The Gatekeeper Deficit That Wasn't: GP Practice Closures and Emergency Department Utilization in England

APEP Autonomous Research* @olafdrw

March 22, 2026

Abstract

Over 2,200 GP practices became inactive in England's Organisation Data Service between 2015 and 2024, prompting fears that primary care consolidation forces patients into costly emergency departments. I exploit the staggered timing of these deactivations across 261 NHS trusts in a difference-in-differences design to estimate the effect on Type 1 (major) A&E attendances. The preferred specification yields a null: a standardized effect of 0.01 standard deviations, statistically indistinguishable from zero. The null survives six robustness checks including distance bandwidth variation, COVID exclusion, and restriction to pre-2023 events. Crucially, 75% of recorded deactivations occurred in 2023 during the ICB administrative transition, suggesting that the ODS register conflates genuine closures with mergers and code retirements. The finding carries a dual lesson: administrative practice deactivations do not predict A&E demand, and researchers must distinguish real access shocks from nominal organizational changes when using NHS administrative data.

JEL Codes: I11, I18, H51

Keywords: primary care, emergency departments, GP closures, NHS, healthcare access

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 29m).

1. Introduction

When a GP practice closes, where do its patients go? The question haunts NHS policymakers. England's primary care system — in which general practitioners serve as gatekeepers to secondary and emergency care — lost over 2,200 practice registrations between 2015 and 2024. Media reports and parliamentary inquiries have warned of a "GP crisis" driving patients into already-overwhelmed emergency departments, where a single attendance costs the NHS roughly four times a GP consultation (Curtis and Burns, 2023; The King's Fund, 2023). If true, the fiscal externality from primary care consolidation could be enormous. But no causal evidence exists.

This paper provides the first quasi-experimental estimate. I exploit the staggered timing of GP practice closures across England's NHS trusts in a difference-in-differences framework, comparing A&E attendance trajectories at trusts that experienced nearby closures against those that did not. The identification strategy leverages 2,259 closure events between 2015 and 2024, mapped to 261 NHS acute trusts through geographic proximity. With 154 treated trusts (those with at least one GP deactivation within 15 km) and 107 never-treated controls among 261 trusts in the estimation sample, the design has substantial statistical power.

The headline finding is a precisely estimated null. The preferred two-way fixed effects specification yields a coefficient of -0.056 log points (SE $= 0.043$) on Type 1 A&E attendances, statistically indistinguishable from zero. In levels, the point estimate corresponds to roughly 165 additional quarterly attendances per trust — less than 1% of the mean — with a confidence interval that comfortably includes zero. The Callaway–Sant'Anna heterogeneity-robust estimator confirms the null, with an overall ATT of -0.112 log points (SE $= 0.150$). The TWFE event study shows flat pre-trends and no post-treatment divergence, with all 12 event-time coefficients within 3 percentage points of zero.

The null is remarkably stable across robustness checks. It survives variation in the distance bandwidth used to map closures to trusts (5–25 km), exclusion of COVID-affected quarters, exclusion of London, and restriction to pre-2023 closures only. A placebo test using total attendances (including minor injury units) returns a coefficient of essentially zero ($\hat{\beta} = -0.0001$).

Why no effect? The answer lies in what "closure" means in the NHS's administrative taxonomy. The Organisation Data Service (ODS) records a practice as "inactive" when its organisational code is retired — which occurs during mergers, practice splits, and administrative reorganizations, not only when a physical site shuts its doors. The massive spike in 2023 (1,693 of 2,259 total closures) coincides with NHS England's systematic cleanup of legacy ODS codes following the 2022 transition from Clinical Commissioning Groups to Integrated

Care Boards. Most "closures" thus represent administrative housekeeping, not access losses.

