Online Appendix to

The Idle Margin: Why Everyday Conservation Goes Undone

Muhammad Usman Naeem*

13th September 2026

Latest version here

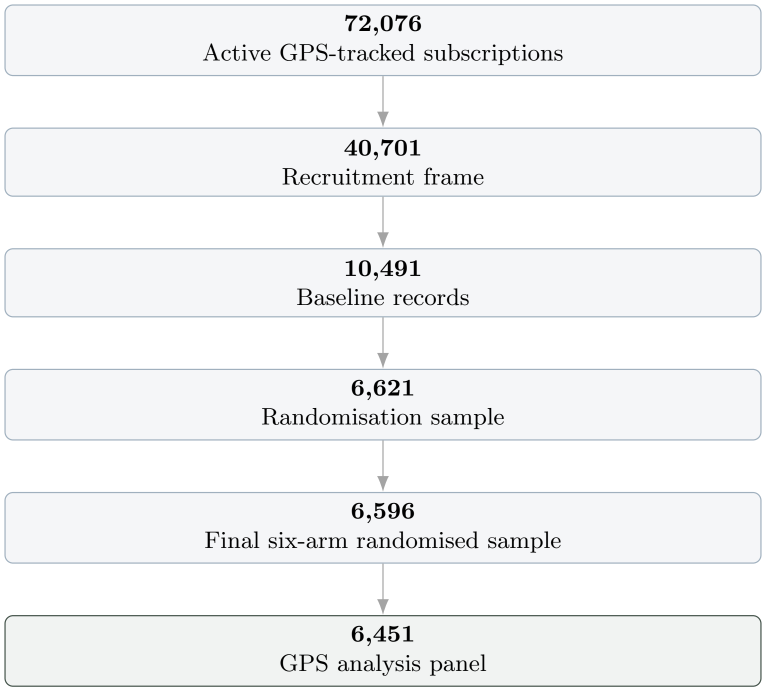
*Tufts University, Department of Economics and The Fletcher School. Email: usman.naeem@tufts.edu.

Contents

A Experimental Design and Implementation

A.1 Sample Recruitment and Construction

Figure A.1: Experimental sample flow and GPS analysis panel construction



Notes: Counts are vehicles. The recruitment frame applies the passenger-four-wheeler, owner-driver, tracking-history, geographic, and non-pilot restrictions to active subscriptions. Baseline records comprise completed baseline surveys before the final respondent and GPS eligibility restrictions. The 25 drivers whose incentivised choices were implemented were removed before the second-stage *Reminder* versus *No-Reminder* randomisation. The GPS analysis panel comprises vehicles in the final six-arm randomised sample observed at least once in the cleaned GPS data between 17 November 2025 and 22 August 2026.

A.2 Treatment Materials

A.2.1 Private Information

Figure A.2: Illustrative *Private Info* treatment message

Falcon-i Weekly Driving Tip

For [PLATE NO.], [MAKE AND MODEL]

Dear [DRIVER NAME],

Early-2000s engines don't need warm-up. Drive sooner to waste less fuel & cut engine wear (US EPA).

You can reduce engine wear & save fuel if you turn your engine off for stops longer than 10 sec while driving (US EPA & DOE).

In recent weeks, you spent 103 min per week (Mon–Sat) with your engine on during stops longer than 30 sec.

Save about Rs 8 on fuel for each min you cut from 103. Max weekly savings are Rs 864. That's Rs 3456 per month.

Notes: The figure reconstructs the participant-facing *Private Info* message from the final implementation script. Identifying fields are shown as placeholders; numerical fields reproduce the values generated for a driver with the median weekly idling benchmark in the treatment-delivery data. Messages were delivered through the partner's mobile app, WhatsApp Business account, and official email platform.

A.2.2 Social Information

Figure A.3: Illustrative *Private & Social Info* treatment message

Falcon-i Weekly Driving Tip

For [Plate No.], [Make and Model]

Dear [Driver Name],

Early-2000s engines don't need warm-up. Drive sooner to waste less fuel, pollute less & cut engine wear (US EPA).

You can reduce engine wear, save fuel & release less CO_2 if you turn your engine off for stops longer than 10 sec while driving (US EPA & DOE).

In recent weeks, you spent 101 min per week (Mon–Sat) with your engine on during stops longer than 30 sec.

Save about Rs 8 on fuel for each min you cut from 101, & help lower environmental harm (e.g., floods, heat) worth about Rs 3. Max weekly savings are Rs 1186. That's Rs 4744 per month.

Engine-on stops emit $PM_{2.5}$, a pollutant that cuts the average Pakistani's life by 3.3 years (Univ of Chicago).

Tailpipe CO_2 adds to flooding

Tailpipe $PM_{2.5}$ shortens lives

Notes: The figure reconstructs the participant-facing *Private & Social Info* message from the final implementation script. Identifying fields are shown as placeholders; numerical fields reproduce the values generated for a driver with the median weekly idling benchmark in the treatment-delivery data. Messages were delivered through the partner's mobile app, WhatsApp Business account, and official email platform.

A.2.3 Payments and *IdleGuard*

Figure A.4: Illustrative *Private & Social Info + Pay + Reminder* treatment message

Falcon-i Weekly Driving Tip

For [PLATE NO.], [MAKE AND MODEL]

Dear [DRIVER NAME],

Early-2000s engines don't need warm-up. Drive sooner to waste less fuel, pollute less & cut engine wear (US EPA).

You can reduce engine wear, save fuel & release less CO_2 if you turn your engine off for stops longer than 10 sec while driving (US EPA & DOE).

In recent weeks, you spent 100 min per week (Mon–Sat) with your engine on during stops longer than 30 sec.

Save about Rs 8 on fuel for each min you cut from 100, & help lower environmental harm (e.g., floods, heat) worth about Rs 4. Max weekly savings are Rs 1179. That's Rs 4716 per month.

For each min you cut from 100 this week (Mon–Sat), we can pay you Rs 4. Max payment for this week: Rs 400. Paid next Mon.

If you forget, *IdleGuard* can remind. It beeps at 30 sec, then every 30 sec, to remind you to turn off the engine if your stop is still over 30 sec.

Engine-on stops emit $PM_{2.5}$, a pollutant that cuts the average Pakistani's life by 3.3 years (Univ of Chicago).

Tailpipe CO_2 adds to flooding

Tailpipe $PM_{2.5}$ shortens lives

Notes: The figure reconstructs the participant-facing week-one *Private & Social Info + Pay + Reminder* message from the final implementation script. Identifying fields are shown as placeholders; numerical fields reproduce the values generated for a driver with the median weekly idling benchmark in the treatment-delivery data. Messages were delivered through the partner's mobile app, WhatsApp Business account, and official email platform.

B Data Construction

B.1 Engine Status and Movement Classification

Ping-level idling classification. In the raw ping-level data, I initially classify a ping as candidate idling when the vehicle's ignition status is on and its recorded speed is zero. I then address two potential measurement concerns. First, because ignition status records whether the vehicle's ignition circuit is powered rather than directly whether the engine is running, I validate it using tracker-recorded battery voltage. Second, because a recorded speed of zero does not guarantee that the vehicle remained stationary between consecutive pings, I use GPS displacement to identify and reclassify false positives arising from slow movement.

Battery-voltage validation. The GPS tracker records a binary ignition-status field indicating whether it detects the vehicle's ignition circuit as active. I use this field as a proxy for whether the engine is running. A potential concern is that the ignition circuit could remain active while the engine is off, for example when a vehicle is in accessory mode. I validate the interpretation of the ignition-status field using battery voltage recorded independently by the tracker.

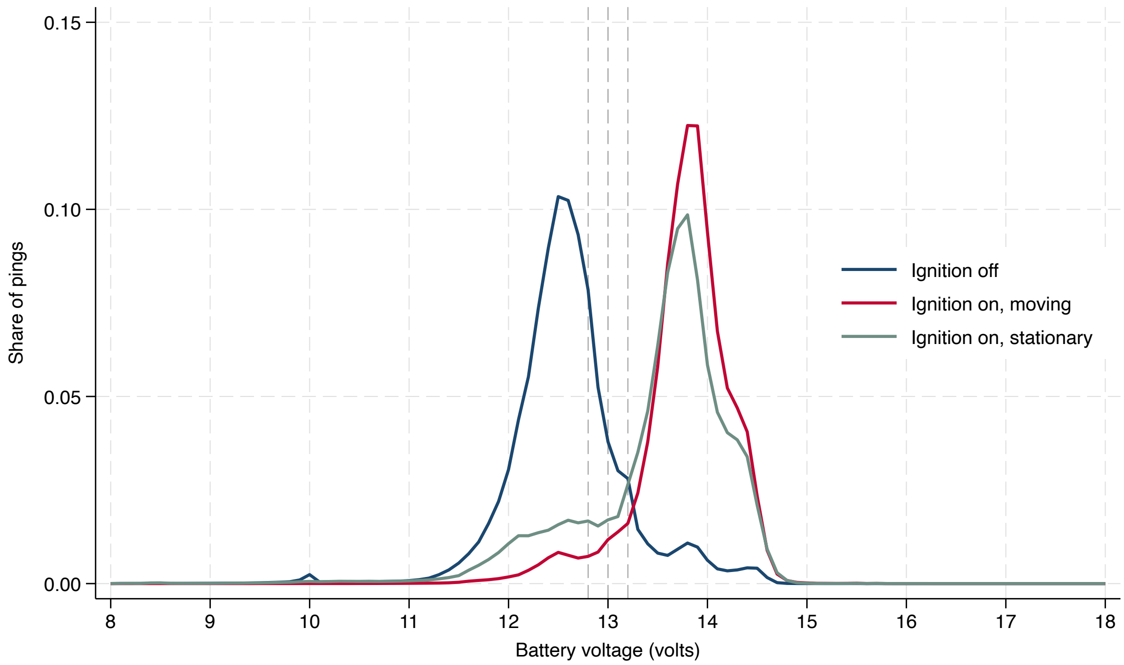
Figure B.1 compares battery-voltage distributions across three mutually exclusive tracker states: ignition off, ignition on while moving, and ignition on while stationary. Moving ignition-on pings provide a benchmark in which the engine is known to be running. The voltage distribution for stationary ignition-on pings closely resembles that of moving ignition-on pings and differs sharply from the ignition-off distribution. In particular, both ignition-on distributions are concentrated in the higher charging-voltage range, whereas ignition-off observations are concentrated at lower voltage.

