Deterrence Beyond Borders: Violence Reduction Units and Knife Crime Spillovers in England and Wales

APEP Autonomous Research* @olafdrw

March 14, 2026

Abstract

Since 2019, the UK Home Office has invested over £254 million in Violence Reduction Units across 20 of 43 police force areas. I exploit this rollout to estimate effects on knife crime and test whether crime displaces to neighboring jurisdictions. Using ONS data for 42 forces over 15 years, I classify untreated forces as "boundary" (adjacent to a VRU force) or "interior" to decompose spillovers. The direct effect on VRU forces is unidentified: pre-trends are violated because forces were selected on high violence. Boundary forces show a 4.39 per 100,000 knife crime reduction under conventional inference ($p = 0.032$), suggesting deterrence rather than displacement, though this is fragile under randomization inference ($p = 0.298$). The results illustrate how selection on pre-treatment outcomes and few-cluster inference complicate evaluation of place-based crime programs.

JEL Codes: K42, H75, R58

Keywords: knife crime, Violence Reduction Units, displacement, deterrence, spillovers, difference-in-differences

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 34m).

1. Introduction

Between March 2018 and March 2019, hospitals in England and Wales recorded over 5,000 admissions for assault by a sharp object — a figure that had nearly doubled in four years and that placed knife violence squarely at the center of domestic policy debate (Allen and Audickas, 2019). The government's response was substantial: the Home Office established Violence Reduction Units in 18 police force areas, backed by £35.5 million in the first year and scaling to £64 million annually, with two additional forces added in 2022 (Home Office, 2019). By 2025, cumulative spending through the Serious Violence Fund had exceeded £254 million. Yet despite this investment, credible evidence on whether VRUs actually reduce violence remains thin. The Home Office's own evaluation relied on a quasi-experimental comparison that the authors acknowledged could not rule out secular trends (Home Office, 2022). More fundamentally, even a well-identified estimate of the direct effect on treated forces would miss a first-order question: does place-based policing reduce total crime, or merely push it across jurisdictional boundaries?

This paper addresses that question by estimating both the direct effect of VRU funding on knife crime and the spatial spillovers to neighboring police force areas. I exploit the fact that the 18 forces in the 2019 cohort were selected on the basis of serious violence rates during 2016–2018, creating a treatment-control contrast across all 43 forces in England and Wales. I classify the 22 untreated forces into "boundary" forces (those sharing a geographic border with at least one VRU force) and "interior" forces (those sharing no such border). If VRUs displace crime, boundary forces should see increases; if the deterrence effect of heightened enforcement radiates outward, boundary forces should see decreases. The sign of the boundary coefficient thus directly discriminates between the two leading theoretical channels in the spatial crime literature (Guerette and Bowers, 2009; Johnson et al., 2014; Weisburd et al., 2006).

The distinction matters quantitatively. VRU forces collectively serve 67.6 percent of England and Wales's population, boundary forces 30.9 percent, and interior forces just 1.5 percent. Even modest spillovers, whether positive (displacement) or negative (deterrence), therefore affect a population comparable in size to that of the directly treated forces. Ignoring them produces a misleading estimate of the program's total impact — a point emphasized by Angelucci et al. (2015) in the context of development interventions and by Blattman et al. (2021) for place-based policing.

My main findings are threefold. First, the direct effect of VRUs on knife crime in treated forces is not robustly estimated. A standard TWFE specification yields a small, statistically insignificant coefficient ($\hat{\beta} = 1.41$, SE $= 3.25$). The Callaway–Sant'Anna estimator, designed for staggered adoption settings, produces a large negative estimate (ATT $= -13.67$, $p < 0.01$), but a formal pre-trends test decisively rejects the parallel trends assumption ($p = 0.000$). Because treatment was assigned to the highest-crime forces, pre-treatment trends are mechanically contaminated by mean reversion and differential trajectories — precisely the selection-on-levels problem that Roth et al. (2023) warn about.

Second, among untreated forces, boundary forces experienced a reduction of 4.39 knife crimes per 100,000 relative to the sole interior force ($p = 0.032$ under conventional cluster-robust inference). This point estimate is consistent with deterrence: enforcement intensity in VRU forces appears to depress crime in adjacent jurisdictions rather than displacing it. However, with only 22 untreated forces and a single interior comparator, the inference is fragile. Randomization inference, which does not require large-sample asymptotics, yields $p = 0.298$; a cluster bootstrap produces $p = 0.303$. The conventional result is thus suggestive but not definitive.