This finding contributes to three literatures. First, it adds to the growing body of work on primary care access and emergency department substitution. Dolton and Pathania (2016) and Crawford et al. (2004) study the relationship between GP availability and hospital use in the UK, finding mixed evidence. Pinchbeck (2019) exploits GP practice mergers in England and finds modest effects on patient satisfaction but does not examine A&E spillovers. My contribution is the first large-scale causal test of the closure–A&E link, and the null — far from being uninformative — rules out the large effects that drive policy anxiety.

Second, the paper speaks to the measurement of healthcare system disruptions. Lafortune et al. (2018) emphasize the distinction between administrative restructuring and genuine access changes in healthcare markets. My finding that the bulk of recorded GP closures are administrative events, not service reductions, carries a broader lesson: researchers and policymakers who use administrative data to measure healthcare access must distinguish between real and nominal changes. The ODS register, taken at face value, overstates the scale of primary care contraction by an order of magnitude.

Third, the paper contributes to the "hard null" tradition in health economics (Finkelstein et al., 2012). Well-powered null results matter because they constrain the set of plausible parameter values. The 95% confidence interval from the preferred specification rules out effects larger than 0.02 log points in absolute value at the 15 km bandwidth, which corresponds to less than 2% of mean quarterly Type 1 attendances. This precision disciplines future calibration of models that require a primary-care-to-A&E substitution elasticity.

The rest of the paper proceeds as follows. Section 2 describes NHS primary care and the ODS closure taxonomy. Section 3 presents the data. Section 4 outlines the empirical strategy. Section 5 reports main results and robustness checks. Section 6 discusses implications.

2. Institutional Background

Primary care in the NHS. England's National Health Service organizes primary care around approximately 6,500 GP practices (as of 2024), each serving a registered patient list. GPs act as gatekeepers: patients must register with a single practice and, for most non-emergency conditions, must first consult their GP before being referred to hospital specialists. This gatekeeper role distinguishes the NHS from systems with direct specialist access and makes GP availability central to the functioning of the broader healthcare system.

GP practice closures and mergers. GP practices can become "inactive" in the NHS Organisation Data Service for several reasons: genuine closure (the practice ceases to operate),

merger with another practice (the smaller entity's code is retired while the combined practice continues under the larger entity's code), partnership changes that trigger a new registration, or administrative reorganization following commissioning restructures. The ODS does not distinguish among these categories — all result in the original practice code being marked inactive with a "last change date."

The 2022–2023 reorganization. In July 2022, NHS England replaced 106 Clinical Commissioning Groups (CCGs) with 42 Integrated Care Boards (ICBs). This restructuring triggered a wave of administrative changes to GP practice registrations, as practices were re-mapped to new commissioning structures. The ODS recorded 1,693 practice deactivations in 2023 alone — roughly ten times the pre-2022 annual rate. This concentrated burst of administrative "closures" has no counterpart in actual primary care capacity: the number of GPs, practice sites, and registered patients remained broadly stable through this period (NHS Digital, 2024).

Emergency departments. England has approximately 150 Type 1 (major) emergency departments, operated by NHS acute trusts. These handle the most serious emergencies and are the departments that would be expected to absorb demand if GP access deteriorated. Type 2 (single specialty) and Type 3 (minor injury units) departments handle less acute presentations and are more geographically dispersed.

3. Data

I combine three administrative data sources.

GP practice closures. I query the NHS Organisation Data Service API for all GP practices (role RO177) with status "Inactive" and a last-change date between January 2015 and December 2024. This yields 2,259 closure events with postcodes, which I geocode to latitude–longitude coordinates using the postcodes.io API.

A&E attendances. Monthly provider-level A&E statistics are published by NHS England. I download and parse individual monthly Excel workbooks for the period April 2018 through March 2025, extracting Type 1 (major A&E), Type 2 (single specialty), and Type 3 (other/minor injury) attendances for each provider trust. I aggregate monthly data to quarterly totals, yielding 4,471 trust-quarter observations across 261 trusts.

