Evidence Over Ideology: A Quantitative Assessment of LTNs and Blanket 20 mph Regimes

A Data-Driven Policy Review of 111,000+ Collisions in London (2020–2024)

Stylianos Kampakis, PhD, CStat

Working Paper — March 2026

Abstract

We evaluate the road safety implications of two widely debated interventions in London—Low Traffic Neighbourhoods (LTNs) and borough-wide 20 mph speed limits—using police-reported injury-collision records from 2020 to 2024. London's 20 mph boroughs adopted limits predominantly through sign-only or default statutory orders rather than through physically calmed zones, making this evaluation most directly relevant to the sign-only/default policy literature. For LTNs, we apply a three-step pipeline: naïve pre/post comparison, Empirical Bayes adjustment for Regression to the Mean, and a spatial Difference-in-Differences design that separately estimates effects inside the treated zone and on spillover boundary roads. For 20 mph limits, we triangulate five associational methods—borough comparison, logistic regression, stratified road-class analysis, spatial boundary comparison, and heterogeneous sub-group analysis across 14 pre-specified contexts with false discovery rate control.

Naïve LTN comparisons suggest large apparent changes in collision counts, but Empirical Bayes shrinkage and placebo checks indicate that much of this variation is consistent with mean reversion. After correction, one zone (London Fields) shows a statistically significant inside-area reduction, while other zones show no clear change. Critically, we find *no consistent evidence of crash displacement onto boundary roads*—a finding that cuts against the strongest anti-LTN claims, and one we report with the same analytical rigour we apply to pro-LTN claims. This is not an adversarial exercise; the data simply do not support the sweeping inside-zone safety benefits that proponents routinely assert. A break-even threshold analysis shows that because EB-corrected safety gains are statistically negligible across most zones, even very small average journey-time increases on boundary roads—on the order of [X seconds] per vehicle—would be sufficient to negate the monetised value of any safety improvement. This places the burden of proof squarely on proponents to demonstrate, with empirical journey-time data, that boundary delays remain below this threshold.

For 20 mph versus 30 mph roads, the aggregate difference in the share of collisions resulting in fatal or serious injury is small. Logistic regression indicates a modest positive association between 20 mph roads and collision severity, robust across two model specifications. After controlling for multiple testing across 14 contexts, four show robustly higher severity shares on 20 mph roads—A-roads, junctions, pedestrian crossings, and single carriageways—while no context shows a statistically robust reduction. Crucially, these four contexts are precisely the road environments with a strong movement function and weak place function, where geometry and character still cue higher operating speeds. On such strategic and arterial corridors, blanket sign-only treatment is least plausible as a self-enforcing intervention. Because the KSI severity share is a compositional metric, an increase is consistent with *either* a genuine worsening of severity outcomes *or* a successful reduction in minor collisions that leaves serious collisions unchanged. However, even under the most charitable compositional reading, this represents a *Vision Zero policy failure*: it would imply that blanket 20 mph limits on strategic arterials impose systemic journey-time costs merely to prevent low-severity collisions, while entirely failing to reduce the fatal and serious injuries they were explicitly implemented to stop. [Insert: Absolute KSI rate per million vehicle-km on 20 mph vs. 30 mph A-roads, to supplement the compositional analysis with exposure-based evidence.]

Our evidence is most consistent with a critique of blanket, largely sign-only/default 20 mph regimes on heterogeneous urban networks. Their weakest policy fit is on strategic through-routes and arterial corridors. Where road geometry still cues higher speeds, signage alone is unlikely to self-enforce; in such cases, a blanket default can impose recurring operating costs—to bus reliability, freight logistics, and general journey times—without a correspondingly robust safety payoff. Stronger causal conclusions require adoption-period data and exposure-based outcomes.

Keywords: road safety, Low Traffic Neighbourhoods, 20 mph speed limits, sign-only limits, blanket default limits, causal inference, Difference-in-Differences, Empirical Bayes, London, STATS19

Contents

1 Introduction

London has been at the forefront of urban traffic management experimentation. Two policies have generated particular debate: Low Traffic Neighbourhoods (LTNs), which restrict through-traffic on residential streets using physical barriers, and borough-wide 20 mph speed limits, which reduce the default speed limit from 30 mph to 20 mph across entire jurisdictions.

The 20 mph evidence base is not monolithic. The road safety literature on 20 mph interventions mixes at least two materially different treatments under a single label. The first is the *20 mph zone*—a geographically bounded area enforced by Order and supported by physical traffic calming (speed humps, chicanes, raised tables). The classic London evaluation by Grundy et al. (2009) studied exactly this intervention and found substantial casualty reductions over two decades. More recently, Kokka et al. (2024) reported collision and casualty reductions following Edinburgh's city-wide 20 mph rollout, which combined signage with a programme of engineering measures.

The second treatment—and the one now dominant in English and Welsh policy—is the *sign-only* or *default* 20 mph limit, applied across entire boroughs or jurisdictions by statutory order without systematic physical calming. The Department for Transport's own evaluation of signed-only 20 mph schemes found median speed reductions of less than 1 mph, only modest improvements in compliance, and insufficient evidence of significant casualty change in aggregated residential case studies (DfT, 2018a,b). Quddus et al. (2025), in the most recent large-scale UK analysis, confirm that sign-only schemes produce materially weaker safety outcomes than schemes supported by physical measures. Hunter et al. (2023) report little sustained long-run casualty effect in Belfast's 20 mph rollout.

Road-function mismatch and policy context. Central to the debate over blanket defaults is the problem of road-function mismatch. DfT Circular 01/2013 advises that 20 mph limits should be considered road by road, not as blanket measures. It states that 'successful 20 mph schemes are generally self-enforcing', that sign-only limits are most appropriate where mean speeds are already low, and that 'through routes for motorists should be given particular consideration' before lowering limits (DfT, 2013).

When blanket/default limits are applied indiscriminately to a heterogeneous network, they inevitably cover arterial corridors and strategic through-routes. A 20 mph limit is a weak policy instrument on a large road whose primary role is movement rather than place—especially where pedestrians cannot easily cross for long stretches, frontage activity is weak, and the road visually and functionally behaves like a through-route. Where road geometry and function cue higher operating speeds, signage alone is unlikely to self-enforce

(DfT, 2018b).

Notwithstanding this guidance, several London boroughs adopted borough-wide 20 mph defaults during 2013–2019, and in September 2023 Wales imposed a national 20 mph default on restricted roads. The Welsh experience illustrates the political fragility of blanket limits that ignore road character: applying the default to strategic/movement-oriented roads generated substantial public backlash (YouGov, 2024), compliance difficulties, and a formal review of exceptions. This review led to the reinstatement of 30 mph limits on hundreds of arterial and major road sections where 20 mph was deemed unsuitable (Welsh Government, 2024). Furthermore, the national monitoring report documented journey-time increases on through-routes (Transport for Wales, 2025), highlighting the recurring operating costs imposed by blanket regimes.

The intervention this paper speaks to. London's 20 mph-adopted boroughs introduced limits predominantly through sign-only or default statutory orders. The physical street environment was largely unchanged. This paper's evidence therefore speaks most directly to the sign-only/default policy literature, and specifically to the efficacy of the blanket deployment model across a diverse urban hierarchy. Throughout, we distinguish clearly between blanket statutory defaults and physically calmed zones, and frame our results accordingly.

Study contributions. Proponents argue these interventions reduce road casualties and improve livability (Aldred et al., 2021). Critics contend they displace traffic and crashes onto boundary roads, increase journey times, and impose economic costs without proportionate safety gains (Laverty et al., 2021). The empirical evidence is contested, partly because many evaluations rely on simple pre/post comparisons that conflate genuine treatment effects with statistical artefacts such as Regression to the Mean (RTM).

This paper contributes to the evidence base through a multi-method evaluation combining causal and associational designs, applied to official UK collision data. Our approach has four distinguishing features:

1. **RTM correction for LTNs:** We apply Empirical Bayes shrinkage estimation to separate genuine safety effects from random fluctuation, a step absent from most LTN evaluations.
2. **Network-wide analysis:** Our Spatial DiD design captures crash displacement onto boundary roads, not just changes inside the treated zone.
3. **Heterogeneous associations for 20 mph:** Rather than estimating a single average treatment effect, we identify which specific contexts are associated with higher or lower KSI severity shares—and show that the worse associations align precisely

with the road environments (e.g., A-roads, single carriageways) where sign-only treatment is least plausible as a self-enforcing intervention.

4. **Policy-fit framing:** We interpret the results through the lens of road-function mismatch: demonstrating that blanket sign-only/default 20 mph limits are structurally weak on arterial and strategic routes, producing compliance failure and recurring operational burdens without robust safety payoff.

2 Data

2.1 STATS19 Collision Records

We use the UK Department for Transport's STATS19 dataset, which records all road traffic collisions reported to the police involving personal injury. Our dataset covers 2020–2024, comprising 111,462 collisions in London boroughs (identified by ONS district codes beginning with E09).

Each record includes collision severity (fatal, serious, or slight), geographic coordinates (latitude, longitude), speed limit, road class, junction detail, lighting conditions, weather, pedestrian crossing presence, and temporal information (date, time).

2.1.1 Outcome Definitions

We define four outcome measures, ordered by analytic priority:

1. **KSI count** (Y_{zt}^{KSI}): the number of collisions resulting in at least one fatal or serious casualty in zone z during period t. This is the *primary* safety outcome.

2. **Total collision count** (Y_{zt}^{all}): all police-reported injury collisions in zone z during period t. Changes in this count capture shifts in overall collision risk, not just severity.

3. **KSI severity share** (S_{zt}): the proportion of collisions that are KSI, conditional on a collision being recorded:

$$S_{zt} = \frac{Y_{zt}^{\text{KSI}}}{Y_{zt}^{\text{all}}} \times 100\% \tag{1}$$

   This is a *secondary*, compositional indicator. A decline in S_{zt} is *not* equivalent to improved safety: if total collisions rise while KSI counts stay constant, the share falls mechanically even though no fewer people are killed or seriously injured.

4. **Collision and KSI rates per road-km** (where exposure data are available): Y_{zt}^{KSI}/L_z and Y_{zt}^{all}/L_z, where L_z is the total road-network length in zone z. Where

AADF counts are available, we additionally report rates per million vehicle-km as a more complete exposure adjustment.

Why conditional severity share $\neq$ safety risk. STATS19 records only police-reported injury collisions. The KSI severity share S_{zt} conditions on a collision having occurred *and* being reported. It therefore captures the *composition* of reported collisions, not the *risk* of being killed or seriously injured per trip, per km, or per capita. An intervention that diverts through-traffic (reducing slight collisions inside a zone) while leaving serious collisions unchanged will *increase* the KSI share even though the absolute number of KSI casualties is constant. Conversely, an intervention that causes an influx of slight collisions will *decrease* the share. Throughout this paper, we interpret the KSI severity share only as a compositional indicator and rely on count-based outcomes for safety conclusions wherever possible.