Table B.1 examines whether fixed battery-voltage cut-offs provide a viable alternative engine-state classification. Battery voltage is observed for 99.5% of pings classified as idling. Depending on the cut-off, voltage screening retains 78.8–85.6% of idling pings and 78.0–84.9% of measured idling minutes. These excluded shares cannot, however, be interpreted as accessory-mode contamination. The same cut-offs also exclude 4.8–10.5% of moving pings, even though the engine must be running while the vehicle is moving. Low voltage therefore does not provide a reliable stand-alone indicator that the engine is off.

Taken together, the battery-voltage evidence validates the ignition-status field as the engine-on measure. Stationary ignition-on observations closely resemble known engine-running moving observations and differ sharply from ignition-off observations. Because universal voltage cut-offs also exclude genuine engine-running observations, I use battery voltage to validate rather than replace the ignition-status measure.

GPS displacement and slow-movement correction. A recorded speed of zero does not necessarily imply that a vehicle remained stationary between consecutive pings. In slow-moving

Figure B.1: Battery voltage by ignition and movement state



Notes: The figure uses ping-level GPS tracker data from the twelve pre-treatment study weeks. Each distribution is normalised separately and reports the share of pings within 0.1-volt bins. Ignition on, moving consists of pings for which ignition status equals one and recorded speed is positive. Ignition on, stationary consists of pings classified as idling under the GPS-adjusted idling definition. The vertical dashed lines mark the alternative battery-voltage cut-offs of 12.8, 13.0, and 13.2 volts examined in Table B.1.

Table B.1: Validation of ignition status using battery voltage

Minimum battery voltage	Idling pings retained	Idling minutes retained	Moving pings excluded
12.8 volts	85.6%	84.9%	4.8%
13.0 volts	82.6%	81.9%	6.3%
13.2 volts	78.8%	78.0%	10.5%
Battery-voltage availability among idling pings		99.5%	

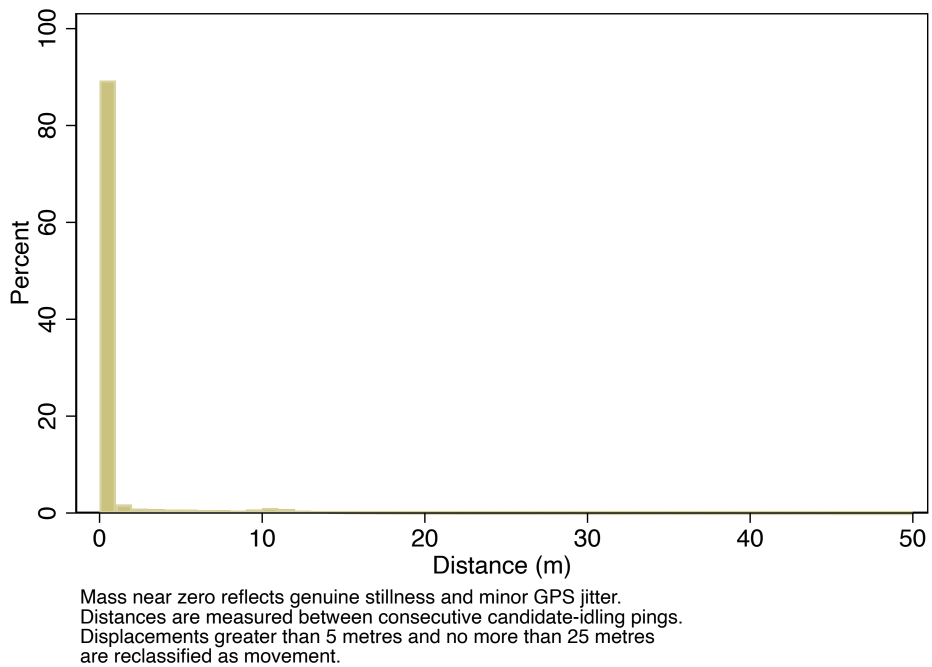
Notes: The table uses ping-level GPS tracker data from the twelve pre-treatment study weeks. Idling pings retained is the share of voltage-observed pings classified as idling whose recorded battery voltage is at or above the indicated cut-off. Idling minutes retained is the corresponding share of measured idling duration. Moving pings excluded is the share of ignition-on pings with positive recorded speed whose battery voltage falls below the indicated cut-off. Because the vehicle is moving in these observations, they provide a benchmark in which the engine is known to be running. Battery-voltage availability is the share of all pings classified as idling with a recorded voltage between 8 and 18 volts.

traffic, for example, the tracker may report zero speed even when the vehicle moves between observations. I therefore use GPS displacement to diagnose and correct potential creeping within the speed-based idling classification.

To calibrate this correction, I use the first complete week of high-frequency GPS data, 17–22 November 2025. For each pair of consecutive candidate-idling pings, I calculate the haversine distance d between their recorded coordinates. Candidate-idling pairs are retained for this diagnostic only when both pings have ignition status on and recorded speed equal to zero, belong to the same vehicle and calendar date, and have valid GPS coordinates. This restriction excludes comparisons that cross vehicle or date boundaries or rely on invalid location records.

Figure B.2 shows the distribution of displacement up to 50 metres. The sharp mass at $d = 0$ indicates that most candidate-idling pairs are genuinely stationary. Small positive displacements are also common and may reflect ordinary GPS jitter rather than actual vehicle movement. I therefore retain observations with displacement of no more than five metres as idling. Displacements satisfying $5 < d \leq 25$ metres are reclassified as movement because they are more consistent with slow rolling or creeping. Values above 25 metres are not automatically reclassified: such large jumps are less plausibly generated by creeping between consecutive candidate-idling pings and may instead reflect GPS or telemetry anomalies.

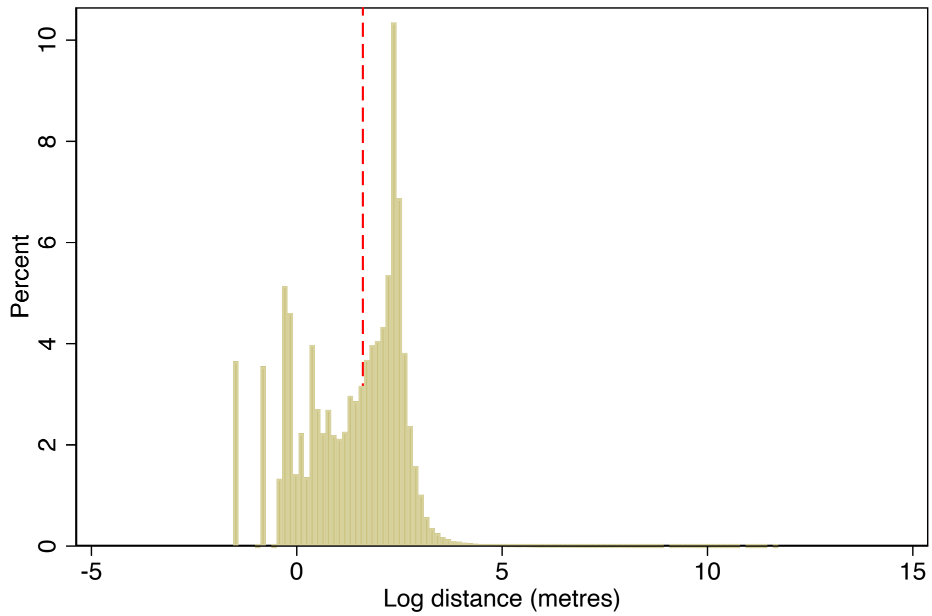
Figure B.2: GPS displacement between consecutive candidate-idling pings



Notes: The figure uses ping-level GPS tracker data from 17–22 November 2025 and reports haversine distance between consecutive candidate-idling pings up to 50 metres. Candidate-idling pings have ignition status on and recorded speed equal to zero. Consecutive pairs are retained only when both pings belong to the same vehicle and calendar date and have valid GPS coordinates. The distribution is measured before applying the GPS-based slow-movement correction. Displacements of no more than five metres are retained as idling, while displacements greater than five and no more than 25 metres are reclassified as movement.

Figure B.3 presents the full displacement distribution on a logarithmic scale. The transformation makes both the dense near-zero region and the much thinner upper tail visible. The figure reinforces the distinction between a narrow GPS-jitter band near zero and larger displacement values that are more consistent with movement. The bounded correction uses this distinction without treating all large GPS jumps as plausible creeping.