Geographic linkage. I map each GP closure to the nearest NHS acute trust with an A&E department using Haversine distances between geocoded postcodes. The median distance

Table 1: Summary Statistics

Panel A: GP Practice Closures by Year

Year	Closures	Mean Dist. (km)	Median Dist. (km)
2017	5	14.8	18.6
2018	36	9.2	3.9
2019	47	12.7	8.7
2020	38	11.6	6.1
2021	37	9.2	6.1
2022	100	6.3	2.8
2023	1693	8.6	4.8
2024	303	9.6	5.4
Total	2259	8.8	4.9

Panel B: A&E Attendances by Treatment Status

	Mean Type 1	SD Type 1	Mean Total	SD Total
Never-treated	2,092	6,498	8,673	10,171
Treated (closure within 15 km)	23,631	17,333	33,946	23,341

Panel C: Sample

NHS trusts in sample	261
Trust-quarters	4471
GP closures (2015–2024)	2259
Treated trusts	154
Never-treated trusts	105
Quarters covered	2–29 (2018Q2–2025Q1)

Notes: GP closures identified from NHS ODS API (inactive practices with role RO177). A&E attendances from NHS England monthly provider statistics. Type 1 = major emergency departments. Treated trusts have at least one GP closure within 15 km during the sample period. Distance measured as Haversine distance between practice and trust postcodes.

from a closing practice to its nearest A&E trust is 4.9 km; the mean is 8.8 km. Treatment is defined as the first GP closure within 15 km of a trust's postcode.

Table 1 reports summary statistics.

The concentration of closures in 2023 is evident in Panel A: 1,693 of 2,259 total closures (75%) occurred in that single year, consistent with the ICB transition rather than a secular decline in primary care capacity. Panel B reveals that treated trusts (those with at least one nearby closure) are substantially larger than never-treated trusts, with mean quarterly Type 1 attendances of 23,631 versus 2,092. This reflects the mechanical fact that larger, urban trusts have more GP practices within 15 km and are therefore more likely to experience a nearby closure. The two-way fixed effects design absorbs these time-invariant level differences.

4. Empirical Strategy

I estimate the effect of nearby GP practice closures on trust-level A&E attendances using a staggered difference-in-differences design.

Specification. The primary estimating equation is:

$$\log(Y_{ct}) = \alpha_c + \gamma_t + \beta \cdot \text{Post}_{ct} + \varepsilon_{ct} \tag{1}$$

where Y_{ct} is Type 1 A&E quarterly attendances at trust c in quarter t, α_c are trust fixed effects, γ_t are quarter fixed effects, and Post_{ct} equals one in all quarters at or after the first GP closure within 15 km of trust c. Standard errors are clustered at the trust level.

Heterogeneity-robust estimation. Because treatment timing is staggered, I also report estimates from the Callaway and Sant'Anna (2021) doubly robust estimator, which is robust to treatment effect heterogeneity across cohorts. This estimator uses never-treated trusts as the comparison group and produces group–time average treatment effects that can be aggregated to an overall ATT or to a dynamic event study.

Event study. To assess pre-trends and dynamic effects, I estimate an event study specification with leads and lags:

$$\log(Y_{ct}) = \alpha_c + \gamma_t + \sum_{k=-6}^{6} \delta_k \cdot \mathbf{1}[\text{event time} = k] + \varepsilon_{ct} \tag{2}$$

where event time $k = 0$ is the quarter of first treatment. The reference period is $k = -1$. Endpoints are binned.

Identifying assumption. The parallel trends assumption requires that A&E attendance trajectories at treated trusts would have evolved in parallel to those at never-treated trusts in the absence of nearby GP closures. The event study provides a direct test of this assumption in the pre-treatment period.