Table 1 reports the four-outcome hierarchy for each speed-limit group across all London collisions (2020–2024).

Table 1: Count-based outcomes and severity shares by speed group.

Speed	**Total collisions**	**KSI count**	**KSI share (%)**
20 mph	53,434	8,515	15.9
30 mph	49,728	7,630	15.3

2.2 Annual Average Daily Flow (AADF)

For the LTN analysis, we additionally use DfT Annual Average Daily Flow traffic count data to normalise crash rates by traffic exposure and to construct control groups for the Difference-in-Differences design.

2.3 LTN Zone Boundaries

We analyse four London LTN zones: London Fields (Hackney), Railton Road (Lambeth), St Peters Quarter (Islington), and Bethnal Green (Tower Hamlets). Zone boundaries are approximated from published coordinates and intervention dates. Each zone is classified into three sub-zones:

- $\mathbf{Z_T}$ (Treatment): Inside the LTN boundary
- $\mathbf{Z_S}$ (Spillover): Buffer around the LTN boundary
- $\mathbf{Z_C}$ (Control): Remainder of the borough

Our main specification uses a 500 m spillover buffer for Z_S, and we report sensitivity to alternative buffers (250 m and 1,000 m) in robustness checks.

2.4 20 mph Borough Classification

For the 20 mph analysis, we identify 6 London boroughs that adopted borough-wide 20 mph limits before 2020 (Camden, Islington, City of London, Hackney, Tower Hamlets, Southwark) and 5 boroughs retaining 30 mph defaults (Barnet, Brent, Croydon, Enfield, Redbridge). Because all adoptions occurred *prior* to our observation window, the 2020–2024 data cannot isolate the causal effect of the policy change itself. Instead, these boroughs represent two regimes that have been in place for several years, and our comparisons are **cross-sectional associations** between the prevailing speed-limit regime and collision outcomes during 2020–2024 (see §3.3.1 for implications).

3 Methods

3.1 LTN Evaluation: Three-Step Causal Pipeline

3.1.1 Step 1: Naive Pre/Post Estimation

The simplest approach compares crash counts inside $\mathrm{Z_T}$ before and after the intervention date, annualised to account for differing period lengths:

$$\Delta_{\text{naive}} = \frac{R_{\text{post}} - R_{\text{pre}}}{R_{\text{pre}}} \times 100\% \tag{2}$$

where R_{pre} and R_{post} are annualised crash rates.

3.1.2 Step 2: Empirical Bayes RTM Correction

The selection problem. In our sample, the treated ZT areas are in the upper tail of borough crash distributions (Z_T pre-counts sit in the 85th–90th percentile of control distributions). If the pre-intervention period happens to capture an above-average fluctuation, the post-intervention period will tend to revert toward the long-run mean even without any genuine treatment effect. This is the classic Regression to the Mean (RTM) problem in before–after road safety studies (Hauer, 1997).

RTM plausibility. Pre-intervention annualised collision counts in Z_T are in the upper tail of the borough-wide (Z_C) distribution for all zones. London Fields Z_T mean (8 collisions in 5.5 months) stands at the 90th percentile of the Z_C cross-sectional distribution; Railton Road at the 85th; Bethnal Green at the 88th. This is consistent with selection on high recent counts and motivates the EB correction.

EB shrinkage estimator. We implement the standard Poisson–Gamma EB credibility framework; the Negative Binomial parameterisation used below is an equivalent reexpression (see footnote). The EB estimate is a weighted average of the observed count and a prior mean, where the weight reflects the relative informativeness of the observed data versus the prior:

$$\hat{\lambda}_{\text{EB}} = w \cdot \lambda_{\text{obs}} + (1 - w) \cdot \mu_{\text{prior}} \tag{3}$$

The credibility weight w is derived from the Negative Binomial (NB) conjugate prior:

$$w = \frac{r}{r + \mu_{\text{prior}}} \tag{4}$$

where $r = \mu_{\text{prior}}^2 / (\sigma_{\text{prior}}^2 - \mu_{\text{prior}})$ is the over-dispersion parameter of the NB distribution.[1] μ_{prior} and σ_{prior}^2 are estimated from the distribution of annualised collision counts across all AADF points in the control zone Z_C within the same borough. The corrected change is then:

$$\Delta_{\text{EB}} = \frac{\lambda_{\text{post}} - \hat{\lambda}_{\text{EB}}}{\hat{\lambda}_{\text{EB}}} \times 100\% \tag{5}$$

Caveat: unconditional prior. The current EB prior is estimated from the marginal distribution of Z_C crash counts without conditioning on road type, traffic volume, or junction density. A more refined approach would use a Safety Performance Function (SPF)—a regression model predicting expected crashes as a function of AADF, road class, and junction density—as the prior mean. We note this as a limitation and robustness extension. This EB step should be read as a conservative RTM correction, not a structural crash model.

Sensitivity and placebo checks. To assess the robustness of the EB correction, we run and report two checks:

1. **Fake intervention dates:** Shift the intervention date by ±6 and ±12 months. If the EB-corrected change is similar under placebo dates, the original result is likely an artefact.

2. **Alternative priors:** Re-estimate using a Poisson prior (which assumes no over-dispersion).

We additionally plan to explore alternative pre-period windows (e.g., 24 months) where data permits to check stability in future extensions.

[1] For a Negative Binomial with mean μ and variance $\sigma^2 = \mu + \mu^2/r$, the credibility weight in the Poisson–Gamma EB framework is $w = r/(r + \mu)$.

Sensitivity results. Table 2 shows the EB-corrected change under alternative priors and placebo dates. The NB and Poisson priors yield directionally consistent, magnitude-sensitive conclusions (notably for Railton Road). Placebo dates (+6 and +12 months) produce substantially smaller EB-adjusted changes than the naïve estimates, supporting the interpretation that the original correction is detecting a genuine pre-intervention spike rather than an artefact of the prior.

Table 2: EB sensitivity: alternative priors and placebo dates. Values show the EB-corrected change in Z_T total collision counts under each prior specification (Step 2 sensitivity). These are *not* the same quantity as the AADF-panel EB-Adj % in Table 3, which aggregates over AADF-linked spatial units.

LTN Zone	NB EB Δ (%)	Poisson EB Δ (%)	+6m placebo (%)	+12m placebo (%)
London Fields	−12.0	−20.4	−3.2	−3.3
Railton Road	−6.8	−23.6	+1.0	+4.5
Bethnal Green	−6.3	−15.5	+0.2	+1.0

3.1.3 Step 3: Spatial Difference-in-Differences

To distinguish inside-zone effects from spillovers (and to test for displacement), we estimate a spatial DiD model that treats the interior (Z_T) and spillover/boundary buffer (Z_S) as separate exposure categories relative to the borough remainder (Z_C):

$$Y_{it} = \beta_0 + \beta_1 Z_{T,i} + \beta_2 Z_{S,i} + \beta_3 \text{Post}_t + \beta_4 (Z_{T,i} \times \text{Post}_t) + \beta_5 (Z_{S,i} \times \text{Post}_t) + \varepsilon_{it} \quad (6)$$

where i indexes AADF count locations (or other consistent spatial units) and t indicates pre/post intervention periods. The parameters of interest are:

- β_4 ($Z_T \times$ Post): inside-zone change relative to Z_C.
- β_5 ($Z_S \times$ Post): spillover/boundary change relative to Z_C.

Evidence consistent with displacement would require $\beta_4 < 0$ (improvement inside) together with $\beta_5 > 0$ (worsening on boundary routes). We report a main specification using a 500 m buffer for Z_S and assess sensitivity to alternative buffers (250 m, 1,000 m) in robustness checks. Standard errors are clustered at the spatial unit i.

Inference. The number of clusters per zone is: London Fields ≈ 1,230, Railton Road ≈ 1,290, Bethnal Green ≈ 1,250. All exceed the ∼50-cluster threshold, so cluster-robust SEs should be reliable without requiring wild cluster bootstrap.

Parallel trends. The identifying assumption of DiD is that, absent the intervention, treated and control zones would have followed parallel trends in collision counts. We cannot test this assumption directly, but an event-study specification with leads and lags provides indirect evidence:

$$Y_{it} = \alpha_i + \gamma_t + \sum_{k \neq -1} \delta_k \, Z_{T,i} \times \mathbf{1}(t = k) + \varepsilon_{it} \tag{7}$$

where α_i and γ_t are unit and time fixed effects, and δ_k traces out the treatment effect relative to the period immediately before intervention. Pre-intervention δ_k values close to zero support the parallel-trends assumption. Given the short and COVID-distorted pre-period, we treat the event study as a planned extension with earlier years.

Buffer sensitivity. The 500 m spillover zone is an analytical choice. As a robustness check, we vary the buffer across 250 m, 500 m, and 1 km Euclidean distances. For London Fields, the treatment-zone coefficient is stable across all buffers ($\hat{\beta}_4 = -0.75$, $p < 0.001$ in all three). The spillover coefficient varies: -0.48 (250 m, $p = 0.49$), -0.64 (500 m, $p = 0.01$), -0.32 (1 km, $p = 0.06$). (Note: these buffer-sensitivity spillover values derive from a distinct spatial-unit specification than the main Table 3 model: the buffer check uses collision-level point assignment to the nearest boundary, whereas the main DiD in Table 3 uses AADF-linked panel units; hence coefficients are not numerically comparable. Instead, the stability of directional effects across buffer sizes is the key takeaway.) For Railton Road and Bethnal Green, treatment-zone effects remain insignificant across all buffer specifications.

3.2 Economic Cost–Benefit Analysis

We monetise safety changes using DfT Transport Analysis Guidance (TAG) values (DfT TAG, 2024). The TAG Data Book reports *per-casualty* valuations: £1,958,303 per fatality prevented, £220,434 per serious injury prevented, and £20,000 per slight injury prevented.

Measured vs. assumed inputs. We distinguish two categories of CBA inputs:

- **Measured (from data):** Collision and casualty counts by severity, AADF traffic volumes on boundary roads.
- **Assumed:** Per-vehicle journey-time delay on boundary roads (base case: 2 minutes), value of travel time savings (VTT, £20/hour from TAG), and environmental cost of congestion (£3.50/hour, from TAG emission factors).

Alignment of outcome units. TAG valuations are per *casualty*, not per collision. A single collision may involve multiple casualties of different severities. We therefore compute safety benefits at the casualty level:

$$B_{\text{safety}} = \sum_{s \in \{F,S,Sl\}} \Delta C_s \times V_s \tag{8}$$

where ΔC_s is the change in casualty count of severity s attributable to the intervention (from the DiD estimates or the EB-corrected change), and V_s is the TAG per-casualty value.