Figure B.3: Full distribution of GPS displacement between consecutive candidate-idling pings



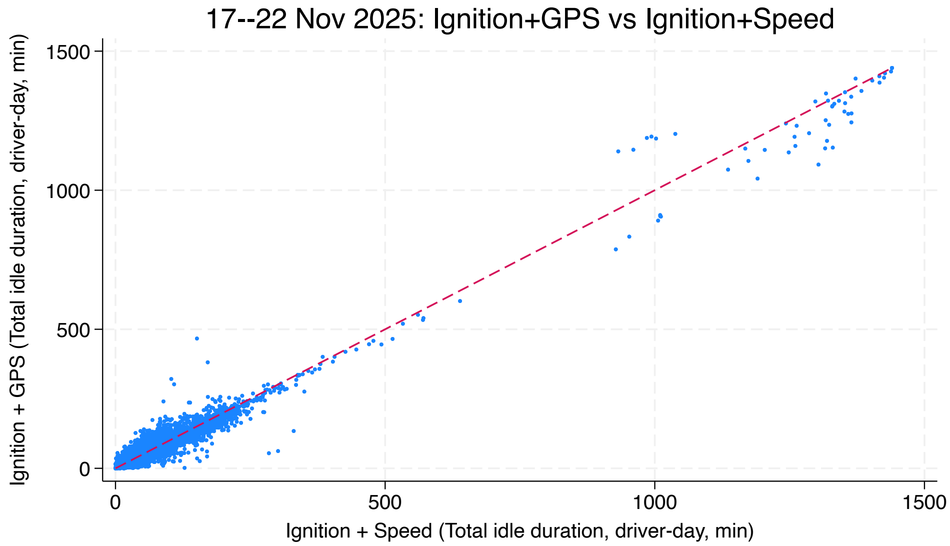
Notes: The figure uses ping-level GPS tracker data from 17–22 November 2025 and reports the log-transformed haversine distance between consecutive candidate-idling pings before applying the GPS-based slow-movement correction. Candidate-idling pings have ignition status on and recorded speed equal to zero. Consecutive pairs are retained only when both pings belong to the same vehicle and calendar date and have valid GPS coordinates. The bounded cleaning rule reclassifies displacements greater than five and no more than 25 metres as movement; larger values are not automatically reclassified.

Comparison with alternative idling definitions. I next assess whether the GPS-based correction materially changes measured idling. Using the same diagnostic week, I construct three alternative vehicle-day measures of total idling duration. The first uses the baseline ignition-and-speed definition, under which ignition-on pings with recorded speed equal to zero are classified as idling. The second uses an ignition-and-GPS benchmark based on realised displacement rather than recorded speed. The third is the preferred cleaned measure, which begins with the ignition-and-speed definition and reclassifies only candidate-idling observations with displacement satisfying $5 < d \leq 25$ metres.

I compare total vehicle-day idling rather than average spell duration because alternative definitions may generate different numbers of detected spells. Vehicle-day totals therefore provide a common outcome that is not mechanically affected by variation in the number of spells.

Figure B.4 compares the ignition-and-GPS benchmark with the baseline ignition-and-speed definition. Observations are concentrated close to the 45-degree line, indicating that the speed-based measure is highly aligned with the displacement-based benchmark. The GPS benchmark nevertheless produces modestly lower idling totals on average, consistent with the ignition-and-speed rule classifying some slow movement as idling.

Figure B.4: Vehicle-day idling under ignition-and-GPS and ignition-and-speed definitions



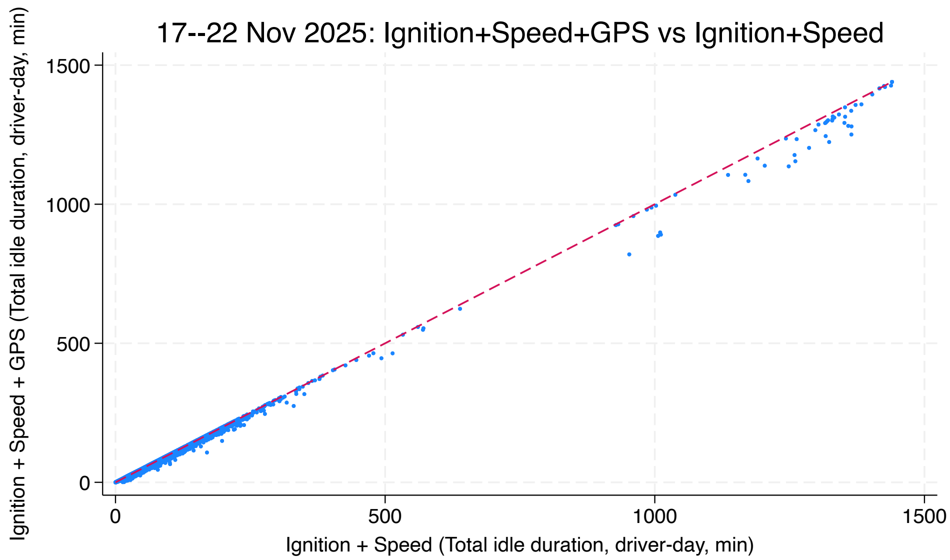
Notes: The figure uses vehicle-day GPS tracker data from 17–22 November 2025. The horizontal axis reports total daily idling duration under the baseline ignition-and-speed definition. The vertical axis reports total daily idling duration under the ignition-and-GPS benchmark. The 45-degree line denotes equality between the two measures.

Figure B.5 compares the preferred cleaned measure with the baseline ignition-and-speed definition. The two measures remain tightly aligned, showing that the GPS correction leaves the baseline measure largely intact while applying a limited downward adjustment to vehicle-days containing likely creeping.

During the diagnostic week, mean vehicle-day total idling is 39.8 minutes under the baseline ignition-and-speed definition, 35.9 minutes under the ignition-and-GPS benchmark, and 38.0 minutes under the preferred cleaned definition. These comparisons support two conclusions. First, the baseline ignition-and-speed measure is broadly valid because it closely tracks the GPS-based benchmark. Second, the bounded correction is not overly aggressive: it removes likely creeping while preserving most of the original ignition-and-speed measure.

The figures in this subsection use one diagnostic week to illustrate and calibrate the measurement correction. The resulting rule is not limited to that week. In the construction of each weekly GPS file, I separately calculate displacement between consecutive candidate-idling pings and reclassify observations with $5 < d \leq 25$ metres as movement. The same correction is therefore applied uniformly throughout the study panel.

Figure B.5: Vehicle-day idling before and after the GPS-based slow-movement correction



Notes: The figure uses vehicle-day GPS tracker data from 17–22 November 2025. The horizontal axis reports total daily idling duration under the baseline ignition-and-speed definition. The vertical axis reports total daily idling duration under the preferred cleaned definition, which reclassifies candidate-idling observations with displacement greater than five and no more than 25 metres as movement. The 45-degree line denotes equality between the two measures.

B.2 Idling Spell Construction

I define a trip as the sequence of pings between an ignition-on event and the subsequent ignition-off event within a vehicle-day. Within each vehicle-day and trip, I group consecutive pings classified as idling into idling spells.

For each idling ping i, observed at time t_i, I assign the elapsed time until the next ping, $t_{i+1} - t_i$, provided that the next ping belongs to the same vehicle-day and trip. The next ping may record either idling or movement because the interval is attributed to the state observed at the current ping. I sum these elapsed-time contributions across the idling pings in each spell to obtain the spell's duration.

I do not assign the interval $t_{i+1} - t_i$ to idling when consecutive pings within the same vehicle-day and trip are separated by more than 30 minutes. Such gaps may reflect interruptions in telemetry rather than continuous idling. This rule does not define the tracker reporting frequency: trackers report at approximately 10-second intervals. It instead prevents a telemetry blackout from being counted as continuous idling.

B.3 Vehicle-Day Panel Construction

I aggregate idling spells within each vehicle-day to construct daily idling duration and spell counts. The primary outcome is daily idling duration from spells longer than 30 seconds. The treatment informed drivers that shutting off the engine during stops longer than 10 seconds generally saves fuel and reduces emissions, while personalised feedback and payments were based on idling spells longer than 30 seconds (U.S. EPA, 2025; U.S. DOE, 2026; Gaines and Rask, 2012). I also construct outcomes using the 10-second threshold and total idling without a minimum spell-duration threshold.

I exclude vehicle-days with more than 18 recorded hours of idling. This distribution-based screen is intended to capture cases in which a tracker appears to remain stuck in a single ignition state.

I manually audit observations in the far-right tail of the resulting vehicle-day idling distribution. The remaining extreme values appear consistent with genuine idling and are retained in the primary measure. As a conservative robustness check, I also winsorise daily idling at the 99.9th percentile; the results are unchanged.

C Experimental Validity and Data Quality

C.1 Additional Balance and Pre-Treatment Comparability

The first-stage randomisation assigned the 6,621-vehicle randomisation sample to five groups. Table C.1 assesses balance across baseline characteristics plausibly related to idling, drawn from the baseline survey and partner administrative data. For each treatment group, the table reports differences relative to *Control*, together with a p-value from a joint test that all displayed differences are zero. The estimated differences are generally small, and none of the joint tests rejects balance at conventional significance levels.

Table C.1: Pre-randomisation balance across first-stage assignment groups

Variable	Private Info	+ Social Info	+ Pay	+ Tech	N	Control mean
1. Age (years)	0.36 (0.49)	0.49 (0.50)	-0.53 (0.48)	0.31 (0.41)	6621	40.86
2. Bachelor's degree or higher (=1)	-0.01 (0.02)	-0.01 (0.02)	0.00 (0.02)	-0.00 (0.02)	6614	0.67
3. Engine size $\leq$1000cc (=1)	-0.04 (0.02)	-0.01 (0.02)	-0.02 (0.02)	-0.02 (0.02)	6505	0.51
4. Stop duration very important for shutoff decision (=1)	0.03 (0.02)	0.02 (0.02)	0.02 (0.02)	0.01 (0.01)	6621	0.19
5. Believes warm-up increases engine life (=1)	0.01 (0.02)	0.03 (0.02)	0.02 (0.02)	0.02 (0.02)	6621	0.75
6. Believes warm-up harms environment (=1)	-0.00 (0.02)	0.01 (0.02)	0.04 (0.02)	0.02 (0.02)	6621	0.38
7. Fuel-saving shutoff threshold $\leq$30 seconds (=1)	0.00 (0.01)	-0.01 (0.01)	-0.00 (0.01)	-0.00 (0.01)	6490	0.07
8. Environmental-harm shutoff threshold $\leq$30 seconds (=1)	0.00 (0.02)	0.01 (0.02)	0.02 (0.02)	0.00 (0.01)	6476	0.15
9. WWF-Pakistan donation out of Rs. 100	-2.91 (1.24)	-3.04 (1.24)	-1.15 (1.19)	-1.84 (0.99)	6504	91.27
Joint p-value	0.219	0.193	0.289	0.462		

Notes: The unit of observation is the vehicle. The sample comprises the 6,621 vehicles in the first-stage randomisation. Treatment columns report differences relative to *Control*, with robust standard errors in parentheses. A plus sign indicates that the named component is added to the preceding package. The *Private & Social Info + Pay + Tech* column denotes the original technology group before the subsequent *Reminder* versus *No-Reminder* randomisation. Regressions adjust for the vehicle-age and main-driver-certainty assignment strata. Indicators equal zero or one. The fuel-saving and environmental-harm indicators equal one for reported shutoff thresholds of 30 seconds or less. WWF-Pakistan donation is the amount donated from a Rs. 100 endowment. The final row reports the p-value from a joint test that all displayed differences equal zero.

