.5 | 11.8 | 11.6 |
| Northern Africa and Western Asia | 4.5 | 4.7 | 4.8 |
| Oceania | 2.7 | 2.7 | 2.7 |
| Europe and Northern America | 3.6 | 3.4 | 3.1 |
| Eastern and South-Eastern Asia | 1.4 | 0.7 | 0.6 |
| World | 5.8 | 5.1 | 5.0 |

Estimated victims per 100,000 population, based on UNODC (2025), Figure 7. These regional aggregates contain no country-level GDP observations.

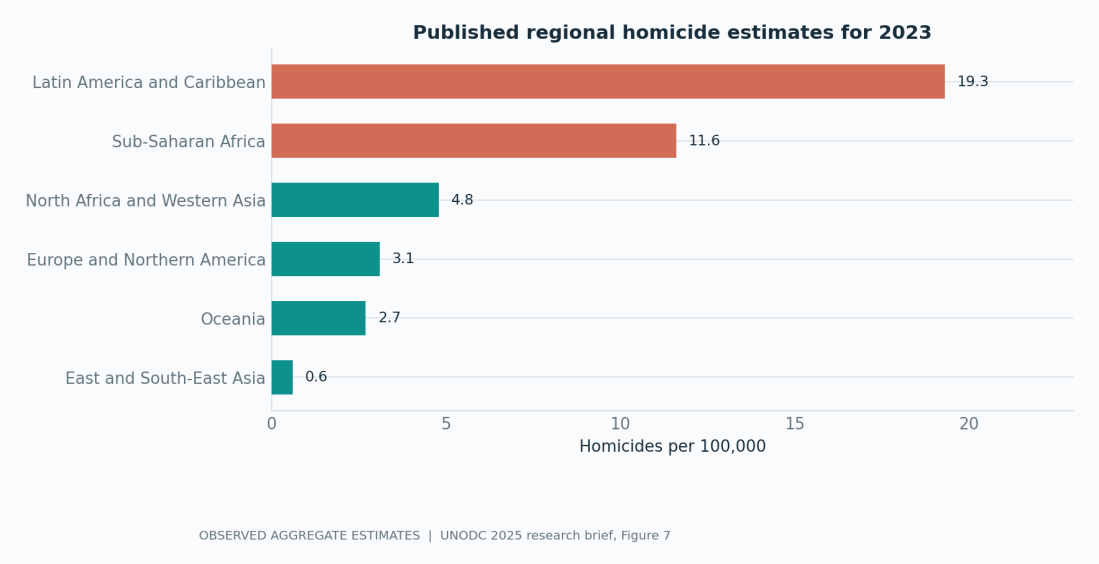
*Figure 2. Published homicide rates at three benchmark years*




The plotted lines connect three published observations for each series. Between those dates, no annual values have been inserted. The world estimates decline at the reported benchmarks, although the regional paths are not uniform. Northern Africa and Western Asia, for instance, have a higher reported

rate in 2023 than in 2010. Neither the connecting segments nor their apparent slopes should be read as estimated annual trends. The evidence describes three snapshots, not a continuous longitudinal model. Connecting 2010 with 2020 and 2023 is useful for locating the published estimates, but the resulting segments do not recover intervening fluctuations. The reported fall in the global rate is supported by the endpoints. A claim about uninterrupted annual decline would require all intervening observations and a continuity audit. Changes in source coverage can also complicate regional comparisons across time. These limits do not make the UNODC figures uninformative. They remain a useful description of homicide burden at the published dates. What they cannot provide is the covariance between national income and homicide that a within–between comparison would need.

***Figure 3. Regional intentional homicide rates in 2023***




Reported 2023 homicide rates differ by more than an order of magnitude across the six regions displayed. The values are legitimate descriptive aggregates, yet none contains the matched national GDP observations required for the income regression examined later. The bars neither rank individual countries nor establish why one region reports more lethal violence than another. Their purpose is to preserve the distinction between published homicide evidence and the generated units used for the methodological experiment.

Regional means can obscure large differences among the countries that compose them. Kadir (2026e) similarly documented cross-national differences in the recorded lethal share of serious violence, while cautioning that this ecological ratio was not an individual-level escalation probability. The UNODC values presented here cannot be unpacked into national observations, nor do they identify how much each country contributes to a regional total. Consequently, an explanation fitted to a region would address a different unit of analysis from one fitted to countries. Spatial work on homicide and political legitimacy, including Nivette (2012), illustrates why aggregation choices matter. The figure establishes only the published ordering of regional homicide rates. Its high and low endpoints should not be turned into a claim about the income characteristics of particular countries.

A national prosperity–homicide association would require observed income and homicide for the same countries and years, together with a rule for selecting those records. A review of adjusted homicide rankings by Kadir (2026j) likewise distinguished recorded-rank movement from the unknown true ranking of lethal risk. A longitudinal claim would also require a model that isolates change inside each country. The regional table meets neither condition. It supplies background rather than a test of economic theory. Research into welfare arrangements and homicide suggests mechanisms that might matter in a genuine country panel (Rogers and Pridemore 2017, Rudolph and Starke 2020). Here those studies inform the substantive motivation, but their findings are not used to manufacture missing observations or to infer an income coefficient from regional bars.

**A Known Generating Mechanism with Opposing Associations**

The simulation fixes an otherwise unobservable distinction before analysis begins. Differences in average income among units have a negative coefficient in the generating equation, while deviations from a unit's own average have a positive coefficient. A shared temporal shift is added and later

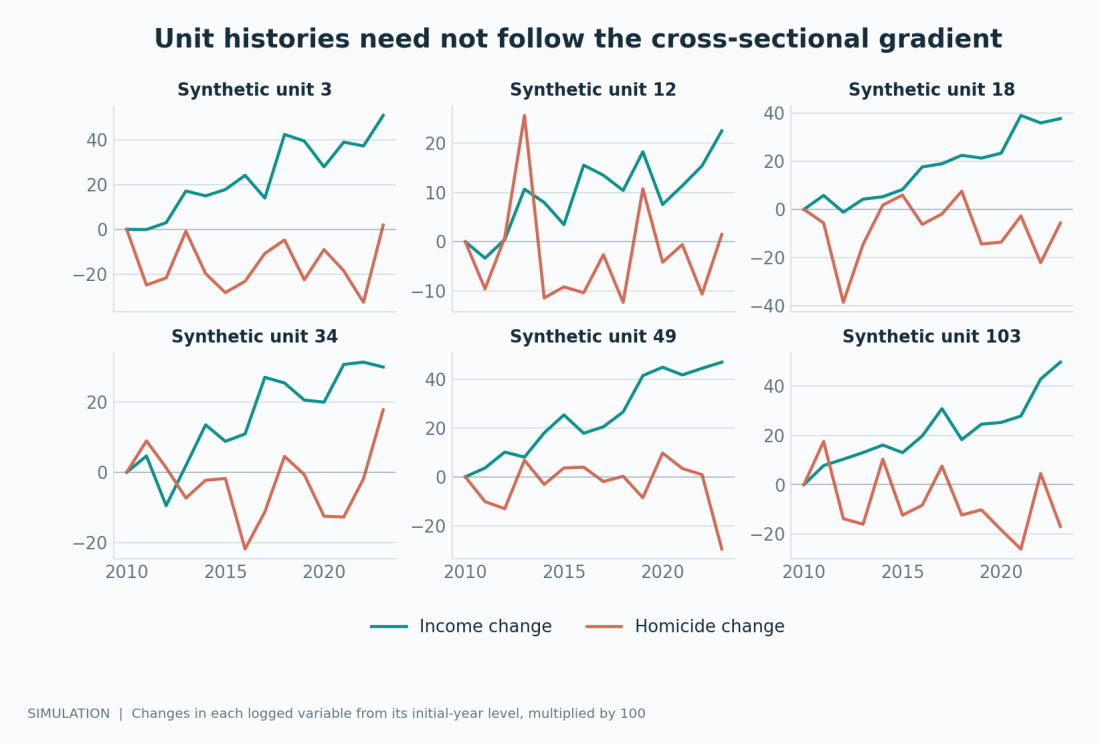
absorbed through year effects. Stable, unexplained unit differences and random disturbances prevent a mechanically perfect scatterplot. Table 2 records the parameters used. They are experimental choices rather than estimates of homicide determinants. Calling the outcome homicide-like keeps the subject matter recognisable without suggesting that the generated records reproduce the causes or recording practices of actual countries. Changing the relative dispersion of income between and within units would change the contribution each component makes to the pooled estimate.

*Table 2. Parameters and structure of the synthetic experiment*

| Component | Design setting | Interpretive role |
|---|---|---|
| Synthetic units and periods | 120 units × 14 periods | Balanced reference panel |
| Between-income parameter | −0.65 | Imposed cross-unit gradient |
| Within-income parameter | +0.45 | Imposed within-unit relationship |
| Baseline log-income mean | 9.65 | Location of generated x |
| Income mean dispersion | 0.84 standard deviation | Cross-unit income heterogeneity |
| Persistent unit disturbance | 0.18 standard deviation | Unexplained stable differences |
| Outcome disturbance | 0.11 standard deviation | Transitory uncertainty |
| Common time shift | −0.017 per period | Shared trend controlled by year effects |
| Main experiment seed | 4937 | Reference-panel reproducibility |
| Country-block bootstrap | 700 resamples | Contrast uncertainty |

All quantities refer to the data-generating process, not characteristics measured for any real country.

*Figure 4. Income and homicide-like trajectories for six synthetic units*

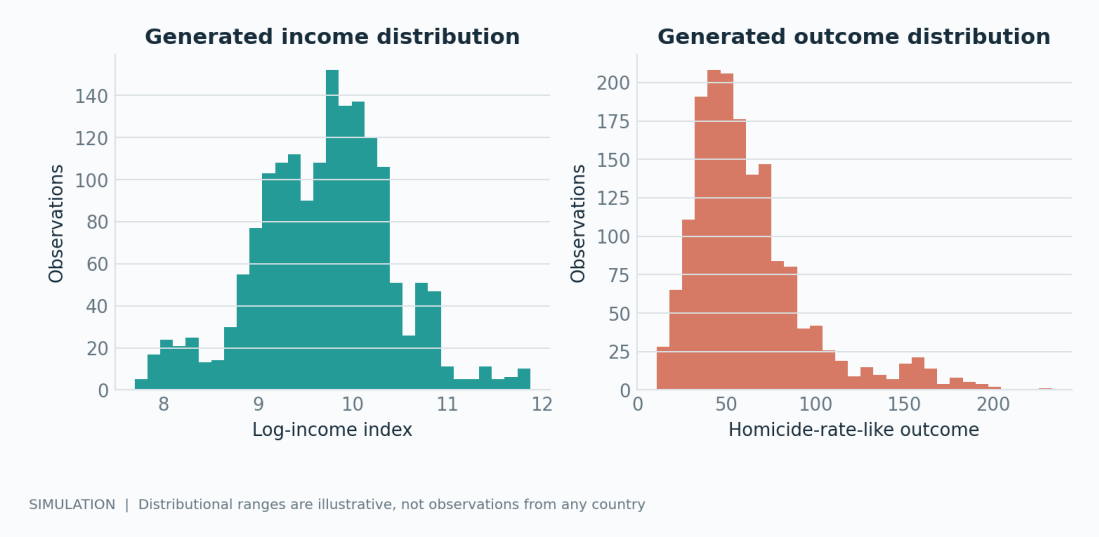



Six units are followed through all 14 simulated periods. Their income paths tend upward, but the homicide-like values fluctuate rather than tracing smooth responses. The chosen identifiers were fixed before the trajectories were inspected. Such examples reveal the noise surrounding a shared generating relation without being selected for dramatic appearance. The figures also warn against treating a visible co-movement in a single unit as a causal estimate. A unit history is part of the sample, not an independent validation of the mechanism.

