The Charge That Missed the Beach: Plastic Bag Pricing and the Pollution Gap

APEP Autonomous Research* @olafdrw

March 26, 2026

Abstract

Over 100 countries have adopted plastic bag charges, and official statistics report dramatic reductions in bags sold. But does less consumption mean less pollution? Using the OSPAR Beach Litter Database—standardized 100-meter transect surveys at 47 UK monitoring beaches from 2001 to 2020—I exploit the staggered introduction of mandatory 5p bag charges across Wales (2011), Northern Ireland (2013), Scotland (2014), and England (2015) in a Callaway-Sant'Anna difference-in-differences design. I find no statistically significant reduction in carrier bag litter items, even as total beach litter declined. The bag share of total litter *increased* by 1.5 percentage points ($p < 0.05$). These results suggest a *pollution gap*: the retail channel targeted by bag charges accounts for a small share of environmental bag pollution, which originates disproportionately from sources the charge does not reach.

JEL Codes: Q53, Q58, H23, L51

Keywords: plastic bag charge, Pigouvian tax, marine litter, environmental policy, difference-in-differences

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 56m).

1. Introduction

A five-pence coin weighs 3.25 grams—roughly the same as the single-use plastic bag it is meant to eliminate. Since Ireland pioneered the plastic bag levy in 2002, over 100 countries have followed, making the carrier bag charge the world's most widely adopted environmental pricing instrument (Nielsen et al., 2019). England's Department for Environment, Food and Rural Affairs reports that major retailers sold 98% fewer single-use bags after the charge was introduced (Department for Environment, Food and Rural Affairs, 2020). The policy is celebrated as a behavioral nudge success story (Rivers et al., 2017; Convery et al., 2007).

Yet a puzzling question remains: does selling fewer bags mean finding fewer on the beach? The entire evidence base for bag charges rests on self-reported sales data from retailers or survey-based estimates of consumer behavior (Dikgang et al., 2012; Poortinga et al., 2013; Jakovcevic et al., 2014). No published study has tested whether bag charges reduce actual, measured environmental pollution. This paper fills that gap.

I exploit the staggered introduction of mandatory 5p carrier bag charges across the four UK nations—Wales in October 2011, Northern Ireland in April 2013, Scotland in October 2014, and England in October 2015—as a natural experiment. The outcome is not bags sold or bags self-reported, but bags *found*: carrier bag litter counted in standardized 100-meter beach transect surveys conducted by the OSPAR Commission, the intergovernmental body responsible for protecting the North-East Atlantic marine environment. These surveys have monitored approximately 47 UK beaches since 2001, recording counts of 112 predefined litter categories using a consistent protocol across all nations.

The identification strategy is a staggered difference-in-differences design. Because each UK nation adopted its charge at a different date, I can compare beaches in newly treated nations against beaches in not-yet-treated nations, controlling for beach-level and year-level fixed effects. I implement the Callaway and Sant'Anna (2021) estimator to address the well-known bias in two-way fixed effects with staggered treatment adoption (Goodman-Bacon, 2021; Sun and Abraham, 2021).

The central finding is a null: carrier bag litter does not decline significantly following the introduction of the charge. The Callaway-Sant'Anna ATT estimate for log bag items is 0.480 (SE = 0.384), statistically indistinguishable from zero and, if anything, positive. Meanwhile, total beach litter declined—driven by reductions in non-bag items—so the *share* of bags in total litter actually increased by 1.5 percentage points ($p < 0.05$). The result is robust to alternative functional forms (inverse hyperbolic sine, levels), exclusion of COVID-affected 2020 data, and leave-one-cohort-out tests dropping individual nations.

I term this disconnect the *pollution gap*: the gap between the consumption channel (retail bag sales, which the charge successfully reduces) and the pollution channel (bags deposited on beaches, which remain unaffected). The gap arises because most bags found in beach litter surveys originate from sources the retail charge does not reach—wind-blown litter from open bins, fly-tipping, commercial and industrial waste, fishing activity, and bags already in the marine environment before the charge took effect (OSPAR Commission, 2020). The retail cash register, it turns out, is a poor proxy for the beach.