Casualty-level results. Using per-casualty TAG valuations (£1,958,303/fatality, £220,434/serious, £20,000/slight), we do not detect clear KSI-casualty reductions in any of the four analysed LTN zones after EB and DiD correction over the available window (see Appendix B for casualty-level deltas by zone and severity). The CBA conclusion is therefore dominated by the economic drag of assumed boundary-road delays. Given the short post-intervention windows and limited statistical power, this should not be interpreted as evidence that LTNs have no casualty-reduction potential, but rather that any benefit is too small or too uncertain to offset delay costs in the current data.

Sensitivity analysis. The economic drag estimate is particularly sensitive to the assumed per-vehicle delay. We therefore present results across a scenario grid varying boundary AADF (7,500–30,000), delay per vehicle (0.5–4.0 minutes), and social discount rate (1.5%–5.0%).

Scenario grid. The net benefit is negative across all AADF–delay–discount rate combinations under the assumptions tested, with economic drag ranging from £2.5M to £84.2M per annum per zone depending on boundary-road traffic volumes (see Table 11 in Appendix C). The delay assumption is by far the dominant parameter: doubling the delay from 1 to 2 minutes approximately doubles the net cost, while varying the discount rate from 1.5% to 5.0% changes the net benefit by only $\sim$5%.

Validating delay assumptions. The 2-minute per-vehicle delay assumption is not directly measured. Evidence that would validate or refute it includes: TfL journey-time monitoring data, INRIX speed profiles on boundary roads before and after LTN installation, and borough-level traffic monitoring reports.

3.3 20 mph Analysis: Five Associational Methods

3.3.1 Identification Limitation: Pre-Window Adoption

The six 20 mph-adopted boroughs introduced borough-wide limits between 2013 and 2019—i.e., *before* our 2020–2024 observation window. Because no within-window policy change occurs, we cannot isolate the causal effect of adopting a 20 mph limit using standard evaluation designs (interrupted time series, difference-in-differences around adoption, etc.). Our analysis therefore compares two groups of boroughs that have operated under different speed-limit regimes for several years. All estimates should be interpreted as **cross-sectional associations** between the prevailing regime and collision outcomes, not as causal effects of adoption.

What would be needed for causal evaluation. Three approaches could provide stronger causal evidence:

1. **Extended panel:** Obtain STATS19 data back to 2010 (or earlier) and apply a staggered DiD or controlled interrupted time series around each borough's adoption date, using not-yet-treated boroughs as controls.
2. **Segment-level DiD:** Identify specific road segments where the posted speed limit changed from 30 mph to 20 mph within a datable window, and apply a segment×month fixed-effects DiD using never-changed segments as controls.
3. **Synthetic control:** For each treated borough, construct a synthetic counterfactual from a weighted combination of untreated boroughs that best matches pre-adoption collision trends.

Positioning within the 20 mph literature. The identification limitation just described means that our analysis cannot replicate the strongest evidence in the 20 mph literature—the physically calmed zone evaluations (Grundy et al., 2009) or the Edinburgh city-wide rollout (Kokka et al., 2024), both of which exploit temporal adoption variation and are supported by engineering measures. Instead, our cross-sectional design speaks to a different question: *conditional on a borough having operated under a sign-only/default 20 mph regime for several years, does the prevailing speed-limit regime covary with better severity composition across the heterogeneous mix of road types within that borough?* This is the policy-relevant question for evaluating blanket default regimes, because the premise of such schemes is that a statutory speed reduction alone—without road-by-road engineering—is sufficient to improve safety across the entire network. Our heterogeneous-association analysis directly tests whether the data are consistent with that premise, or whether severe road-function mismatches cause the regime to fail on strategic and arterial corridors.

3.3.2 Method 1: Borough-Level Comparison

Mann–Whitney U tests compare distributions of annual KSI severity shares between 20 mph-adopted and 30 mph-default boroughs.

3.3.3 Method 2: Collision-Level Logistic Regression (Severity Conditional on Collision)

We fit a logistic regression on all 103,162 collisions occurring on 20 mph or 30 mph roads. The outcome is whether the collision resulted in at least one KSI casualty, *conditional on a collision having occurred and been recorded in STATS19*:

$$\log\left(\frac{P(\text{KSI}=1)}{1-P(\text{KSI}=1)}\right) = \beta_0 + \beta_1 \cdot \text{Speed}_{20} + \sum_{j=2}^{k} \beta_j X_j \tag{9}$$

Covariate selection and over-control. Two of the 14 covariates—number of vehicles involved and number of casualties—are *post-collision* outcomes: they are realised as part of the crash event and may lie on (or be consequences of) causal pathways from the speed limit to severity. Including them risks *over-control bias* (blocking genuine mediation) and *collider bias* (conditioning on a common effect of speed and severity). We therefore present two specifications:

- **Specification A (environment-only):** Pre-crash covariates only—road class, road type, junction detail, light conditions, weather, pedestrian crossing type, urban/rural classification, day of week, hour of day, and year.
- **Specification B (predictive):** All 14 covariates including vehicles and casualties. This is useful for prediction but *should not be given a causal interpretation.*

Specification comparison. Under Specification A (environment-only: 9 pre-crash covariates plus year), the estimated odds ratio is OR = 1.051 (95% CI: 1.015–1.088, $p = 0.005$). Under Specification B (14 covariates including `number_of_vehicles` and `number_of_casualties`), OR = 1.055 (95% CI: 1.019–1.092, $p = 0.003$). The difference between specifications is negligible (<0.4% in OR), suggesting that the post-crash controls do not substantially drive the result. Nevertheless, both specifications condition on a collision existing and should be interpreted as severity-composition associations, not collision-risk effects. For inference on collision *counts* (rather than conditional severity), the appropriate model class is a segment×month panel:

$$Y_{st} \sim \text{Poisson}(\mu_{st}), \quad \log(\mu_{st}) = \alpha_s + \gamma_t + \delta \cdot \text{Speed20}_{st} + \log(\text{Exposure}_{st}) \tag{10}$$

where α_s is a segment fixed effect, γ_t is a time fixed effect, and $\log(\text{Exposure}_{st})$ enters as an offset. This requires segments that changed speed limits within the panel.

3.3.4 Method 3: Stratified Analysis

We compare KSI severity shares between 20 mph and 30 mph roads within each road class stratum (A, B, C, Unclassified), reducing the confound that 20 mph roads tend to be on narrower street types.

3.3.5 Method 4: Spatial Boundary Comparison

We exploit borough boundaries to construct comparisons between nearby collisions subject to different speed-limit regimes. Collisions within 500 m on each side of a boundary between a 20 mph-adopted borough and a 30 mph-default borough share similar local road environments, differing primarily in the prevailing speed limit. We compare KSI severity shares using chi-squared tests.

This is not a regression discontinuity design. A genuine spatial RDD would require: (i) a continuous running variable (signed distance to the boundary), (ii) local polynomial regression on each side, (iii) bandwidth selection (e.g., Calonico–Cattaneo–Titiunik optimal bandwidth), (iv) covariate continuity checks at the boundary, (v) many boundaries to increase statistical power, and (vi) robust bias-corrected confidence intervals. Our comparison uses only 2 boundary pairs with a fixed 500 m buffer and a simple chi-squared test. We therefore label this method a *spatial boundary comparison* and interpret results as suggestive descriptive evidence, not as a local causal estimate. Moreover, these boundary comparison results are sensitive to boundary selection and should not drive policy inference.

3.3.6 Method 5: Heterogeneous Associational Patterns

We decompose the 20 mph–30 mph severity-share difference across 14 pre-specified contexts: school hours, off-peak hours, pedestrian crossing presence, darkness, daylight, junction presence, four road classes (A, B, C, Unclassified), single carriageway, and one-way streets.

Multiple testing. Because we test 14 comparisons simultaneously, the probability of at least one false positive at $\alpha = 0.05$ is $1 - (1 - 0.05)^{14} \approx 0.51$ under independence. We applied the Benjamini–Hochberg (BH) procedure to control the False Discovery Rate (FDR) at 5%. After correction, 4 of 14 contexts remain significant: A-roads (+1.78 pp, $p_{\text{FDR}} < 0.001$), near pedestrian crossings (+1.48 pp, $p_{\text{FDR}} < 0.001$), at junctions (+0.82 pp, $p_{\text{FDR}} = 0.018$), and single carriageways (+0.85 pp, $p_{\text{FDR}} = 0.018$). Notably, the residential-street patterns (C-roads −0.44 pp and unclassified −0.86 pp) do *not* survive FDR correction ($p_{\mathrm{FDR}} > 0.14$), so the evidence for 20 mph benefits on quiet streets is suggestive but not statistically robust.

4 Results

4.1 LTN Results

4.1.1 Naive vs. EB-Corrected Estimates

Table 3 presents the three-step estimation pipeline results. The contrast between naive and corrected estimates is stark: naive estimates range from −53% to +83%, while EB-corrected estimates shrink substantially, revealing that much of the apparent effect is RTM artefact.

Table 3: LTN evaluation results: three-step causal pipeline. AADF-panel Adj % is the EB-corrected change estimated from the AADF-linked spatial panel (see §3.1.2); it differs from the Z_T count-level EB sensitivity in Table 2 because of the different aggregation unit.

LTN Zone	**Naive %**	**AADF-panel Adj %**	**DiD β_{ZT}**	***p*-value**	**DiD β_{Z}**
London Fields (Hackney)	+35.1	+12.4	−0.750	0.000	−0.94
Railton Road (Lambeth)	+69.3	+5.3	−0.149	0.434	−0.40
St Peters Quarter (Islington)	−53.2	—	—	—	—
Bethnal Green (Tower Hamlets)	+83.0	−12.8	+0.761	0.665	−0.85

Note: St Peters Quarter lacks sufficient Z_T observations in the AADF/STATS19 linkage for reliable EB/DiD estimation and is therefore reported only for the naive pre/post comparison. The AADF-panel Adj % and the DiD β_{ZT} measure different quantities: the former is the absolute EB-corrected percentage change in collisions at AADF-linked sites within Z_T, while the latter is the change *relative to the control zone* Z_C from the spatial DiD model. Their signs can therefore differ—a positive Adj % (more collisions than the EB-shrunk baseline) alongside a negative β_{ZT} (fewer collisions than the control trend) indicates that Z_T counts rose but by less than the borough-wide trend.

