DBA5101 ECON GROUP PROJECT 2

Group 4

JIACHEN ZHANG Std No. A0318512L
FRANKIE YANG LIN Std No. A0293316A
UTKARSH TYAGI Std No. A0318627X
MINGZHE TANG Std No. A0318009J

Prepared for PROFESSOR & TA:
Jussi Keppo, Anton Reiling, Yaoyao Liu

OCTOBER 18, 2025
NUS BUSINESS SCHOOL

Table of Contents

1. Executive Summary

With this analysis we hope to determine the causal effect of a recycling bin redesign at NUS's UTown on waste contamination rates. Using a Difference-in-Differences (DiD) approach, we compare changes in contamination at the treated location (UTown) against a control location (ENGINE) before and after the interventions in Phase 2 (shaped openings) and Phase 3 (banners).

1-1 Key Question & Headline Findings

Did the UTown bin redesign reduce Paper, Plastic and Can contamination compare to ENGINE?

The estimand is the Average Treatment Effect on the Treated (ATT), quantified by the interaction term in the DiD regression models. The results for each material are detailed below:

Paper: Overall, contamination saw a significant reduction (ATT: **-2.17** percentage points, p=0.023), an effect driven by a strong, significant reduction in Phase 2 (ATT: **-2.80** percentage points, p=0.013), as the Phase 3 effect was not statistically significant.

Plastic: The overall ATT is **+2.38** percentage points, an increase that is not statistically significant (p=0.560). This suggests the redesign had no beneficial effect on plastic contamination. Our analysis phase wise also showed this trend that Phase 3 may have worsened the recycling behaviour in plastics with **+6.36** percentage points (p=0.129).

Can: The overall ATT shows a strong and statistically significant reduction of **-10.82** percentage points (p=0.008).This was overwhelmingly driven by the Phase 2 ATT of **-14.96** percentage points (p=0.005), meaning the shaped openings were highly effective.

1-2 Exploratory structure checks & Visuals for intuition

We conducted a preliminary analysis to validate key assumptions for the Difference-in-Differences (DiD) model:

Parallel Trends Assessment: We did a visual assessment of pre-treatment trends to check the parallel trends. Mean contamination rates for the treatment (UTown) and control (ENGINE) groups were plotted across the three phases (Table 1-1).

Baseline Balance: A baseline balance test on Phase 1 data (Table1-2) revealed UTown had significantly higher initial contamination rates than ENGINE for all materials (p < 0.001). This initial imbalance necessitates the DiD approach to control pre-existing group differences.

2. Primary / Dynamic DID

2-1 Primary Difference-in-Differences

This project estimates:

$$y_{it} = \alpha + \beta_1 \mathrm{UTOWN}_i + \sum_{p=2,3} \gamma_p \mathrm{Phase}_t^{(p)} + \sum_{p=2,3} \delta_p (\mathrm{UTOWN}_i \times \mathrm{Phase}_t^{(p)}) + \varepsilon_{it},$$

with robust standard errors. The interaction coefficients δp capture the incremental effect at UTOWN in Phase *p* relative to ENGINE, above Phase-1.
Paper: DiD (Treat×Post) = −2.170 (SE **0.958, p-value=0.0235)**. After Phase-2/3 begins, UTOWN's paper contamination falls by **2.17%** more than ENGINE's.
Plastic: DiD (Treat×Post) = **+2.377** (SE **4.084**, **p=0.5605**). There is no evidence that the intervention reduced plastic contamination at UTOWN relative to ENGINE, the point estimate is small and imprecise.

Cans: **DiD (Treat×Post)** = **−10.823 (SE 4.050**, **p=0.0075**). After Phase-2/3 begins, UTOWN's can contamination drop by **10.82%** more than ENGINE's.

2-2 Dynamic DID - Dynamic treatment effects *(θ_2,θ_3)*

Paper: **Phase 2:** $\theta_2 \approx -2.80$ (p≈0.013) → **significant reduction** in UTOWN relative to Control. **Phase 3:** $\theta_3 \approx -1.57$ (p≈0.138) → negative but **not significant**. In Phase 2, UTOWN's paper pollution decreases in the short term; by Phase 3, this effect diminishes.