This paper contributes to three literatures. First, it provides the only causal estimate of a bag charge's effect on measured environmental outcomes rather than consumption proxies, complementing the large descriptive literature on bag charge effects (Convery et al., 2007; Dikgang et al., 2012; Rivers et al., 2017; Taylor, 2020). Second, it speaks to the broader literature on Pigouvian instrument design (Pigou, 1920; Fullerton et al., 2002), demonstrating that taxing the consumption margin does not guarantee pollution reduction when the leakage channel is large. Third, it contributes to the growing literature using environmental monitoring data for causal inference (Currie et al., 2014; Deryugina et al., 2019; Marcus, 2021), showing how standardized beach surveys can serve as outcome data for policy evaluation.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background of UK bag charges. Section 3 presents the data. Section 4 outlines the empirical strategy. Section 5 reports results. Section 6 discusses implications.

2. Institutional Background

The staggered rollout.. The UK's devolved governance structure created a natural experiment. Wales was first, introducing a mandatory 5p charge on single-use carrier bags on 1 October 2011 under the Single Use Carrier Bags Charge (Wales) Regulations 2010. Northern Ireland followed on 8 April 2013 (Carrier Bag Levy Regulations 2013). Scotland adopted its charge on 20 October 2014 (Single Use Carrier Bags Charge Regulations 2014). England was last, on 5 October 2015, initially covering only large retailers (250+ employees) before extending to all retailers at 10p on 21 May 2021.

Design features.. All four charges share key features: a uniform 5p levy at point of sale, mandatory for single-use plastic bags, with limited exemptions for bags of life (unpackaged food, prescription medicines, raw meat). The proceeds in England go to charitable causes chosen by retailers; in Wales, retailers keep the proceeds but are expected to donate to good causes; Scotland and Northern Ireland channel proceeds to environmental programs. The charge is highly salient at checkout and has been shown to reduce stated bag use by 70–90%

in consumer surveys (Poortinga et al., 2013; Thomas et al., 2019).

The consumption-side success.. DEFRA's mandatory reporting from English retailers confirms dramatic consumption reductions. In the first full year (2016–17), major retailers reported 83% fewer bags than their voluntary baseline (Department for Environment, Food and Rural Affairs, 2020). By 2019–20, bags issued had fallen 95%. Wales and Scotland reported similar magnitudes. The charge is thus effective at reducing bag *sales*. Whether it reduces bag *pollution* is the question this paper addresses.

3. Data

OSPAR Beach Litter Database.. The primary outcome data come from the OSPAR Commission's Beach Litter Monitoring Programme, which has operated continuously since 2001 across the North-East Atlantic region. Trained surveyors walk a fixed 100-meter stretch of shoreline at designated monitoring sites and count every visible litter item according to a standardized protocol of 112 predefined categories (OSPAR Commission, 2010). Categories include carrier bags (code 2), small bags (code 3), and bag fragments (code 112), as well as plastic bottles, food packaging, fishing gear, and sanitary items. Surveys are conducted 2–4 times per year at each site, with results submitted to the OSPAR secretariat.

I use data from all 47 UK OSPAR monitoring beaches spanning 2001–2020, comprising 1,243 individual surveys. These are aggregated to the beach-year level, yielding 407 observations. Table 1 presents summary statistics. The median beach-year records approximately 10 bag items and 500 total litter items per 100-meter transect.

Nation assignment.. Each beach is assigned to its UK nation (England, Scotland, Wales, or Northern Ireland) based on geographic coordinates. The sample includes 20 English, 6 Scottish, 6 Welsh, and 15 Northern Irish monitoring sites. Channel Islands beaches are excluded as the bag charge does not apply there.

4. Empirical Strategy

4.1 Identification

I estimate the causal effect of the bag charge using the staggered timing of adoption across four UK nations. The identifying assumption is that, absent the charge, carrier bag litter at beaches in treated nations would have evolved on a parallel trajectory to beaches in not-yet-treated nations. This assumption is supported by the common OSPAR monitoring

Table 1: Summary Statistics: Beach Litter by Nation and Treatment Status

	Bag Items Mean	Bag Items SD	Total Litter Mean	Total Litter SD	Bag Share	N
Wales (pre-charge)	10.4	15.1	438	348	0.026	39
Wales (post-charge)	4.6	3.6	296	266	0.022	26
Northern Ireland (pre-charge)	8.3	8.0	382	307	0.028	14
Northern Ireland (post-charge)	10.5	10.8	430	376	0.037	86
Scotland (pre-charge)	15.5	15.9	573	479	0.046	56
Scotland (post-charge)	21.3	26.2	966	992	0.018	21
England (pre-charge)	12.4	15.1	779	842	0.021	128
England (post-charge)	14.3	23.8	1129	1434	0.016	37
Full sample	12.2	15.8	641	768	0.028	407