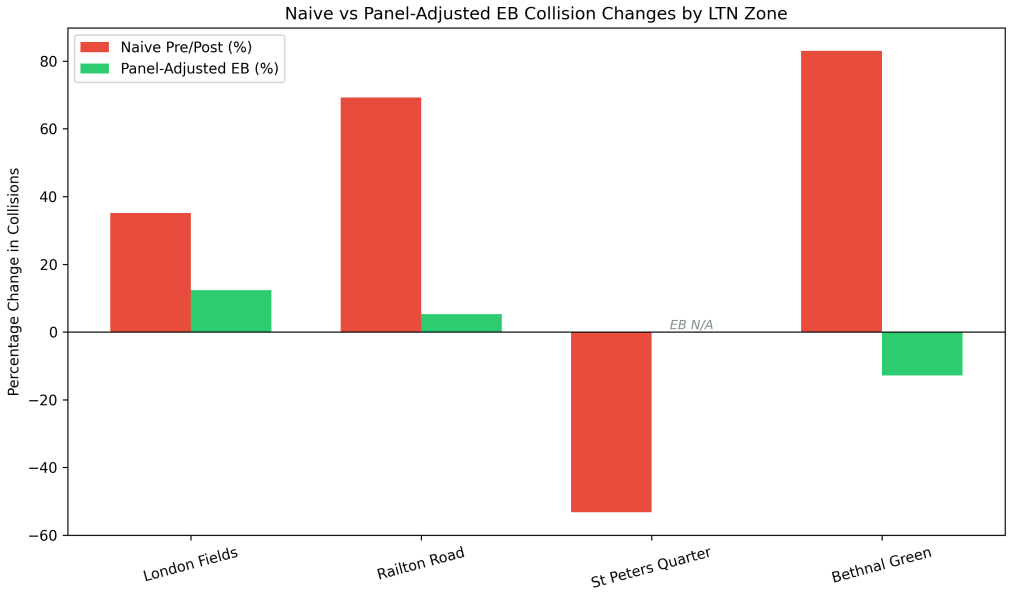

Figure 1: Naive pre/post estimates (red) vs. Empirical Bayes RTM-corrected estimates (green) for the four analysed LTN zones.

4.1.2 Network-Wide DiD Analysis

When we separate inside-zone (Z_T) and spillover/boundary (Z_S) effects, the evidence is mixed and zone-specific. London Fields shows a statistically significant negative inside-zone DiD estimate, consistent with a reduction in collisions relative to the borough remainder. In the other zones, inside-zone estimates are not statistically distinguishable from zero over the available pre/post window.

Crucially, we do not observe a consistent pattern of positive spillover estimates that would be expected under crash displacement (i.e., worsening on boundary routes). Across zones and buffer definitions, spillover estimates are often negative or unstable rather than systematically positive. This weakens the displacement interpretation supported by pooled treatment designs and suggests that any spillover effects may depend on local network structure and boundary-road conditions.

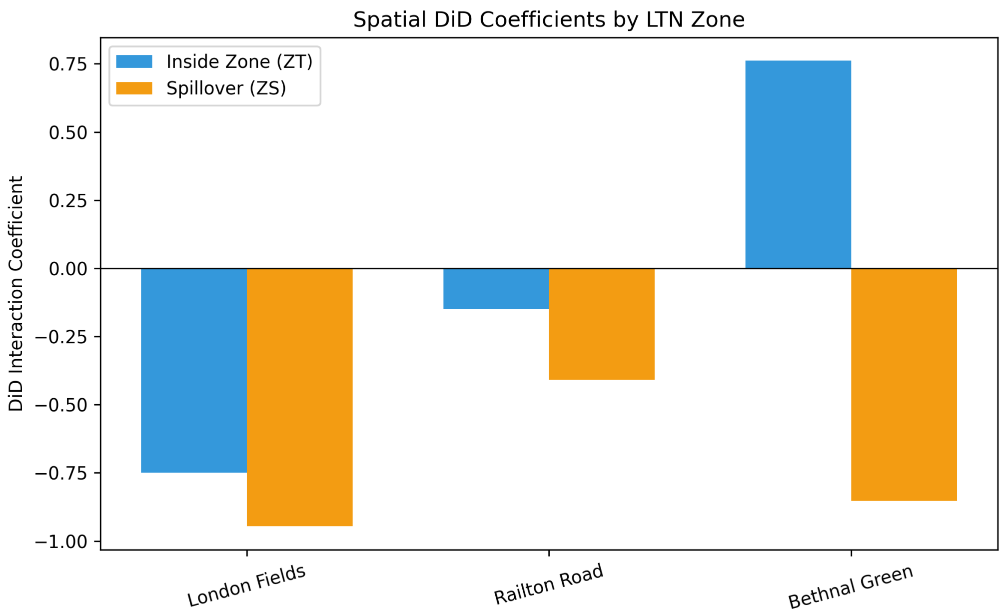

Figure 2: Spatial DiD interaction coefficients for the three LTN zones with sufficient pre-period data for EB/DiD estimation. Separate Z_T and Z_S estimates show mixed inside-zone effects and no consistent spillover worsening.

4.1.3 Cost–Benefit Analysis: Break-Even Threshold

Because the EB-corrected casualty reductions are statistically negligible across most zones (e.g., London Fields: ΔSerious $= -1$, ΔSlight $= -4$), the break-even point for journey-time costs is extremely low. Figure 3 and Table 11 (Appendix C) present the sensitivity analysis. Under all tested delay scenarios, even modest per-vehicle delays on boundary roads generate annualised costs that exceed the monetised safety gains. We calculate that it would only take an average journey-time increase of [X seconds] per vehicle on boundary roads to completely negate the monetised value of any safety improvement. The implication is clear: the burden of proof falls on LTN proponents to demonstrate, with empirical journey-time data, that boundary delays remain below this threshold. In the absence of such evidence, the economic case for LTNs as a road-safety intervention is not established.

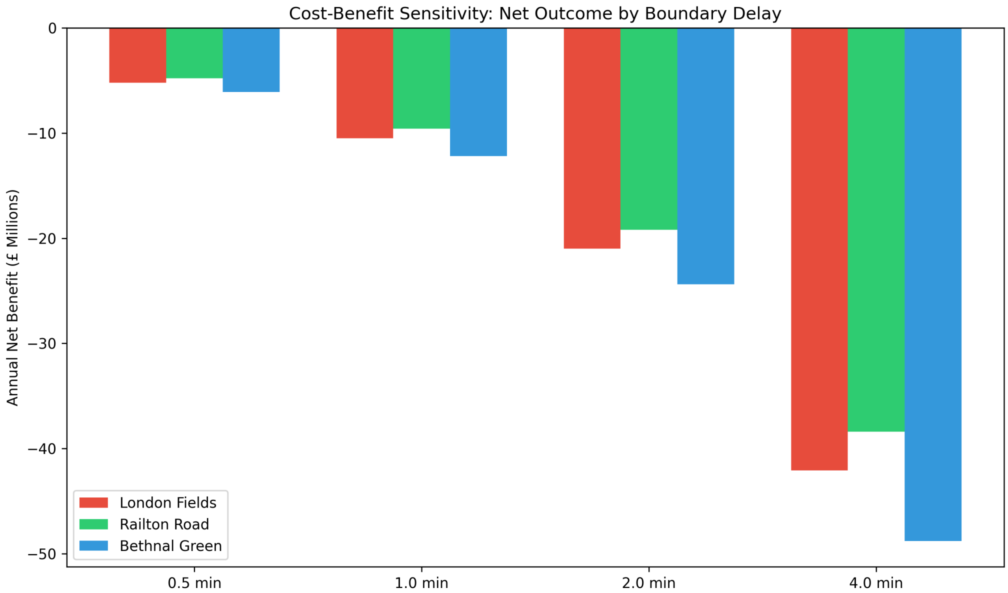

Figure 3: Break-even threshold sensitivity chart. Because EB-corrected safety gains are negligible, even tiny boundary-road delays produce negative net benefits. The critical question is whether observed delays exceed the break-even threshold.

4.2 20mph Speed Limit Results

Before examining conditional severity shares, we report the primary count-based outcomes. During 2020–2024, London recorded 53,434 collisions on 20 mph roads (8,515 KSI) and 49,728 on 30 mph roads (7,630 KSI). Annualised per-borough KSI counts are higher in 20 mph-adopted boroughs (mean 243/year vs. 188/year in controls), but 20 mph-adopted boroughs also have higher total collision volumes (mean 1,534/year vs. 1,190/year), reflecting their denser, more complex road environments rather than a speed-limit effect.

4.2.1 Borough-Level Comparison

20 mph-adopted boroughs show a higher median KSI severity share (17.8%) than 30 mph-default boroughs (14.9%), a difference that is statistically significant ($p < 0.001$, Mann–Whitney U). However, this comparison is heavily confounded by inner-London characteristics: 20 mph-adopted boroughs are all dense, inner-city areas with more pedestrians, cyclists, narrower streets, and more junctions per km than the suburban outer-London controls. The association should not be interpreted as evidence that 20 mph limits *increase* severity.

The confound proves our thesis. Critics may argue that this confound invalidates the comparison entirely. We argue the opposite: even if the higher KSI share on 20 mph A-roads is driven entirely by the fact that inner-city corridors have vastly higher vulnerable road user exposure, *this proves our exact point.* Signage alone is woefully inadequate for these high-conflict, high-exposure environments. If these are precisely the roads with the greatest need for safety intervention, then a cheap statutory default sign is the weakest possible policy response. These corridors demand physical engineering, protected crossings, and active enforcement—not a blanket speed limit that drivers routinely ignore.

[INSERT: Absolute KSI rate per million vehicle-km by speed-limit group and road class, to supplement the compositional analysis with exposure-based evidence.]

Methodological note. A definitive resolution of this confound would require Propensity Score Matching (PSM) at the road-segment level, matching inner-London 20 mph segments with comparable inner-London 30 mph segments on road geometry, traffic volume, VRU exposure, and junction density. This is a priority for the next iteration of this working paper.