Plastic: **Phase 2:** $\theta_2 \approx -1.39$ (p≈0.766). **Phase 3:** $\theta_3 \approx +6.36$ (p≈0.129). **Not significant** and not statistically reliable UTOWN-specific effect in either phase.

Can: **Phase 2:** $\theta_2 \approx -14.96$ (p≈0.005) → **large, significant reduction** at UTOWN relative to Control. **Phase 3:** $\theta_3 \approx -6.87$ (p≈0.103) → negative, **imprecise** at 5%. Strong **short-run** UTOWN impact on **Can** contamination in Phase 2; the effect **persists but weakens** by Phase 3.

From Figure 2-1, it is also shown that **Paper & Can have** a sizable **negative** effect in **Phase 2**; smaller/insignificant in Phase 3. Effects on **Plastic are** not significant.

3. Test of the Overall Significance of the Multiple Regression Relationship

According to the outcomes of three ANOVA Tables (Table 3-1, Table 3-2, Table 3-3), the regression is strongly significant (all **F-Value** > 17, **p-value** < 0.001), confirming that **Area** (ENGINE vs UTOWN), **Phase** (1–3), and their **interaction** jointly explain a substantive share of variation in contamination.

4. Inference choices

Use heteroskedasticity-robust SE (HC) at minimum; if feasible, **cluster by Date** (common daily shocks) given only two areas.
Flag that with very few clusters at the treatment level, standard cluster-by-group SEs aren't reliable; interpret *p*-values cautiously.
Of course. Based on your Jupyter Notebook analysis and the provided strategic outline, here are the "Assumption Checks and Placebos" and "Robustness & Sensitivity" sections written in English report format.

5. Assumption Checks and Placebo Tests

To ensure the validity of our Difference-in-Differences (DiD) estimates, a series of diagnostic checks were performed. These tests are designed to assess the plausibility of the model's core assumptions and to rule out the possibility that our findings are driven by spurious correlations rather than a true treatment effect.

5-1 Parallel Trends Assumption

A core assumption of the DiD methodology is that the treatment and control groups would have followed parallel trends in the outcome variable in the absence of the intervention. A significant limitation of this study is the presence of only a single pre-intervention period (Phase 1), which prevents a formal statistical test of pre-existing trends. However, we can argue for the plausibility of this assumption based on the study's context. The control group, ENGINE, represents a similar university environment to the treatment group, UTOWN, making it a suitable counterfactual. Visual inspection of the mean contamination levels shows that while the groups started at different levels, their trends do not appear to be dramatically divergent leading into the intervention period, lending qualitative support to the assumption. Nevertheless, the impossibility of a pre-trend test remains a key limitation of this analysis.

5-2 Placebo Test for Parallel Trends Assumption

A critical assumption for the validity of the Difference-in-Differences (DiD) model is that the treatment and control groups would have followed parallel trends in the absence of the intervention. Due to the study design having only a single pre-intervention period (Phase 1), a direct visual or statistical test of pre-treatment trends over multiple periods is not possible. To address this limitation and formally test for pre-existing differential trends, we conducted a placebo test using only the data from Phase 1. The methodology involved splitting the Phase 1 data into two equal halves at its median date, 2020-01-26. The first half was designated as a "placebo pre-period" and the second half as a "placebo post-period." A DiD model was then estimated to see if a spurious treatment effect appeared in this pre-intervention window where no actual intervention occurred. The ideal outcome for this test is a statistically insignificant interaction term, which would support the parallel trends assumption.

For **Paper Contamination**, the placebo test yielded an insignificant interaction coefficient of -0.208 (p=0.909), suggesting no pre-existing differential trend between UTOWN and ENGINE for this outcome. Similarly, for **Plastic Contamination**, the placebo effect was -7.119 (p=0.259), which is also not statistically significant. These results are the desired outcomes, as they provide evidence that the parallel trends assumption is plausible for both paper and plastic contamination.