The final six experimental arms were formed only after the second-stage randomisation conducted during treatment briefing. Table C.2 therefore assesses their comparability in GPS observability and idling duration from stops longer than 30 seconds over the two clean weeks immediately preceding treatment launch. The table reports differences relative to *Control*, together with group-specific joint-test p-values. The estimated differences are generally small, and none of the joint tests rejects comparability at conventional significance levels.

Table C.3 additionally tests for differential daily dynamics across the final six assignment groups during this pre-treatment window. Table C.4 reports the corresponding focused test for the second-

Table C.2: Pre-main-treatment GPS comparability across assignment groups

Variable	Private Info	+ Social Info	+ Pay	+ Reminder	+ No-Reminder	N	Control mean
1. Share of days observed	0.00 (0.01)	0.02 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.03 (0.01)	6596	0.87
2. Idling minutes from stops >30s	1.58 (1.30)	1.92 (1.36)	1.79 (1.35)	0.69 (1.08)	-0.07 (0.99)	5846	19.45
Joint p-value	0.443	0.186	0.289	0.705	0.153		

Notes: The unit of observation is the vehicle. The sample comprises the 6,596 vehicles in the final six-arm randomised sample. Treatment columns report differences relative to *Control*, with robust standard errors in parentheses. The table uses the two clean immediate pre-main-treatment GPS weeks, starting 23 and 30 March 2026. Share of days observed is computed over all possible vehicle-days, with unobserved days coded as zero; idling minutes from stops longer than 30 seconds are averaged over observed GPS vehicle-days. The *Reminder* and *No-Reminder* columns split the original first-stage technology group; the 25 vehicles whose incentivised choices were implemented are excluded. Regressions adjust for the vehicle-age and main-driver-certainty assignment strata. The final row reports the p-value from a joint test that both displayed differences equal zero.

stage *IdleGuard* versus *No-Reminder* comparison. Neither table provides evidence of differential pre-treatment dynamics.

Table C.3: Immediate pre-main-treatment daily dynamics across assignment groups

Outcome	Vehicles	Vehicle-days	Joint p-value
Observed on day (=1)	6596	79152	0.395
Idling minutes from stops >30s	5840	68980	0.564

Notes: The unit of observation is the vehicle-day. The table uses the two clean immediate pre-main-treatment GPS weeks, starting 23 and 30 March 2026, comprising 12 study days before treatment began on 6 April. Observed on day equals one if the vehicle has at least one retained GPS ping and zero otherwise; idling minutes from stops longer than 30 seconds are computed over observed GPS vehicle-days. For each outcome, the reported p-value tests whether all assignment-group-by-date interactions equal zero in a regression absorbing vehicle and calendar-date fixed effects. *Control* and the first date are omitted. The final six groups incorporate the second-stage *Reminder* versus *No-Reminder* randomisation; the 25 vehicles whose incentivised choices were implemented are excluded. Standard errors are clustered by vehicle.

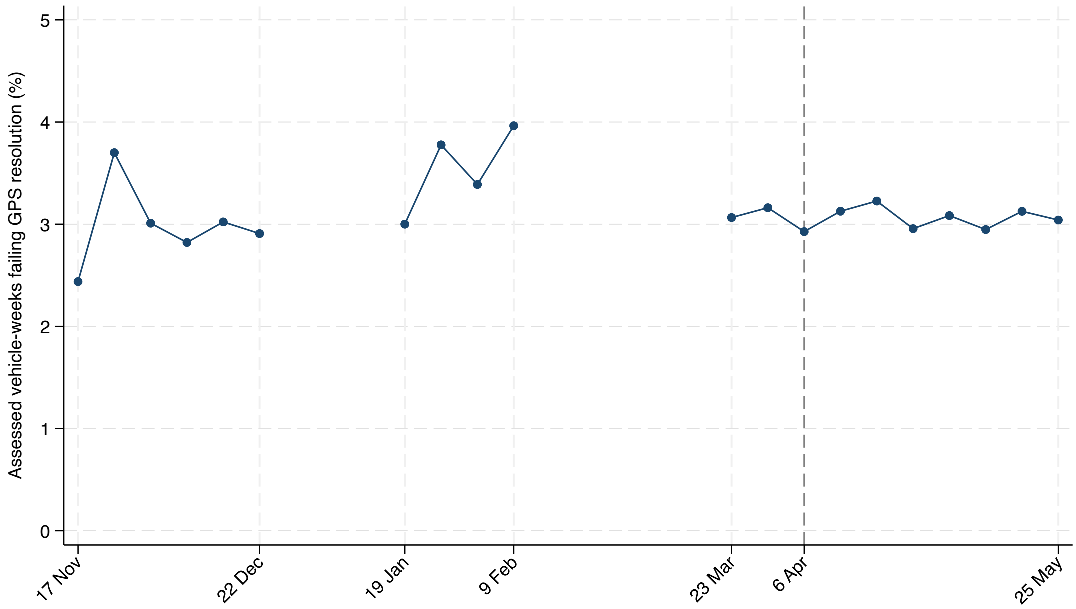
Table C.4: Immediate pre-main-treatment daily dynamics: *IdleGuard* versus *No-Reminder*

Outcome	Vehicles	Vehicle-days	Joint p-value
Observed on day (=1)	2,403	28,836	0.977
Idling minutes from stops >30s	2,198	25,962	0.767

Notes: The unit of observation is the vehicle-day. The sample comprises the *IdleGuard* arm and matched *No-Reminder* arm during the two clean immediate pre-main-treatment GPS weeks, starting 23 and 30 March 2026. Observed on day equals one if the vehicle has at least one retained GPS ping and zero otherwise; idling minutes from stops longer than 30 seconds are computed over observed GPS vehicle-days. For each outcome, the reported p-value tests whether all *IdleGuard*-by-date interactions equal zero in a regression absorbing vehicle and calendar-date fixed effects. *No-Reminder* and the first date are omitted. Standard errors are clustered by vehicle.

C.2 GPS Reporting and Resolution Diagnostics

Figure C.1: Weekly share of assessed vehicle-weeks failing the GPS resolution rule



Notes: The figure reports the weekly share of assessed vehicle-weeks failing the GPS resolution rule, defined as having a 95th-percentile idling ping gap greater than 15 seconds. Assessed vehicle-weeks are observed vehicle-weeks with a non-missing 95th-percentile gap. Disconnected segments denote excluded calendar periods. The dashed vertical line marks the start of weekly main treatment delivery on 6 April 2026.

Figure C.1 shows that the weekly failure rate remains broadly stable around treatment launch. Table C.5 formally tests for differential post-launch changes across assignment groups. The estimated treatment-arm-by-post differences are small, and the joint test does not reject equality across treatment groups relative to the control arm.

Behavioural outcomes remain missing during periods of tracker non-reporting rather than being coded as zero. Differential changes in GPS availability after treatment launch could therefore induce differential selection into the observed vehicle-day sample. Following the GPS data-quality attrition check in Kreindler (2024), Table C.6 examines this possibility using two complementary vehicle-week measures.

The first is a stringent indicator equal to one only if the partner's internal monitoring system never classified the tracker as non-reporting during the week. The second is the share of study days in the week for which the cleaned GPS data contain at least one retained ping. The estimated treatment-arm-by-post differences are small for both measures, and none is individually statistically significant. The omnibus tests likewise fail to reject the joint null of no differential post-launch change in reporting or observability across treatment groups relative to the control arm.

Table C.5: Post-launch treatment-arm differences in GPS resolution failure

	Resolution failure
Private Info × Post	0.001
	(0.005)
Private & Social Info × Post	0.010
	(0.005)
Private & Social Info + Pay × Post	-0.000
	(0.005)
Private & Social Info + Pay + Reminder × Post	-0.001
	(0.005)
Private & Social Info + Pay + No-Reminder × Post	0.000
	(0.005)
Control post-period mean	0.027
Observations	114,170
Vehicles	6,389
Vehicle FE	Yes
Week FE	Yes
Joint p-value	0.354

Notes: The unit of observation is the vehicle-week. The outcome equals one if the vehicle-week's 95th-percentile idling ping gap exceeds 15 seconds. The sample comprises observed vehicle-weeks with a non-missing 95th-percentile gap. The table reports treatment-arm-by-post coefficients from a regression absorbing vehicle and resolution-week fixed effects, with the control arm omitted. Post covers the eight main-treatment resolution weeks, from 6 April through 25 May 2026. Standard errors are clustered by vehicle and shown in parentheses. The joint p-value tests that all displayed coefficients equal zero.

