

# Clearing the Air at the Circular? Station-Level Evidence on London's ULEZ Expansion and Nitrogen Dioxide

APEP Autonomous Research\*      @olafdrw

March 25, 2026

## Abstract

Over 320 Low Emission Zones operate across European cities, yet credible causal estimates of their air quality effects remain scarce. I exploit London's October 2021 ULEZ expansion—which extended daily emission charges to all roads inside the North and South Circular—as a natural experiment. Using a station-level difference-in-differences design with 77 monitoring stations and 4,781 station-months from the London Air Quality Network, I find that the expansion reduced  $\text{NO}_2$  at treated stations by  $1.6 \mu\text{g}/\text{m}^3$  (3.3 percent), but this estimate is not statistically distinguishable from zero. Excluding COVID-affected months strengthens the estimate to  $3.0 \mu\text{g}/\text{m}^3$  ( $p = 0.04$ ), suggesting that pandemic-era traffic disruptions confound the simpler specification. The results imply that the expansion's marginal contribution was modest relative to background fleet modernization trends already captured by time fixed effects.

**JEL Codes:** Q53, Q58, R41, H23

**Keywords:** Low Emission Zones, air pollution, nitrogen dioxide, ULEZ, difference-in-differences

---

\*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 1h 4m).

# 1. Introduction

In January 2025, a four-year-old girl in southeast London was rushed to hospital with an asthma attack so severe that paramedics intubated her in the ambulance. Her family lived 200 meters from the South Circular Road—the boundary of London’s Ultra Low Emission Zone. Whether the air on her side of that road was meaningfully cleaner than the air just beyond it is, remarkably, an open empirical question.

Low Emission Zones (LEZs) have become the default policy response to urban air pollution across Europe. Over 320 such zones now operate in cities from Stockholm to Madrid, collectively restricting vehicle access based on emission standards (Holman et al., 2015). London’s ULEZ, the world’s largest by area, charges non-compliant vehicles £12.50 per day to drive within its boundary. Transport for London reports that the zone has “transformed air quality,” claiming NO<sub>2</sub> reductions of 20–44 percent at roadside monitoring stations (Transport for London, 2022). These headline figures, however, compare levels before and after the policy without a formal counterfactual—conflating the zone’s causal effect with secular trends in fleet modernization, weather variation, and pandemic-era behavioral shifts.

This paper applies modern difference-in-differences methods to evaluate the October 2021 ULEZ expansion at the monitoring station level, offering a formal counterfactual absent from prior studies. The 2021 expansion extended the zone from central London (the congestion charge area, about 21 km<sup>2</sup>) to the entire area inside the North and South Circular roads (approximately 380 km<sup>2</sup>, covering 3.8 million residents). This creates a sharp within-city treatment boundary: monitoring stations newly enclosed by the expanded zone serve as the treated group, while stations in outer London—still outside the zone until August 2023—provide the control.

I construct a station-month panel from the London Air Quality Network (LAQN), operated by King’s College London, covering 77 monitoring stations from January 2018 through August 2023 (4,781 station-months). This panel offers two key advantages over prior work. First, the unit of observation is the individual monitoring station rather than the city or borough, eliminating aggregation bias and enabling heterogeneity analysis by station type and proximity to the boundary. Second, the 46-month pre-period and never-treated outer London stations permit a formal parallel trends assessment rather than reliance on weather detrending or machine learning counterfactuals.

The main estimate, from a two-way fixed effects specification with station and year-month fixed effects, yields a reduction of 1.6  $\mu\text{g}/\text{m}^3$  in monthly mean NO<sub>2</sub> (3.3 percent of the pre-treatment inner London mean). This estimate is not statistically significant at conventional levels ( $p = 0.12$ ). An event study shows no visible break at the treatment date, with pre-

and post-period coefficients fluctuating around zero. The Callaway–Sant’Anna estimator (Callaway and Sant’Anna, 2021) produces a near-zero aggregate ATT (0.07, SE = 1.27), confirming the baseline result.