The six histories share features of the generating equation, so they cannot be treated as six independent case studies. Some movements come from common time effects, while others reflect unit trends and random disturbances. A unit may display a rising or falling sequence that differs from the

average conditional association. The fixed-effects estimate pools information from all 120 units after defined adjustments. It does not describe every trajectory equally well. With observed homicide, local shocks and changes in recording can create further departures from a smooth national pattern. The small panels make the distinction visible without identifying any such historical mechanism.

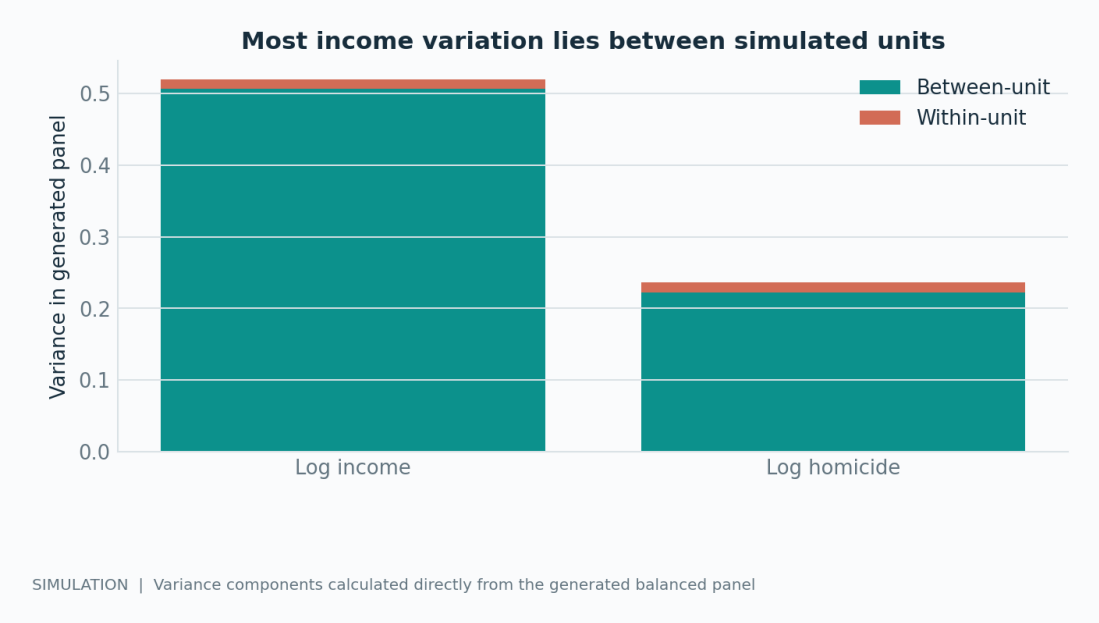
***Figure 5. Distributions of the generated income and homicide-like variables***




Generated income covers a relatively wide range because its unit means have deliberately high dispersion. The homicide-like measure is more right-skewed after the logarithmic outcome is transformed back to its original scale. Neither distribution was estimated from UNODC national data. Their shapes depend on the parameters in Table 2, including the intercept variation and random disturbances. The display mainly documents the scale on which the regression exercise operates, thereby making the choice of transformations and the influence of high generated values easier to inspect.

A log transformation changes what a slope measures and how extreme values influence fitting. In this experiment the outcome is created on the log-one-plus-rate scale, so regression on that same scale is intentionally well matched to the generating equation. That convenience cannot justify applying the transformation without inspection to a real homicide dataset. Observed rates may contain zeros, population denominators vary greatly, and countries can differ in data quality. Counts fitted with a population offset would answer a related but not identical question. The present result should be interpreted in generated transformed units, with no direct translation into a universal percentage change in actual homicide.

*Figure 6. Between-unit and within-unit variation in generated income*




Stable differences in mean income account for most of its variance in the generated panel. Departures around each unit mean are smaller, so the pooled estimate gives greater weight to the negative cross-unit component. This composition was chosen by design. It does not establish the relative importance of these components in world GDP records. The decomposition is nevertheless useful because it exposes the mathematical route through which a pooled coefficient can resemble the between slope even when the within slope points in the opposite direction (Bollen and Brand 2010, Curran and Bauer 2011).

In a balanced panel, total covariance can be partitioned into the component arising among unit means and the component arising from deviations within units. The corresponding income variances determine their influence on the pooled coefficient. Between-unit variation dominates the generated data, making the pooled slope negative. Within-unit covariance is positive, but its contribution carries less weight. The pooled result is therefore not a summary of the temporal association. This conclusion follows from the structure of the synthetic variables rather than from a claim about economic theory. A researcher can calculate the same pooled number correctly and still misdescribe what it says about changes in homicide.

**Pooled, Between and Within Estimates**

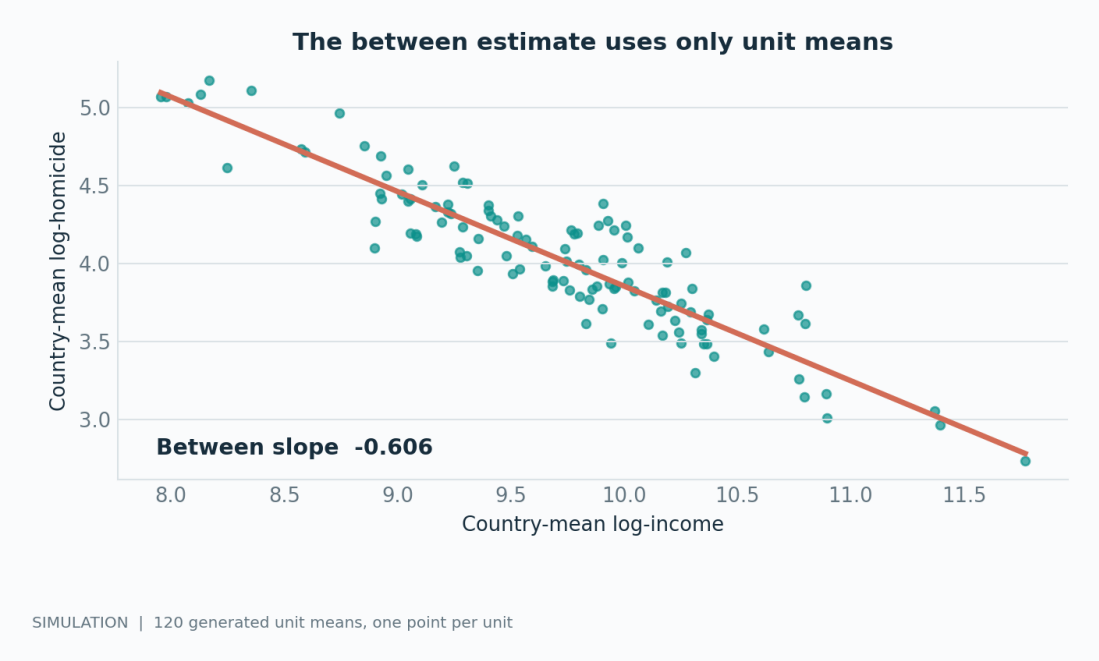
The reference panel yields a between estimate of −0.606 and a within estimate of +0.516. The pooled coefficient is −0.590, close to the between slope because the synthetic income variation is dominated by differences in unit averages. The within-minus-between contrast is +1.122. These are fitted coefficients from the generated observations, with uncertainty assessed as reported in Table 3. Their opposite signs were not discovered in external homicide data. Instead, the generating mechanism was designed with coefficients of opposite signs so that the ability of each estimator to recover its target could be examined.

***Table 3. Reference-panel coefficients and unit-bootstrap intervals***

| Estimand | Estimate | Bootstrap 95% interval | Analytical unit |
|---|---|---|---|
| Pooled | -0.590 | -0.631 to -0.545 | 1,680 generated records |
| Between | -0.606 | -0.650 to -0.561 | 120 generated unit means |
| Within | +0.516 | +0.446 to +0.579 | 1,680 generated records |
| Within minus between | +1.122 | +1.041 to +1.203 | Country-block bootstrap |

Percentile intervals from 700 resamples of complete synthetic units. They quantify conditional sampling uncertainty within the simulated design.