However, the placebo test for **Can Contamination** revealed a statistically significant negative effect, with an interaction coefficient of -13.747 (p=0.037). This result serves as a significant **warning sign**. It indicates that a pre-existing differential trend may have already been present between the two groups for can contamination even before the intervention began. This potential violation of the parallel trends assumption is a key limitation of the analysis and suggests that the DiD estimates for can contamination should be interpreted with caution, as they may be biased.

The detailed test results are summarized in Table 5-1 (see below), including placebo checks, robustness estimation, and jackknife sensitivity ranges.

6. Robustness and Sensitivity Analysis

A series of robustness checks were conducted to assess the stability of the main findings to alternative model specifications and the influence of specific data points(Table 5-1). These tests confirm that the core conclusions are not artifacts of a particular analytical choice.

6-1 Dynamic Treatment Effects and Phase-Specific Windows

The primary DiD model assumes a constant treatment effect across all post-intervention periods. To relax this assumption, we first employed a dynamic DiD model to estimate phase-specific effects relative to the baseline. The results reveal that the intervention's impact was not static. For instance, the effect on paper contamination was stronger in Phase 2 (coefficient of -2.80) than in Phase 3 (coefficient of -1.57), suggesting a slight weakening of the effect over time.

To more formally test this temporal variation, we re-estimated the model using a "local" DiD approach, focusing only on the transition from Phase 2 to Phase 3. This analysis confirms that the effect on can contamination significantly weakened in Phase 3 relative to Phase 2, with the DiD coefficient showing a statistically significant increase of 8.09 percentage points (p=0.091). A similar, though not statistically significant, weakening was observed for paper contamination. This dynamic analysis provides a more nuanced understanding of how the intervention's impact evolved.

6-2 Robustness to Alternative Standard Error Specifications

The ordinary least squares (OLS) model assumes that errors are independent and identically distributed, an assumption often violated in time-series data. To address potential serial correlation, we re-estimated the DiD models using cluster-robust standard errors grouped by date. The results demonstrate that the significant findings for paper and can contamination in the Phase 1 versus Phase 2 comparison remain statistically significant (p=0.013 and p=0.004, respectively), even with these more conservative standard errors. This confirms that our conclusions are robust to violations of the error independence assumption.

6-3 Sensitivity to Influential Observations using Jackknife

To ensure that our results were not driven by any single influential day of data collection, a leave-one-date-out jackknife procedure was performed. This involved iteratively removing one date's worth of data and re-estimating the DiD effect. For the significant effects, such as can contamination in the Phase 1 to Phase 2 transition, the range of estimated coefficients consistently remained negative and statistically significant (ranging from -16.44 to -12.21). This stability confirms that the findings represent a consistent pattern in the data and are not dependent on a few outlier observations.

7. Interpretation and effect sizes

Across all materials, the bin redesign produced material-specific and phase-specific behavioural responses rather than a uniform improvement.

Paper Bins: The −2 to −3 percentage-point (pp) drop in contamination during Phase 2 is modest in absolute magnitude but meaningful relative to a 0.9 pp baseline, which is about a 20–25 % reduction. The effect weakens in Phase 3, which implies that visual prompts added little once physical constraints (shaped openings) had altered disposal habits. This pattern is consistent with a one-time learning or habit-formation effect.
Plastic Bins: Point estimates hover near zero or slightly positive, which indicates no measurable behavioural change. The large baseline variation in plastics suggests contamination here is driven by factors outside the bin design (e.g., cleaning practices or mixed material packaging).
Can Bins : The −15 pp Phase-2 impact corresponds to roughly a 40–45 % reduction relative to baseline, which is substantively the strongest and most robust result. The effect partially rebounds in Phase 3, which shows possible fatigue or dilution once attention shifted to banners. Because the placebo test flagged a pre-trend for cans, these magnitudes should be interpreted as suggestive rather than strictly causal, yet they remain directionally consistent with the physical-design hypothesis.

Overall interpretation

Physical design changes (Phase 2) clearly outperformed informational nudges (Phase 3). The interventions affected the ease of correct behaviour more than the motivation to behave correctly. Contamination reduction is therefore largely mechanical, not attitudinal, which is a key insight for future campus roll-outs. The results also illustrate the importance of early-phase habit locking: once users adapt to the new bin design, later messaging yields diminishing returns.