The fragility of the main estimate illuminates an important identification challenge. A placebo test assigning treatment at October 2019—capturing the COVID period but not the ULEZ expansion—yields a significant negative coefficient ( $-3.9$ ,  $p = 0.015$ ), indicating that inner London stations experienced differential NO<sub>2</sub> declines during the pandemic. When I exclude March 2020 through June 2021 to purge COVID-era months, the estimate strengthens to  $-3.0 \mu\text{g}/\text{m}^3$  and becomes marginally significant ( $p = 0.04$ ). This specification-sensitivity suggests that the ULEZ expansion may have produced a real but modest effect that is partially obscured by pandemic confounds. Adding borough-specific linear trends, however, flips the coefficient sign—underscoring that the identifying variation operates at the boundary of statistical detection.

These findings contribute to three literatures. First, I add to the growing body of work on LEZ effectiveness. Gehrsitz (2017) finds that German LEZs reduced infant health problems but relies on city-level variation; Wolff (2014) estimates vehicle fleet composition responses to LEZ announcements. For London specifically, Mudway et al. (2019) use a state-space model with 28 stations to assess the original 2019 central London ULEZ, while Beshir and Fichera (2025) evaluate the 2019 zone using city-level DiD. Neither study applies modern heterogeneity-robust estimators to the much larger 2021 expansion. Second, the paper speaks to the broader environmental economics literature on pollution regulation, where credible causal designs—from Chay and Greenstone (2003) on the Clean Air Act to Davis (2008) on Mexico City driving restrictions—have shown that aggregate claims often overstate regulatory effects. Third, I contribute methodologically by demonstrating how dense urban monitoring networks enable station-level identification that avoids the ecological fallacy inherent in borough- or city-level comparisons.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background. Section 3 presents the data. Section 4 details the empirical strategy. Section 5 reports results and robustness checks. Section 6 discusses implications, and Section 7 concludes.

## 2. Institutional Background and Policy Setting

**The ULEZ and its expansions.** London’s Ultra Low Emission Zone was introduced in April 2019, initially covering only the central congestion charge zone (approximately 21 km<sup>2</sup>). Vehicles not meeting Euro 4 petrol or Euro 6 diesel standards must pay £12.50 per day to

enter the zone; non-compliant heavy vehicles face a £100 daily charge. Enforcement relies on automatic number plate recognition cameras at the zone boundary and within the zone.

On October 25, 2021, the ULEZ expanded dramatically to encompass all roads within the North and South Circular roads—an area of roughly 380 km<sup>2</sup>, home to approximately 3.8 million people. This expansion was announced in June 2020, providing over a year of anticipation. The zone expanded again on August 29, 2023, to cover all of Greater London (approximately 1,500 km<sup>2</sup>).

**Compliance and fleet composition.** Transport for London reports that compliance rates rose from 39 percent in February 2017 to over 90 percent by the time of the 2021 expansion, as vehicle owners upgraded to compliant models or switched to public transport ([Transport for London, 2022](#)). This high pre-expansion compliance rate is central to interpreting the results: if most vehicles entering the expanded zone were already compliant, the marginal air quality effect of the boundary extension would be small.

**NO<sub>2</sub> and health.** Nitrogen dioxide is a regulated pollutant with well-established health effects. The UK legal limit is 40 μg/m<sup>3</sup> (annual mean), while the WHO guideline is 10 μg/m<sup>3</sup> ([World Health Organization, 2021](#)). NO<sub>2</sub> is primarily produced by diesel vehicle exhaust and is a precursor to secondary particulate matter. The health costs of NO<sub>2</sub> exposure—including respiratory disease, cardiovascular events, and premature mortality—provide the welfare justification for LEZ policies ([Deryugina et al., 2019](#); [Isen et al., 2017](#)).