Third, the pre-COVID restriction illuminates an important dynamic. When I limit the sample to financial years ending through 2020 — capturing only the first eleven months of VRU operation before lockdowns altered crime patterns entirely — the treated-force coefficient is positive and significant ($\hat{\beta} = 10.25$, $p = 0.008$). This finding is consistent with VRU forces having been selected on rising violence: in the immediate post-treatment period, the upward pre-treatment trajectory dominates any nascent treatment effect.

This paper contributes to three literatures. Most directly, it provides the first econometric evaluation of the UK's VRU program using methods designed for causal inference, complementing the government's descriptive evaluation (Home Office, 2022). Prior work on the Serious Violence Strategy has been largely qualitative (IRISS, 2020; Wong et al., 2019). Second, it advances the empirical study of crime displacement and deterrence spillovers. The classic framework of Cornish and Clarke (1987) predicts that rational offenders respond to localized enforcement by relocating, while Weisburd et al. (2006) document that in many policing experiments, crime drops in treated areas without rising in adjacent ones. My boundary/interior decomposition builds on the concentric-zone approach of Bowers and Johnson (2003) and the spatial DiD strategies in Draca et al. (2011) and Di Feliciantonio and Ferrara (2022), extending them to a national-scale, multi-force setting. Third, the paper illustrates the inferential challenges of evaluating programs assigned to high-risk jurisdictions — a common feature of place-based policies (Busso et al., 2013; Kline and Moretti, 2013). The divergence between TWFE, Callaway–Sant'Anna, and pre-COVID results underscores the fragility of causal claims when treatment is selected on pre-treatment outcomes, a problem formalized by Callaway and Sant'Anna (2021) and Sun and Abraham (2021).

The remainder of the paper proceeds as follows. Section 2 describes the VRU program and the institutional features of policing in England and Wales. Section 3 introduces the data and defines the boundary/interior classification. Section 4 presents the empirical strategy. Section 5 reports results, and Section 7 concludes.

2. Institutional Background

The Serious Violence Strategy. In April 2018, the Home Office published the Serious Violence Strategy, identifying knife crime, gun crime, and county-lines drug trafficking as the primary threats to public safety in England and Wales (Home Office, 2018). The strategy adopted a "public health" approach to violence, emphasizing multi-agency partnerships between police, health services, local government, education, and community organizations — an approach inspired by Scotland's Violence Reduction Unit, established in 2005, and by similar models in Glasgow and the United States (Deuchar, 2018; Butts et al., 2015).

Selection into treatment. The Home Office allocated Serious Violence Fund grants to 18 police force areas in 2019, selecting forces based on hospital admissions for assault with a sharp object and police-recorded serious violence during March 2016–March 2018 (Home Office, 2019). This selection mechanism is consequential for identification: treated forces are, by construction, those with the highest pre-treatment violence rates. The 18 forces include every metropolitan force (Metropolitan Police, Greater Manchester, West Midlands, Merseyside, West Yorkshire), all of the large urban forces, and several smaller forces with concentrated violence problems. In 2022, Cleveland and Humberside were added, bringing the total to 20. Funding ranged from £880,000 for smaller forces to £7 million for the Metropolitan Police in the first year.

What VRUs do. Each VRU is required to produce a local "strategic needs assessment" identifying the drivers of serious violence and to coordinate a multi-agency response plan. In practice, VRUs fund school-based mentoring, hospital "teachable moment" interventions for knife-wound patients, community outreach workers, enhanced intelligence-sharing between police and health services, and targeted enforcement operations. The intervention bundle varies across forces, making the "treatment" inherently heterogeneous — a common challenge in evaluating place-based programs (Kline and Moretti, 2013).

Policing geography. England and Wales are divided into 43 territorial police force areas, each headed by an operationally independent Chief Constable. Force boundaries are fixed administrative units that have remained stable since 1974 (with minor exceptions). Critically, offenders do not face additional legal penalties for crossing force boundaries, and police cooperation across boundaries, while facilitated by regional organized crime units, is operationally more difficult than within-force coordination. This institutional structure makes force boundaries a natural unit for measuring spatial spillovers: if enforcement in one force raises the perceived risk of crime, offenders may shift to adjacent forces where enforcement is lighter.

3. Data

3.1 Crime Data

I use the ONS Police Force Area Data Tables, which report annual counts of recorded crime by offence category and police force area for financial years ending 2011 through 2025 (Office for National Statistics, 2024).[1] The primary outcome is knife crime ("offences involving a knife or sharp instrument"), a category the ONS has published consistently since the financial year ending 2011. To construct crime rates, I divide counts by mid-year population estimates from ONS Table P3, expressing all outcomes as offences per 100,000 population.