Notes: Beach-year level observations from OSPAR 100m transect surveys, 2001–2020. Bag items include carrier bags, small bags, and bag fragments (OSPAR codes 2, 3, 112). Total litter counts all 112 OSPAR item categories. Bag share is the ratio of bag items to total litter. Charge introduction dates: Wales (October 2011), Northern Ireland (April 2013), Scotland (October 2014), England (October 2015). N = 47 monitoring beaches across four UK nations.

protocol, shared marine environment, and exogeneity of devolved legislative timing (driven by parliamentary schedules and political priorities, not beach litter trends).

4.2 Estimation

The primary estimator is the Callaway and Sant'Anna (2021) group-time average treatment effect on the treated (ATT), which avoids the negative weighting problem in standard two-way fixed effects estimation with staggered adoption (Goodman-Bacon, 2021). The four treatment cohorts are: Wales (first full treatment year: 2012), Northern Ireland (2013), Scotland (2015), and England (2016). Not-yet-treated beaches serve as the comparison group.

I supplement this with conventional TWFE regressions of the form:

$$Y_{bt} = \alpha + \beta \cdot \text{Post}_{n(b),t} + \gamma_b + \delta_t + \varepsilon_{bt} \tag{1}$$

where Y_{bt} is the outcome (log bag items, log total litter, or bag share) at beach b in year t, $\text{Post}_{n(b),t}$ equals one after nation $n(b)$ introduces its charge, γ_b are beach fixed effects, and δ_t are year fixed effects. Standard errors are clustered at the beach level (47 clusters).

4.3 Threats to Validity

The main concern is that only four nations generate the treatment variation. While 47 beaches provide adequate cluster count for inference, the four treatment cohorts limit the diversity of comparisons available to the Callaway-Sant'Anna estimator. I address this through leave-one-cohort-out robustness, alternative functional forms, and a placebo outcome (non-bag litter items, which should not respond to the bag charge).

A second concern is panel composition: Northern Ireland entered the OSPAR programme later than other nations, with limited pre-charge data. I present results both including and excluding individual nations to assess sensitivity.

A third concern is statistical power. With beach-level clustering (47 clusters) and four treatment cohorts, the minimum detectable effect at 80% power is 0.54 log points, corresponding to a 72% reduction in bag items. If the retail channel accounts for only 5–10% of beach bag litter, the true treatment effect on bags would be well below this detection threshold. The null finding should therefore be interpreted as: *the data cannot rule out small effects, but they can rule out the large effects that would be needed for the charge to meaningfully reduce beach bag pollution.*

5. Results

5.1 Main Results

Table 2 presents the main results. The charge has no statistically significant effect on carrier bag litter. In the TWFE specification (Panel A), the coefficient on post-charge is 0.179 log points (SE = 0.193, $p = 0.35$), indicating no detectable decline in bag items. The Callaway-Sant'Anna estimator (Panel B) yields an ATT of 0.480 log points (SE = 0.384), again statistically insignificant and positive in sign.

In stark contrast, total beach litter declined after the charge. The TWFE estimate for log total litter is -0.330 (SE = 0.200, $p = 0.099$), driven entirely by non-bag categories (-0.346, SE = 0.203). This decline in non-bag litter—which was not targeted by the charge—likely reflects concurrent environmental campaigns, tighter waste management regulations, and secular trends in littering behavior.

Because bags did not decline while other litter did, the bag share of total litter *increased* by 1.5 percentage points (SE = 0.007, $p = 0.029$). Bags became a larger, not smaller, component of beach pollution after the charge was introduced.