Table C.6: Post-launch treatment-arm differences in tracker reporting and GPS observability

	Partner tracker reporting	Cleaned GPS observability
Private Info × Post	-0.007	-0.008
	(0.008)	(0.007)
Private & Social Info × Post	-0.003	-0.006
	(0.007)	(0.007)
Private & Social Info + Pay × Post	-0.001	-0.010
	(0.008)	(0.007)
Private & Social Info + Pay + Reminder × Post	-0.008	0.003
	(0.007)	(0.006)
Private & Social Info + Pay + No-Reminder × Post	-0.010	0.001
	(0.008)	(0.007)
Control post-period mean	0.929	0.843
Observations	105536	65960
Vehicles	6596	6596
Vehicle FE	Yes	Yes
Week FE	Yes	Yes
Joint p-value	0.694	0.325

Notes: The unit of observation is the vehicle-week. Column 1 equals one if the partner's monitoring system never classified the tracker as non-reporting during the week; Column 2 is the share of possible study days with at least one retained GPS ping. Each column reports treatment-arm-by-post coefficients from a regression absorbing vehicle and week fixed effects, with the control arm omitted. Post begins on 6 April 2026. Column 1 uses partner tracker-status data from 5 February through 24 May; Column 2 uses the two clean immediate pre-main-treatment weeks and the available main-treatment weeks through 25 May. Both samples exclude the 25 vehicles whose incentivised choices were implemented. The control post-period mean is reported separately by outcome. Standard errors are clustered by vehicle and shown in parentheses. The joint p-value tests that all displayed coefficients in the corresponding column equal zero.

C.3 Fuel-Survey Availability and Attrition

I assess differential attrition in the weekly chatbot-based fuel-expenditure survey using two indicators. The first equals one if the respondent answers "Yes" to the weekly consent question and proceeds to the fuel-expenditure question, and zero following an explicit refusal or no response. The second equals one if the subsequent response yields a valid cleaned fuel-expenditure outcome, including a valid report of zero expenditure.

Table C.7 shows that drivers in the three payment-containing arms were more likely than those in the control arm to provide valid fuel-expenditure outcomes during the seven-week survey period. This pattern may reflect confusion among drivers receiving anti-idling payments about whether responding to the fuel survey was connected to those payments, despite the survey being separately incentivised. Alternatively, receiving anti-idling payments may have made the separate fuel-survey incentives more credible.

I therefore do not report fuel-expenditure effects for these arms relative to the control arm or the incremental effect of payments. Table C.8 finds no evidence of differential availability for *Private Info* relative to the control arm or for the incremental effect of *IdleGuard*.

Table C.7: Treatment-arm differences in weekly fuel-survey availability

	Any response	Valid outcome
Private Info	-0.007	-0.010
	(0.011)	(0.011)
Private & Social Info	-0.013	-0.013
	(0.011)	(0.011)
Private & Social Info + Pay	0.045***	0.039***
	(0.012)	(0.012)
Private & Social Info + Pay + Reminder	0.043***	0.037***
	(0.011)	(0.011)
Private & Social Info + Pay + No-Reminder	0.044***	0.037***
	(0.012)	(0.012)
Control mean	0.133	0.122
Vehicle-weeks	46144	46144
Vehicles	6592	6592
Survey-week FE	Yes	Yes
Assigned-day FE	Yes	Yes
Randomisation-stratum FE	Yes	Yes
Joint p-value	0.000	0.000

Notes: The unit of observation is the vehicle-week. Any response equals one if the respondent answers "Yes" to the weekly consent question and proceeds to the fuel-expenditure question, and zero otherwise. Valid outcome equals one if the subsequent response yields a usable cleaned daily fuel-expenditure measure, including a valid report of zero. Each column reports treatment-arm coefficients from a linear probability model with the control arm omitted. Regressions include survey-week, assigned survey-day, and original randomisation-stratum fixed effects. Survey day was randomised across Tuesday through Sunday within treatment group and survey week. The control mean is calculated in the corresponding regression sample. Standard errors are clustered by vehicle and shown in parentheses. The joint p-value tests that all five displayed coefficients equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table C.8: Fuel-survey availability for experimental fuel-effect estimands

	Any response	Valid outcome
Private Info	-0.007	-0.010
	(0.011)	(0.011)
Social information	-0.006	-0.003
	(0.011)	(0.010)
Payments	0.058***	0.052***
	(0.012)	(0.011)
IdleGuard	-0.001	-0.001
	(0.012)	(0.012)
Vehicle-weeks	46144	46144
Vehicles	6592	6592
Survey-week FE	Yes	Yes
Assigned-day FE	Yes	Yes
Randomisation-stratum FE	Yes	Yes
Joint p-value	0.000	0.000

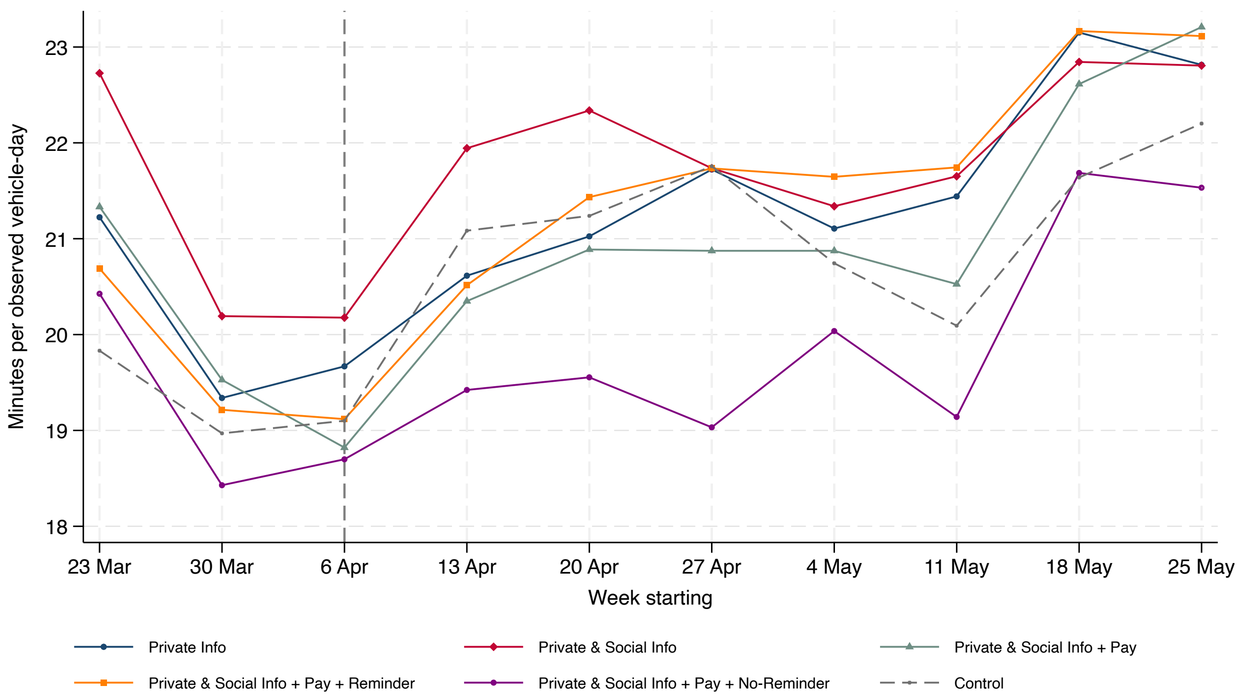
Notes: The unit of observation is the vehicle-week. Outcomes are defined as in Appendix Table C.7. *Private Info* compares that assignment group with the control arm; social information compares *Private & Social Info* with *Private Info*; payments compare *Private & Social Info + Pay* with *Private & Social Info*; and *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. Estimates come from the corresponding six-arm linear probability models in Appendix Table C.7. Regressions include survey-week, assigned survey-day, and original randomisation-stratum fixed effects. Standard errors are clustered by vehicle. The joint p-value tests that all four displayed effects equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

D Additional Experimental Results

D.1 Raw Weekly Trends

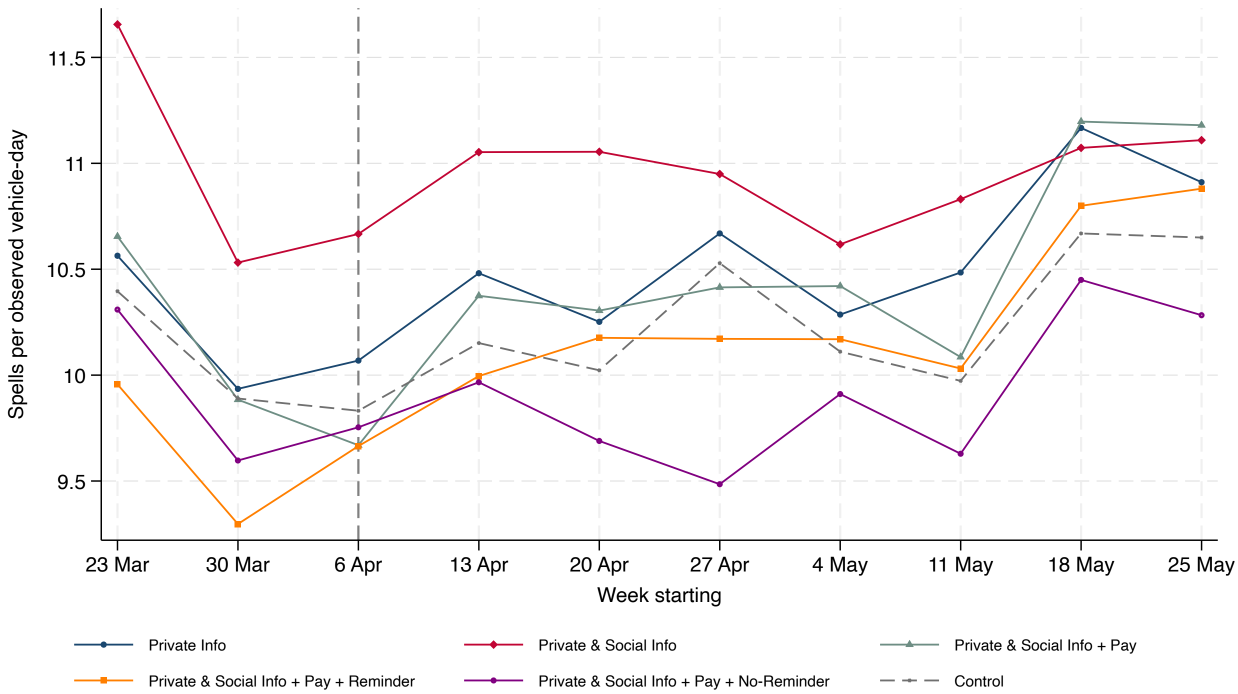
Figures D.1 and D.2 report unadjusted weekly means across assignment groups for daily idling minutes and spell counts from stops longer than 30 seconds.