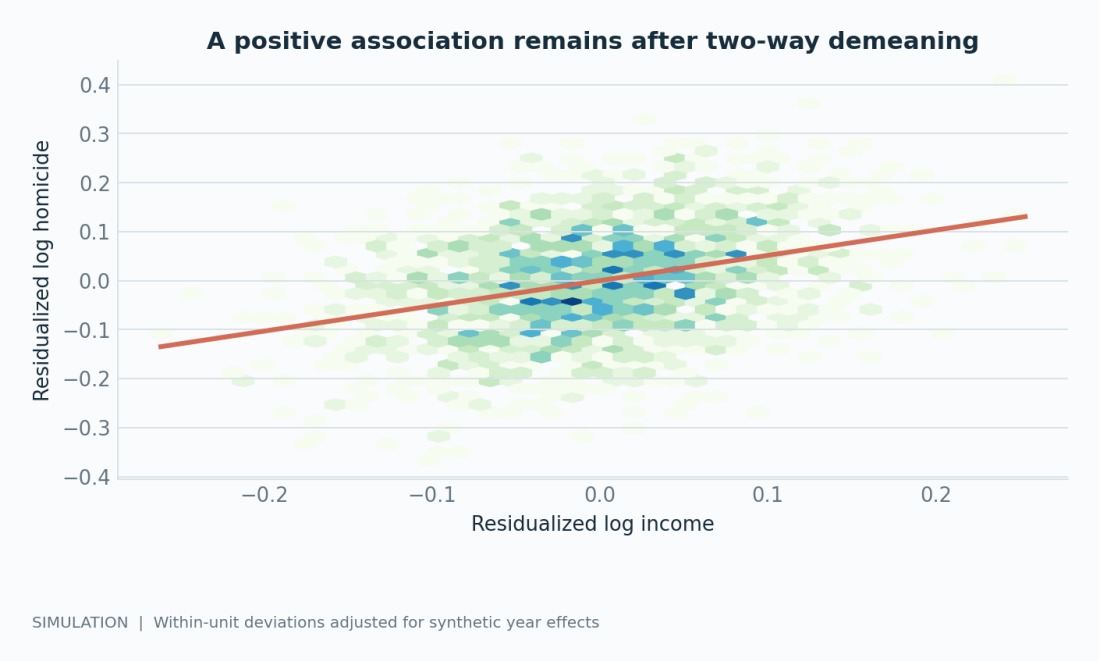
***Figure 7. Between-unit association based on synthetic unit means***




The regression in Figure 7 uses one mean observation from each of 120 generated units. Its negative slope describes differences among those unit averages, not movements inside any individual unit. The unobserved within coefficient cannot be recovered from this scatterplot unless assumptions are introduced that connect the two levels of variation. A cross-sectional association may be informative for a cross-sectional question. Problems arise when an account of temporal change is attached to it without temporal evidence. Nivette (2011) reviewed the varied structural predictors used in cross-national crime research.

Unexplained stable differences remain in the generated unit means, so the between regression is neither deterministic nor a uniquely causal description of income. Observed nations would carry many additional sources of persistent heterogeneity. Criminal justice administration, inequality and political legitimacy could all vary systematically with prosperity. Nivette and Eisner (2013) examined legitimacy, while Messner, Raffalovich and Sutton (2010) questioned the validity of constructs used in international homicide comparisons. Their work does not validate the synthetic slope, but it gives substantive reasons not to equate a between coefficient with the effect of income alone. The coefficient here describes the relation imposed and estimated across artificial units.

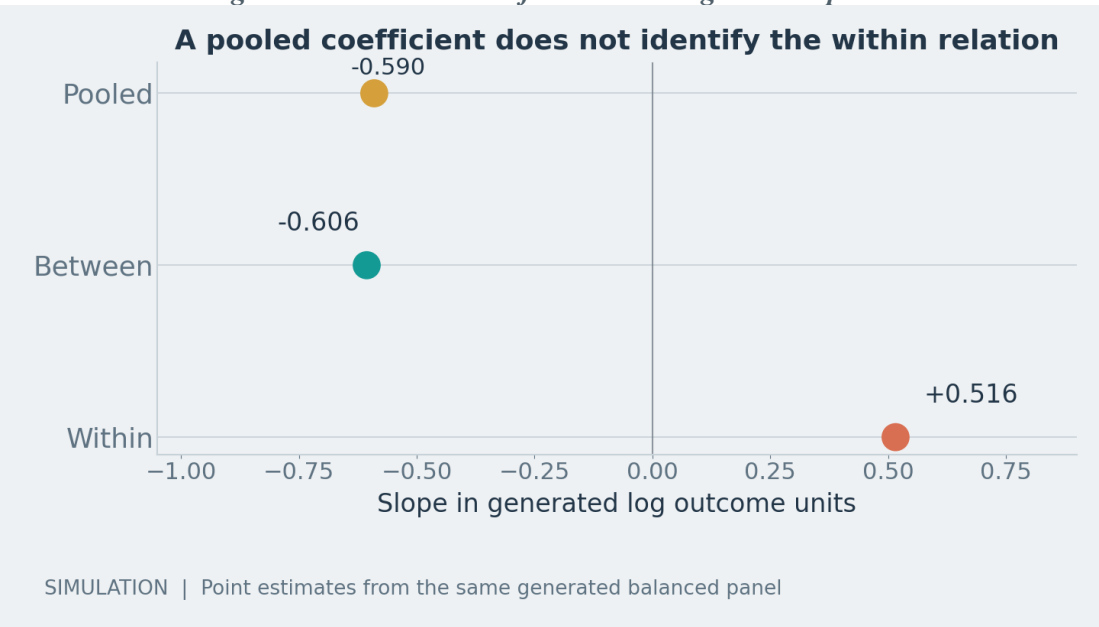
*Figure 8. Within-unit association after unit and period adjustment*




The residualised plot retains all 1,680 generated observations. Income and the homicide-like outcome are adjusted for unit and period means, leaving the variation from which the two-way fixed-effects slope is calculated. Its fitted line is positive, unlike the negative regression on unit averages. The plot therefore visualises the information used by the within model. It cannot establish causality in observational homicide research, where changing confounders may remain even after fixed effects are included (Imai and Kim 2019).

Removing unit means eliminates additive features that remain constant within each unit, and year effects remove shocks shared by every unit in a given period. Neither operation controls for a factor that varies within units and jointly affects income and homicide. Those confounders were not introduced into the main experiment, which tests coefficient recovery under a known generating equation. Real data would demand a separate account of timing, mechanisms and possible reverse causality. Adding observed controls may improve interpretation, but it may also change the question or condition on variables influenced by income. Fixed effects specify a comparison. They do not supply a causal identification strategy on their own.

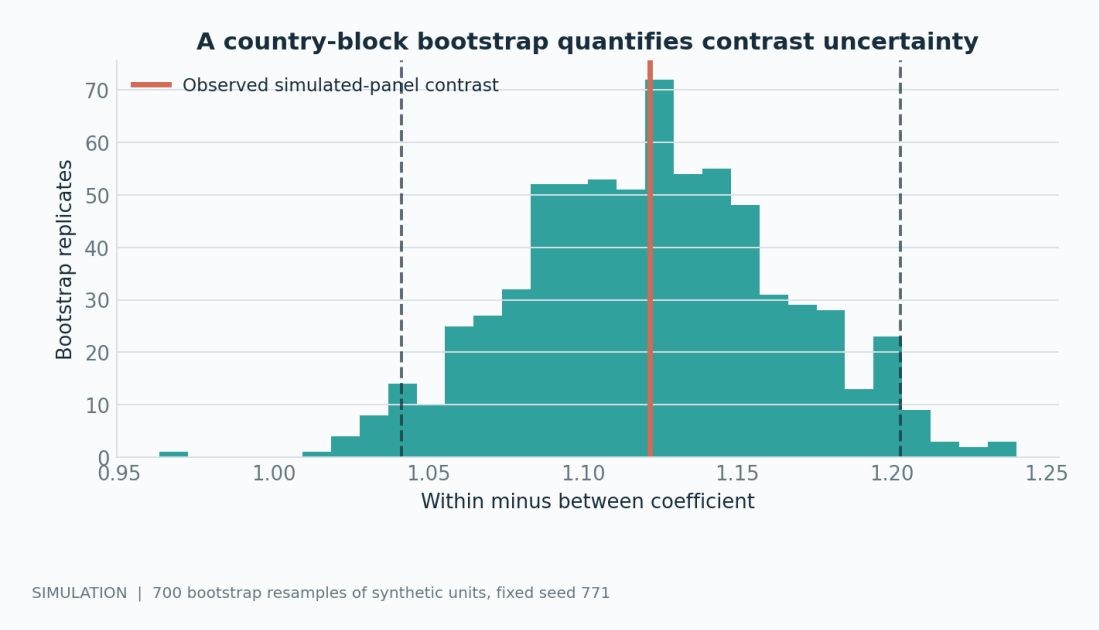
*Figure 9. Three estimates from the same generated panel*



The pooled slope is negative even though the fitted within slope is positive. Because the generated income variance is concentrated between units, the pooled covariance largely follows stable differences in their averages. There is no numerical error in the pooled regression. The interpretive error would arise if its coefficient were reported as an estimate of the change associated with income growth inside a unit. Bell, Fairbrother and Jones (2019) argue for choosing longitudinal models in light of the question rather than a conventional fixed-versus-random preference.

Pooling repeated observations does not automatically make a coefficient longitudinal in meaning. In this simulation, the pooled fit answers a question about all unit-period records jointly, and the high dispersion of unit means dominates that fit. The within model answers a question about deviations from those means after common-time adjustment. A small standard error on the pooled regression would not correct the mismatch if the paper's claim concerned economic change. Statistical precision is distinct from correspondence between the estimand and the proposition. This is why readers should examine the source of variation before treating significance as evidence for a particular crime theory.

***Figure 10. Bootstrap uncertainty for the within–between contrast***




The fitted within-minus-between contrast is +1.122. Across 700 resamples that draw complete synthetic units, the percentile 95% interval spans +1.041 to +1.203 and remains above zero. Resampling at the unit level retains the dependence among records from the same generated unit. The calculation is appropriate to the sampling structure used in this experiment, though other sources of dependence would demand additional thought in observational settings. Cameron and Miller (2015) and Abadie and colleagues (2023) discuss principles of cluster-based inference. The displayed interval is not a confidence interval for any observed GDP–homicide relationship.

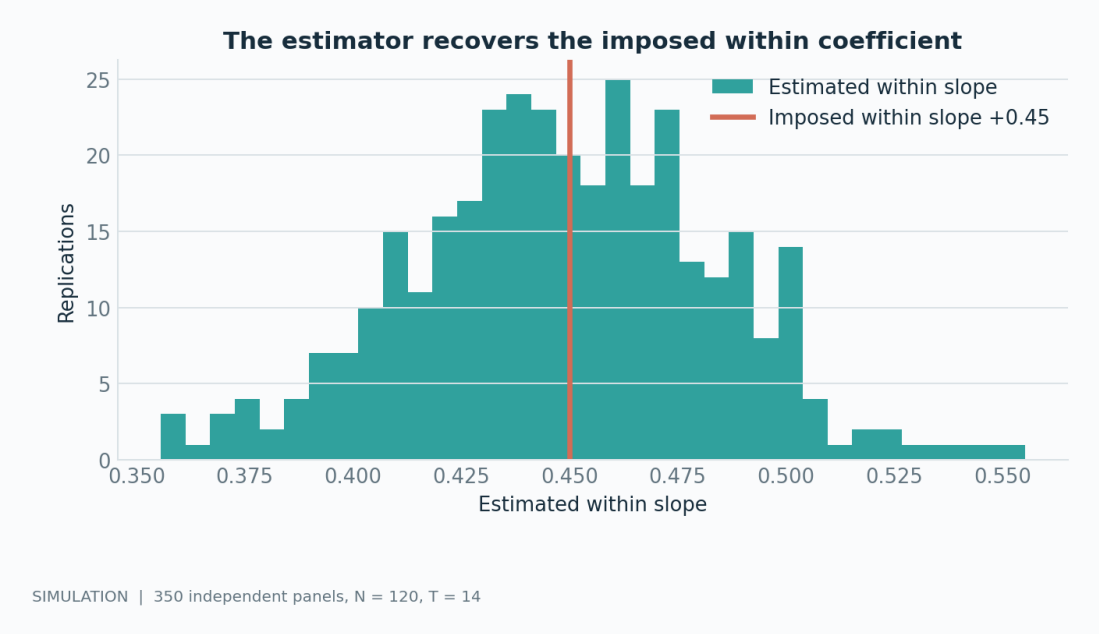
A narrow bootstrap interval is informative about the reference design and no more. Resampling the 120 synthetic units approximates variability under repeated draws resembling the simulated population. The generating parameters, including the opposed signs, remain fixed throughout. The interval therefore cannot express uncertainty about whether actual countries have those relationships. Seven hundred replications are adequate for this demonstration's reported precision, but the number is not offered as a general rule. In an observed panel, clustering would need to match its dependence structure, and confidence intervals could not repair mismatched definitions of homicide or poorly justified sample selection.