I exclude the British Transport Police, which has no fixed territorial jurisdiction, leaving 42 geographic forces. With 15 financial years, the analysis sample comprises 630 force-year observations. For some specifications, I also examine firearm offences ("offences in which firearms were used") as an alternative outcome.

3.2 Force Classification

I classify forces into three mutually exclusive groups based on treatment status and geographic contiguity:

- **VRU forces** (20): received Serious Violence Fund grants in 2019 (18 forces) or 2022 (Cleveland, Humberside).
- **Boundary forces** (21): never received VRU funding but share a geographic border with at least one VRU force.
- **Interior forces** (1): never received VRU funding and share no border with any VRU force.

[1] An alternative data source is the street-level crime records from `data.police.uk`, which would provide monthly granularity. However, each monthly archive exceeds 1.6 GB, making the full panel (∼150 GB) impractical for this analysis. The annual ONS data compensates with a longer pre-period (9 years), which strengthens tests of parallel trends.

The near-total coverage of VRU forces across England and Wales means that almost every untreated force is adjacent to at least one treated force. Only one force — Dyfed-Powys, in rural mid-Wales — qualifies as "interior." This extreme imbalance (21 boundary vs. 1 interior) is itself informative: the VRU program is so geographically extensive that clean untreated comparators barely exist. It also creates an inferential challenge that I address through randomization inference and cluster bootstrap methods.

3.3 Summary Statistics

Table 1 reports pre-treatment summary statistics by force classification. VRU forces have a mean knife crime rate of 58.8 per 100,000, nearly double that of boundary forces (31.8) and more than double interior forces (26.1). This gap confirms the selection mechanism: treatment was assigned to the highest-crime forces. The standard deviation within VRU forces (30.0) is also roughly twice that of boundary forces (15.4), reflecting the heterogeneity from including both the Metropolitan Police (with rates exceeding 100 per 100,000) and smaller urban forces.

Table 1: Pre-Treatment Summary Statistics by Force Classification

Force Type	Mean Knife Crime Rate	SD	Mean Pop. (000s)	Forces	Force-Years
VRU	58.8	30.0	2088	20	180
Boundary	31.8	15.4	910	21	189
Interior	26.1	20.4	940	1	9

Notes: Knife crime rate per 100,000 population. Pre-treatment period: financial years ending 2011–2019. VRU forces received Home Office Serious Violence Fund grants beginning April 2019. Boundary forces are untreated forces sharing a geographic border with at least one VRU force. Interior forces share no borders with VRU forces. Source: ONS Police Force Area Data Tables.

4. Empirical Strategy

4.1 Estimating Equation

The primary specification is a two-way fixed effects model:

$$Y_{ft} = \alpha_f + \gamma_t + \beta_1(\text{VRU}_f \times \text{Post}_t) + \beta_2(\text{Boundary}_f \times \text{Post}_t) + \varepsilon_{ft} \tag{1}$$

where Y_{ft} is the knife crime rate per 100,000 in force f and financial year t; α_f and γ_t are force and year fixed effects; VRU_f indicates forces receiving VRU funding; and Boundary_f indicates untreated forces adjacent to at least one VRU force. The omitted category is interior forces. The parameter β_1 captures the direct effect of VRU funding; β_2 captures the spillover to boundary forces. Standard errors are clustered at the police force area level.

The sign of β_2 is the key test. Under displacement, $\beta_2 > 0$: crime pushed out of VRU forces raises crime in adjacent areas. Under deterrence, $\beta_2 < 0$: enforcement externalities reduce crime beyond the treated jurisdiction. If $\beta_2 \approx 0$, the VRU effect is spatially contained.

4.2 Staggered Treatment

Because 18 forces were treated in 2019 and two in 2022, standard TWFE may be biased by treatment effect heterogeneity across cohorts (Goodman-Bacon, 2021; de Chaisemartin and d'Haultfœuille, 2020). I therefore complement TWFE with the Callaway–Sant'Anna estimator, which constructs cohort-specific group-time average treatment effects using a doubly robust approach and never-treated forces as the comparison group (Callaway and Sant'Anna, 2021). I also report Sun–Abraham interaction-weighted estimates (Sun and Abraham, 2021). However, because the spillover decomposition requires estimating the boundary coefficient separately from the direct VRU effect, and because the Callaway–Sant'Anna framework does not natively accommodate multi-treatment-arm designs, the spillover analysis relies on TWFE with the full force classification.