Figure D.1: Raw weekly trends in idling minutes from stops longer than 30 seconds



Notes: The figure reports unadjusted assignment-group-by-week means of daily idling minutes from stops longer than 30 seconds, computed over observed GPS vehicle-days. It covers the two clean immediate pre-main-treatment weeks and eight main-treatment weeks, from 23 March through 25 May 2026. The dashed vertical line marks the start of weekly main treatment delivery on 6 April 2026. The sample excludes the 25 vehicles whose incentivised choices were implemented before the second-stage *Reminder* versus *No-Reminder* randomisation.

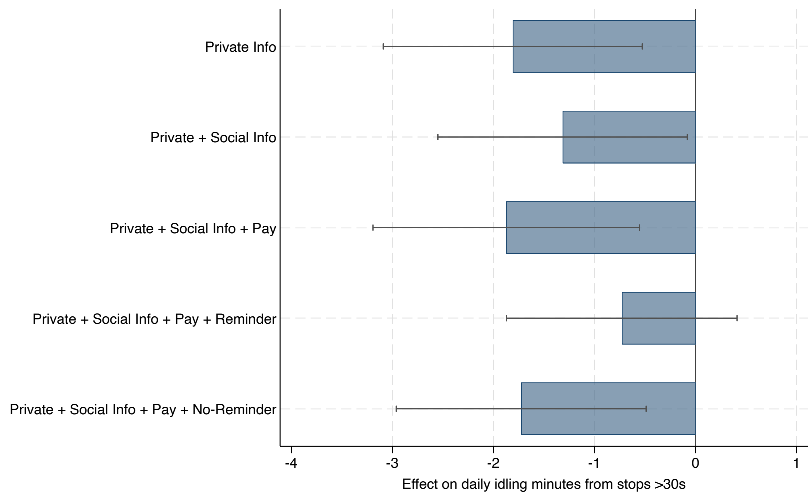
Figure D.2: Raw weekly trends in idling spells from stops longer than 30 seconds



Notes: The figure reports unadjusted assignment-group-by-week means of the daily number of idling spells from stops longer than 30 seconds, computed over observed GPS vehicle-days. It covers the two clean immediate pre-main-treatment weeks and eight main-treatment weeks, from 23 March through 25 May 2026. The dashed vertical line marks the start of weekly main treatment delivery on 6 April 2026. The sample excludes the 25 vehicles whose incentivised choices were implemented before the second-stage *Reminder* versus *No-Reminder* randomisation.

D.2 Additional Main Treatment Effects

Figure D.3: Treatment effects on idling minutes from stops longer than 30 seconds



Notes: The figure reports treatment-arm-by-post coefficients for daily idling minutes from stops longer than 30 seconds, estimated relative to the control arm. Bars show point estimates and capped lines show 95% confidence intervals. Regressions absorb vehicle and calendar-date fixed effects. Post covers the eight main-treatment weeks, from 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle.

Table D.1: Treatment effects across behavioural margins using vehicle-level ANCOVA

	Idling duration (minutes)			Idle spells (count)		Vehicle use	
	Total	>10s	>30s	>10s	>30s	Moving minutes	Trips
Private Info	-2.090**	-2.043**	-1.899***	-0.487	-0.459*	-0.551	-1.256
	(0.827)	(0.808)	(0.737)	(0.940)	(0.237)	(1.674)	(1.174)
Private + Social Info	-0.978	-0.887	-0.687	-0.868	-0.344	-0.003	-1.661
	(0.727)	(0.708)	(0.647)	(0.792)	(0.259)	(1.878)	(1.108)
Private + Social Info + Pay	-1.663**	-1.571**	-1.378**	-0.814	-0.269	-1.337	-1.853
	(0.747)	(0.728)	(0.666)	(0.767)	(0.236)	(1.687)	(1.148)
Private + Social Info + Pay + Reminder	-0.511	-0.415	-0.397	0.702	0.108	-0.330	-1.085
	(0.679)	(0.660)	(0.612)	(1.039)	(0.261)	(1.556)	(1.279)
Private + Social Info + Pay + No-Reminder	-1.910***	-1.857***	-1.694***	-0.901	-0.441**	-2.497	-1.315
	(0.705)	(0.686)	(0.620)	(0.678)	(0.223)	(1.627)	(1.119)
Control treatment-period mean	27.743	26.767	20.986	32.488	10.198	80.207	11.791
Vehicles	5833	5833	5833	5833	5833	5833	5833
Baseline outcome control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Randomisation-stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Vehicle FE	No	No	No	No	No	No	No
Calendar-date FE	No	No	No	No	No	No	No
Joint p-value	0.023	0.020	0.021	0.564	0.192	0.600	0.381

Notes: The unit of observation is the vehicle. Each column reports treatment-arm coefficients from a regression of the vehicle's average outcome during the eight treatment weeks on treatment-arm indicators, its average outcome during the twelve pre-treatment study weeks, and randomisation-stratum fixed effects; the control arm is omitted. Vehicle-level means are computed over observed GPS vehicle-days. The estimation sample in each column includes vehicles with a non-missing outcome in both the pre-treatment and treatment periods. Columns 1–3 report idling duration, Columns 4–5 idling-spell counts, and Columns 6–7 vehicle-use margins; thresholds refer to total spell duration. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. Regressions do not include vehicle or calendar-date fixed effects. Robust standard errors are shown in parentheses. The joint p-value tests whether all displayed treatment-arm coefficients equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.2: Incremental treatment effects across behavioural margins

	Idling duration (minutes)			Idle spells (count)		Vehicle use	
	Total	$>$10s	$>$30s	$>$10s	$>$30s	Moving minutes	Trips
Social information	0.471	0.485	0.493	-0.022	0.053	-0.387	-0.226
	(0.691)	(0.679)	(0.609)	(0.758)	(0.249)	(1.890)	(0.472)
Payments	-0.564	-0.562	-0.557	0.030	0.064	-0.762	-0.228
	(0.712)	(0.698)	(0.629)	(0.768)	(0.244)	(1.940)	(0.476)
IdleGuard	1.050*	1.081*	0.996*	0.891	0.387*	2.075	-0.188
	(0.609)	(0.595)	(0.532)	(0.613)	(0.207)	(1.502)	(1.086)
Control post-period mean	27.721	26.770	20.972	32.590	10.238	80.756	10.567
Observations	701441	701441	701441	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint p-value	0.293	0.251	0.212	0.549	0.291	0.499	0.888

Notes: The unit of observation is the vehicle-day. Social information compares *Private & Social Info* with *Private Info*; payments compare *Private & Social Info + Pay* with *Private & Social Info*; and *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. Regressions absorb vehicle and calendar-date fixed effects. Post covers 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles removed before the second-stage randomisation. Standard errors are clustered by vehicle. The joint p-value tests whether all three displayed effects equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

D.3 Alternative Outcome Definitions and Duration Decompositions

D.3.1 Excess-over-Threshold Outcomes

Appendix Tables D.3–D.4 report alternative duration outcomes that count only time beyond the relevant threshold. The primary outcomes count the full duration of qualifying spells; the excess-over-threshold outcomes instead exclude the first 10 or 30 seconds of each qualifying spell.

Table D.3: Treatment effects using excess-over-threshold idling duration

	All idling	Excess-beyond-threshold duration		Idle spells (count)		Vehicle use	
	Total minutes	$>$10s	$>$30s	$>$10s	$>$30s	Moving minutes	Trips
Private Info × Post	-1.993***	-1.775***	-1.543***	-1.067	-0.532**	-1.321	-0.325
	(0.737)	(0.647)	(0.576)	(0.705)	(0.229)	(1.739)	(0.559)
Private & Social Info × Post	-1.523**	-1.286**	-1.076**	-1.089	-0.479**	-1.707	-0.551
	(0.722)	(0.622)	(0.542)	(0.737)	(0.242)	(1.967)	(0.448)
Private & Social Info + Pay × Post	-2.087***	-1.853***	-1.666***	-1.058	-0.415*	-2.470	-0.779
	(0.756)	(0.667)	(0.601)	(0.716)	(0.223)	(1.793)	(0.563)
Private & Social Info + Pay + Reminder × Post	-0.826	-0.769	-0.697	-0.024	-0.066	-0.470	-0.341
	(0.665)	(0.577)	(0.512)	(0.637)	(0.206)	(1.629)	(1.123)
Private & Social Info + Pay + No-Reminder × Post	-1.876***	-1.701***	-1.499***	-0.915	-0.453**	-2.545	-0.153
	(0.718)	(0.625)	(0.551)	(0.660)	(0.222)	(1.709)	(0.457)
Control post-period mean	27.721	21.339	15.853	32.590	10.238	80.756	10.567
Observations	701441	701441	701441	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint p-value	0.023	0.021	0.024	0.280	0.051	0.568	0.661

Notes: The unit of observation is the vehicle-day. The table repeats the main treatment-arm specification, replacing the duration outcomes in Columns 2–3 with excess-over-threshold measures that exclude the first 10 or 30 seconds of each qualifying spell. Column 1 reports total idling minutes, Columns 4–5 idling-spell counts, and Columns 6–7 vehicle-use margins. Each column reports treatment-arm-by-post coefficients from a regression absorbing vehicle and calendar-date fixed effects; the control arm is omitted. Post covers 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles whose incentivised choices were implemented before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle. The joint p-value tests whether all displayed coefficients equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.4: Incremental treatment effects using excess-over-threshold idling duration