**Repeated Simulations and Sources of Uncertainty**

The 350 independent generated panels return an average within estimate of +0.449, close to the imposed value of +0.45. Every panel in this experiment has opposite fitted signs for the between and within coefficients. The result is expected, since the generating process fixes the opposite directions and the design contains enough units to estimate them precisely. It does not estimate how common reversals are in comparative homicide research. Repeated simulation instead checks whether the

procedures recover coefficients under known assumptions. This gives an internal test of the analysis, rather than an empirical prevalence estimate.

***Figure 11. Recovery of the imposed within coefficient across panels***




The 350 fitted within coefficients cluster around +0.45, the parameter assigned in advance. Their dispersion arises from repeated draws of disturbances and unit characteristics under an otherwise stable design. Recovery near the imposed value indicates that the code and estimator behave as expected in this setting. It says nothing about the correctness of a GDP model fitted to real countries. An empirical panel may contain omitted variables or recording changes that are absent from the generated observations. The distribution is consequently a diagnostic of computation, not external validation.

Estimator recovery is necessary for the simulation to be credible, but it is a limited test. The data were generated so that the additive fixed-effects estimator matches the relevant part of the outcome equation. It would be more surprising if repeated estimates did not centre near the chosen parameter. Real homicide processes could involve feedback, delayed relationships or changes to recording rules, none of which is represented in the baseline mechanism. Arellano and Bond (1991) developed tests for dynamic panels, but such methods are unnecessary for the present question. Their broader relevance is a reminder that identifying assumptions must follow the process being studied.

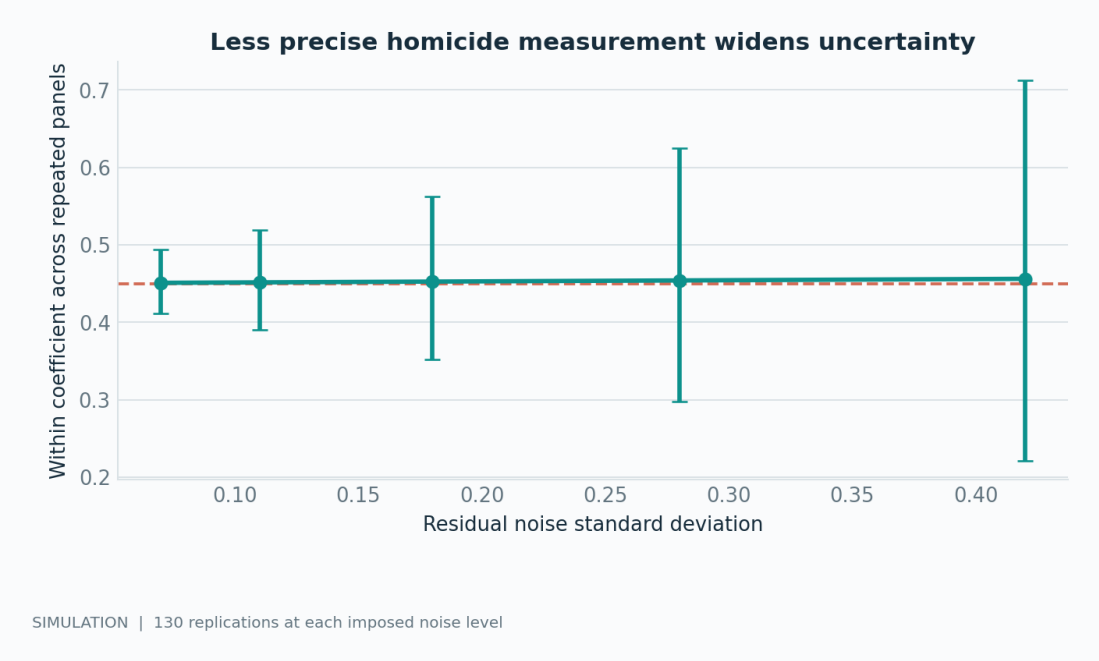
***Figure 12. Detection of opposite signs across imposed within coefficients***



As the imposed within coefficient moves from negative values through zero to positive values, the share of simulations with opposite fitted signs increases. Near zero, noise can produce a positive estimate even when no meaningful within relationship exists. A simple sign count therefore exaggerates certainty around weak effects. The parameter sweep is more useful when read alongside the imposed values and the dispersion of the estimates. It identifies the conditions in this specific simulation under which sign reversal becomes easy to detect, not a general threshold for real homicide research.

A regression can have a positive sign even when the true coefficient is zero. This is especially common if random disturbances are large relative to the available variation. Treating every positive within estimate beside a negative between estimate as a theoretical contradiction would elevate noise into a finding. Here the parameter sweep keeps the between mechanism fixed and changes the strength of the within relation. Readers can therefore distinguish a weak estimated sign from a substantial imposed difference. Simonsohn, Simmons and Nelson (2020) argue for inspecting results across defensible analytic specifications rather than retaining a single attractive estimate. The same care is needed when claims rest on a reversal.

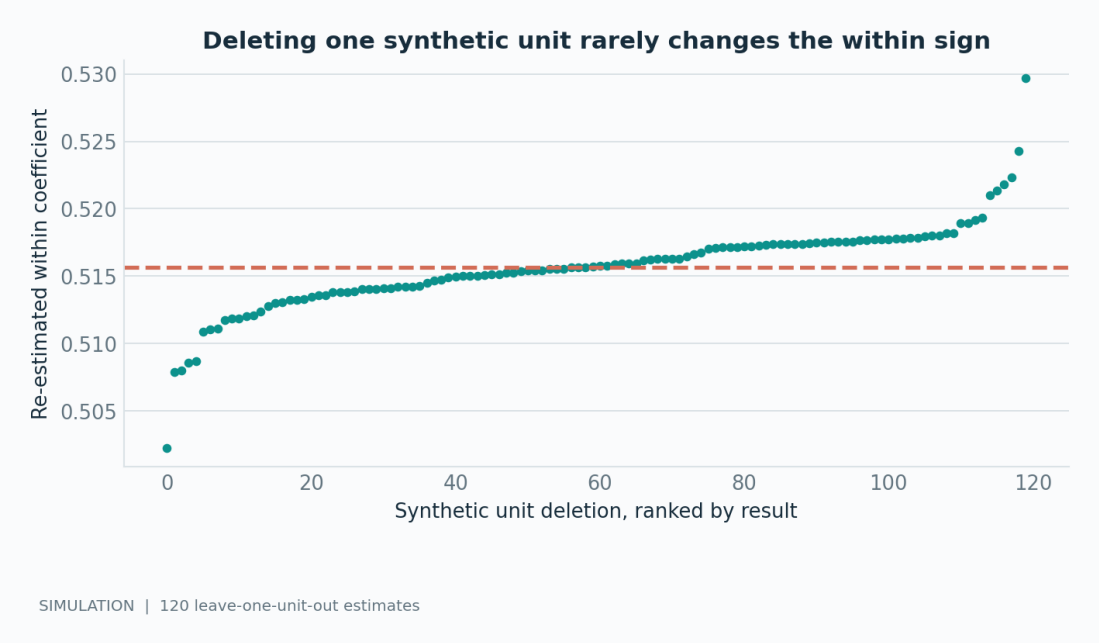
***Figure 13. Sampling dispersion as outcome noise increases***




Increasing the variance of the generated outcome noise widens the distribution of estimated within slopes. Their mean remains close to +0.45 because the added disturbances are independent of income in this experiment. The difference between a stable average estimate and reduced precision is visible in the plotted intervals. Real homicide recording errors may be related to state capacity or economic conditions instead, so they need not behave like independent noise. Rogers and Pridemore (2023) document why alternative homicide measures can lead to different cross-national inferences.

The noise experiment isolates random disturbance rather than systematic measurement error. Its errors are independent of the income mechanism, so increasing their size mainly changes precision. Actual homicide counts may be affected by incomplete death registration, changes in classification and administrative practice. These processes can vary with development and can affect coefficient estimates, not simply widen their uncertainty. Buil-Gil, Pina-Sánchez and Aebi (2026) quantify the relevance of counting rules for international homicide comparison. The generated experiment cannot assign a numerical correction for those mechanisms. It shows the consequences of one specified source of noise while leaving more difficult forms of measurement bias for empirical study.