4.3 Identification Assumptions and Threats

The identifying assumption for β_1 is parallel trends between VRU forces and interior forces in the absence of treatment. Given that VRU forces were selected on high pre-treatment violence, this assumption is suspect. I present three pieces of evidence on its plausibility: (i) the Callaway–Sant'Anna pre-trends test, (ii) visual inspection of pre-treatment trajectories, and (iii) the pre-COVID restriction.

For the spillover coefficient β_2, identification requires that boundary and interior forces would have followed parallel trends absent VRU spillovers. This assumption is more plausible than for the direct effect: neither boundary nor interior forces received VRU funding, and the classification is based on geography rather than crime levels. The pre-treatment rates in Table 1 are more similar across these two groups (31.8 vs. 26.1) than between VRU and non-VRU forces.

Inference with few clusters. With 42 forces and only 22 in the untreated sample for the spillover analysis, conventional cluster-robust standard errors may over-reject. I supplement conventional inference with two alternatives: randomization inference, which permutes the VRU treatment indicator across forces (1,000 permutations) and computes the empirical distribution of the test statistic under the sharp null of no effect (Fisher, 1935; Young, 2019); and the wild cluster bootstrap, which resamples at the force level with 999 replications (Cameron et al., 2008; Roodman et al., 2019).

5. Results

5.1 Direct Effect on VRU Forces

Table 2 presents the main results. Column 1 reports the Callaway–Sant'Anna aggregated ATT: a reduction of 13.67 knife crimes per 100,000 ($p < 0.01$). Taken at face value, this would represent a 23 percent decline relative to the pre-treatment VRU mean of 58.8. However, the formal pre-trends equivalence test rejects the null of parallel trends ($p = 0.000$), rendering this estimate unreliable. VRU forces were on steeper upward trajectories before treatment than comparison forces — precisely because they were selected on high and rising violence.

Table 2: Effect of Violence Reduction Units on Knife Crime

	(1)	(2)	(3)	(4)	(5)
VRU × Post	-13.67***	1.41	10.25***	-0.19**	2.34
	(4.14)	(3.25)	(3.68)	(0.09)	(3.22)
Observations	630	630	420	624	615
Forces	42	42	42	42	41
Treated Forces	20	20	20	20	19
Estimator	CS-DiD	TWFE	TWFE	TWFE	TWFE
Pre-COVID Only	No	No	Yes	No	No
Dep. Var.	Level	Level	Level	Log	Level
Dep. Var. Mean	58.8				

Notes: Dependent variable: knife crime rate per 100,000 population (columns 1–3, 5) or log(rate + 1) (column 4). Column 1 reports the Callaway and Sant'Anna (2021) aggregated ATT. Columns 2–5 report TWFE estimates with force and year fixed effects. Column 3 restricts to financial years ending 2011–2020 (pre-COVID). Column 5 excludes the Metropolitan Police. Standard errors clustered at the police force area level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Column 2 reports the standard TWFE estimate: $\hat{\beta}_1 = 1.41$ (SE = 3.25, $p = 0.665$). The positive but insignificant point estimate is consistent with either no effect or a small effect masked by residual confounding from differential pre-trends. Column 4 uses the log transformation, yielding a 19.1 percent reduction ($p = 0.049$), but this marginal significance vanishes under randomization inference. Column 5 excludes the Metropolitan Police — the largest force, with knife crime rates several times the national average — and produces a similarly null result ($\hat{\beta} = 2.34$, $p = 0.467$).

Pre-COVID dynamics. Column 3 restricts the sample to financial years ending through 2020, capturing only the first year of VRU operation. The coefficient is positive and strongly significant ($\hat{\beta} = 10.25$, $p = 0.008$). This counterintuitive finding has a straightforward explanation: VRU forces were selected on rising violence, and the first post-treatment year simply extends the pre-existing upward trajectory. The intervention had at most a few months to operate before COVID-19 lockdowns disrupted all crime patterns. This result underscores the difficulty of separating treatment effects from selection dynamics in short post-treatment windows.

The honest conclusion is that the data cannot identify the direct effect of VRUs on knife crime in treated forces. Selection on pre-treatment violence levels and trends confounds all estimators. This finding echoes the Home Office's own acknowledgment that its evaluation design "cannot confirm a causal relationship between VRU activity and changes in serious violence" (Home Office, 2022).

5.2 Spillover Analysis

The more informative — and more credibly identified — analysis concerns spillovers. Table 3 decomposes the untreated-force effect into boundary and interior components using Equation (1).