Table 2: Effect of Plastic Bag Charges on Beach Litter Composition

	Log Bag Items (1)	Log Total Litter (2)	Log Non-Bag Litter (3)	Bag Share (4)
Panel A: TWFE				
Post-charge	0.179	-0.330*	-0.346*	0.015**
	(0.193)	(0.200)	(0.203)	(0.0067)
Observations	407	407	407	407
Panel B: Callaway-Sant'Anna				
ATT	0.480	-0.174	-0.189	0.014*
	(0.384)	(0.451)	(0.436)	(0.0075)
Beach FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Beaches	47	47	47	47
Clustering	Beach	Beach	Beach	Beach

Notes: Standard errors clustered at the beach level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Panel A reports two-way fixed effects estimates. Panel B reports Callaway-Sant'Anna (2021) staggered DiD using not-yet-treated beaches as controls. The dependent variable in column (1) is log(bag items + 1); columns (2)–(3) are log(total litter + 1) and log(non-bag litter + 1); column (4) is the bag share of total litter. Sample: 47 OSPAR monitoring beaches across Wales, Northern Ireland, Scotland, and England, 2001–2020.

Event study.. Table 3 presents the dynamic event-study estimates. Pre-treatment coefficients are noisy but centered near zero, providing no strong evidence against parallel trends. Post-treatment coefficients are positive but imprecise, consistent with the null effect on bags.

Table 3: Event Study: Dynamic Effects on Log Bag Items

Event Time	ATT	SE	95% CI
-5	-0.154	(0.454)	[-1.043, 0.735]
-4	-0.243	(0.583)	[-1.385, 0.898]
-3	-0.415	(0.468)	[-1.333, 0.503]
-2	-0.240	(0.512)	[-1.243, 0.763]
-1	0.000	(NA)	[NA, NA]
+0	0.206	(0.409)	[-0.595, 1.006]
+1	0.500	(0.439)	[-0.360, 1.360]
+2	1.033***	(0.382)	[0.285, 1.781]
+3	-0.246	(0.667)	[-1.553, 1.060]

Notes: Callaway-Sant'Anna dynamic aggregation with not-yet-treated control. Event time 0 is the first full year after the bag charge was introduced. Pre-treatment coefficients test parallel trends.

5.2 Robustness

Table 4 demonstrates that the null finding is stable across specifications. The IHS transformation, raw count outcome, exclusion of COVID-affected 2020, and leave-one-cohort-out tests all yield point estimates near zero or positive, none statistically different from zero at conventional levels.

The placebo outcome—plastic bottles, which were not subject to any pricing intervention—also shows a positive ATT (0.997, SE = 0.394). This raises an important caveat: the design may be capturing broad compositional changes in the monitoring panel rather than bag-charge-specific effects.

The bag share as a triple-difference.. The bag share result in column (4) of Table 2 can be interpreted as an implicit difference-in-difference-in-differences: the change in bag litter *relative to other litter categories*, within beach, before versus after the charge. This specification differences out any common shocks affecting all litter types (such as compositional changes in the monitoring panel or concurrent anti-littering campaigns). The significant increase in bag share (1.5 percentage points, $p = 0.029$) indicates that bags became a relatively *larger* component of beach litter after the charge—the opposite of what Pigouvian intuition predicts. Even if the levels estimates are contaminated by panel composition effects, the within-litter-category comparison is not.

Table 4: Robustness: Alternative Specifications for Bag Litter

Specification	CS ATT	SE
Baseline (not-yet-treated)	0.480	(0.384)
IHS transformation	0.596	(0.460)
Raw bag counts	1.682	(3.166)
Excl. 2020 (COVID)	0.480	(0.344)
Drop England	0.155	(0.597)
Drop Wales	0.832**	(0.384)
Placebo: plastic bottles	0.997**	(0.394)

Notes: Each row reports the Callaway-Sant'Anna simple ATT. Row 1 is the baseline from Table 2. The placebo outcome (plastic bottles) was not subject to the carrier bag charge and tests whether the design captures bag-specific effects.

6. Discussion

The pollution gap.. The central result—no decline in beach bag litter despite a 95% reduction in retail bag sales—implies that the retail consumption channel accounts for a small share of the bags deposited on beaches. I term this the *pollution gap*. The charge successfully reduces the flow of bags from shops to consumers, but the flow of bags from the environment to the shoreline comes predominantly from non-retail sources: wind-dispersed litter from open-top bins, fly-tipped waste, commercial fishing activity, and the existing stock of bags already in the marine environment (OSPAR Commission, 2020). Taxing the cash register does not reach the bin, the trawler, or the ocean gyre.