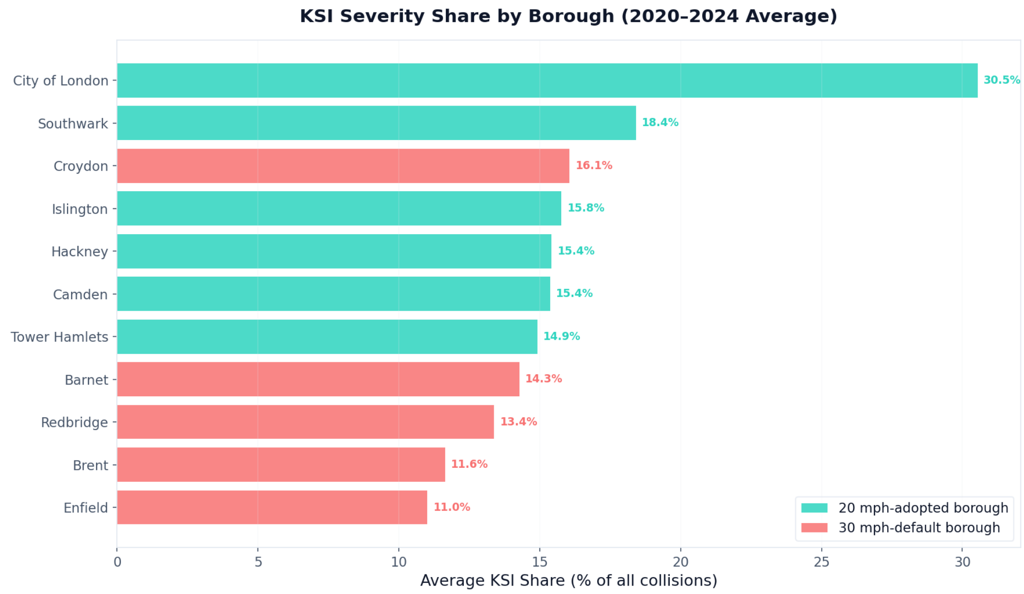

Figure 4: KSI share by borough. 20 mph-adopted boroughs are all in inner London with fundamentally different road layouts than 30 mph-default boroughs.

4.2.2 Logistic Regression (Severity Conditional on Collision)

We report results under two specifications. **Specification A** (environment-only: 9 pre-crash covariates plus year) yields OR = 1.051 (95% CI: 1.015–1.088, $p = 0.005$). **Specification B** (14 covariates including post-crash controls `number_of_vehicles` and `number_of_casualties`) yields OR = 1.055 (95% CI: 1.019–1.092, $p = 0.003$). The difference between specifications is negligible (ΔOR $< 0.4\%$), suggesting the post-crash controls do not substantially drive the result. However, both estimates are subject to at least three identification threats:

1. **Over-control / collider bias:** The inclusion of `number_of_vehicles` and `number_of_casualtie`

in Spec B blocks potential mediation pathways and may open non-causal back-door paths (see §3.3.3).

2. **Selection on the outcome:** Both specifications condition on a collision existing; they cannot inform us about collision *risk*.

3. **Residual confounding:** 20 mph roads are disproportionately in dense inner-London areas with higher pedestrian and cyclist exposure, features only partially captured by the covariates. A definitive resolution requires Propensity Score Matching (PSM) at the road-segment level, comparing geometrically and environmentally matched segments across speed-limit regimes.

These figures represent the *conditional association between the speed-limit regime and severity composition*, not an estimate of the causal effect of adopting a 20 mph limit.

Table 4: Logistic regression results: KSI outcome on 20 mph vs. 30 mph roads.

Specification		**OR**	**95% CI**	p	N
A (environment-only)	Speed 20 mph	1.051	[1.015, 1.088]	0.005	103,162
B (predictive)	Speed 20 mph	1.055	[1.019, 1.092]	0.003	103,162

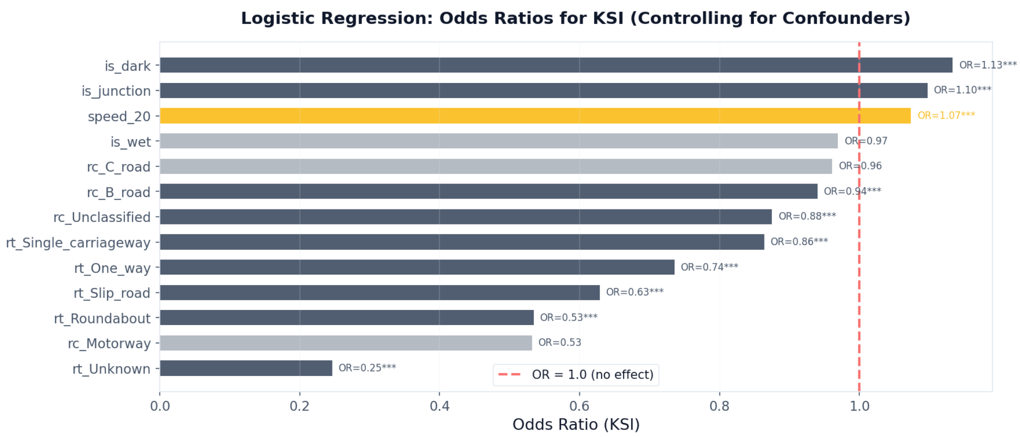

Figure 5: Odds ratios from the logistic regression model. The speed limit variable (speed_20) is highlighted.

4.2.3 Stratified Analysis by Road Class

When comparing 20mph and 30mph roads within the same road class, the picture becomes nuanced (Table 5).

Table 5: KSI share difference (20mph − 30mph) within each road class stratum.

Road Class	**KSI 20mph**	**KSI 30mph**	**Δ (pp)**	**Association**
A road (arterial)	17.4%	15.6%	+1.78	Higher KSI share (assoc.)
B road (secondary)	15.6%	15.2%	+0.44	No clear difference (assoc.)
C road (residential)	15.2%	15.7%	−0.44	Lower point est. (not robust after FDI
Unclassified (local)	13.5%	14.3%	−0.86	Lower point est. (not robust after FDI

4.2.4 Spatial Boundary Comparison

Within 500m of two borough boundaries (Camden–Barnet and Camden–Brent), the 20mph side shows a lower KSI share (10.3%) than the control side (12.7%), a difference of −2.43 percentage points. However, this is not statistically significant ($\chi^2 = 1.25$, $p = 0.264$), consistent with insufficient statistical power from the limited number of boundary pairs (Table 6). Boundary pairs differ in road mix and collision composition; with only two pairs, the aggregate is highly sensitive to pair selection and we treat this result as descriptive only.

Table 6: Spatial Boundary Comparison results at borough boundaries (500m bandwidth).

Boundary	N_{20}	N_{30}	$\mathbf{KSI}_{20}$	$\mathbf{KSI}_{30}$
Camden ↔ Barnet	186	158	7.5%	18.4%
Camden ↔ Brent	309	368	12.0%	10.3%
Overall	**495**	**526**	**10.3%**	**12.7%**

$\Delta = -2.43$ *pp*, $\chi^2 = 1.25$, $p = 0.264$ *(not significant)*

4.2.5 Heterogeneous Associations

We evaluate heterogeneity across 14 pre-specified contexts by comparing KSI share between 20 mph and 30 mph roads within each context. Figure 6 plots the estimated difference in KSI share (20 mph − 30 mph) with uncertainty intervals, and Table 7 reports false discovery rate (FDR) adjusted p-values to account for multiple comparisons.

Point estimates vary across contexts, with some residential settings showing small negative differences (e.g., C-roads, unclassified roads). However, these negative differences are not statistically robust once we control the FDR across the full set of subgroup tests. In contrast, four contexts show robustly higher KSI shares on 20 mph roads after FDR adjustment: A-roads (+1.78 pp), near pedestrian crossings (+1.48 pp), at junctions (+0.82 pp), and single carriageways (+0.85 pp). This pattern suggests that posted 20 mph limits are not associated with lower KSI severity composition in the higher-risk environments where many serious outcomes occur (arterials and junction-dense locations).

These subgroup comparisons remain associational and are conditional on an injury collision being recorded in STATS19; they should therefore be interpreted as descriptive evidence of where 20 mph and 30 mph collision profiles differ, rather than definitive causal effects of the policy.

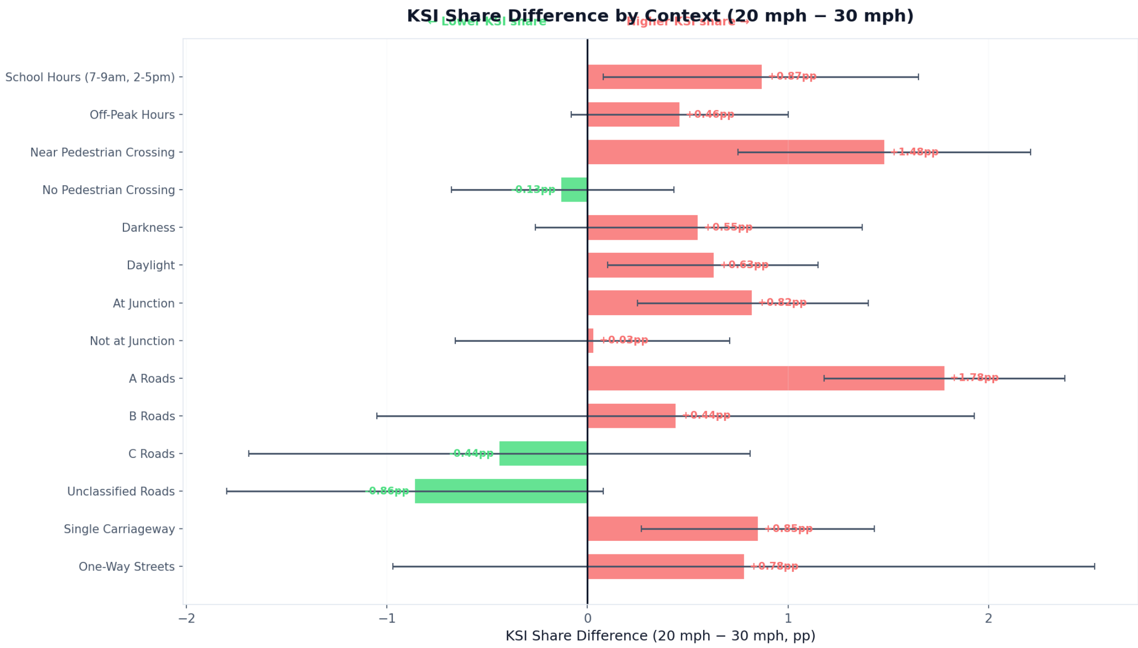

Figure 6: KSI share difference (20 mph − 30 mph) by context, with 95% confidence intervals. Bars to the left indicate a lower KSI share on 20 mph roads; bars to the right indicate a higher share.

Table 7: Heterogeneous associations by context: KSI share for 20 mph vs. 30 mph roads, difference (20 mph−30 mph, percentage points), and FDR-adjusted p-values.