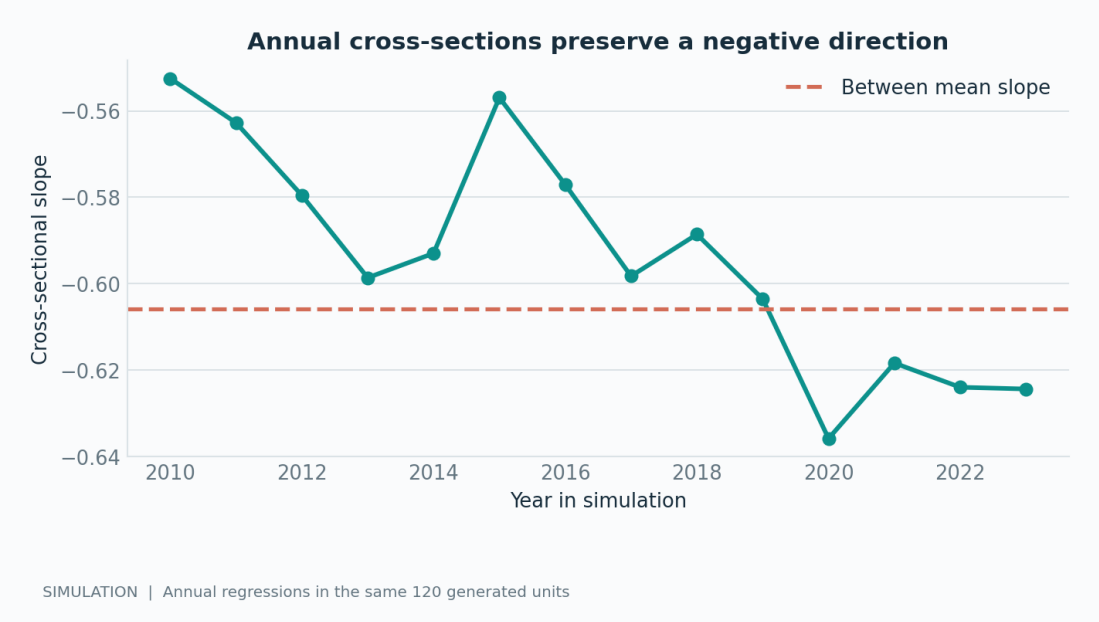
*Figure 14. Within estimates after deleting each synthetic unit*




Removing one unit at a time yields within estimates between +0.502 and +0.530. All 120 leave-one-unit-out fits retain the positive sign. The main result is therefore not driven by any single synthetic unit in this reference sample. That is a limited robustness finding. It does not rule out dependence on groups of units or on a different selection procedure. Influence diagnostics remain useful because a coefficient can be stable on average yet sensitive to a small set of high-leverage observations (King and Roberts 2015).

A leave-one-unit-out check asks a narrow question. Would the fitted coefficient change appreciably if a particular unit were absent? It does not test what happens when whole types of units are missing. Nor does it capture recording failures that affect selected periods within many countries. The reference simulation was not constructed with these more complex departures. Its stable positive slope after each deletion should be reported without treating it as proof of general robustness. Rubin (1976) and Little (1988) describe reasons why the process governing missing observations deserves separate attention from the precision of an estimate computed on the remaining sample.

*Figure 15. Annual cross-sectional coefficients in the synthetic panel*

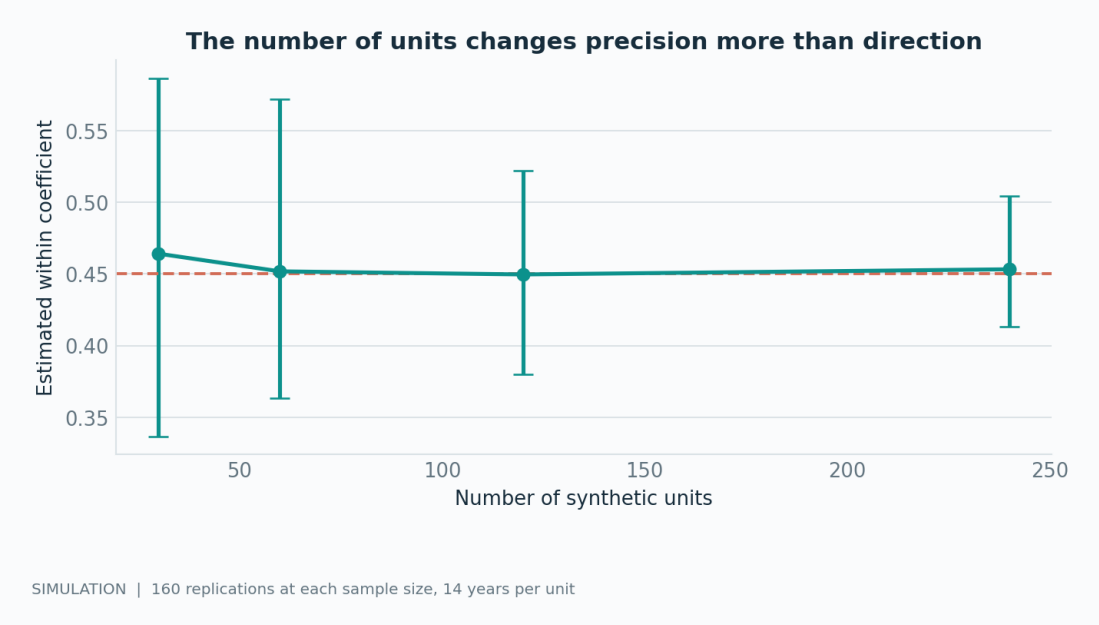



Each annual cross-sectional regression compares different synthetic units in one period. The negative slopes persist across the displayed periods because income differences among unit means

remain influential. They are not repeated within-unit estimates. Even fourteen negative annual coefficients leave open whether income growth inside a unit is positively, negatively or weakly associated with its homicide-like outcome. The figure therefore illustrates why repeated snapshots cannot replace a longitudinal decomposition. Its apparent consistency reflects persistence of the cross-sectional composition, not fourteen independent confirmations of a causal mechanism.

Repeated annual rankings can create a false sense of replication. The same unit characteristics contribute to each snapshot, so successive negative slopes are far from independent discoveries. The simulation makes that persistence easy to see. A positive within coefficient coexists with a negative slope in every annual cross-section, because the underlying comparisons differ. Historical country panels, such as those examined by LaFree and Tseloni (2006), allow questions that annual rankings cannot resolve on their own. The interpretation should begin with whether a study compares countries at a point in time or follows changes within them. Both may be useful, but their findings need not coincide.

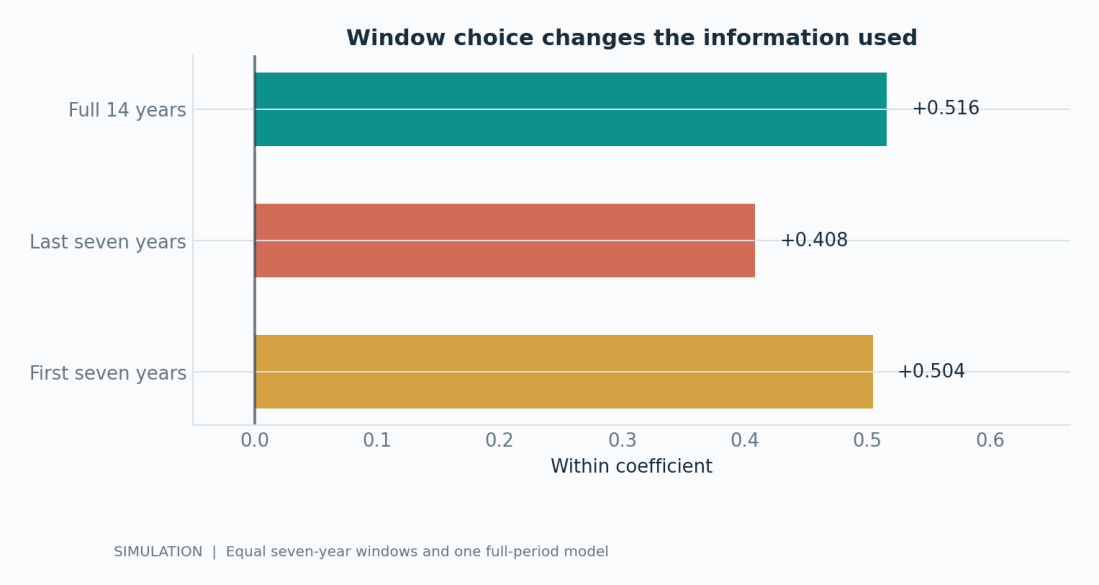
***Figure 16. Precision of within estimates under different sample sizes***




Larger simulated panels produce narrower distributions of the within coefficient, while average estimates remain close to the imposed +0.45. Figure 16 compares repeated experiments with different numbers of units. The reduction in sampling variability is a property of the specified design, not a recommendation to include every country regardless of data quality. Additional observations cannot, by themselves, repair systematic under-recording or incompatible homicide definitions. Precision and validity are distinct considerations, as the comparative-data concerns discussed by Harrendorf (2018) make clear.

An increase from 30 to 240 generated units improves precision because the simulation draws observations from a common, correctly specified mechanism. That assumption is unlikely to hold unchanged across every country in a real global panel. Including countries with different reporting systems could increase coverage while also altering the population being described. The figure does not establish an optimal sample size for homicide research. It makes one familiar statistical consequence visible under controlled conditions, then leaves the more difficult issue of selecting comparable countries unresolved. A larger sample can produce a tighter confidence interval around a misleading estimand.

*Figure 17. Within estimates from two non-overlapping time windows*




The first seven synthetic periods produce a within estimate of +0.504, compared with +0.408 for the final seven. The reference estimate using all periods is +0.516. The generating coefficient does not change between these windows. Their difference therefore arises from finite-sample variation and the information retained in each subset. Because the period labels are illustrative, the split should not be described as a historical break in 2017. A similar divergence in observed data would require separate examination of economic conditions and data continuity.

A change in a fitted slope between two windows does not necessarily imply a turning point in the underlying process. Shorter spans contain less information, and their income variation and random disturbances need not be comparable. The stable mechanism used here rules out a true structural break by construction. Real homicide trends, in contrast, may respond to policies or changing conditions that an ordinary fixed-effects model does not identify. Tuttle (2018) studied social protection and lethal violence in a cross-national setting, illustrating why timing can have substantive meaning. Our split-sample result is a sensitivity check, not a claim about a historical change in economic violence.