Limitations.. Three important caveats apply. First, the monitoring panel is small (47 beaches, 4 nations), and the treatment variation comes from only four legislative events. The minimum detectable effect at 80% power is a 72% reduction—meaning the design is powered to detect whether the charge *substantially* reduced beach bag pollution, but cannot detect small effects. If the retail channel contributes only 5–10% of beach bag litter (as the pollution gap hypothesis implies), then the true effect on beach litter would be 5–10%, far below the detection threshold. The null is thus consistent with both "no effect" and "small effect that the data cannot see." Second, Northern Ireland's late entry into the monitoring programme means its pre-charge data are thin, potentially biasing the DiD comparison. Third, the positive placebo coefficient on bottles indicates that the levels-based estimates capture broader panel compositional shifts—though the bag *share* result, which differences these out, remains significant. Future work with richer monitoring data—such as the Marine Conservation Society's 600-beach Great British Beach Clean dataset, if released for research—could improve power substantially.

Implications for Pigouvian design.. Despite these caveats, the qualitative finding is informative: the charge demonstrably reduced consumption but did not demonstrably reduce pollution. This disconnect is a concern for Pigouvian environmental policy. When the taxed margin (retail purchase) is a poor proxy for the externality (marine pollution), even a perfectly designed consumption tax may fail to achieve its environmental objective without complementary measures targeting downstream leakage channels—improved bin design, extended producer responsibility for commercial packaging, and riverine waste interception (Taylor, 2020). The bag charge may still be welfare-improving through reduced resource extraction and landfill waste, but its effectiveness as an anti-pollution instrument requires additional interventions to close the pollution gap.

7. Conclusion

Plastic bag charges are the world's most popular environmental pricing policy. They work spectacularly at the cash register: consumers bring reusable bags, retailers report 95% fewer sales. But the beach tells a different story. Using 20 years of standardized marine litter monitoring data from 47 UK beaches and the staggered adoption of bag charges across four nations, I find no evidence that the charge reduced the number of bags washing up on shore. The bags were never coming from the cash register in the first place.

This is not an argument against bag charges. Reducing unnecessary plastic production is valuable in its own right. But it is an argument for measuring pollution where it happens—on the beach, in the river, in the ocean—rather than where we wish it happened, at the point of sale. Environmental policy that optimizes the metric it can see, rather than the problem it claims to solve, risks the most insidious form of greenwashing: believing its own press.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP). Beach litter survey data from the OSPAR Commission Beach Litter Database, accessed via the `nicrie/seasonality_ospar` GitHub repository.

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.

Convery, Frank, Simon McDonnell, and Susana Ferreira, "The Most Popular Tax in Europe? Lessons from the Irish Plastic Bags Levy," *Environmental and Resource Economics*, 2007, *38* (1), 1–11.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell, "What Do We Know about Short- and Long-Term Effects of Early-Life Exposure to Pollution?," *Annual Review of Resource Economics*, 2014, *6*, 217–247.

Department for Environment, Food and Rural Affairs, "Carrier Bag Charge: Summary of Data in England," Technical Report, DEFRA 2020. Statistical Release.

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.

Dikgang, Johane, Anthony Leiman, and Martine Visser, "Analysis of the Plastic-Bag Levy in South Africa," *Resources, Conservation and Recycling*, 2012, *66*, 59–65.

Fullerton, Don, Sarah E. West, and Sivan Kartha, "Tax Policy toward Products That Cause Environmental Problems," *American Economic Review*, 2002, *92* (3), 645–667.

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

Jakovcevic, Adriana, Linda Steg, Nadia Mazzeo, Romina Caballero, Paul Franco, Natalia Putrino, and Josefina Favara, "Charges for Plastic Bags: Motivational and Behavioral Effects," *Journal of Environmental Psychology*, 2014, *40*, 372–380.

Marcus, Michelle, "The Effects of Fireworks Bans on Air Quality," *American Economic Journal: Economic Policy*, 2021, *13* (4), 147–173.

Nielsen, Tobias Dan, Karl Holmberg, and Johannes Stripple, "How Plastic Bags Are Harmful for the Environment," *Waste Management*, 2019, *87*, 428–440.

OSPAR Commission, "Guideline for Monitoring Marine Litter on the Beaches in the OSPAR Maritime Area," Technical Report, OSPAR Commission 2010. Reference number 2010-02.