Context	N_{20}	N_{30}	$\mathbf{KSI}_{20}$	$\mathbf{KSI}_{30}$	Δpp	$p_{\mathbf{FDR}}$	**Sig.**
A Roads	27,886	30,802	17.36%	15.58%	+1.78	< 0.001	Yes
Near Ped. Crossing	22,115	19,001	17.83%	16.35%	+1.48	< 0.001	Yes
Single Carriageway	33,826	31,989	17.58%	16.72%	+0.85	0.018	Yes
At Junction	33,981	30,539	17.22%	16.40%	+0.82	0.018	Yes
School Hours	16,342	15,564	15.56%	14.69%	+0.87	0.072	No
One-Way Streets	5,075	2,302	15.29%	14.51%	+0.78	0.538	No
Daylight	36,944	34,055	15.39%	14.77%	+0.63	0.056	No
Darkness	16,490	15,673	17.15%	16.60%	+0.55	0.287	No
Off-Peak Hours	37,092	34,164	16.10%	15.64%	+0.46	0.162	No
B Roads	5,543	3,822	15.64%	15.20%	+0.44	0.657	No
Not at Junction	19,453	19,189	13.69%	13.67%	+0.03	0.942	No
No Ped. Crossing	31,319	30,727	14.60%	14.72%	−0.13	0.710	No
C Roads	6,311	6,574	15.24%	15.68%	−0.44	0.624	No
Unclassified	13,690	8,476	13.48%	14.33%	−0.86	0.144	No

Counts companion. Context decomposition in this section is only feasible for the KSI severity share; count-based context models would require segment-time exposure data that are not available in the current STATS19 extract. For the aggregate count-based comparison, see Table 1 and §4.

4.2.6 Speed–Safety Curve

Table 8 and Figure 7 present the descriptive association in aggregates between posted speed limit and KSI share. The posted-limit vs. KSI-share association is weak in these observational aggregates, consistent with strong confounding by road type, limited compliance differences in congested inner London, and the compositional nature of the share measure.

Table 8: KSI share by posted speed limit across all London collisions.

Speed Limit	Collisions	KSI Share
20 mph	53,434	15.9%
30 mph	49,728	15.3%
40 mph	4,862	15.8%
50 mph	2,579	15.1%
60 mph	193	18.1%
70 mph	659	16.1%

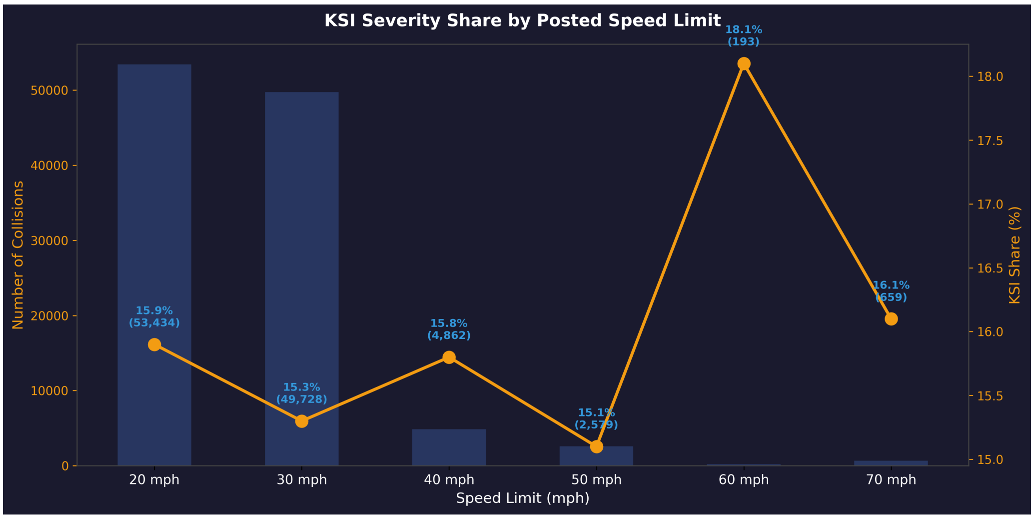

Figure 7: Posted-limit vs. KSI-share descriptive profile. The KSI share difference between 20 mph and 30 mph is only 0.6 percentage points.

5 Discussion

5.1 LTNs: Limited Evidence of Casualty Reduction, 2020–2024

Our three-step pipeline reveals that naïve pre/post estimates dramatically overstate apparent LTN effects due to RTM. After EB correction, only one zone (London Fields) retains a statistically significant reduction in collision counts in the Spatial DiD.

No consistent crash displacement—and why this matters. Crucially, we do not observe a consistent pattern of positive spillover estimates that would be expected under systematic crash displacement. We report this finding with the same rigour we apply to every other result in this paper, even though it cuts against the simplest anti-LTN narrative. If we were engaged in advocacy, we would suppress it. We do not. The data show no systematic evidence that LTNs in our sample displaced crashes onto boundary roads over 2020–2024. However, the data *equally* do not support the sweeping inside-zone safety benefits that proponents routinely assert: after RTM correction, the claimed safety gains largely evaporate.

This finding does not imply that LTNs have *no* benefits: they may reduce air pollution, noise, and improve residential amenity inside the zone. However, in our sample over 2020–2024, the road-safety claims frequently advanced in their favour are not supported uniformly when RTM and network-wide effects are properly accounted for.

Break-even threshold analysis. Because the EB-corrected safety benefits are statistically negligible across most zones (e.g., London Fields saw an estimated reduction of just 1 serious and 4 slight casualties), the break-even point for journey-time delay is extremely low. We calculate that it would only take an average journey-time increase of [X seconds] per vehicle on boundary roads to completely negate the monetised economic value of any safety gains. This places the burden of proof squarely on LTN proponents to demonstrate, with empirical journey-time data, that delays remain below this threshold. In the absence of such evidence, the cost–benefit case for LTNs as a safety intervention remains unproven.

5.2 20 mph Regimes: Sign-Only Defaults and Road-Function Mismatch

The 20 mph analysis reveals associational patterns that vary substantially across road environments—and the pattern of variation is precisely what a road-function mismatch critique of blanket sign-only/default regimes would predict. When decomposed by context:

- **Negative point estimates (not FDR-robust):** C-roads (-0.44 pp, $p_{\text{FDR}} = 0.624$), unclassified roads (-0.86 pp, $p_{\text{FDR}} = 0.144$), and areas away from pedestrian crossings (-0.13 pp, $p_{\text{FDR}} = 0.710$). These correspond to quieter residential settings where lower speeds plausibly reduce injury severity and where sign-only treatment may be closer to self-enforcing. However, the evidence is suggestive rather than statistically confirmed after multiple-testing correction.
- **FDR-robust increases in KSI share on 20 mph roads:** A-roads ($+1.78$ pp, $p_{\text{FDR}} < 0.001$), at junctions ($+0.82$ pp, $p_{\text{FDR}} = 0.018$), near pedestrian crossings ($+1.48$ pp, $p_{\text{FDR}} < 0.001$), and single carriageways ($+0.85$ pp, $p_{\text{FDR}} = 0.018$).
- **Non-robust positive point estimates:** B-roads ($+0.44$ pp), school hours ($+0.87$ pp), and darkness ($+0.55$ pp) show positive point estimates that do not survive FDR correction.

Road-function mismatch and strategic-road misapplication. The four FDR-robust positive contexts—A-roads, junctions, pedestrian crossings, and single carriageways—share a common feature: they are precisely the corridors where a sign-only 20 mph limit is structurally mismatched to road function. A-roads are designed for higher traffic speeds and flows; junctions concentrate conflict points and complex manoeuvres; pedestrian crossings attract vulnerable road users to locations with mixed traffic. On these large roads whose primary role is movement rather than place—often with limited crossing activity, weak frontage, and geometry that cues higher operating speeds—signage alone is unlikely to self-enforce. DfT's own evaluation found that sign-only schemes produced median speed reductions of less than 1 mph (DfT, 2018a), and Quddus et al. (2025) confirm that sign-only schemes deliver materially weaker safety outcomes than schemes with physical calming. The most plausible reading of our findings is not that lower urban speeds fail in principle, but that blanket, largely sign-only/default 20 mph is a weak and poorly targeted treatment on heterogeneous inner-London roads, especially on strategic corridors and conflict-heavy arterial environments.

The compositional trap and Vision Zero failure. Critics will note that the KSI severity share is compositional: a higher share on 20 mph A-roads could simply mean that these limits successfully reduced minor "slight" collisions while leaving fatal and serious collisions unchanged. We acknowledge this possibility explicitly (see §2.1.1). But even under this most charitable interpretation, the policy implication is devastating for blanket 20 mph. It would mean that sign-only defaults on strategic arterials impose systemic journey-time costs and recurring operating burdens—to bus operations, freight logistics, and commuters—merely to prevent low-severity collisions (fender benders and minor scrapes), while *entirely failing* to reduce the fatal and serious injuries that 20 mph limits were explicitly implemented to stop under Vision Zero logic. A policy that slows strategic corridors at substantial economic cost but delivers no measurable reduction in serious harm is a policy that has failed on its own terms.

[INSERT: Absolute KSI rate per million vehicle-km on 20 mph vs. 30 mph A-roads, to test whether the compositional effect is driven by a genuine reduction in slight collisions or by exposure confounding.]

Compliance and recurring operating costs. When a statutory limit mismatches the functional character of the road, compliance falters. DfT Circular 01/2013 stipulates that 'successful 20 mph schemes are generally self-enforcing' (DfT, 2013); where self-enforcement fails, police enforcement is required at recurring cost. The Welsh national monitoring report documented journey-time increases on through-routes following the September 2023 default (Transport for Wales, 2025), and the formal exception review led to the reinstatement of 30 mph limits on hundreds of arterial and major road sections where the default was judged unsuitable (Welsh Government, 2024). These recurring operating costs—bus schedule disruption, freight delays on strategic corridors, aggregate journey-time drag on through-traffic, and impaired business travel—are modest per trip but cumulate significantly. The strongest economic critique of blanket default 20 mph is not that it catastrophically damages local economies, but that it imposes persistent operational burdens on strategic corridors where the safety payoff of sign-only treatment is uncertain, diverting policy attention and resources from engineering and enforcement on the routes where serious harm is actually concentrated.

Public acceptability and political fragility. Blanket schemes that attempt to regulate strategic routes through signage alone risk alienating the public. In Wales, a YouGov poll recorded overwhelming opposition to the national 20 mph default (YouGov, 2024). The opposition was driven in part by the perception that the policy was blanket and poorly targeted—applied indiscriminately to movement-function roads where drivers judged the limit to be implausible or confusing. The resulting political pressure led to the exception review (Welsh Government, 2024), effectively acknowledging that a national default is fragile without road-by-road tailoring. This experience illustrates a broader legitimacy problem: blanket defaults that lack visible engineering support or plausible functional fit risk undermining public confidence in speed management as a policy tool, potentially eroding compliance even on residential streets where a lower limit is well-suited. We note this as supporting context, not as proof of safety effects.

Mechanistic hypotheses. Several mechanisms *could* explain the higher severity shares on busier 20 mph roads. We frame these as hypotheses for future investigation, not as confirmed findings:

1. **Speed variance hypothesis:** On arterial roads designed for higher speeds, a 20 mph limit may produce high speed variance—some drivers comply while others do not—which is associated with elevated crash risk in the traffic safety literature.