The boundary coefficient is -4.39 per 100,000 (SE $= 2.03$, $p = 0.032$). The sign is negative, pointing toward deterrence rather than displacement: knife crime fell more in untreated forces adjacent to VRU forces than in the single interior comparator. The magnitude implies a 13.8 percent reduction relative to the boundary-force pre-treatment mean of 31.8, or roughly one-third of a standard deviation.

However, the inferential basis for this result is thin. Randomization inference yields $p = 0.298$, and the cluster bootstrap produces $p = 0.303$. The discrepancy between conventional and permutation-based inference is driven by the extreme control-group imbalance: with 21 boundary forces and 1 interior force, the effective comparison is between 21 units and a single unit. Conley and Taber (2011) show that inference in such settings requires care; the RI p-value is the more reliable guide.

Interpretation. If one takes the conventional estimate seriously, the results suggest that VRU forces' enforcement spillovers reduced knife crime in adjacent jurisdictions. This is consistent with the "diffusion of benefits" hypothesis documented in the hot-spots policing

Table 3: Displacement vs. Deterrence: Spillovers at Force Boundaries

	(1) All Forces	(2) Untreated Only
Panel A: Direct Effect		
VRU × Post	-2.78	
	(2.64)	
Panel B: Spillover		
Boundary × Post	-4.39**	-4.39**
	(1.98)	(2.03)
Observations	630	330
Forces	42	22
RI *p*-value (spillover)	0.298	0.298
Bootstrap *p*-value	0.303	0.303

Notes: Dependent variable: knife crime rate per 100,000. Column 1 includes all 42 forces; column 2 restricts to untreated forces only (22 forces). Boundary forces share a geographic border with at least one VRU force. If VRUs displace crime, the boundary coefficient should be positive; if VRUs deter crime regionally, the coefficient should be negative. Randomization inference permutes VRU assignment across forces (1,000 permutations). Cluster bootstrap resamples forces with replacement (999 replications). All specifications include force and year fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

literature (Weisburd et al., 2006; Braga et al., 2019) and with theoretical models in which deterrence operates through offender expectations of apprehension that are not perfectly localized (Chalfin et al., 2022). The effect would be economically meaningful: 4.39 fewer knife crimes per 100,000 across 21 forces serving roughly 19 million people implies approximately 834 fewer knife offences per year in boundary forces alone.

If instead one privileges the RI results, the finding is a well-powered null: the data are consistent with VRU spillovers being zero, meaning that whatever effects VRUs have (or do not have) are spatially contained within treated forces. Under this interpretation, the program neither displaces nor deters crime across borders. This ambiguity is partly structural: the boundary coefficient is effectively a comparison between 21 forces and one outlier (Dyfed-Powys), and any idiosyncratic shock to that single comparator could drive the result. Future work with finer geographic data — for example, distance-based gradients using crime records geocoded to LSOA level — could provide a more credible decomposition by exploiting within-force variation near boundaries.

5.3 Robustness

Table 4 consolidates the robustness analysis. The direct effect is fragile across specifications: TWFE, Callaway–Sant'Anna, log, and pre-COVID estimates disagree on sign, magnitude, and significance, confirming that the direct effect is not identified. The spillover estimate is stable in point estimate (-4.39 across all variants) but sensitive to the inference method. Firearm offences — a less common outcome that should also be affected if VRUs operate through general deterrence — show a precisely estimated null ($\hat{\beta} = -0.12$, SE $= 1.26$, $p = 0.926$), consistent with VRU interventions being knife-specific or with the outcome being too rare for detection at the force-year level.

6. Discussion

Why the direct effect is unidentified. The fundamental challenge is that the Home Office selected VRU forces on the very outcome the program targets. This is sensible policy — allocate resources where the problem is worst — but it is devastating for causal inference. Selection on levels induces differential trends through mean reversion (Ashenfelter, 1978), through ceiling effects, and through the mechanical correlation between high initial levels and subsequent changes. The divergence between the Callaway–Sant'Anna estimate (strongly negative) and the TWFE estimate (null) reflects different approaches to this problem, neither of which resolves it fully. Future evaluations would benefit from the government committing to randomized or regression-discontinuity-based allocation, as Heller et al. (2017) advocate

Table 4: Robustness of Spillover Estimate

Specification	Estimate	SE	p-value
Direct effect (VRU forces)			
TWFE	1.41	3.25	0.665
Callaway-Sant'Anna	-13.67	4.14	0.001
Log(rate + 1)	-0.191	0.094	0.042
Pre-COVID only	10.25	3.68	0.005
Excl. Met Police	2.34	3.22	0.467
Randomization inference	1.41	—	0.668
Spillover (boundary forces)			
TWFE	-4.39	2.03	0.030
Randomization inference	-4.39	—	0.298
Cluster bootstrap	-4.39	2.62	0.303

Notes: Dependent variable: knife crime rate per 100,000 (levels) or log(rate + 1). Direct effect estimates use all 42 forces; spillover estimates use 22 untreated forces only. Randomization inference permutes treatment assignment (1,000 draws); cluster bootstrap resamples forces with replacement (999 draws).

for US anti-violence programs.