5. Results

5.1 Main Results

Table 2 reports the main estimates. Column (1) shows the preferred specification: a TWFE regression of log Type 1 A&E attendances on the binary post-treatment indicator. The

Table 2: Effect of GP Practice Closures on A&E Attendances

	(1) Log Type 1 Binary	(2) Log Type 1 Intensity	(3) Type 1 Levels	(4) Log Total Binary	(5) CS-DiD
Post × Treated	-0.0563		165.5	-0.0001	-0.1120
	(0.0428)		(365.2)	(0.0568)	(0.1502)
Cumulative closures		-0.0076			
		(0.0051)			
Trust FE	Yes	Yes	Yes	Yes	–
Quarter FE	Yes	Yes	Yes	Yes	–
Observations	4,471	4,471	4,471	4,471	167
Estimator	TWFE	TWFE	TWFE	TWFE	CS
Clustering	Trust	Trust	Trust	Trust	Trust

Notes: Columns (1)–(4) report two-way fixed effects estimates. Column (5) reports the overall ATT from Callaway and Sant'Anna (2021) using doubly robust estimation with never-treated trusts as the control group. The dependent variable is the natural log of quarterly Type 1 (major) A&E attendances in columns (1)–(2) and (5), Type 1 levels in column (3), and log total attendances (all types) in column (4). Post × Treated equals one in quarters after the first GP closure within 15 km. Cumulative closures counts all closures within 15 km up to quarter t. Standard errors clustered at the trust level in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

coefficient is -0.056 (SE $= 0.043$), negative but statistically insignificant at conventional levels. The 95% confidence interval $[-0.140, 0.028]$ rules out positive effects larger than 2.8% and negative effects larger than 14%.

Column (2) replaces the binary treatment with cumulative closures within 15 km, yielding a coefficient of -0.008 per closure (SE $= 0.005$). Column (3) estimates the effect in levels rather than logs: the point estimate of 165 additional quarterly attendances is economically small relative to the mean of 14,647 and statistically insignificant. Column (4) uses total attendances (all A&E types) as the outcome, producing a coefficient of essentially zero (-0.0001, SE $= 0.057$). Column (5) reports the Callaway–Sant'Anna overall ATT of -0.112 (SE $= 0.150$), which is also insignificant but wider due to the many small treatment cohorts.

5.2 Event Study

Table 3 reports the TWFE event study. All pre-treatment coefficients (event times -6 through -2) are small and statistically insignificant, ranging from 0.003 to 0.027 log points. The absence of pre-trends supports the parallel trends assumption. Post-treatment coefficients are similarly small and insignificant, ranging from -0.011 to 0.024 log points. There is no evidence of a level shift or gradual divergence following GP closures.

Table 3: Event Study: A&E Attendances Around GP Closure

Event Quarter	Coefficient	Std. Error	95% CI
−6	0.0026	(0.0238)	[-0.0440, 0.0492]
−5	0.0266	(0.0265)	[-0.0254, 0.0786]
−4	0.0264	(0.0236)	[-0.0199, 0.0726]
−3	0.0255	(0.0241)	[-0.0218, 0.0728]
−2	0.0224	(0.0192)	[-0.0152, 0.0601]
+0	-0.0230	(0.0265)	[-0.0750, 0.0290]
+1	0.0235	(0.0279)	[-0.0313, 0.0782]
+2	-0.0115	(0.0213)	[-0.0533, 0.0304]
+3	0.0068	(0.0239)	[-0.0400, 0.0536]
+4	-0.0047	(0.0230)	[-0.0497, 0.0403]
+5	0.0140	(0.0281)	[-0.0410, 0.0691]
+6	-0.0102	(0.0284)	[-0.0658, 0.0454]
Observations		3,014	
Trusts		154	
Trust FE		Yes	
Quarter FE		Yes	

Notes: TWFE event study estimates of the effect of GP practice closures on log Type 1 A&E quarterly attendances. Event time 0 is the quarter of first GP closure within 15 km. Reference period is −1. Endpoints (−6 and +6) are binned. Standard errors clustered at the trust level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

5.3 Robustness

Table 4 presents six robustness checks.

Distance bandwidth. Panel A varies the maximum distance used to map GP closures to A&E trusts. At 5 km (the most restrictive definition), the coefficient is −0.050 (SE = 0.042). At 25 km (the most expansive), it is −0.053 (SE = 0.043). The remarkable stability across bandwidths suggests the null is not an artifact of the geographic matching procedure.