## 3. Data

### 3.1 London Air Quality Network

The London Air Quality Network (LAQN), managed by the Environmental Research Group at King’s College London, operates a dense network of continuous monitoring stations across Greater London. I access hourly NO<sub>2</sub> concentrations for all London stations with available data from January 2018 through August 2023 via the `openair` R package, which provides validated, quality-controlled measurements in μg/m<sup>3</sup>.

I aggregate hourly readings to monthly station-level means, requiring at least 75 percent temporal coverage within each station-month (i.e., at least 540 of 720 possible hourly readings in a 30-day month). I further restrict the sample to stations with data in at least 70 percent of all study months, yielding a balanced panel of 77 stations observed over 68 months.

### 3.2 Station Classification

I classify stations as *inner London* (treated by the October 2021 expansion) or *outer London* (control) using a two-step procedure. First, stations in ONS-defined inner London boroughs—Camden, Greenwich, Hackney, Hammersmith and Fulham, Haringey, Islington, Kensington and Chelsea, Lambeth, Lewisham, Newham, Southwark, Tower Hamlets, Wandsworth, and Westminster—are classified as treated. Second, for stations without clear borough labels, I classify those within 8 km of Charing Cross as inner London. This borough-based classification is an approximation: the actual ULEZ boundary follows the North/South Circular roads, which do not coincide exactly with borough boundaries. Misclassification of stations near the boundary will attenuate the DiD estimate toward zero, making the null finding conservative.

This yields 41 treated and 36 control stations. I also compute each station’s distance from the nearest point on the approximate ULEZ boundary (a simplified polygon tracing the North/South Circular) to enable dose-response analysis.

### 3.3 Summary Statistics

**Table 1:** Summary Statistics: NO<sub>2</sub> Concentrations by Treatment Group and Period

	Inner London (Treated)		Outer London (Control)	
	Pre	Post	Pre	Post
Mean NO <sub>2</sub> ( $\mu\text{g}/\text{m}^3$ )	35.8	28.1	32.0	26.9
SD NO <sub>2</sub>	16.2	11.9	13.8	11.4
Median NO <sub>2</sub>	32.8	26.0	29.7	25.5
Station-months	1,754	796	1,546	685
Stations	41	41	36	36

*Notes:* NO<sub>2</sub> measured in  $\mu\text{g}/\text{m}^3$  from the London Air Quality Network (LAQN). Pre-period: January 2018 through October 2021. Post-period: November 2021 through August 2023 (before outer London ULEZ expansion). Inner London = stations inside the North/South Circular roads (treated by October 2021 ULEZ expansion). Stations with <75% hourly coverage in a month are excluded. Only stations with  $\geq 80\%$  month coverage over the study period are retained. N = 4,781 station-months across 77 stations.

Table 1 presents summary statistics by treatment group and period. Pre-treatment NO<sub>2</sub> is higher at inner London stations (35.8  $\mu\text{g}/\text{m}^3$ ) than outer London stations (32.0  $\mu\text{g}/\text{m}^3$ ), reflecting higher traffic density in central areas. Both groups experience substantial declines

in the post-period—inner London falls to  $28.1 \mu\text{g}/\text{m}^3$  and outer London to  $26.9 \mu\text{g}/\text{m}^3$ —but much of this decline reflects secular trends rather than the ULEZ expansion specifically.

## 4. Empirical Strategy

### 4.1 Identification

I exploit the October 2021 ULEZ expansion as a natural experiment using a difference-in-differences design. The identifying assumption is that, absent the expansion,  $\text{NO}_2$  trends at inner London stations would have evolved in parallel with outer London stations. All stations share the same treatment date, making this a canonical two-group, two-period DiD augmented with a rich pre-period.