**Sensitivity to Alternative Generating Mechanisms and Samples**

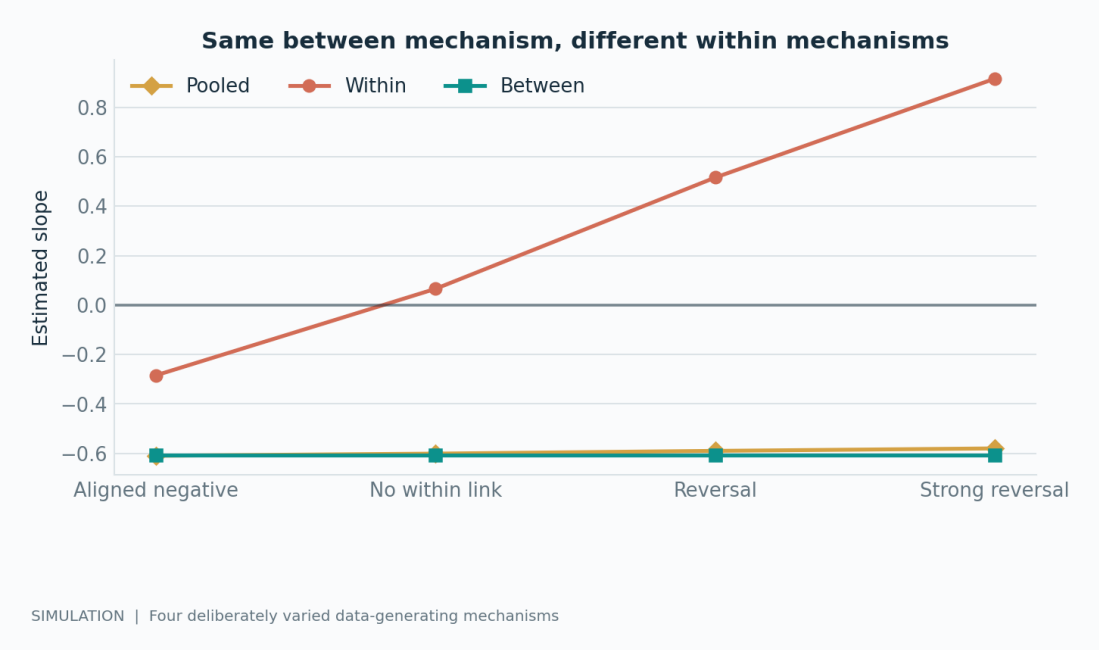
Four alternative generating mechanisms keep the between coefficient negative while changing the within parameter. One within relation is also negative, another is exactly zero, and two are positive. The fitted between slopes remain close because they come from the same cross-unit structure. Pooled slopes change only slightly, whereas within estimates follow the imposed mechanism. This controlled comparison asks whether a single pooled summary can distinguish different relations within units. In these constructed examples it cannot do so reliably. The finding is conditional on the dispersion and covariance parameters in Table 2, which were held stable across scenarios.

*Table 4. Estimates under four deliberately different within mechanisms*

| Scenario | True within | Between estimate | Within estimate | Pooled estimate |
|---|---|---|---|---|
| Aligned negative | -0.35 | -0.606 | -0.284 | -0.610 |
| No within link | +0.00 | -0.606 | +0.066 | -0.601 |
| Reversal | +0.45 | -0.606 | +0.516 | -0.590 |
| Strong reversal | +0.85 | -0.606 | +0.916 | -0.580 |

All scenarios retain an imposed between coefficient of −0.65. Their within coefficients are selected by design.

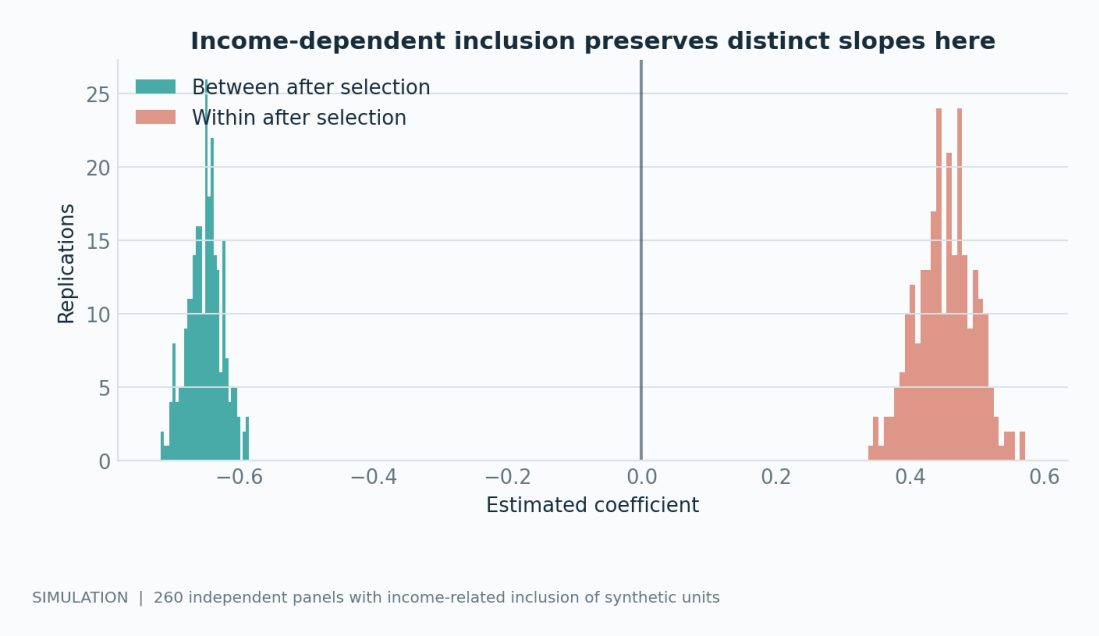
*Figure 18. Similar pooled slopes under different within mechanisms*




Across four imposed mechanisms, the pooled coefficient stays near −0.60 although the within parameter ranges from −0.35 to +0.85. A similar pooled summary thus arises from datasets with markedly different within-unit relations. This result is constructed rather than discovered in the world, but it is consequential for interpretation. A pooled coefficient may describe the joint data correctly while supplying weak evidence about the change process that a theory seeks to explain. The panel-data arguments of Mundlak (1978) and Bell and Jones (2015) establish why these estimands should be separated.

The scenarios make the ambiguity tangible. A negative cross-unit slope can coexist with a negative within association, a zero within parameter or a positive within association. The negative pooled summary barely changes because stable differences dominate its covariance. The analyst cannot infer the within mechanism from that pooled number unless additional restrictions connect variation across levels. Substantive theories might justify such restrictions, but they must be stated and assessed. Comparative research on welfare and homicide (Rogers and Pridemore 2017) and governance (Neumayer 2003) addresses mechanisms richer than the single artificial income component used here. Their presence in the literature does not make the synthetic parameter a real-world estimate.

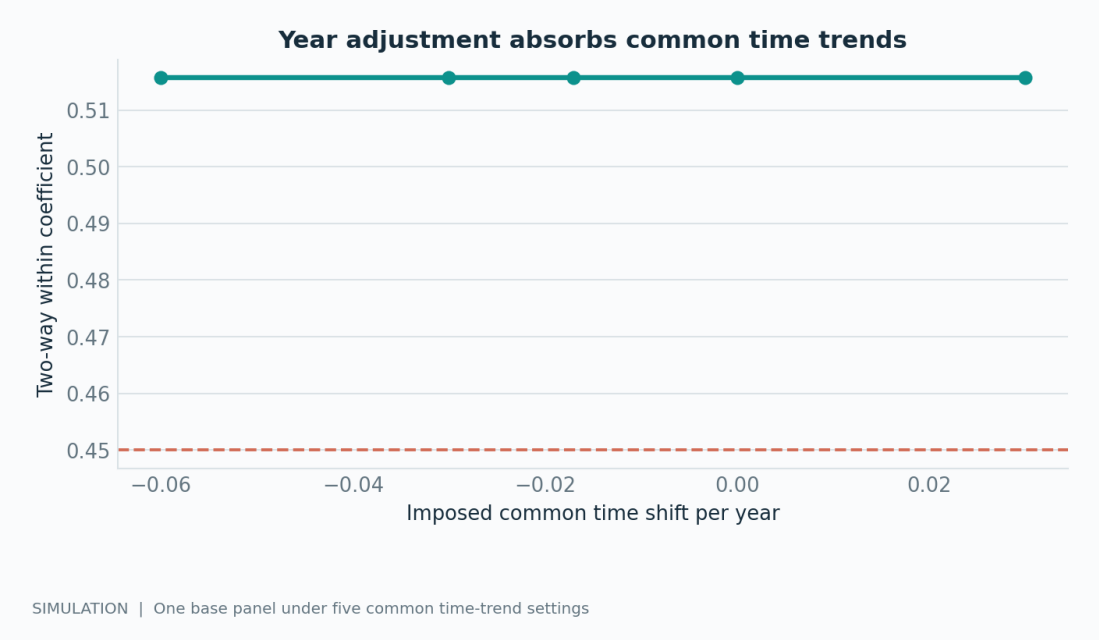
*Figure 19. Results after income-related selection of complete units*




Income-related selection retains an average of 94.0 of the 120 generated units across 260 independent panels. The average within slope is +0.453, while the between slope averages −0.650. Opposite signs persist under this particular complete-unit selection rule. The exercise does not show that missingness is harmless. It removes whole units with probabilities related to their income, not isolated periods or observations tied to violent shocks. Other selection processes could change both estimates. The missing-data arguments developed by Rubin (1976) and Little (1988) remain pertinent to any real country panel.

The selection rule favours higher-income synthetic units, leaving fewer lower-income units in a typical retained sample. It removes whole records rather than individual years. Those conditions differ from an international database in which homicide reports disappear after conflict, institutional disruption or a change in statistical practice. The stability found here is therefore tied to the imposed selection process. It cannot establish the direction of selection bias in UNODC data. A stronger observational claim would require information about excluded country-years and the reasons they were not recorded. The simulation shows a possible robustness pattern, not the absence of missing-data bias.

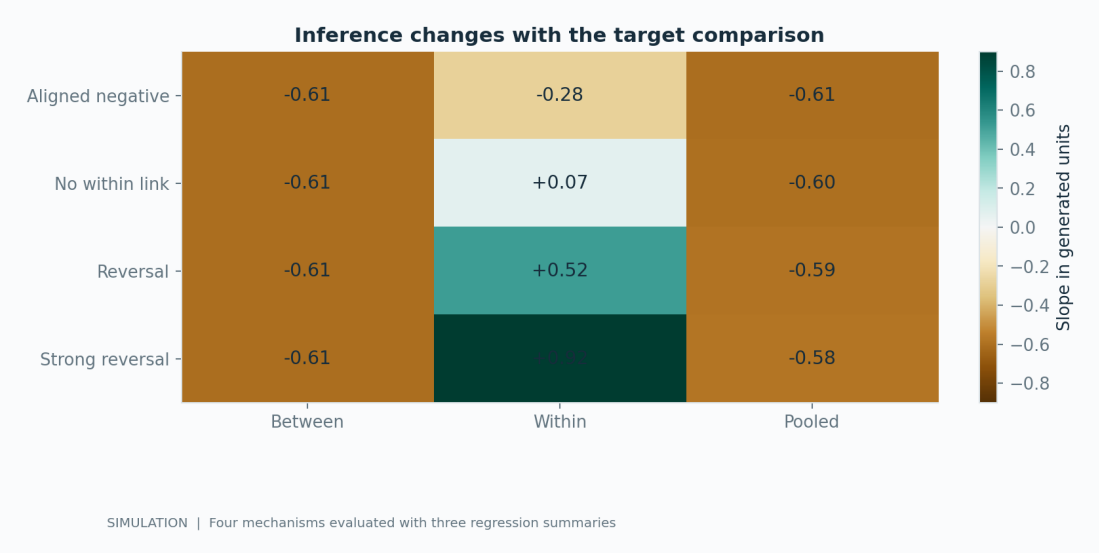
*Figure 20. Invariance to an additive common time trend*



Changing the common synthetic time trend leaves the two-way fixed-effects estimate almost unchanged. The model absorbs an additive period effect shared by all units, so the result follows from how it was specified. This diagnostic does not test whether country-specific trends are benign or whether changing confounders affect income and homicide together. Those problems remain outside the simulation. Beck and Katz (1995) describe wider dependence concerns in time-series cross-sectional data. The figure verifies a property of the estimator rather than revealing a new regularity in observed homicide.