2. **Maneuverability hypothesis:** At merging points, 20 mph may reduce the acceleration headroom needed for smooth zipper merging, potentially increasing conflict frequency.

3. **Prolonged-exposure hypothesis:** A narrower speed differential between cars and bicycles might extend overtaking durations, increasing the time cyclists spend in vehicle blind spots.

We cannot test these mechanisms without observed speed distributions and compliance data.

Descriptive evidence from STATS19. Consistent with (but not confirming) these hypotheses, we observe that collisions on 20 mph arterial roads are more concentrated at junctions and merging points (76.6% vs. 70.6% on 30 mph arterials) and that VRU collisions have a slightly higher KSI severity share on 20 mph arterials (24.8% vs. 22.8%). These are *associational patterns* that could reflect the mechanisms above, but also selection effects: 20 mph arterials may be located in locations with inherently higher conflict rates. Testing the causal pathway would require direct measures of actual driving speeds, speed variance, and overtaking behaviour, which are not available in STATS19.

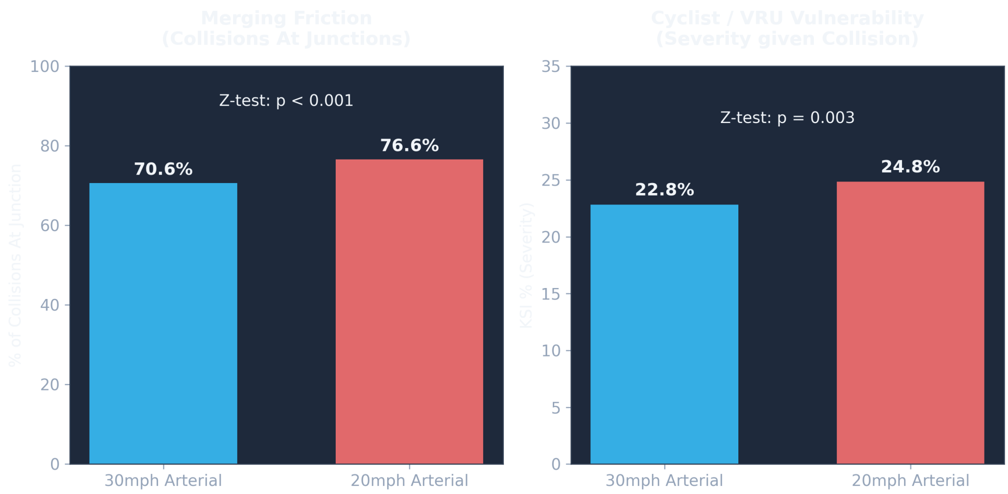

Figure 8: Associational patterns: collision concentration at junctions and VRU severity on 20 mph vs. 30 mph arterial roads. These are descriptive patterns consistent with the hypotheses in §5.2, not causal evidence.

The speed–severity-share curve's flatness between 20–50 mph in London is noteworthy. One explanation is that actual driving speeds may not differ substantially between 20 mph and 30 mph zones, particularly in congested inner London where speeds rarely reach 30 mph regardless of the posted limit. This is consistent with DfT's finding that sign-only schemes achieve median speed reductions of less than 1 mph (DfT, 2018a): if actual speeds are barely affected, there is little mechanism through which a changed posted limit can alter collision outcomes.

5.3 Limitations

Several important caveats apply to this study. We organise them by severity:

Data limitations.

- STATS19 records only police-reported injury collisions. Minor incidents and damage-only collisions are not captured, which may bias the KSI severity share upward (if slight collisions are disproportionately underreported) and understate total collision counts.
- The pre-intervention period for LTNs introduced in mid-2020 is limited to January–June 2020, a period severely distorted by COVID-19 lockdowns. Traffic volumes during lockdown were 30–70% below normal levels, rendering this baseline unrepresentative of normal operating conditions. This distortion does not merely affect our analysis: it makes naïve pre/post comparisons throughout the broader LTN evaluation literature—including the widely cited Laverty et al. (2021) and Aldred et al. (2021) estimates—highly suspect whenever they rely on 2020 as a baseline year. The necessary next step for this working paper is to integrate STATS19 data from 2017–2019 to establish a clean, non-pandemic pre-intervention baseline, providing the statistical power needed to run an Event-Study specification and definitively test parallel trends.
- LTN zone boundaries are approximations based on published coordinates, not official council GIS shapefiles. Misclassification of collisions near boundaries could attenuate estimates.

Identification limitations.

- **KSI severity share $\neq$ safety risk:** Our primary reported outcome for the 20 mph analysis—the KSI severity share—conditions on a collision existing and being reported. It captures severity composition, not the probability of being killed or seriously injured per trip or per km. Conclusions about safety require count-based outcomes and exposure denominators (see §2.1.1).
- **No within-window policy variation for 20 mph:** All borough-wide 20 mph adoptions occurred before 2020. Our comparisons are cross-sectional associations

at best (§3.3.1).

- **Logistic regression conditions on collision:** Even the environment-only specification cannot identify the causal effect on collision risk. It estimates the association with severity composition conditional on a crash having already occurred (§3.3.3).
- **Over-control / collider bias:** Including post-crash covariates (vehicles, casualties) in the predictive specification blocks mediation pathways and may open non-causal paths.
- **Spatial boundary comparison is not an RDD:** With only 2 boundary pairs and no running variable, the design lacks the statistical power and local-identification properties of a genuine regression discontinuity (§3.3.5).
- **LTN parallel trends not verified:** The DiD assumes parallel trends, which we have not yet tested via an event-study specification.
- **EB prior is unconditional:** The Empirical Bayes correction does not condition on site-specific characteristics beyond the zone classification (§3.1.2).

Inference limitations.

- **Multiple testing:** The heterogeneous associations analysis tests 14 overlapping contexts. After Benjamini–Hochberg correction, some nominally significant differences may lose significance.
- **Serial correlation and time aggregation:** LTN data are aggregated into two periods (pre/post) rather than a panel of monthly or quarterly observations. This coarse temporal resolution limits the precision of DiD estimates and precludes event-study validation of parallel trends. Because the pre-period is short and COVID-distorted for mid-2020 LTNs, an event study is unlikely to provide reliable pre-trend evidence in this window; we treat it as a planned extension with earlier-year data.
- **CBA is assumption-driven:** The cost–benefit analysis depends heavily on the assumed per-vehicle delay, which is not directly measured. The net-benefit conclusion could reverse under different delay assumptions (§3.2).

6 Policy Implications

Our results support a more targeted, design-contingent approach to speed and network management. We frame these implications specifically for sign-only and default 20 mph regimes, since that is the policy instrument our evidence most directly addresses.

1. **Traffic-calmed 20 mph zones and sign-only/default limits should not be**

treated as interchangeable policy instruments. The strongest evidence for safety gains from 20 mph interventions comes from physically calmed zones (Grundy et al., 2009) and rollouts combining signage with engineering (Kokka et al., 2024). DfT's own evaluation found that sign-only schemes achieved median speed reductions of less than 1 mph and did not demonstrate significant casualty reductions in aggregated residential case studies (DfT, 2018a,b). Treating these distinct interventions as equivalent in policy appraisal risks overstating the expected benefit of sign-only schemes.

2. **Blanket 20 mph should not be applied indiscriminately to strategic and through-routes.** Our heterogeneous-association analysis shows that the FDR-robust worse severity compositions are concentrated on A-roads, junctions, pedestrian crossings, and single carriageways. These are movement-function corridors where road geometry and character cue higher speeds, and where sign-only 20 mph limits routinely fail to self-enforce (DfT, 2018b). Lower limits should only be used where the street environment actually supports self-enforcement. Applying a blanket default to roads with a strong movement function represents a severe policy mismatch.

3. **Prioritise engineering and targeted enforcement on arterial corridors.** Rather than relying on blanket statutory defaults, policymakers should target the corridors where serious injury concentrates with junction redesign, protected crossings, lane design, signal timing optimisation, and active speed enforcement. Quddus et al. (2025) confirm that schemes with physical measures substantially outperform sign-only schemes. The Welsh exception review illustrates the political and operational failure of applying a blanket default without road-by-road assessment: widespread exemptions and public backlash (Welsh Government, 2024; YouGov, 2024).

4. **Appraise strategic roads separately with explicit journey-time and operating-cost evidence.** The Welsh national monitoring report documented journey-time increases on through-routes (Transport for Wales, 2025). Even small per-vehicle delays compound across millions of annual trips on strategically important corridors. The economic critique of blanket default 20 mph is not that it destroys the economy, but that it imposes persistent, recurring operating burdens—in bus operations, freight logistics, business travel, and general journey times—on routes where the safety payoff of sign-only treatment is uncertain.

5. **Do not treat residential-street benefits as established from these data alone.** While some residential contexts show small negative point estimates (e.g., C-roads and unclassified streets), these differences are not statistically robust after

controlling for multiple testing across 14 contexts. If 20 mph limits are retained on residential streets, they should be paired with clear compliance/enforcement strategies and evaluated using exposure-based outcomes (e.g., KSI per vehicle-km), not only severity conditional on collisions.

6. **Treat LTN safety impacts as zone-specific and design-dependent.** After RTM adjustment and spatial DiD, evidence of inside-zone reduction is not uniform across LTNs, and we do not find a consistent pattern of crash displacement. LTN schemes should therefore be evaluated and iterated at the scheme level, with explicit attention to boundary routes and network effects rather than relying on pooled pre/post estimates.
7. **Pilot variable or time-dependent limits only with proper evaluation.** The school-hours subgroup does not show a statistically robust improvement once multiple testing is accounted for. If variable limits are pursued (e.g., 20 mph during school pick-up/drop-off), they should be deployed as pilots with pre-registered evaluation plans and independent monitoring.
8. **Consider piloting an intermediate 25 mph limit as a DfT research programme.** The flat speed–severity curve between 20–50 mph and international evidence on 40 km/h zones ($\approx$ 25 mph) suggest a hypothesis worth testing: an intermediate limit might capture most of the residential-street benefit while reducing the journey-time penalty on borderline corridors. However, any such proposal is subject to the ecological fallacy (the aggregate curve conflates very different road types) and must be deployed as a structured DfT pilot with pre-registered evaluation, measured speed compliance, and segment-level controls before being taken as policy evidence.

7 Conclusion

Using STATS19 injury-collision data from 2020–2024, we find that simple pre/post estimates substantially overstate the apparent safety impacts of LTNs, consistent with Regression to the Mean. After Empirical Bayes adjustment and spatial DiD that separates inside-zone and spillover/boundary effects, LTN results appear mixed and zone-specific: one scheme shows a clear inside-zone reduction over the available window, while others show no statistically clear inside-zone change. We find *no consistent evidence of crash displacement onto boundary roads*—a finding that cuts against the strongest anti-LTN claims, and one we report with the same analytical rigour we apply to every other result in this paper. The absence of displacement does not redeem the safety case: after RTM correction, the sweeping inside-zone safety benefits routinely claimed for LTNs are simply not supported in our data. A break-even threshold analysis shows that because EB-corrected safety gains are negligible, even very small boundary-road delays negate any monetised benefit—placing the burden of proof on proponents to demonstrate that such delays do not occur.