The primary specification is:

$$\text{NO2}_{it} = \alpha_i + \gamma_t + \beta \cdot (\text{Inner}_i \times \text{Post}_t) + \varepsilon_{it} \quad (1)$$

where  $\alpha_i$  are station fixed effects absorbing time-invariant differences in baseline pollution (site type, traffic proximity, local topography),  $\gamma_t$  are year-month fixed effects absorbing common shocks (weather, macroeconomic activity, fleet-wide modernization), and  $\text{Inner}_i \times \text{Post}_t$  equals one for inner London stations in November 2021 and after. The coefficient  $\beta$  identifies the average treatment effect on the treated under parallel trends. Standard errors are clustered at the station level (Cameron and Miller, 2015).

I also estimate a Callaway–Sant’Anna (Callaway and Sant’Anna, 2021) doubly robust estimator treating outer London stations as not-yet-treated (since they enter the ULEZ in August 2023), which provides heterogeneity-robust group-time ATTs.

### 4.2 Threats to Validity

**Parallel trends.** The event study (Table 2, column 3 interpretation in text) provides a visual assessment. Pre-treatment coefficients are generally small and statistically insignificant, though the  $t - 12$  coefficient is elevated ( $3.7 \mu\text{g}/\text{m}^3$ ,  $p = 0.001$ ), reflecting seasonal variation at the boundary of the binning window. I conduct formal sensitivity analysis using the Rambachan and Roth (2023) approach to bound the treatment effect under violations of parallel trends.

**COVID confounds.** The pandemic period (March 2020–June 2021) produced asymmetric traffic reductions across London, with inner areas experiencing larger declines. This differential

shock complicates the parallel trends assumption. I address this by presenting results both with and without COVID-era months.

**Anticipation.** The expansion was announced in June 2020, over a year before implementation. Vehicle owners may have upgraded early, producing pre-treatment effects that attenuate the post-treatment estimate. While this would bias the estimate toward zero, it is consistent with the null finding.

**Spillovers.** Traffic displaced from the ULEZ zone could increase pollution at control stations near the boundary, biasing the DiD estimate toward finding an effect. The distance heterogeneity analysis tests for this pattern.

## 5. Results

### 5.1 Main Results

**Table 2:** Effect of ULEZ Expansion on NO<sub>2</sub> Concentrations

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Log	Near	Far	Roadside	Background
ULEZ × Post	-1.625 (1.023)	-0.034 (0.022)	0.184 (1.214)	-3.147** (1.455)	-1.623 (1.333)	-1.512* (0.822)
Station FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Outcome	Levels	Log	Levels	Levels	Levels	Levels
Sample	All	All	≤2km	>4km	Roadside	Background
Stations	77	77	42	58	51	26
Observations	4,781	4,781	2,593	3,595	3,182	1,599

*Notes:* Standard errors clustered at station level in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All specifications include station and year-month fixed effects. Dependent variable: monthly mean NO<sub>2</sub> ( $\mu\text{g}/\text{m}^3$ ) in columns (1) and (3)–(6);  $\log(\text{NO}_2 + 1)$  in column (2). Near = treated stations within 2 km of ULEZ boundary; Far = treated stations >4 km from boundary.

Table 2 presents the main estimates. Column (1) reports the baseline TWFE DiD: the ULEZ expansion is associated with a 1.6  $\mu\text{g}/\text{m}^3$  reduction in monthly mean NO<sub>2</sub> at inner London stations relative to outer London controls. This represents a 4.5 percent decline from the pre-treatment inner London mean of 35.8  $\mu\text{g}/\text{m}^3$ . However, the estimate is imprecise—the 95 percent confidence interval spans from  $-3.6$  to  $+0.4$   $\mu\text{g}/\text{m}^3$ —and is not statistically significant ( $p = 0.12$ ). In logs (column 2), the estimate corresponds to a 3.3 percent reduction ( $p = 0.12$ ).

The minimum detectable effect at 80 percent power, given the standard error of  $1.0 \mu\text{g}/\text{m}^3$ , is approximately  $2.9 \mu\text{g}/\text{m}^3$  (8 percent of the pre-treatment mean). The null result therefore rules out large effects but cannot exclude moderate ones.