Displacement versus deterrence in context. The suggestive negative spillover aligns with a growing body of evidence that place-based policing does not simply redistribute crime. Weisburd et al. (2006) found that hot-spots policing in Jersey City produced a "diffusion of benefits" to surrounding areas in 11 of 12 experimental sites. Draca et al. (2011) exploited the redeployment of London police after the July 2005 bombings and found crime reductions in redeployed areas with no offsetting increase elsewhere. Blattman et al. (2021) randomized policing intensity across sectors in Bogotá and Medellín, finding crime reductions that persisted after the intervention ended and did not spill to neighboring sectors. My boundary/interior decomposition provides a complementary national-scale test of the same hypothesis, though at a coarser geographic resolution.

Limitations. Three limitations deserve emphasis. First, the "interior" control group consists of a single force (Dyfed-Powys), making the boundary coefficient effectively a comparison of 21 forces against one. This is not an artifact of research design but a structural feature of VRU coverage: the program is so geographically extensive that nearly every untreated force borders a treated one. Second, the annual frequency of the data limits the ability to trace dynamic effects or to isolate the pre-COVID treatment window precisely. Monthly data, which the Home Office publishes for aggregate crime categories but not for knife crime specifically at the force level, would permit sharper event-study designs. Third, the treatment is a bundle — VRUs fund many distinct activities — and I cannot decompose which components drive any observed effect, a challenge common to evaluations of complex programs (Ludwig et al., 2011).

Policy implications. The most robust conclusion is methodological: evaluating place-based crime programs assigned on the basis of pre-treatment violence is extremely difficult, and the UK government's VRU allocation design does not permit credible causal inference about the direct program effect. This does not mean VRUs are ineffective — it means the current evidence base cannot tell us whether they are. The suggestive evidence against displacement is more policy-relevant than the direct-effect estimates: even if we cannot quantify how much VRUs reduce crime where they operate, the data provide no evidence that they merely push crime elsewhere.

7. Conclusion

This paper has asked whether Britain's £254 million investment in Violence Reduction Units reduces knife crime or displaces it. The honest answer is that the direct effect cannot be credibly identified, because VRU forces were selected on the basis of pre-treatment violence. The more interesting finding concerns the spatial margin: conventional inference suggests that untreated forces adjacent to VRU forces experienced lower knife crime than more distant comparators, consistent with deterrence rather than displacement. This result, however, is fragile under randomization inference, reflecting the extreme geographic coverage of the VRU program — which leaves almost no untreated comparator that does not border a treated force.

The broader lesson is one of institutional design. If policymakers wish to know whether violence-prevention programs work, they must build evaluability into the allocation mechanism. Randomized phase-ins, regression-discontinuity thresholds, or stratified rollouts that preserve credible comparison groups are not luxuries for academic researchers — they are prerequisites for knowing whether public money is well spent. Until then, the question of whether VRUs save lives or merely move violence remains, uncomfortably, unanswered.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: https://github.com/SocialCatalystLab/ape-papers

Contributors: @olafdrw

First Contributor: https://github.com/olafdrw

References

Allen, Grahame and Lukas Audickas, "Knife Crime in England and Wales," Briefing Paper No. SN04304, House of Commons Library, London 2019.

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman, "Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco," *American Economic Journal: Applied Economics*, 2015, *7* (1), 151–182.

Ashenfelter, Orley, "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics*, 1978, *60* (1), 47–57.

Blattman, Christopher, Donald P. Green, Daniel Ortega, and Santiago Tobón, "Place-Based Interventions at Scale: The Direct and Spillover Effects of Policing and City Services on Crime," *Journal of the European Economic Association*, 2021, *19* (4), 2221–2272.

Bowers, Kate J. and Shane D. Johnson, "Measuring the Geographical Displacement and Diffusion of Benefit Effects of Crime Prevention Activity," *Journal of Quantitative Criminology*, 2003, *19* (3), 275–301.

Braga, Anthony A., Brandon Turchan, Andrew V. Papachristos, and David M. Hureau, "Hot Spots Policing of Small Geographic Areas Effects on Crime," *Campbell Systematic Reviews*, 2019, *15* (3), e1046.