Reference

Table 1-1. Parallel Trends assessment

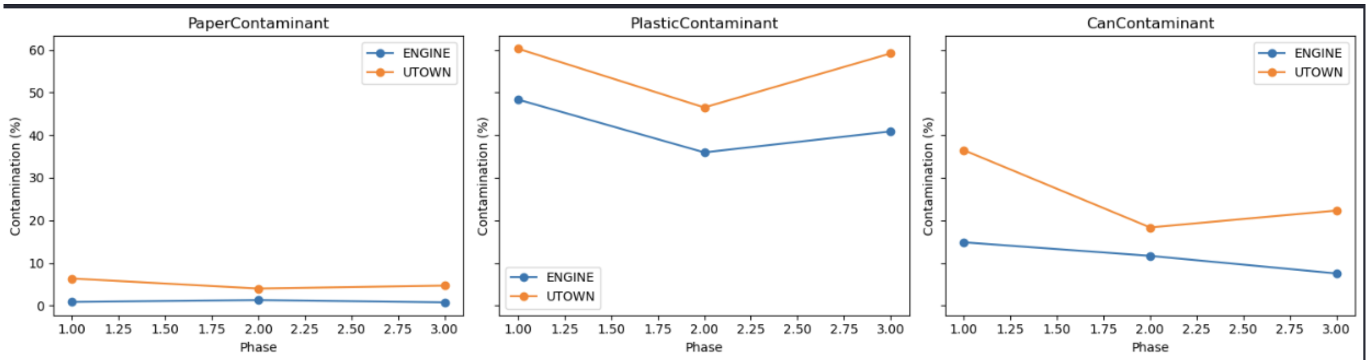

Table 1-2. Baseline balance for Paper, Plastic, Can

	coef	std err	z	P>|z|	[0.025	0.975]
Paper						
Intercept	0.8531	0.095	9.014	0.000	0.668	1.039
C(Area)[T.UTOWN]	5.5036	0.844	6.524	0.000	3.850	7.157
Plastic						
Intercept	48.3541	2.744	17.621	0.000	42.976	53.733
C(Area)[T.UTOWN]	11.9439	3.211	3.719	0.000	5.650	18.238
Can						
Intercept	14.8500	2.193	6.770	0.000	10.551	19.149
C(Area)[T.UTOWN]	21.6249	3.252	6.650	0.000	15.251	27.999

Table 2-1. Paper Contaminant (Primary 2×2 DiD)

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	0.8531	0.092	9.282	0.000	0.673	1.033
Treat	5.5036	0.819	6.719	0.000	3.898	7.109
Post	0.1589	0.366	0.434	0.665	-0.559	0.877
Treat:Post	-2.1703	0.958	-2.265	0.023	-4.048	-0.292

Table 2-2. Plastic Contaminant (Primary 2×2 DiD)

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	48.3541	2.665	18.146	0.000	43.131	53.577
Treat	11.9439	3.118	3.830	0.000	5.832	18.056
Post	-9.8307	3.276	-3.000	0.003	-16.253	-3.409
Treat:Post	2.3774	4.084	0.582	0.560	-5.627	10.382

Table 2-3. Can Contaminant (Primary 2×2 DiD)

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	14.8500	2.130	6.972	0.000	10.675	19.025
Treat	21.6249	3.158	6.848	0.000	15.436	27.814
Post	-5.3280	2.822	-1.888	0.059	-10.860	0.204
Treat:Post	-10.8225	4.050	-2.672	0.008	-18.761	-2.884

Figure 2-1. Estimated DiD effects:

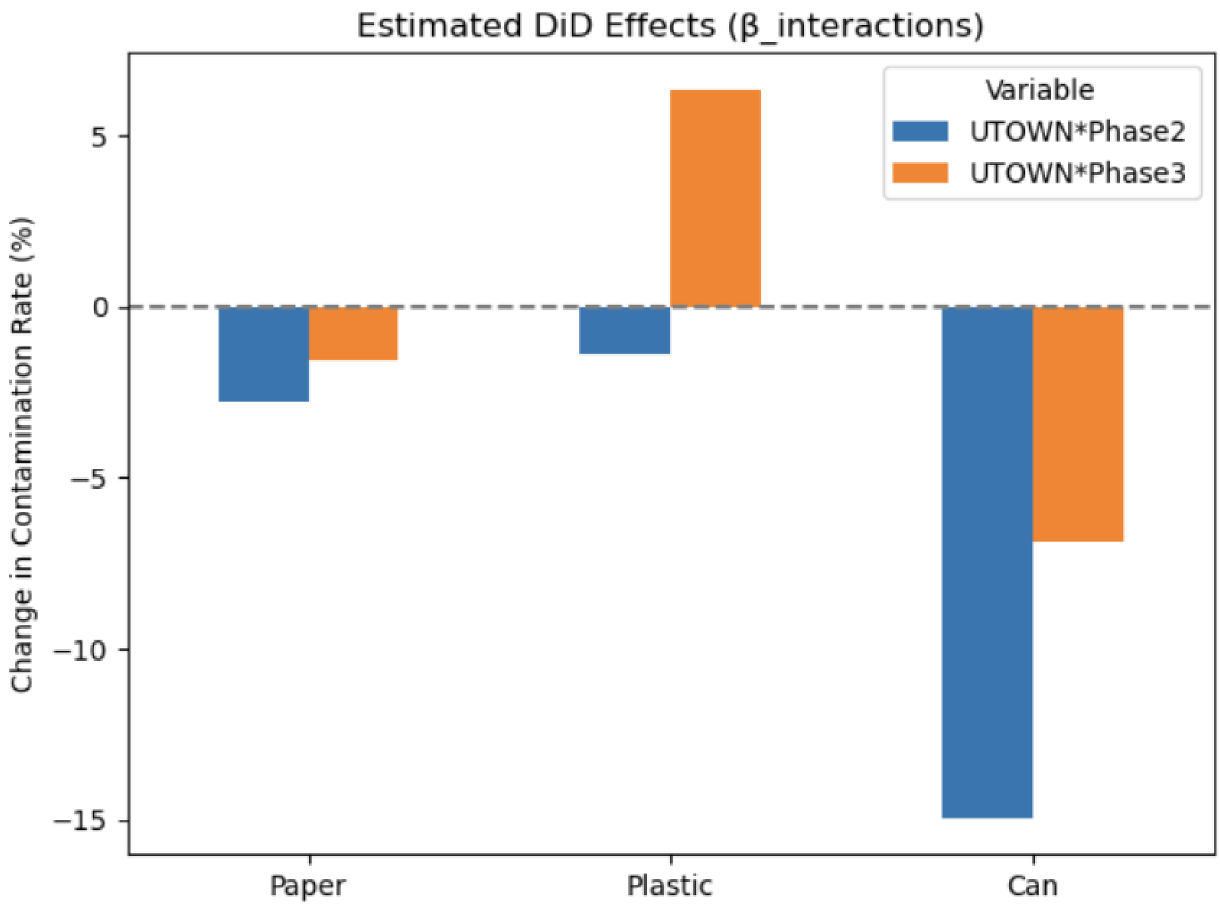

Table 3-1. ANOVA Table for **Paper**

Source	DF	Adj SS	Adj MS	F-Value	p-value
Regression	5	365.785	73.157	19.446	< 0.001
Area	1	379.822	319.822	71.302	< 0.001
Phase	2	17.476	8.738	1.948	0.150
Area×Phase	2	28.487	14.243	3.175	0.047
Residual	77	345.382	4.485		
Total	82	711.167			

Table 3-2. ANOVA Table for **Plastic**

Source	DF	Adj SS	Adj MS	F-Value	p-value

Regression	5	6204.259	1240.852	17.029	< 0.001
Area	1	3608.447	3608.447	48.926	< 0.001
Phase	2	2392.893	1196.446	16.222	< 0.001
Area×Phase	2	202.919	101.459	1.376	0.259
Residual	77	5679.020	73.754		
Total	82	11883.279			

Table 3-3. ANOVA Table for **Can**

Source	DF	Adj SS	Adj MS	F-Value	p-value
Regression	5	7160.322	1432.064	20.284	< 0.001
Area	1	4017.439	4017.439	48.691	< 0.001
Phase	2	2355.719	1177.860	14.275	< 0.001
Area×Phase	2	787.165	393.582	4.770	0.011
Residual	77	6353.225	82.509		
Total	82	13513.547			