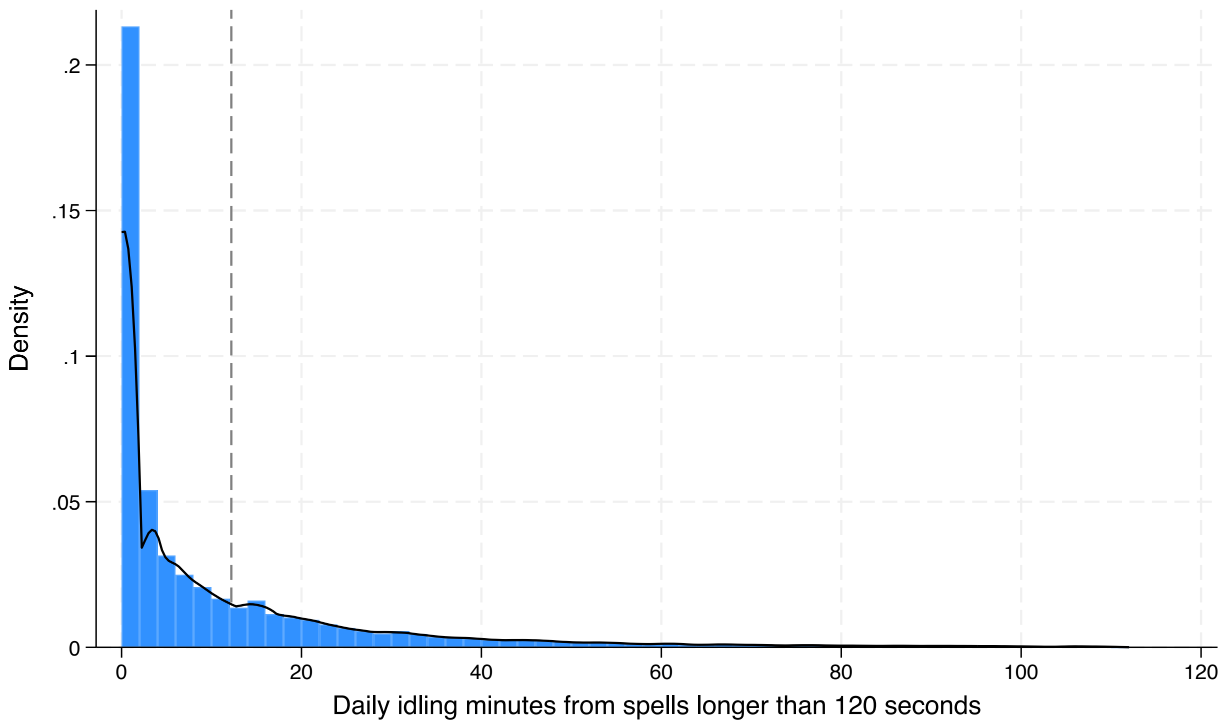
	All idling	Excess-beyond-threshold duration		Idle spells (count)		Vehicle use	
	Total minutes	>10s	>30s	>10s	>30s	Moving minutes	Trips
Social information	0.471	0.489	0.467	-0.022	0.053	-0.387	-0.226
	(0.691)	(0.600)	(0.518)	(0.758)	(0.249)	(1.890)	(0.472)
Payments	-0.564	-0.567	-0.589	0.030	0.064	-0.762	-0.228
	(0.712)	(0.621)	(0.546)	(0.768)	(0.244)	(1.940)	(0.476)
IdleGuard	1.050^{*}	0.932^{*}	0.802^{*}	0.891	0.387^{*}	2.075	-0.188
	(0.609)	(0.527)	(0.457)	(0.613)	(0.207)	(1.502)	(1.086)
Control post-period mean	27.721	21.339	15.853	32.590	10.238	80.756	10.567
Observations	701441	701441	701441	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint p-value	0.293	0.243	0.213	0.549	0.291	0.499	0.888

Notes: The unit of observation is the vehicle-day. The table repeats Appendix Table D.2, replacing the duration outcomes in Columns 2–3 with excess-over-threshold measures that exclude the first 10 or 30 seconds of each qualifying spell. Social information compares *Private & Social Info* with *Private Info*; payments compare *Private & Social Info + Pay* with *Private & Social Info*; and *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. Positive coefficients indicate an increase in the outcome relative to the preceding arm or matched no-reminder arm. The estimation sample, periods, outcome construction, fixed effects, and clustered standard errors follow Appendix Table D.3. The joint p-value tests whether all three displayed incremental effects equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

D.3.2 Duration-Bin and Idling-Context Decompositions

Figure D.4 shows the baseline distribution of daily idling minutes contributed by idling spells longer than 120 seconds. The distribution is highly right-skewed, with a large mass of low-idling vehicle-days and a long right tail.

Figure D.4: Baseline distribution of daily idling minutes from spells longer than 120 seconds



Notes: The figure reports the distribution of daily idling minutes contributed by idling spells longer than 120 seconds, computed over observed GPS vehicle-days. Bars show the histogram, the solid line shows the kernel density, and the dashed vertical line marks the sample mean. The sample uses the two clean immediate pre-main-treatment weeks, starting 23 and 30 March 2026. For readability, the figure displays observations through the 99th percentile of the baseline distribution; the underlying outcome is not winsorised. The sample excludes the 25 vehicles whose incentivised choices were implemented before the second-stage *Reminder* versus *No-Reminder* randomisation.

Table D.5: Treatment effects across idling-spell duration bins

	Idling duration bins (minutes)				
	0–10s	10–30s	30–60s	60–120s	>120s
Private Info × Post	-0.041	-0.143	-0.159*	-0.218**	-1.431**
	(0.033)	(0.131)	(0.085)	(0.104)	(0.563)
Private & Social Info × Post	-0.055*	-0.152	-0.137	-0.241**	-0.937*
	(0.032)	(0.135)	(0.088)	(0.115)	(0.523)
Private & Social Info + Pay × Post	-0.058*	-0.156	-0.105	-0.127	-1.641***
	(0.034)	(0.133)	(0.083)	(0.099)	(0.596)
Private & Social Info + Pay + Reminder × Post	-0.053	-0.043	-0.034	-0.072	-0.624
	(0.040)	(0.117)	(0.077)	(0.092)	(0.502)
Private & Social Info + Pay + No-Reminder × Post	-0.022	-0.128	-0.114	-0.176*	-1.435***
	(0.030)	(0.120)	(0.080)	(0.100)	(0.537)
Control post-period mean	0.950	5.799	3.699	4.220	13.053
Observations	701441	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes
Joint p-value	0.413	0.745	0.344	0.182	0.026

Notes: The unit of observation is the vehicle-day. The columns decompose total daily idling minutes into mutually exclusive bins defined by the total duration of the underlying spell. Each column reports treatment-arm-by-post coefficients from a regression absorbing vehicle and calendar-date fixed effects; the control arm is omitted. Post covers 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles whose incentivised choices were implemented before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle. The joint p-value tests whether all displayed coefficients equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.6: Incremental treatment effects across idling-spell duration bins

	Idling duration bins (minutes)				
	0–10s	10–30s	30–60s	60–120s	>120s
Social information	-0.014	-0.008	0.022	-0.023	0.495
	(0.029)	(0.141)	(0.091)	(0.120)	(0.495)
Payments	-0.002	-0.004	0.033	0.114	-0.704
	(0.030)	(0.143)	(0.089)	(0.116)	(0.532)
IdleGuard	-0.031	0.085	0.081	0.104	0.811*
	(0.035)	(0.112)	(0.075)	(0.095)	(0.442)
Control post-period mean	0.950	5.799	3.699	4.220	13.053
Observations	701441	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes
Joint p-value	0.769	0.899	0.671	0.488	0.144

Notes: The unit of observation is the vehicle-day. Each column reports incremental effects constructed from the corresponding treatment-arm-by-post coefficients in Appendix Table D.5. Social information compares *Private & Social Info* with *Private Info*; payments compare *Private & Social Info + Pay* with *Private & Social Info*; and *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. Positive coefficients indicate increased idling relative to the preceding arm or matched no-reminder arm. The columns decompose total daily idling minutes into mutually exclusive bins defined by the total duration of the underlying spell. The estimation sample, periods, outcome construction, fixed effects, and clustered standard errors follow Appendix Table D.5. The joint p-value tests whether all three displayed incremental effects equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.7: Treatment effects by idling context

	All idling minutes				Idling minutes from stops >30s			
	Total	Warming	Traffic	Waiting	Total	Warming	Traffic	Waiting
Private Info × Post	-1.993***	-0.671**	-0.702**	-0.621*	-1.809***	-0.670**	-0.510**	-0.629*
	(0.737)	(0.296)	(0.347)	(0.341)	(0.654)	(0.283)	(0.246)	(0.341)
Private & Social Info × Post	-1.523**	-0.320	-0.492	-0.711**	-1.316**	-0.299	-0.309	-0.708**
	(0.722)	(0.243)	(0.355)	(0.342)	(0.629)	(0.221)	(0.254)	(0.341)
Private & Social Info + Pay × Post	-2.087***	-0.678**	-0.572*	-0.837**	-1.873***	-0.644**	-0.399*	-0.830**
	(0.756)	(0.310)	(0.338)	(0.367)	(0.673)	(0.294)	(0.232)	(0.366)
Private & Social Info + Pay + Reminder × Post	-0.826	-0.015	-0.204	-0.607*	-0.730	-0.064	-0.059	-0.607*
	(0.665)	(0.270)	(0.305)	(0.317)	(0.582)	(0.243)	(0.214)	(0.316)
Private & Social Info + Pay + No-Reminder × Post	-1.876***	-0.502*	-0.555*	-0.819**	-1.725***	-0.507**	-0.399*	-0.819**
	(0.718)	(0.265)	(0.322)	(0.338)	(0.631)	(0.245)	(0.227)	(0.337)
Control post-period mean	27.721	5.866	13.874	7.981	20.972	4.504	8.939	7.528
Observations	701441	701441	701441	701441	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint p-value	0.023	0.053	0.269	0.197	0.017	0.049	0.157	0.198