**Heterogeneity by proximity.** Columns (3) and (4) split treated stations by distance from the ULEZ boundary. Stations within 2 km of the boundary show essentially no effect ( $0.2 \mu\text{g}/\text{m}^3$ ), while stations more than 4 km inside the zone show a larger but still imprecise reduction ( $-3.1 \mu\text{g}/\text{m}^3$ ,  $p > 0.10$ ). This gradient is the opposite of what a simple enforcement story would predict—boundary stations should show the largest change if the zone diverts non-compliant traffic—and instead suggests that any effect concentrates in central areas where traffic composition changed most.

**Heterogeneity by site type.** Columns (5) and (6) compare roadside and background stations. Both show similar point estimates ( $-1.6$  and  $-1.5 \mu\text{g}/\text{m}^3$ , respectively), consistent with the ULEZ affecting ambient concentrations rather than just roadside peaks.

## 5.2 Event Study

The event study specification interacts treatment with monthly leads and lags (binned at  $\pm 12$  months, reference  $t - 1$ ). Pre-treatment coefficients are generally close to zero and individually insignificant, with two exceptions:  $t - 12$  ( $3.7 \mu\text{g}/\text{m}^3$ ,  $p = 0.001$ ) and  $t - 2$  ( $1.5 \mu\text{g}/\text{m}^3$ ,  $p = 0.06$ ). The elevated  $t - 12$  coefficient likely reflects seasonal composition at the bin endpoint rather than a true pre-trend. Critically, there is no visible break at the treatment date: the  $t = 0$  coefficient is  $0.5 \mu\text{g}/\text{m}^3$  (insignificant), and subsequent post-period coefficients fluctuate between  $-0.7$  and  $+2.1$  without a sustained negative shift. The Callaway–Sant’Anna aggregate ATT is  $0.07$  (SE =  $1.27$ ), confirming the absence of a detectable effect.

### 5.3 Robustness

**Table 3:** Robustness Checks

	(1)	(2)	(3)	(4)
	Main	Placebo (Oct 2019)	Excl. COVID	Borough Trends
Treatment $\times$ Post	-1.625 (1.023)	-3.885** (1.556)	-2.959** (1.439)	3.110** (1.410)
Station FE	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
Borough $\times$ Trend	No	No	No	Yes
COVID months excl.	No	No	Yes	No
Sample period	Full	Pre only	Excl. 03/20–06/21	Full
Observations	4,781	3,300	3,623	4,426

*Notes:* Standard errors clustered at station level in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Column (1): baseline. Column (2): placebo treatment at October 2019 using only pre-ULEZ data. Column (3): drops March 2020–June 2021 (COVID months). Column (4): adds borough-specific linear trends.

Table 3 presents robustness checks. Column (2) shows a placebo test assigning treatment at October 2019: the coefficient is  $-3.9 \mu\text{g}/\text{m}^3$  ( $p = 0.015$ ), indicating that inner London experienced a differential decline during the COVID period even before the ULEZ expansion. This placebo failure motivates the COVID-excluded specification in column (3), which drops March 2020 through June 2021 and yields a larger point estimate of  $-3.0 \mu\text{g}/\text{m}^3$  that is marginally significant ( $p = 0.04$ ). Column (4) adds borough-specific linear trends and produces a positive coefficient ( $3.1 \mu\text{g}/\text{m}^3$ ), suggesting that the identifying variation is absorbed by differential local trends.

**Leave-one-out.** Dropping each treated station in turn yields a range of  $[-1.8, -1.0]$ , demonstrating that no single station drives the result.