The near-flat line is expected because common year shocks are removed before the within coefficient is calculated. The test confirms that this adjustment works for additive shocks of the sort included in the data-generating equation. It provides no protection against a change in one unit that is unrelated to the others. Economic crises or policy shifts can affect countries differently, and an observational model would need to address those patterns directly. Presenting the flat line as evidence that time effects solve temporal confounding would overstate the result. Its value is diagnostic, with a deliberately narrow target.

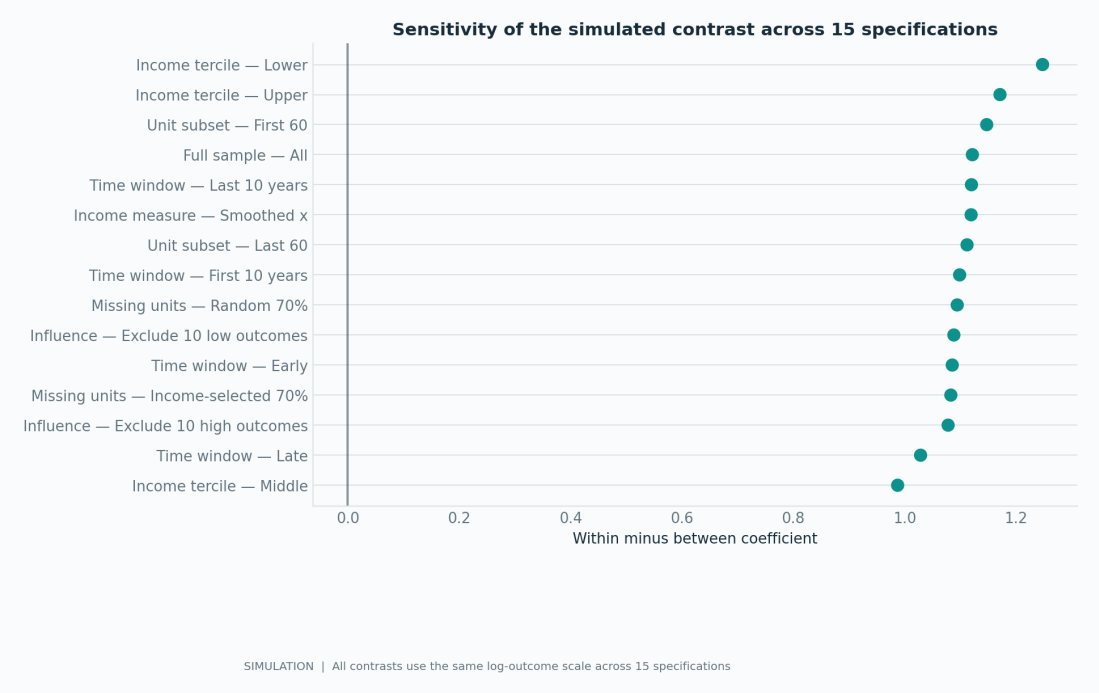
***Figure 21. Fitted signs across four imposed mechanisms***




The heatmap contrasts imposed within parameters with fitted coefficient signs under four generated mechanisms. A small positive estimate appears even when the true within parameter is zero. That sign should not be given the substantive meaning of a positive income effect. Larger imposed positive parameters produce larger fitted within slopes while the pooled summary remains negative. The image is useful for spotting differences in what a model estimates, although the underlying numbers in Table 4 remain necessary for judging their size and meaning.

A fitted sign is least informative near zero. In the scenario with no within relation, a small positive estimate comes from random error. In the two positive scenarios, the fitted slopes reflect the parameters introduced into the generating equation. Across all four, the between estimate remains negative because that part of the mechanism was unchanged. Colour alone can make these differences look more definite than they are. The accompanying numbers allow readers to distinguish a negligible estimate from a sizable one. This is also why sign-based claims in observational homicide research require intervals and an explicit account of the comparison being estimated.

***Figure 22. Within–between contrasts from fifteen sensitivity checks***




Fifteen variants of the reference analysis retain positive within-minus-between contrasts, although their sizes differ. The alternatives include time windows, unit subsets, income groups and a smoothed income measure. All estimates use the same transformed outcome, so their numerical contrasts are comparable on that scale. The result is not a comprehensive multiverse of plausible homicide models. It describes robustness to a limited menu of changes within the synthetic system. Steegen and colleagues (2016) and Simonsohn, Simmons and Nelson (2020) explain why the defensibility of the menu matters as much as the plotted curve.

A specification curve cannot be judged solely by the number of estimates displayed. The researcher must explain why each analytic choice was included and which plausible choices remain absent. Our fifteen specifications retain the generated outcome and omit conflict, social structure and measurement differences that might be consequential in observed homicide research. Their positive contrasts are unsurprising under the opposed generating parameters. The curve nevertheless shows whether the code and substantive interpretation remain consistent after several transparent perturbations. Its purpose is diagnostic. It should not become a substitute for a full robustness analysis once actual country data are available.

**Table 5. Selected sensitivity estimates on a common outcome scale**

| Specification | Between | Within | Contrast |
|---|---|---|---|
| Full sample — All | -0.606 | +0.516 | +1.122 |
| Time window — Early | -0.581 | +0.504 | +1.085 |
| Time window — Late | -0.620 | +0.408 | +1.029 |
| Income tercile — Lower | -0.661 | +0.586 | +1.248 |
| Income tercile — Upper | -0.659 | +0.512 | +1.171 |
| Missing units — Random 70% | -0.594 | +0.501 | +1.094 |
| Missing units — Income-selected 70% | -0.596 | +0.487 | +1.083 |

The full set of fifteen variants is shown in Figure 22. All estimates are based on generated observations.

**Implications for Comparative Homicide Theory and Measurement**

Comparative explanations can shift too easily from describing social settings to predicting the consequences of changing one feature of them. Income is a particular risk because average differences between countries may be large while annual changes within a country are comparatively modest. The generated data were designed to reproduce precisely that imbalance. They do not demonstrate its magnitude in world GDP records. Chon (2012) examined population heterogeneity alongside inequality, while Nivette (2014) developed an account of legitimacy in cross-national criminology. Their work offers substantive reasons to question what a single income gradient represents. A paired-country comparison by Kadir (2026d) likewise found no uniform homicide pattern that could be attributed to religious-majority status. An adequate empirical test would need to measure such processes rather than infer them from the GDP coefficient.

If matched observed records produced a negative between-country coefficient and a positive within-country coefficient, the result would undermine any assertion that the two associations are interchangeable. It would not establish that growth itself increases homicide or that prosperity is unimportant. Changes in income could coincide with demographic shifts, institutional reform or unmeasured shocks, while stable differences between countries could reflect other historical conditions. Lederman, Loayza and Menéndez (2002) examined social capital, Levchak (2019) discussed modernization, and Santos, Testa and Weiss (2018) investigated poverty in cross-national homicide research. These are possible explanatory pathways, not mechanisms measured by the synthetic panel. Kadir (2026f) treated post-shock homicide persistence as conditional rather than a universal response to economic contraction. These explanations may vary across income levels and historical periods, possibilities that this experiment cannot adjudicate.

Recorded homicide depends on more than an underlying lethal event. Police investigation, death certification and legal classification may shape the figures released to researchers. Rogers and Pridemore (2023) found that source selection affects cross-national results, and Buil-Gil, Pina-Sánchez and Aebi (2026) demonstrate the importance of counting rules. The random noise in this experiment does not reproduce a change in classification that happens to correlate with development. Kadir (2026i) demonstrated in a separate simulation how differences between criminal-justice and public-health recording systems can distort the apparent persistence of homicide series. More observations can reduce sampling uncertainty without correcting systematic measurement error. An observed test of cross-level divergence would therefore need to examine source comparability alongside the GDP coefficients.

A negative association among countries can still be informative if it is described as such. Kadir (2026h) also distinguished victim–offender relationships and situational contexts that national homicide totals cannot reveal. It may reveal patterned differences among national settings and prompt hypotheses about institutions or social structure. A within coefficient answers another question, one concerned with deviations inside units after specified adjustments. Neither estimate acquires the assumptions of the other. Fajnzylber, Lederman and Loayza (2002a) and Paré and Felson (2014) used different treatments of inequality and poverty to investigate comparative crime relationships. Their work illustrates why the theoretical statement should determine the relevant comparison before a coefficient is given substantive meaning. Perceived inequality adds a further measurement question to this relationship (Rogers and Pridemore 2022).

An estimand is a question expressed in statistical terms. Fixed effects can remove time-invariant additive differences, but they cannot settle reverse causality or changing confounding. Correlated random-effects approaches can estimate within and between components jointly where the specification supports that interpretation (Bell and Jones 2015, Schunck 2013). Robust standard errors address an aspect of uncertainty, not the meaning of a coefficient or the quality of an outcome measure (King and Roberts 2015). The practical implication is straightforward. Researchers should decide whether their theory concerns national differences, within-country change or both, and then show the estimate that answers that question.

There are clear limits to what was established. UNODC's published figures support statements about regional homicide levels at three dates, but they supply no GDP pairing. The other twenty figures come from a documented synthetic data-generating process. They cannot establish a reversal among actual countries in 2010–2023 or identify its historical cause. Empirical testing would require a harmonised country-year panel, rules for sample inclusion, scrutiny of data sources and direct estimates of the between and within associations. A separate review by Kadir (2026b) described how cleared-

case selection restricts the population represented in offender-specific homicide research. That is a distinct selection problem, but it reinforces the need to align conclusions with the records actually observed. The simulation provides a diagnostic for future national-panel analysis, not a substitute for it.

## CONCLUSION

The generated panel produced a negative between-unit coefficient of −0.606 and a positive within-unit coefficient of +0.516, while the pooled estimate of −0.590 remained close to the between-unit association. The difference between the two slopes was +1.122, with a 95% unit-bootstrap interval ranging from +1.041 to +1.203. Across 350 independently generated panels, the within estimator recovered the prespecified positive parameter on average, and all fitted panels retained opposing coefficient signs. The contrast also remained positive across fifteen sensitivity specifications, although its magnitude varied with sample composition and observation periods. These results establish how a pooled association can obscure the direction of change within units under the conditions imposed by the simulation. They do not establish that similar reversals occur in observed national homicide data.