Busso, Matias, Jesse Gregory, and Patrick Kline, "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 2013, *103* (2), 897–947.

Butts, Jeffrey A., Caterina G. Roman, Lindsay Bostwick, and Jeremy R. Porter, "Cure Violence: A Public Health Model to Reduce Gun Violence," *Annual Review of Public Health*, 2015, *36*, 39–53.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Chalfin, Aaron, Jacob Kaplan, and Michael LaForest, "Crime Reduction Spillovers of a Place-Based Policing Strategy," *Criminology*, 2022, *60* (2), 243–271.

Conley, Timothy G. and Christopher R. Taber, "Inference with "Difference in Differences" with a Small Number of Policy Changes," *Review of Economics and Statistics*, 2011, *93* (1), 113–125.

Cornish, Derek B. and Ronald V. Clarke, "Understanding Crime Displacement: An Application of Rational Choice Theory," *Criminology*, 1987, *25* (4), 933–948.

de Chaisemartin, Clément and Xavier d'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

Deuchar, Ross, *Gangs and Spirituality: Global Perspectives*, Cham: Palgrave Macmillan, 2018.

Draca, Mirko, Stephen Machin, and Robert Witt, "Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks," *American Economic Review*, 2011, *101* (5), 2157–2181.

Feliciantonio, Cesare Di and Eliana Ferrara, "Organized Crime and Economic Growth: Evidence from Municipalities Infiltrated by the Mafia," *Italian Economic Journal*, 2022, *8* (1), 1–30.

Fisher, Ronald A., *The Design of Experiments*, Edinburgh: Oliver and Boyd, 1935.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Guerette, Rob T. and Kate J. Bowers, "Assessing the Extent of Crime Displacement and Diffusion of Benefits: A Review of Situational Crime Prevention Evaluations," *Criminology*, 2009, *47* (4), 1331–1368.

Heller, Sara B., Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A. Pollack, "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago," *Quarterly Journal of Economics*, 2017, *132* (1), 1–54.

Home Office, "Serious Violence Strategy," Policy Paper, HM Government, London 2018.

— , "Serious Violence Fund: Guidance for Police Force Areas," Guidance Document, HM Government, London 2019.

— , "Violence Reduction Units: Year Ending March 2022 Evaluation Report," Research Report, HM Government, London 2022.

IRISS, "Violence Reduction Units: A Review of the Evidence," Technical Report, Institute for Research and Innovation in Social Services, Glasgow 2020.

Johnson, Shane D., Rob T. Guerette, and Kate J. Bowers, "Crime Displacement: What We Know, What We Don't Know, and What It Means for Crime Reduction," *Journal of Experimental Criminology*, 2014, *10* (4), 549–571.

Kline, Patrick and Enrico Moretti, "Place Based Policies with Unemployment," *American Economic Review*, 2013, *103* (3), 238–243.

Ludwig, Jens, Jeffrey R. Kling, and Sendhil Mullainathan, "Mechanism Experiments and Policy Evaluations," *Journal of Economic Perspectives*, 2011, *25* (3), 17–38.

Office for National Statistics, "Crime in England and Wales: Police Force Area Data Tables," Technical Report, Office for National Statistics, Newport 2024.

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb, "Fast and Wild: Bootstrap Inference in Stata Using Boottest," *Stata Journal*, 2019, *19* (1), 4–60.

Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe, "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Weisburd, David, Laura A. Wyckoff, Justin Ready, John E. Eck, Joshua C. Hinkle, and Frank Gajewski, "Does Crime Just Move Around the Corner? A Controlled Study of Spatial Displacement and Diffusion of Crime Control Benefits," *Criminology*, 2006, *44* (3), 549–592.

Wong, Kevin, Kris Christmann, Lin Meadows, and Matthew Maycock, "Can a "Public Health" Approach Reduce Knife Crime?," *British Journal of Criminology*, 2019. Forthcoming.

Young, Alwyn, "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *Quarterly Journal of Economics*, 2019, *134* (2), 557–598.

A. Data Appendix

A.1 Data Sources

The primary data source is the ONS Police Force Area Data Tables, published annually as supplementary tables to the "Crime in England and Wales" statistical bulletin. I use Table P1 ("Police recorded crime by offence and police force area") for knife crime counts and Table P3 for mid-year population estimates. Both are available from the ONS website at https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice.

Knife crime is defined as "offences involving a knife or sharp instrument" and includes the following offence groups: homicide, attempted murder, threats to kill, assault with injury and assault with intent to cause serious harm, robbery, rape, sexual assault, and selected other offences. The ONS notes that recording practices for knife-flagged offences have improved over time, with the National Crime Recording Standard audit process enhancing consistency from 2014 onward.