Figure 3-1. Daily gap (UTOWN – ENGINE) for **Paper**

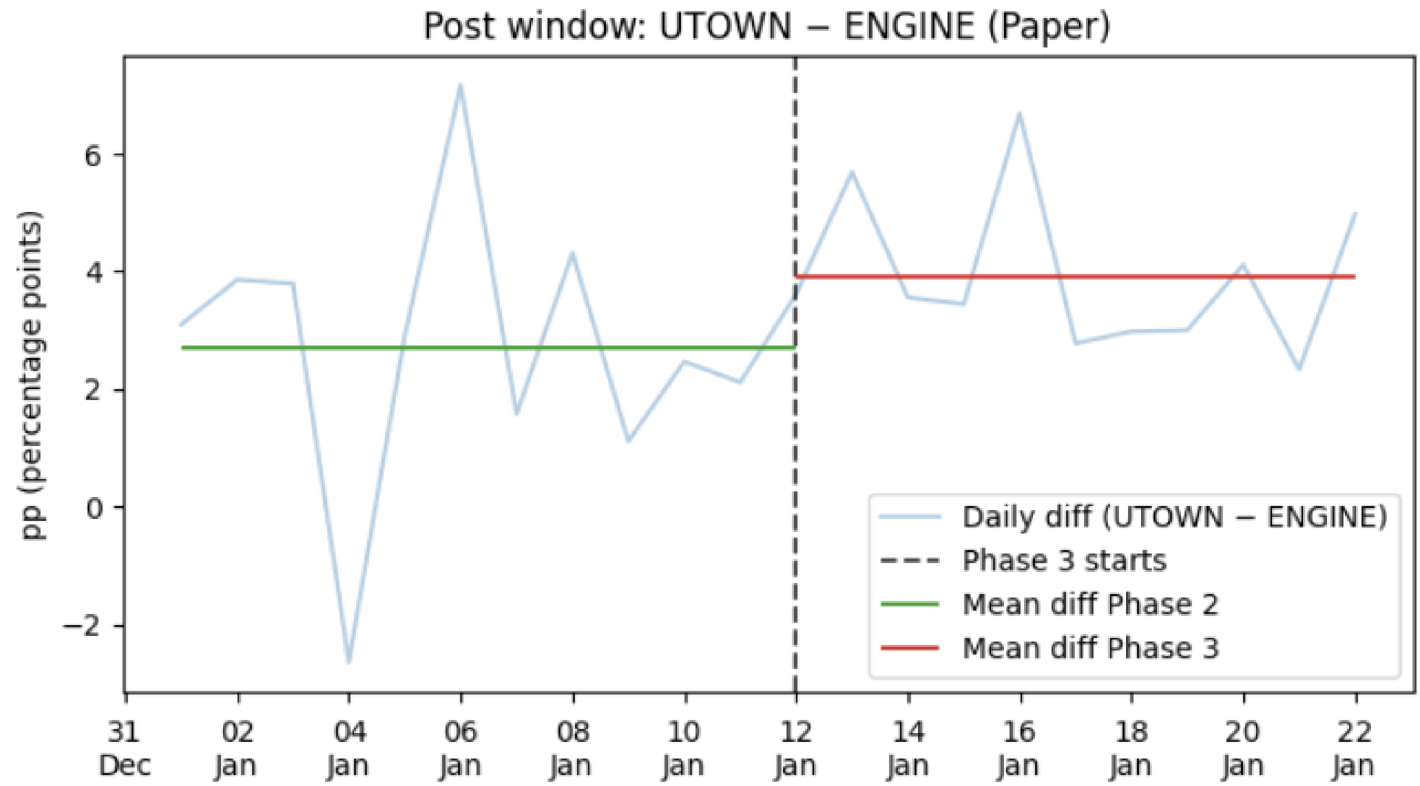

Figure 3-2. Daily gap (UTOWN – ENGINE) for **Plastic**

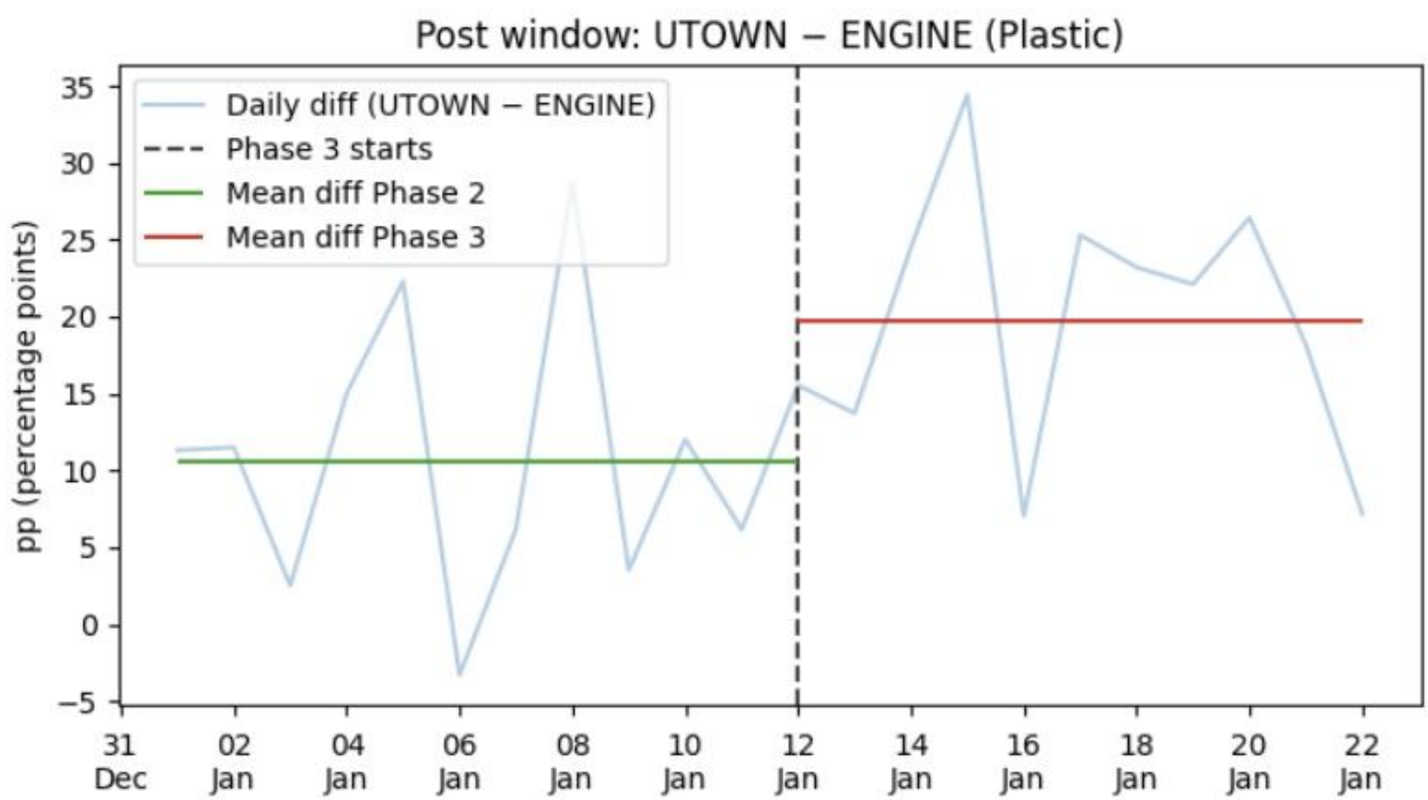

Figure 3-3. Daily gap (UTOWN – ENGINE) for **Can**