Cross-national differences in economic conditions cannot be treated as interchangeable with economic changes within individual countries. The distinction matters because both relationships may be statistically valid while supporting different interpretations of a proposed crime theory. The simulated results establish the conditions under which such inferential disagreement can arise, although they provide no evidence that GDP growth causes homicide to increase or decline. Published UNODC statistics document substantial geographic variation in homicide rates, but those figures cannot independently validate the simulated income relationships. Comparative criminological research consequently requires greater precision in identifying the source of variation underlying its statistical claims. Whether the same divergence occurs in actual homicide patterns remains an empirical question that requires comparable country-year observations and direct estimation of both associations.

## REFERENCES

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? Quarterly Journal of Economics, 138(1), 1–35. doi.org/10.1093/qje/qjac038

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data — Monte Carlo evidence and an application to employment equations. Review of Economic Studies, 58(2), 277–297. doi.org/10.2307/2297968

Beck, N., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. American Political Science Review, 89(3), 634–647. doi.org/10.2307/2082979

Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models — Making an informed choice. Quality & Quantity, 53, 1051–1074. doi.org/10.1007/s11135-018-0802-x

Bell, A., & Jones, K. (2015). Explaining fixed effects — Random effects modeling of time-series cross-sectional and panel data. Political Science Research and Methods, 3(1), 133–153. doi.org/10.1017/psrm.2014.7

Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects — A structural equations approach. Social Forces, 89(1), 1–34. doi.org/10.1353/sof.2010.0072

Buil-Gil, D., Pina-Sánchez, J., & Aebi, M. F. (2026). The effect of counting rules on cross-national comparisons of homicide. Social Indicators Research, 181, Article 44. doi.org/10.1007/s11205-025-03802-8

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317–372. doi.org/10.3368/jhr.50.2.317

Chon, D. S. (2012). The impact of population heterogeneity and income inequality on homicide rates — A cross-national assessment. International Journal of Offender Therapy and Comparative Criminology, 56(5), 730–748. doi.org/10.1177/0306624X11414813

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. Annual Review of Psychology, 62, 583–619. doi.org/10.1146/annurev.psych.093008.100356

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

Fajnzylber, P., Lederman, D., & Loayza, N. (2002a). Inequality and violent crime. Journal of Law and Economics, 45(1), 1–40. doi.org/10.1086/338347

Fajnzylber, P., Lederman, D., & Loayza, N. (2002b). What causes violent crime? European Economic Review, 46(7), 1323–1357. doi.org/10.1016/S0014-2921(01)00096-4

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. Psychological Methods, 20(1), 102–116. doi.org/10.1037/a0038889

Harrendorf, S. (2018). Prospects, problems, and pitfalls in comparative analyses of criminal justice data. Crime and Justice, 47(1), 159–207. doi.org/10.1086/696042

Imai, K., & Kim, I. S. (2019). When should we use unit fixed effects regression models for causal inference with longitudinal data? American Journal of Political Science, 63(2), 467–490. doi.org/10.1111/ajps.12417

Kadir, Z. K. (2026a). Are We Safer, or Just Older? Testing the Demographic Trap Through Cross-National Homicide Rates and Age-Conditional Ranking Instability. Punggawa Global Research – Jurnal Multidisiplin, 1(6). doi.org/10.67707/pgr.v1i6.489

Kadir, Z. K. (2026b). The Cleared-Case Canon – Investigative Selection as a Structural Limit on General Homicide Theory Building. LEXARA – Journal of Law and Legal Studies, 1(2). doi.org/10.67707/lexara.v1i2.197

Kadir, Z. K. (2026c). Criminological Doppelgängers? Similar Starting Homicide Rates and Divergent Change in Six Latin American Countries. LEXARA – Journal of Law and Legal Studies, 1(3). doi.org/10.67707/lexara.v1i3.490

Kadir, Z. K. (2026d). Does Islam Explain Homicide? A Paired Cross-National Comparison of Religion, State Capacity, and Lethal Violence. SIMURGA – Journal of Islamic Studies and Muslim Societies, 1(1). punggawajournal.com/simurga/article/view/89

Kadir, Z. K. (2026e). The Geography of Lethal Violence Escalation – Cross-National Differences in Recorded Homicide and Serious Assault across Structural Contexts. Punggawa Law Review, 1(4). doi.org/10.67707/plr.v1i4.158

Kadir, Z. K. (2026f). The Homicide Hysteresis Problem and Persistent Lethal Violence after Temporary Macroeconomic Shocks. SOCILEXIA – Journal of Social Sciences and Humanities, 1(2). punggawajournal.com/socilexia/article/view/176

Kadir, Z. K. (2026g). The Homicide Visibility Gap – Cross-National Discrepancies between UNODC-Recorded Intentional Homicide and Mortality-Based Violence. Punggawa Global Research – Jurnal Multidisiplin, 1(4). doi.org/10.67707/pgr.v1i4.157

Kadir, Z. K. (2026h). The Social Anatomy of Homicide – Relationships, Situations, and Legal Categories in Indonesia. SOCILEXIA – Journal of Social Sciences and Humanities, 1(1). punggawajournal.com/socilexia/article/view/115

Kadir, Z. K. (2026i). The State Does Not Count the Same Dead – The Politics of Homicide Measurement Across Criminal Justice and Public Health Systems. Punggawa Global Research – Jurnal Multidisiplin, 1(5). doi.org/10.67707/pgr.v1i5.245

Kadir, Z. K. (2026j). The World's Most Dangerous Country May Be a Statistical Illusion – Reassessing the Evidence Behind Homicide Rankings. SOCILEXIA – Journal of Social Sciences and Humanities, 1(4). doi.org/10.67707/socilexia.v1i4.487

King, G., & Roberts, M. E. (2015). How robust standard errors expose methodological problems they do not fix, and what to do about it. Political Analysis, 23(2), 159–179. doi.org/10.1093/pan/mpu015

LaFree, G., & Tseloni, A. (2006). Democracy and crime — A multilevel analysis of homicide trends in forty-four countries, 1950–2000. Annals of the American Academy of Political and Social Science, 605(1), 25–49. doi.org/10.1177/0002716206287169

Lederman, D., Loayza, N., & Menéndez, A. M. (2002). Violent crime — Does social capital matter? Economic Development and Cultural Change, 50(3), 509–539. doi.org/10.1086/342422

Levchak, P. J. (2019). Clarification and refinement of modernization theory — How the global economy impacts homicide rates. Sociological Focus, 52(4), 309–323. doi.org/10.1080/00380237.2019.1668319

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. Journal of the American Statistical Association, 83(404), 1198–1202. doi.org/10.1080/01621459.1988.10478722

Messner, S. F. (1982). Poverty, inequality, and the urban homicide rate — Some unexpected findings. Criminology, 20(1), 103–114. doi.org/10.1111/j.1745-9125.1982.tb00450.x

Messner, S. F. (1989). Economic discrimination and societal homicide rates — Further evidence on the cost of inequality. American Sociological Review, 54(4), 597–611. doi.org/10.2307/2095881

Messner, S. F., Raffalovich, L. E., & Sutton, G. M. (2010). Poverty, infant mortality, and homicide rates in cross-national perspective — Assessments of criterion and construct validity. Criminology, 48(2), 509–537. doi.org/10.1111/j.1745-9125.2010.00194.x

Mundlak, Y. (1978). On the pooling of time series and cross section data. Econometrica, 46(1), 69–85. 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. doi.org/10.1177/00223433030406001

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

Nivette, A. E. (2012). Spatial patterns of homicide and political legitimacy in Europe. International Journal of Comparative and Applied Criminal Justice, 36(3), 155–171. doi.org/10.1080/01924036.2012.670928

Nivette, A. E. (2014). Legitimacy and crime — Theorizing the role of the state in cross-national criminological theory. Theoretical Criminology, 18(1), 93–111. doi.org/10.1177/1362480613499793

Nivette, A. E., & Eisner, M. (2013). Do legitimate polities have fewer homicides? A cross-national analysis. Homicide Studies, 17(1), 3–26. doi.org/10.1177/1088767912452131

Paré, P.-P., & Felson, R. (2014). Income inequality, poverty and crime across nations. British Journal of Sociology, 65(3), 434–458. doi.org/10.1111/1468-4446.12083

Pridemore, W. A. (2011). Poverty matters — A reassessment of the inequality–homicide relationship in cross-national studies. British Journal of Criminology, 51(5), 739–772. doi.org/10.1093/bjc/azr019

Rogers, M. L., & Pridemore, W. A. (2017). How does social protection influence cross-national homicide rates in OECD nations? Sociological Quarterly, 58(4), 576–594. doi.org/10.1080/00380253.2017.1367265

Rogers, M. L., & Pridemore, W. A. (2022). Perceived inequality and cross-national homicide rates. Justice Quarterly, 39(2), 225–251. doi.org/10.1080/07418825.2020.1729392

Rogers, M. L., & Pridemore, W. A. (2023). A review and analysis of the impact of homicide measurement on cross-national research. Annual Review of Criminology, 6, 447–470. doi.org/10.1146/annurev-criminol-030521-102909

Rubin, D. B. (1976). Inference and missing data. Biometrika, 63(3), 581–592. doi.org/10.1093/biomet/63.3.581

Rudolph, M., & Starke, P. (2020). How does the welfare state reduce crime? The effect of program characteristics and decommodification across 18 OECD-countries. Journal of Criminal Justice, 68, 101684. doi.org/10.1016/j.jcrimjus.2020.101684

Santos, M. R., Testa, A., & Weiss, D. B. (2018). Where poverty matters — Examining the cross-national relationship between economic deprivation and homicide. British Journal of Criminology, 58(2), 372–393. doi.org/10.1093/bjc/azx013

Schunck, R. (2013). Within and between estimates in random-effects models — Advantages and drawbacks of correlated random effects and hybrid models. Stata Journal, 13(1), 65–76. doi.org/10.1177/1536867X1301300105

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2020). Specification curve analysis. Nature Human Behaviour, 4, 1208–1214. doi.org/10.1038/s41562-020-0912-z

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. Perspectives on Psychological Science, 11(5), 702–712. doi.org/10.1177/1745691616658637

Tuttle, J. (2018). Specifying the effect of social welfare expenditures on homicide and suicide — A cross-national, longitudinal examination of the stream analogy of lethal violence. Justice Quarterly, 35(1), 87–113. doi.org/10.1080/07418825.2017.1293711

United Nations Office on Drugs and Crime. (2025). Levels and trends of criminal and other intentional killings since the UNTOC was adopted in 2000. Research Brief on Transnational Organized Crime, Special Points of Interest, Figure 7, p. 13. United Nations Office on Drugs and Crime.