A.2 VRU Force List

The 18 forces in the 2019 cohort are: Metropolitan Police, Greater Manchester, West Midlands, West Yorkshire, Merseyside, South Yorkshire, Lancashire, Nottinghamshire, Avon and Somerset, Thames Valley, Hampshire, Kent, Essex, Sussex, Northumbria, Leicestershire, South Wales, and West Mercia. Cleveland and Humberside were added in 2022.

A.3 Contiguity Matrix

Force contiguity is defined by shared land borders using the official force boundary polygons. Island forces (there are none in the mainland classification) and forces connected only by water are not treated as contiguous. The resulting contiguity matrix has 42 rows and columns. Of the 22 untreated forces, 21 share at least one border with a VRU force (boundary forces) and 1 shares no such border (Dyfed-Powys, classified as interior).

A.4 Sample Construction

The analysis sample includes 42 territorial police forces observed over 15 financial years (ending March 2011 through March 2025), yielding 630 force-year observations. The British Transport Police is excluded because it has no fixed geographic jurisdiction. Six observations have missing knife crime counts in the raw data and are coded as zero where the ONS indicates suppression for small counts; results are robust to dropping these observations entirely.

B. Identification Appendix

B.1 Pre-Trends

The Callaway–Sant'Anna pre-trends equivalence test rejects the null of parallel pre-treatment trends between VRU forces and never-treated forces ($p = 0.000$). This is expected given selection on pre-treatment levels: forces with higher baseline violence tend to be on steeper upward trajectories, and the parallel trends assumption is violated even conditional on force fixed effects.

For the spillover analysis, the pre-treatment divergence between boundary and interior forces is smaller (31.8 vs. 26.1 per 100,000), and the pre-treatment trajectories, while not formally testable with a single interior force, appear broadly parallel in levels.

B.2 Randomization Inference

I implement Fisher-type randomization inference by permuting the VRU treatment indicator across forces 1,000 times, holding the number of treated forces fixed at 20. For each permutation, I re-estimate Equation (1) and record the coefficient on the VRU $\times$ Post and Boundary $\times$ Post interactions. The RI p-value is the fraction of permuted coefficients exceeding the observed coefficient in absolute value. This procedure is valid regardless of the number of clusters and does not rely on asymptotic approximations (Young, 2019).

C. Robustness Appendix

C.1 Alternative Outcome: Firearm Offences

To test whether VRU effects extend to other weapon types, I replicate the main specification using firearm offences per 100,000 as the dependent variable. The direct effect is $\hat{\beta} = -0.12$ (SE $= 1.26$, $p = 0.926$), a precisely estimated null. This is consistent with VRU activities being focused primarily on knife crime and with firearm offences being too rare at the force-year level (mean $= 5.2$ per 100,000) for detection.

C.2 Excluding the Metropolitan Police

The Metropolitan Police is an outlier in both population (8.9 million, roughly 15 percent of the England and Wales total) and knife crime rate (exceeding 100 per 100,000 in recent years). Column 5 of Table 2 drops the Metropolitan Police and yields $\hat{\beta} = 2.34$ (SE $= 3.22$, $p = 0.467$), confirming that the null direct effect is not driven by a single influential observation.

C.3 Population Shares

VRU forces serve 67.6 percent of the England and Wales population, boundary forces 30.9 percent, and the single interior force (Dyfed-Powys) 1.5 percent. These shares underscore the near-universal geographic reach of the VRU program and the challenge of finding credible untreated comparators.

D. Standardized Effect Sizes

Table 5: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Direct effect (VRU forces)						
Knife crime	1.41	3.25	30.0	0.047	0.108	Small positive
Firearm offences	-0.12	1.26	6.6	-0.018	0.191	Small negative
Spillover (boundary forces)						
Knife crime	-4.39	2.03	15.4	-0.285	0.132	Large negative

Notes: SDE = $\hat{\beta}$ / SD(Y), where SD(Y) is the pre-treatment standard deviation. Classification: Large ($|SDE| > 0.15$), Moderate ($0.05 < |SDE| \leq 0.15$), Small ($0.005 < |SDE| \leq 0.05$), Null ($|SDE| \leq 0.005$). Classifications refer to magnitude, not statistical significance. Research question: Do Violence Reduction Units reduce or displace knife crime? Data: ONS Police Force Area Data Tables, financial years ending 2011–2025. Method: TWFE DiD. Sample: 42 police force areas × 15 years. Treatment: VRU funding (20 forces from April 2019/2022).