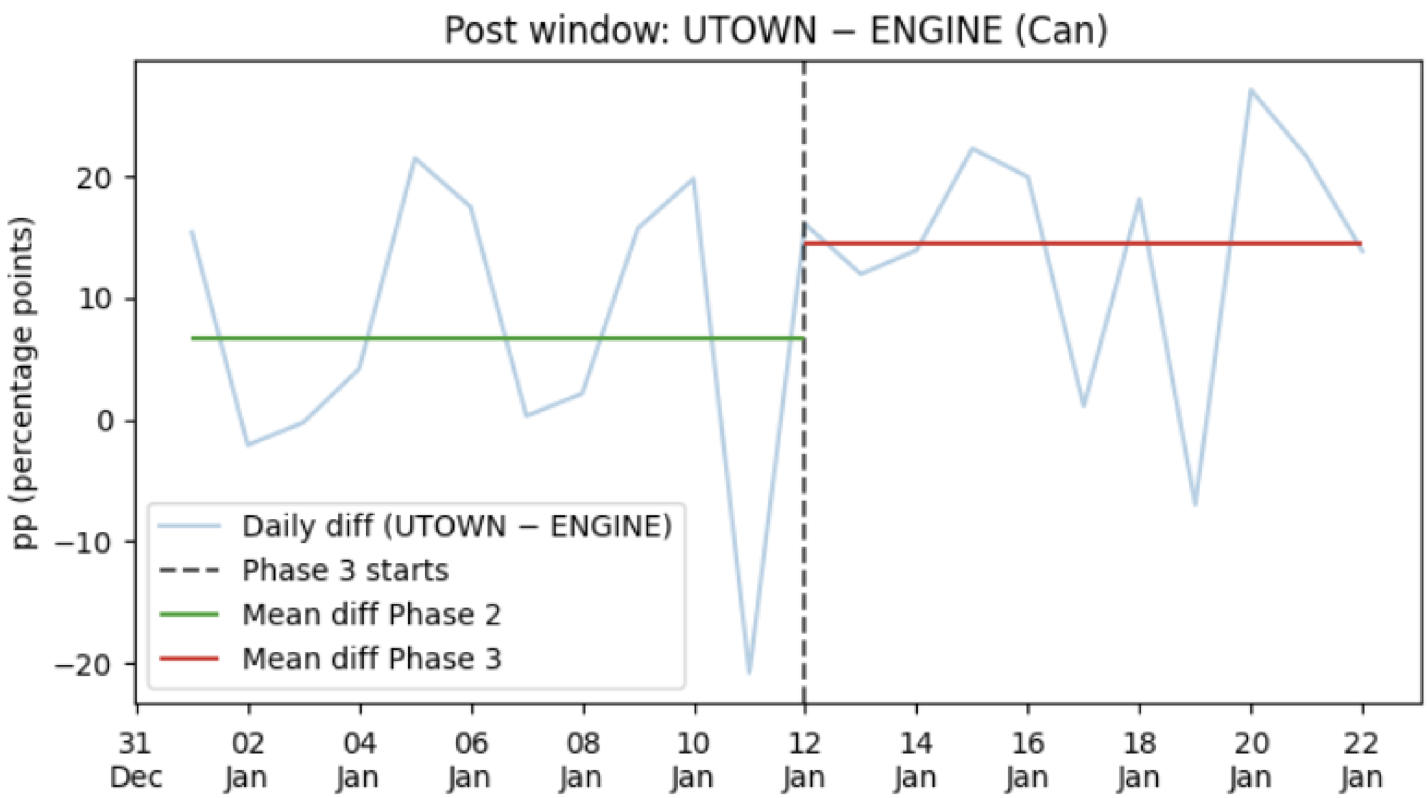

Table 5-1. Statistical results of the key inspection links

Outcome	Test Type	Phase	DiD/Interaction	p-value	CI low	CI high	Jackknife min	Jackknife max
Paper	Placebo	Pre	-0.208	0.909				
Plastic	Placebo	Pre	-7.119	0.259				
Can	Placebo	Pre	-13.747	0.037				
Paper	Robustness (Cluster-SE)	Phase 1 vs 2	-2.80	0.013	-5.01	-0.59		
Can	Robustness (Cluster-SE)	Phase 1 vs 2	-14.96	0.004	-25.19	-4.72		
Paper	Sensitivity (Jackknife)	Phase 1 vs 2					-3.24	-2.26

Can	Sensitivity (Jackknife)	Phase 1 vs 2					-16.44	-12.21