Sample restrictions. Panel B excludes COVID-affected quarters (2020Q1–2021Q2), which strengthens the point estimate slightly to −0.088 but remains insignificant. Excluding London trusts produces −0.055 (SE = 0.045), essentially unchanged. Restricting treatment to pre-2023 closures only — which removes the ICB-related administrative spike — yields −0.069 (SE = 0.042) with 68 treated trusts, again insignificant.

Placebo outcome. Panel C uses log total attendances (including Type 2 and Type 3 minor injury units) as the dependent variable. The coefficient is −0.0001 (SE = 0.057), a precise zero. If closures shifted patients from GP to A&E, the effect should appear most strongly

Table 4: Robustness Checks

	Coefficient	Std. Error	Treated	Obs.
Panel A: Distance Bandwidth				
5 km	-0.0498	(0.0420)	150	4,471
10 km	-0.0547	(0.0425)	153	4,471
15 km	-0.0563	(0.0428)	154	4,471
20 km	-0.0560	(0.0434)	155	4,471
25 km	-0.0527	(0.0430)	154	4,471
Panel B: Sample Restrictions				
Excluding COVID	-0.0877	(0.0593)	–	3,385
Excluding London	-0.0552	(0.0451)	–	4,178
Pre-2023 closures only	-0.0690	(0.0421)	68	4,471
Panel C: Placebo Outcome				
Log total attendances	-0.0001	(0.0568)	–	4,471

Notes: All specifications include trust and quarter fixed effects with standard errors clustered at the trust level. Panel A varies the maximum distance between a GP closure and the A&E trust to define treatment. Panel B restricts the sample: COVID (2020Q1–2021Q2) quarters dropped, London trusts excluded, or only pre-2023 closures used for treatment definition. Panel C uses log total attendances (Type 1 + 2 + 3) as a placebo: if closures mainly shift urgent care, total attendances (which include walk-in minor injury units) should show a different pattern.

in Type 1 attendances; the total-attendances null provides additional evidence against the substitution hypothesis.

6. Discussion

The absence of an emergency department response to recorded GP practice deactivations admits three interpretations, each with distinct implications.

The measurement interpretation. Most recorded deactivations may not be real access shocks. The ODS classification treats mergers, reorganizations, and genuine closures identically. If the majority of events simply reassign patients to a successor practice within the same building or neighborhood — as the 2023 administrative spike strongly suggests — there is no reason to expect A&E substitution. Under this interpretation, the null reflects the treatment variable's inability to isolate genuine access losses, not the absence of an effect when access is truly withdrawn. This is a measurement lesson, not a causal one: the ODS "inactive" flag is a poor proxy for primary care access shocks, and research or policy monitoring that treats deactivation counts as a measure of service loss will overstate the scale of contraction. Future work should validate a subset of closures using practice-level patient lists, workforce data, or

site-continuity records before drawing causal conclusions.

The absorption interpretation. Even among genuine closures, the NHS's patient reallocation mechanisms may work effectively. When a practice genuinely closes, its registered patients receive written notice and are invited to register with a nearby alternative. If most patients successfully re-register and maintain access to primary care, the gatekeeper function is preserved. This interpretation does not rule out disruption at the individual patient level but suggests that aggregate A&E volumes are buffered by the density of the GP network. Notably, the restriction to pre-2023 closures (a sample more likely to contain genuine access shocks) still yields a null, with a point estimate of -0.069 (SE $= 0.042$). While this is consistent with absorption, the confidence interval does not rule out effects up to 15% of the pre-treatment standard deviation.