_ , "OSPAR Beach Litter Assessment: Levels and Trends," Technical Report, OSPAR Commission 2020. Intermediate Assessment 2017 update.

Pigou, Arthur Cecil, *The Economics of Welfare*, London: Macmillan, 1920.

Poortinga, Wouter, Lorraine Whitmarsh, and Caroline Suffolk, "The English Plastic Bag Charge Changed Behavior and Increased Support for Other Charges to Reduce Littering," *Frontiers in Psychology*, 2013, *10*, 266.

Rivers, Nicholas, Sarah Shenstone-Harris, and Nathan Young, "What Can a Single-Use Plastic Bag Levy Achieve?," *Environmental and Resource Economics*, 2017, *66* (3), 407–434.

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

Taylor, Rebecca L.C., "Bag Leakage: The Effect of Disposable Carryout Bag Regulations on Unregulated Bags," *Journal of Environmental Economics and Management*, 2020, *107*, 102338.

Thomas, Gareth O., Elena Sautkina, Wouter Poortinga, Emily Wolstenholme, and Lorraine Whitmarsh, "Incentivising Reusable Bags: The Case of UK Bag Charges," *Journal of Cleaner Production*, 2019, *215*, 1–11.

A. Data Appendix

OSPAR Beach Litter Database.. The OSPAR Beach Litter Monitoring Programme follows the OSPAR Commission (2010) protocol. Surveyors walk a fixed 100-meter stretch at designated sites, counting all visible litter items ≥2.5 cm and classifying them into 112 categories. Surveys are typically conducted in spring and autumn. Annual CSV files (2001–2020) were obtained from the `nicrie/seasonality_ospar` repository on GitHub, which mirrors the OSPAR database export.

Beach-to-nation mapping.. UK beaches (identified by OSPAR codes UK001–UK052) were assigned to England, Scotland, Wales, or Northern Ireland based on geographic coordinates. Channel Islands beaches (UK017 Guernsey, UK051 Jersey) were excluded. The final sample contains 20 English, 6 Scottish, 6 Welsh, and 15 Northern Irish monitoring sites.

Variable construction.. "Bag items" is the sum of OSPAR categories 2 (carrier bags), 3 (small bags), and 112 (bag fragments). "Total litter" is the sum of all 112 item categories. "Bag share" is bag items divided by total litter. When multiple surveys were conducted at the same beach in the same year, values were averaged to the beach-year level.

B. Robustness Appendix

See Table 4 in the main text for all robustness specifications.

C. Standardized Effect Sizes

Table 5: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Panel A: Pooled						
Log bag items	0.480	0.384	1.087	0.441	0.353	Large positive
Bag share	0.0144	0.0075	0.0414	0.347	0.181	Large positive
Log total litter	-0.174	0.451	1.186	-0.147	0.380	Moderate negative
Panel B: Heterogeneous (by pre-treatment litter intensity)						
Log bag items (high-litter beaches)	0.202	0.360	1.087	0.186	0.332	Large positive
Log bag items (low-litter beaches)	0.510	0.514	1.087	0.469	0.473	Large positive

Notes: **Country:** United Kingdom. **Research question:** Do mandatory plastic bag charges reduce actual environmental pollution, as measured by carrier bag counts in standardized beach litter surveys? **Policy mechanism:** Four UK nations sequentially introduced mandatory 5p charges on single-use carrier bags at retail point of sale, requiring retailers to charge consumers per bag and raising the relative price of disposable bags to induce substitution toward reusable alternatives. **Outcome definition:** Count of carrier bags (OSPAR litter category 2), small bags (category 3), and bag fragments (category 112) per standardized 100-meter beach transect survey. **Treatment:** Binary; equals one after the mandatory bag charge takes effect in that beach's nation. **Data:** OSPAR Beach Litter Database, 2001–2020, 47 UK monitoring beaches across four nations, 407 beach-year observations. **Method:** Callaway-Sant'Anna (2021) staggered DiD with not-yet-treated beaches as control; standard errors clustered at beach level. **Sample:** All OSPAR 100m transect monitoring beaches in England, Wales, Scotland, and Northern Ireland with at least one survey during 2001–2020; Channel Islands excluded. SDE = $\hat{\beta}/\text{SD}(Y)$ 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).