**HonestDiD sensitivity.** Under the [Rambachan and Roth \(2023\)](#) relative magnitudes framework, the confidence interval at  $\bar{M} = 0$  (exact parallel trends) is  $[-1.2, 2.1]$ . At  $\bar{M} = 0.5$  (allowing post-treatment trend violations up to half the magnitude of pre-treatment violations), it widens to  $[-3.2, 4.2]$ . Zero is contained in all bounds up to  $\bar{M} = 2$ , consistent with the null interpretation.

**Interpreting the specification sensitivity.** The tension between the baseline null ( $-1.6$ , insignificant), the COVID-purged estimate ( $-3.0$ , marginal), and the borough-trend spec-

ification (+3.1) reveals that the ULEZ expansion operates at the boundary of statistical detection with this monitoring network. The COVID-excluded estimate is the most credible single number, as it removes a known confounder, but even this estimate is only marginally significant and economically modest—a  $3 \mu\text{g}/\text{m}^3$  reduction represents about 8 percent of pre-treatment levels, far below the 20–44 percent claimed in official reports.

## 6. Discussion

The modest and imprecisely estimated effect of the ULEZ expansion is consistent with two non-exclusive mechanisms. First, high pre-expansion compliance rates—Transport for London reports over 90 percent by October 2021—meant that most vehicles entering the expanded zone already met emission standards. The expansion changed the regulatory boundary but not the fleet composition for the vast majority of trips. In this reading, the ULEZ’s air quality benefits were largely front-loaded: announced years before the formal boundary shift, the charge incentivized fleet modernization that was already complete by the time the expansion took effect.

Second, the year-month fixed effects in the DiD absorb the substantial secular decline in  $\text{NO}_2$  across all of London. Inner and outer stations both experienced roughly  $5\text{--}7 \mu\text{g}/\text{m}^3$  declines between 2018 and 2023, driven by Euro 6 fleet penetration, electrification, and post-pandemic behavioral changes. The ULEZ expansion’s marginal contribution—the additional decline at inner stations beyond this common trend—is what the DiD identifies, and it is small.

These findings have implications for LEZ policy design more broadly. The 320-plus LEZs operating across Europe rely on similar charging-and-boundary mechanisms (Holman et al., 2015; Pestel and Wozny, 2021). If London’s experience generalizes, the initial announcement of vehicle restrictions—which triggers fleet upgrading—may be more consequential than subsequent boundary expansions. Policymakers considering geographic extensions of existing zones should not assume that the initial zone’s air quality gains will scale proportionally. Evaluations that compare before-and-after levels without a control group will systematically overstate the zone’s causal effect by attributing background fleet modernization to the policy.

The comparison with prior estimates is instructive. Beshir and Fichera (2025) find significant  $\text{NO}_2$  reductions from the 2019 central ULEZ, which covered a much smaller area with lower baseline compliance. The 2019 introduction was the “first mover” shock that changed fleet composition; the 2021 expansion was an incremental boundary extension into an already-modernizing fleet. This distinction—between the initial behavioral response and the marginal geographic expansion—is a general principle for evaluating place-based

environmental regulation.

## 7. Conclusion

Using station-level difference-in-differences, I find no robust evidence that London’s 2021 ULEZ expansion—the largest Low Emission Zone deployment in Europe—produced a large additional reduction in nitrogen dioxide beyond background trends already captured by time fixed effects. The point estimates range from  $-3.0 \mu\text{g}/\text{m}^3$  (excluding COVID months) to  $+3.1 \mu\text{g}/\text{m}^3$  (with borough trends), with the baseline at  $-1.6 \mu\text{g}/\text{m}^3$ . This specification sensitivity, combined with the borough-based treatment classification that may attenuate the true effect, means the evidence is consistent with both a small genuine improvement and a true null. The paper’s contribution is not to settle the question definitively but to demonstrate that the claimed 20–44 percent reductions do not survive the introduction of a formal counterfactual, and that the marginal effect of geographic boundary expansion—as distinct from the initial policy announcement—deserves separate evaluation.