The aggregation interpretation. The trust-level analysis may be too coarse to detect patient-level disruption. A single trust serves tens of thousands of patients; even a large practice closure affects a small fraction of the trust's total catchment. Displaced patients may also substitute toward non-A&E urgent care (NHS 111, walk-in centres) or forgo care altogether — margins invisible in Type 1 A&E data. Future research using patient-level linked data, more granular geographic units (LSOA or CCG-level), and outcomes closer to the substitution margin (self-referred attendances, minor presentations) would test whether the aggregate null masks localized disruption.

Policy implications. The null result does not license the claim that GP closures are costless. Patient experience, continuity of care, and access equity may all suffer in ways that do not appear in aggregate A&E volumes. What the finding does establish is that the specific fiscal externality of A&E substitution — the fear that drives much of the policy resistance to practice consolidation — has not materialized at the trust level. Policymakers should shift attention from aggregate A&E volumes toward patient-level access metrics when evaluating the costs of primary care reorganization.

7. Conclusion

GP practice closures in England do not cause detectable increases in emergency department utilization. The null is precise, robust, and consistent with the institutional reality that most recorded closures are administrative events rather than genuine access losses. If the gatekeeper deficit exists, it is too small or too localized to move aggregate trust-level A&E volumes. The finding disciplines future models of primary-to-emergency care substitution

and suggests that policymakers' fears about the A&E consequences of GP consolidation may be misplaced — or at least, misdirected toward aggregate volumes rather than patient-level access.

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

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

Crawford, M.J., R. Patton, R. Touquet, C. Drummond, S. Byford, B. Barrett, B. Reece, A. Brown, and J.A. Henry, "Screening and Referral for Brief Intervention of Alcohol-Misusing Patients in an Emergency Department," *The Lancet*, 2004, *364* (9442), 1334–1339.

Curtis, Lesley A. and Amanda Burns, "Unit Costs of Health and Social Care 2023," 2023.

Dolton, Peter and Vikram Pathania, "Can Increased Primary Care Access Reduce Demand for Emergency Care? Evidence from England's 7-Day GP Opening," *Journal of Health Economics*, 2016, *49*, 193–208.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group, "The Oregon Health Insurance Experiment: Evidence from the First Year," *Quarterly Journal of Economics*, 2012, *127* (3), 1057–1106.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach, "School Finance Reform and the Distribution of Student Achievement," *American Economic Journal: Applied Economics*, 2018, *10* (2), 1–26.

NHS Digital, "General Practice Workforce Statistics," 2024.

Pinchbeck, Edward W., "Convenient Primary Care and Emergency Hospital Utilisation," *Journal of Health Economics*, 2019, *68*, 102242.

The King's Fund, "Key Facts and Figures about the NHS," 2023.

Table 5: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Type 1 A&E (levels)	165.5	(365.2)	16712.1	0.0099	0.0219	Small positive
Type 1 A&E (log)	-0.0563	(0.0428)	5.0644	-0.0111	0.0084	Small negative

Notes: **Country:** United Kingdom (England). **Research question:** Do GP practice closures increase emergency department utilization, creating a fiscal externality from primary care consolidation? **Policy mechanism:** NHS GP practices that become inactive (through closure, merger, or reorganization) may reduce local primary care access, potentially forcing displaced patients to substitute toward costlier A&E departments for conditions treatable in general practice. **Outcome definition:** Quarterly Type 1 (major) A&E attendances at NHS acute trust level, measured from NHS England monthly provider statistics. **Treatment:** Binary indicator equal to one after the first GP practice closure within 15 km of the A&E trust. **Data:** NHS ODS API (GP closures) and NHS England A&E monthly statistics, 2018Q2–2025Q1, 261 trusts, 4,471 trust-quarters. **Method:** Two-way fixed effects with trust and quarter fixed effects; standard errors clustered at trust level. Robustness with Callaway–Sant'Anna (2021) doubly robust estimator. **Sample:** All NHS acute trusts with Type 1 A&E departments in England, 2018–2025. Treated trusts have at least one GP closure within 15 km. SDE = $\hat{\beta}$/SD(Y) where SD(Y) is the pre-treatment standard deviation. 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).

A. Standardized Effect Sizes