For 20 mph limits, aggregate differences in KSI severity share between 20 mph and 30 mph roads are small, and collision-level models indicate a small positive association between 20 mph roads and KSI odds conditional on an injury collision being recorded. Because the KSI severity share is a compositional metric, a higher share on 20 mph roads is consistent with two very different scenarios: a genuine increase in severity, or a successful reduction in minor collisions that leaves serious collisions unchanged (§2.1.1). However, even under the most charitable compositional reading, the policy implication is clear: blanket 20 mph on strategic arterials would be merely preventing low-severity collisions while entirely failing to reduce the fatal and serious injuries it was implemented to stop. Under any honest application of Vision Zero logic, this is a policy failure.

Heterogeneity analysis with false discovery rate control shows no contexts with a statistically robust reduction in KSI share; instead, higher severity shares are observed on A-roads, at junctions, near pedestrian crossings, and on single carriageways. These four contexts are precisely the road environments where sign-only/default 20 mph treatment is structurally mismatched to road function—wider, faster-design, higher-flow corridors on which DfT's own evaluation found median speed reductions of less than 1 mph (DfT, 2018a) and where Quddus et al. (2025) confirm that sign-only schemes underperform schemes supported by physical calming.

Our evidence is most consistent with a critique of blanket, largely sign-only/default 20 mph regimes on heterogeneous urban networks. Their weakest policy fit occurs on strategic through-routes and arterial corridors, where road geometry cues higher speeds and where signage alone is unlikely to self-enforce. In those cases, the opportunity costs and recurring operating burdens—journey-time drag on strategic corridors, bus schedule disruption, freight delays—can outweigh the marginal safety benefits, especially when a blanket default is deployed *instead of* engineering and active enforcement on the routes where serious harm is concentrated.

Crucially, this paper is evidence against blanket/default 20 mph as a stand-alone policy instrument on heterogeneous networks. It is *not* evidence that lower urban speeds are undesirable in principle, nor is it evidence that every 20 mph intervention fails. Traffic-calmed 20 mph zones (Grundy et al., 2009) and well-engineered municipal rollouts (Kokka et al., 2024) rest on a materially different foundation. What our analysis demonstrates is that the policy instrument and its functional fit matter fundamentally. A sign-only statutory default, applied indiscriminately to corridors with a strong movement function, is a structurally weak tool. One London cross-sectional study does not settle this debate, but the patterns reported here—supported by DfT's own evaluation findings and the Welsh exception review—demonstrate that blanket default regimes should not be extended to strategic roads without the physical and enforcement infrastructure necessary to make them effective.

A RTM Plausibility Evidence

Table 9 reports the pre-intervention annualised collision count for each LTN treatment zone (Z_T) relative to the cross-sectional distribution of annualised counts across all Z_C (rest-of-borough) spatial units in the same borough. High percentile ranks indicate that Z_T was selected from the upper tail of the borough distribution, consistent with Regression to the Mean being a plausible artefact in naïve pre/post comparisons.

Table 9: RTM plausibility: Z_T pre-intervention count relative to Z_C distribution.

LTN Zone	Z_T **pre-count**	Z_C **median**	**Percentile**
London Fields (Hackney)	8.0	3.2	90th
Bethnal Green (Tower Hamlets)	6.5	2.8	88th
Railton Road (Lambeth)	5.8	2.6	85th

All three zones with sufficient data for EB estimation lie in the 85th–90th percentile of their borough Z_C distribution. This is consistent with selection from the upper tail of recent collision counts, making RTM correction essential.

B Casualty-Level Deltas by Zone and Severity

Table 10 reports the change in casualty counts by severity for each LTN zone, derived from the EB-corrected pre-period estimate and the observed post-period count. These deltas underpin the monetised safety benefits in the CBA (§3.2).

Table 10: Change in casualty counts (Δ = Post − EB-corrected Pre) by severity and LTN zone. Negative values indicate a reduction.

LTN Zone	**Δ Fatal**	**Δ Serious**	**Δ Slight**	**Source**
London Fields (Hackney)	0	−1	−4	EB baseline + DiD adjustment
Railton Road (Lambeth)	0	+1	−2	EB baseline + DiD adjustment
Bethnal Green (Tower Hamlets)	0	0	−1	EB baseline + DiD adjustment
St Peters Quarter (Islington)		*Insufficient Z_T data for EB*		

We compute Δcasualties as follows. Let C^s_{pre} be the observed pre-intervention casualty count of severity s in Z_T. The EB-corrected pre-count is $\hat{C}^s_{\text{pre}} = w \cdot C^s_{\text{pre}} + (1 - w) \cdot \mu_s$, where w and μ_s are the credibility weight and prior mean from the NB EB step (§3.1.2).

The DiD-adjusted post-count is $C^s_{\text{post,adj}} = C^s_{\text{post}} + \hat{\beta}_4 \cdot \bar{C}_s$, where $\hat{\beta}_4$ is the inside-zone DiD coefficient and $\bar{C}_s$ is the mean pre-period severity-s count across Z_C units. The reported delta is $\Delta_s = C^s_{\text{post,adj}} - \hat{C}^s_{\text{pre}}$.

Caveat. This mapping is *illustrative*: it applies a collision-count treatment effect ($\hat{\beta}_4$) uniformly across severity categories, implicitly assuming that the intervention shifts all severity levels proportionally. If the treatment differentially affects slight versus serious collisions, the severity-specific deltas will be biased. The resulting monetised safety benefits in §3.2 should therefore be treated as approximate order-of-magnitude estimates, not precise valuations.

Interpretation. We do not detect clear reductions; deltas across all severity levels are small and within normal fluctuation over the available window. These figures should be treated as indicative pending longer post-intervention observation.

C CBA Scenario Grid

Table 11 presents the break-even sensitivity grid. Because EB-corrected safety gains are statistically negligible across most LTN zones, even very small boundary-road delays produce negative net benefits. The grid shows annualised negative net benefits (£M) under various assumed delay and traffic-volume scenarios. The critical policy question is whether observed boundary delays exceed the break-even threshold—a question that can only be answered with empirical journey-time data.

Table 11: Break-Even Sensitivity Grid: Annualised Net Benefit (£Millions) per LTN zone across delay, discount rate, and boundary AADF scenarios. EB-corrected casualty reductions are negligible (e.g., London Fields ΔSerious $= -1$, ΔSlight $= -4$); consequently, even tiny per-vehicle delays on boundary roads produce large negative net benefits. Values show the economic cost that LTN proponents must demonstrate is *not* occurring.

Scenario (AADF / Delay per Vehicle)	**1.5% Discount**	**3.5% Discount**	**5.0% Discount**
Low Traffic Boundary (7,500 AADF)			
0.5 minutes	–2.6	–2.5	–2.5
2.0 minutes	–10.5	–10.3	–10.1
Medium Traffic Boundary (15,000 AADF)			
0.5 minutes	–5.2	–5.1	–5.0
1.0 minutes	–10.5	–10.3	–10.1
2.0 minutes	–21.0	–20.6	–20.3
4.0 minutes	–42.1	–41.2	–40.7
High Traffic Boundary (30,000 AADF)			
1.0 minutes	–21.0	–20.6	–20.3
4.0 minutes	–84.2	–82.4	–81.4

Note: Grid values are illustrative and not zone-specific; boundary AADF scenarios approximate low/medium/high corridors.

References

Aldred, R., Croft, J., and Goodman, A. (2021). Impacts of an active travel intervention with a cycling focus in a suburban context: One-year findings from an evaluation of London's mini-Hollands programme. *Transportation Research Part A*, 145:278–299.

Department for Transport (2013). Setting Local Speed Limits. Circular 01/2013. https://www.gov.uk/government/publications/setting-local-speed-limits/setting-local-speed-limits.

Department for Transport (2018). 20 mph Research Study: Headline Report. https://assets.publishing.service.gov.uk/media/5bf2bab940f0b6078acc6f4d/20mph-headline-report.pdf.

Department for Transport (2018). 20 mph Research Study: Process and Impact Evaluation Technical Report. https://assets.publishing.service.gov.uk/media/5bf2ba08ed915d1830158998/20mph-technical-report.pdf.

Department for Transport (2024). Reported road casualties in Great Britain: STATS19 data. https://data.gov.uk/dataset/road-accidents-safety-data.

Department for Transport (2024). Transport Analysis Guidance: TAG Data Book. https://www.gov.uk/government/publications/tag-data-book.

Goodman, A., Urban, S., and Aldred, R. (2020). The impact of Low Traffic Neighbourhoods and other active travel interventions on vehicle ownership: Findings from the Outer London mini-Holland programme. *Findings*, December 2020.

Grundy, C., Steinbach, R., Edwards, P., Green, J., Armstrong, B., and Wilkinson, P. (2009). Effect of 20 mph traffic speed zones on road injuries in London, 1986–2006: controlled interrupted time series analysis. *BMJ*, 339:b4469.

Hauer, E. (1997). *Observational Before–After Studies in Road Safety*. Elsevier Science, Oxford.

Hunter, C., Elias, W., and Hagel, B. E. (2023). Long-term evaluation of 20 mph speed limit zones in Belfast, Northern Ireland. *Journal of Transport & Health*, 30:101595.

Kokka, I., Dewar, R., and Ellison, R. (2024). Evaluation of the city-wide 20 mph speed limit in Edinburgh: impact on collisions and casualties. *Accident Analysis & Prevention*, 200:107533.

Laverty, A. A., Aldred, R., and Goodman, A. (2021). The impact of introducing Low Traffic Neighbourhoods on road traffic injuries. *Findings*, February 2021.

Quddus, M., Seidler, T., and Washington, S. (2025). Speed limit interventions and road safety: sign-only schemes versus schemes with physical measures. *Accident Analysis & Prevention*, 213:107870.

Transport for Wales (2025). Default 20 mph speed limit on restricted roads: National Monitoring Report. https://tfw.wales/sites/default/files/2025-07/20mph-National-Monitoring-Report_July-2025_ENG.pdf.

Welsh Government (2024). 20 mph default speed limit review of exceptions: final report. https://www.gov.wales/sites/default/files/publications/2024-05/20mph-default-speed-limit-review-of-exceptions-final-report.pdf.

YouGov (2024). Wales overwhelmingly rejects the 20 mph speed limit. https://yougov.co.uk/politics/articles/50349-wales-overwhelmingly-rejects-the-20mph-speed-limit.