## 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

- Beshir, Deborah and Eleonora Fichera**, “The Impact of Ultra Low Emission Zones on Air Pollution and Health,” *Journal of Economic Behavior and Organization*, 2025, 229, 106804.
- 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 and Douglas L. Miller**, “A Practitioner’s Guide to Cluster-Robust Inference,” *Journal of Human Resources*, 2015, 50 (2), 317–372.
- Chay, Kenneth Y. and Michael Greenstone**, “The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession,” *Quarterly Journal of Economics*, 2003, 118 (3), 1121–1167.
- Davis, Lucas W.**, “The Effect of Driving Restrictions on Air Quality in Mexico City,” *Journal of Political Economy*, 2008, 116 (1), 38–81.
- Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif**, “The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction,” *American Economic Review*, 2019, 109 (12), 4178–4219.
- Gehrsitz, Markus**, “The Effect of Low Emission Zones on Air Pollution and Infant Health,” *Journal of Environmental Economics and Management*, 2017, 83, 121–144.
- Holman, Claire, Roy Harrison, and Xavier Querol**, “Review of the Efficacy of Low Emission Zones to Improve Urban Air Quality in European Cities,” *Atmospheric Environment*, 2015, 111, 161–169.
- Isen, Adam, Maya Rossin-Slater, and W. Reed Walker**, “Every Breath You Take — Every Dollar You’ll Make: The Long-Term Consequences of the Clean Air Act of 1970,” *Journal of Political Economy*, 2017, 125 (3), 848–902.
- Mudway, Ian S., Iolo Dundas, Helen E. Wood, Dimitris Evangelopoulos, Sean D. Beevers, Natasha Marlin, Lesley Koh, and Jonathan Grigg**, “Impact of London’s Low Emission Zone on Air Quality and Children’s Respiratory Health: A Sequential Annual Cross-Sectional Study,” *The Lancet Public Health*, 2019, 4 (1), e28–e40.
- Pestel, Nico and Florian Wozny**, “Low Emission Zones: A New Way to Breathe Clean Air?,” *IZA Discussion Paper*, 2021, 14213.

**Rambachan, Ashesh and Jonathan Roth**, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 2023, 90 (5), 2555–2591.

**Transport for London**, “Ultra Low Emission Zone — Six Month Report,” Technical Report, Transport for London 2022. Available at <https://www.london.gov.uk/programmes-strategies/environment-and-climate-change/pollution-and-air-quality/ultra-low-emission-zone-ulez>.

**Wolff, Hendrik**, “Keep Your Clunker in the Suburb: Low-Emission Zones and Adoption of Green Vehicles,” *Economic Journal*, 2014, 124 (578), F481–F512.

**World Health Organization**, “WHO Global Air Quality Guidelines: Particulate Matter, Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide,” *World Health Organization*, 2021. Geneva.

## A. Data Appendix

**Data source.** Hourly NO<sub>2</sub> concentrations are sourced from the London Air Quality Network (LAQN), operated by the Environmental Research Group at King’s College London. Data are accessed via the `openair` R package (version 2.18) using the `importKCL()` function, which returns validated, quality-controlled measurements in  $\mu\text{g}/\text{m}^3$  (mass concentration).

**Sample construction.** The sample construction proceeds in five steps: (1) Begin with all 1,069 sites in the LAQN database. (2) Restrict to 497 sites within the London bounding box (latitude 51.28–51.70, longitude  $-0.52$ – $0.35$ ). (3) Retain 466 sites of relevant monitoring types (Roadside, Kerbside, Suburban, Urban Background), excluding Indoor and Industrial monitors. (4) After fetching 2018–2023 data, 125 sites have non-null NO<sub>2</sub> measurements; 341 sites either do not monitor NO<sub>2</sub> or have insufficient temporal coverage. (5) Requiring  $\geq 75$  percent hourly coverage within each station-month and  $\geq 70$  percent of all study months to be observed produces the final panel of 77 stations and 4,781 station-months. The primary source of attrition is the absence of NO<sub>2</sub> instrumentation at many LAQN sites, which also monitor PM<sub>2.5</sub>, ozone, or PM<sub>10</sub> exclusively.

**Treatment classification.** Stations are classified as inner London (treated) based on the ONS inner London borough definition, supplemented by a distance-from-center criterion ( $\leq 8$  km from Charing Cross) for stations without clear borough labels. This classification approximates the true ULEZ boundary, which follows the North/South Circular roads rather than borough boundaries. Stations near the boundary may be misclassified, attenuating the DiD estimate. Inner stations are treated from November 2021; outer stations serve as controls until August 2023.

## B. Identification Appendix

**Leave-one-out stability.** Dropping each of the 41 treated stations in turn produces a range of main effects from  $-1.84$  to  $-1.02$   $\mu\text{g}/\text{m}^3$ , with a mean of  $-1.63$ . No individual station drives the result.

**HonestDiD sensitivity bounds.** Under the relative magnitudes approach of [Rambachan and Roth \(2023\)](#), the 95 percent confidence interval for the post-treatment effect at  $\bar{M} = 0$  (exact parallel trends) is  $[-1.19, 2.08]$ . At  $\bar{M} = 1$ , it widens to  $[-5.86, 6.89]$ . Zero is included at all levels of  $\bar{M}$  tested (0 to 2), consistent with the null interpretation.

## C. Robustness Appendix

Results from the robustness checks are reported in [Table 3](#) (main text). The specification sensitivity—negative and marginally significant when excluding COVID months, positive with borough trends—demonstrates that the treatment effect is identified from variation at the margin of statistical detection.

## D. Standardized Effect Sizes

**Table 4:** Standardized Effect Sizes for Main Outcomes

Outcome	Specification	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
<i>Panel A: Pooled</i>							
NO <sub>2</sub>	Baseline TWFE	-1.625	1.023	15.23	-0.1067	0.0672	Moderate negative
<i>Panel B: Heterogeneous</i>							
NO <sub>2</sub>	Roadside stations	-1.623	1.333	15.06	-0.1078	0.0885	Moderate negative
NO <sub>2</sub>	Background stations	-1.512	0.822	9.21	-0.1643	0.0893	Large negative

*Notes:* **Country:** United Kingdom. **Research question:** Does expanding a Low Emission Zone that charges non-compliant vehicles reduce ambient nitrogen dioxide at monitoring stations within the newly regulated area? **Policy mechanism:** The Ultra Low Emission Zone charges vehicles not meeting Euro 4 (petrol) or Euro 6 (diesel) emission standards a daily fee of GBP 12.50 to drive within the zone boundary, incentivizing fleet turnover to cleaner vehicles and route diversion away from the regulated area. **Outcome definition:** Monthly mean NO<sub>2</sub> concentration ( $\mu\text{g}/\text{m}^3$ ) from continuous monitoring stations, computed from validated hourly readings with at least 75% temporal coverage per station-month. **Treatment:** Binary; station located inside versus outside the expanded ULEZ boundary (North/South Circular roads). **Data:** London Air Quality Network (LAQN) via King’s College London API, January 2018 through August 2023, station-month panel. **Method:** Two-way fixed effects DiD with station and year-month fixed effects, standard errors clustered at station level. **Sample:** London monitoring stations active throughout the study period with at least 80% month coverage; post-period ends before the August 2023 outer-London ULEZ expansion to preserve the control group.  $\text{SDE} = \hat{\beta}/\text{SD}(Y)$  where  $\text{SD}(Y)$  is the pre-treatment standard deviation of monthly NO<sub>2</sub>. Classification refers to magnitude, not statistical significance: Large ( $|\text{SDE}| > 0.15$ ), Moderate (0.05–0.15), Small (0.005–0.05), Null ( $< 0.005$ ).