Notes: The unit of observation is the vehicle-day. Columns 1–4 report total daily idling minutes and their decomposition into warming, traffic, and waiting idling; Columns 5–8 report the corresponding outcomes for stops longer than 30 seconds. Each column reports treatment-arm-by-post coefficients from a regression absorbing vehicle and calendar-date fixed effects; the control arm is omitted. Post covers 6 April through 30 May 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles whose incentivised choices were implemented before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle. The joint p-value tests whether all displayed coefficients equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.8: Incremental treatment effects by idling context

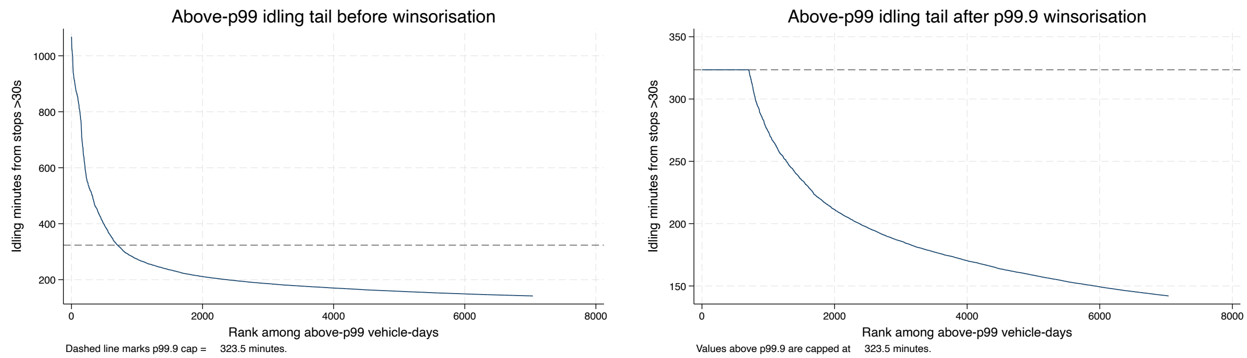
	Idling minutes from stops >30s			
	Total	Warming	Traffic	Waiting
Social information	0.493	0.371	0.201	-0.079
	(0.609)	(0.264)	(0.269)	(0.307)
Payments	-0.557	-0.345	-0.089	-0.123
	(0.629)	(0.276)	(0.256)	(0.335)
IdleGuard	0.996*	0.443*	0.340	0.212
	(0.532)	(0.245)	(0.211)	(0.275)
Control post-period mean	20.972	4.504	8.939	7.528
Observations	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes
Joint p-value	0.212	0.107	0.367	0.812

Notes: The unit of observation is the vehicle-day. The outcomes are daily idling minutes from stops longer than 30 seconds, reported in total and decomposed into warming, traffic, and waiting idling. Social information compares *Private & Social Info* with *Private Info*; payments compare *Private & Social Info + Pay* with *Private & Social Info*; and *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. Positive coefficients indicate increased idling relative to the preceding arm or matched no-reminder arm. The estimation sample, periods, outcome construction, fixed effects, and clustered standard errors follow Appendix Table D.7. The joint p-value tests whether all three displayed incremental effects equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

D.4 Winsorisation and Extreme Values

Figure D.5 shows the far-right tail of the primary outcome before and after winsorising daily idling minutes from stops longer than 30 seconds at the 99.9th percentile. The cap affects only a small number of extreme vehicle-days while leaving the remainder of the distribution unchanged.

Figure D.5: Far-right tail of daily idling minutes from stops longer than 30 seconds



(a) Before winsorisation **(b)** After p99.9 winsorisation

Notes: The figure plots the top 1% of observed vehicle-days ranked by daily idling minutes from stops longer than 30 seconds. The 99th-percentile cut-off, computed over observed vehicle-days in study weeks 1–20, is approximately 142 minutes. Panel A reports the primary outcome before winsorisation. Panel B reports the same vehicle-days after values above the 99.9th percentile are capped at 323.5 minutes. The dashed horizontal line marks this cap.

Tables D.9 and D.10 repeat the main treatment-arm and incremental-effect specifications using outcomes winsorised at their respective 99.9th percentiles. The estimates are similar to the corresponding primary results, showing that the findings are not driven by extreme vehicle-day observations.

Table D.9: Treatment effects across behavioural margins with p99.9-winsorised outcomes

	Idling duration (minutes)			Idle spells (count)		Vehicle use	
	Total	>10s	>30s	>10s	>30s	Moving minutes	Trips
Private Info × Post	-1.729**	-1.687**	-1.523***	-1.197*	-0.482**	-1.262	-0.440
	(0.674)	(0.659)	(0.587)	(0.681)	(0.215)	(1.734)	(0.441)
Private & Social Info × Post	-1.244*	-1.199*	-1.037*	-1.098	-0.449*	-1.669	-0.495
	(0.657)	(0.641)	(0.559)	(0.728)	(0.231)	(1.956)	(0.383)
Private & Social Info + Pay × Post	-1.947***	-1.898***	-1.734***	-1.126	-0.427*	-2.435	-0.583
	(0.704)	(0.689)	(0.619)	(0.711)	(0.222)	(1.786)	(0.458)
Private & Social Info + Pay + Reminder × Post	-0.794	-0.744	-0.689	-0.102	-0.0798	-0.434	-0.431
	(0.608)	(0.593)	(0.526)	(0.624)	(0.203)	(1.625)	(0.463)
Private & Social Info + Pay + No-Reminder × Post	-1.817***	-1.805***	-1.669***	-0.949	-0.477**	-2.484	-0.167
	(0.660)	(0.645)	(0.573)	(0.654)	(0.217)	(1.702)	(0.390)
Control post-period mean	27.560	26.620	20.818	32.576	10.231	80.680	10.414
Observations	701441	701441	701441	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint p-value	0.027	0.024	0.018	0.227	0.055	0.579	0.717

Notes: The unit of observation is the vehicle-day. Each column reports treatment-arm-by-post coefficients from a regression absorbing vehicle and calendar-date fixed effects; the control arm is omitted. Outcomes correspond to the behavioural margins in the main treatment-arm specification, with each winsorised at its 99.9th percentile in the full observed vehicle-day distribution. Post equals one from 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle. The joint p-value tests whether all displayed coefficients equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.10: Incremental treatment effects across behavioural margins with p99.9-winsorised outcomes

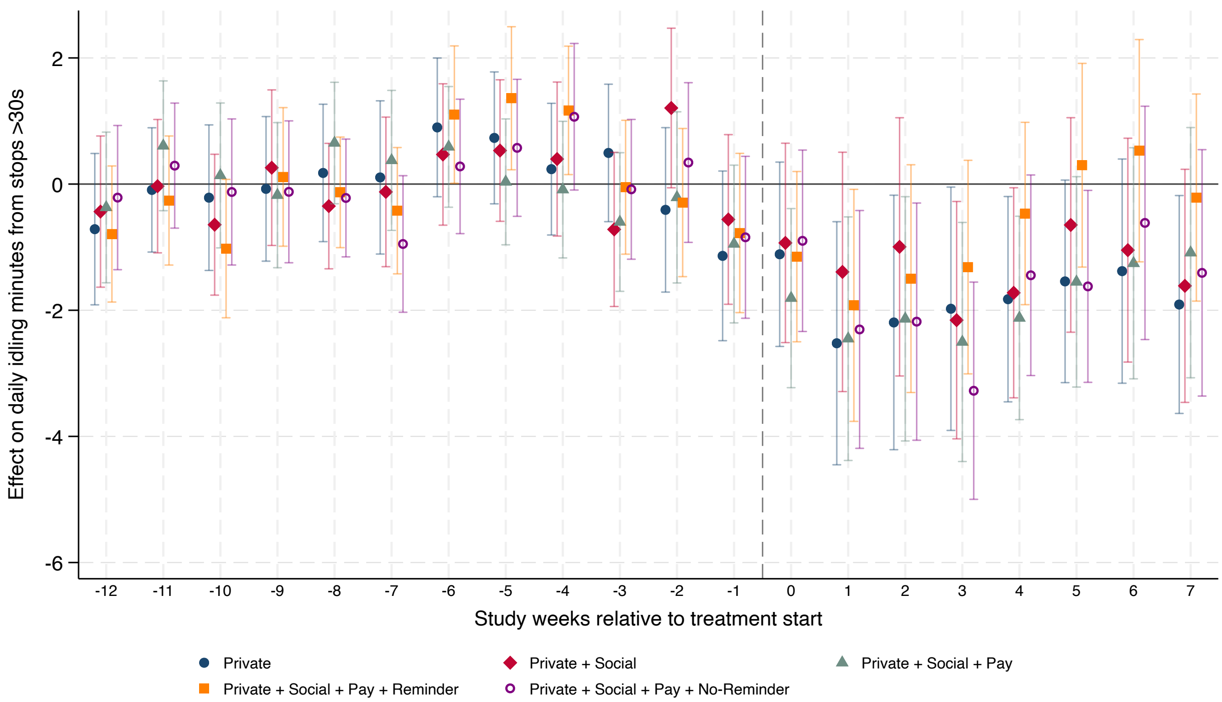
	Idling duration (minutes)			Idle spells (count)		Vehicle use	
	Total	>10s	>30s	>10s	>30s	Moving minutes	Trips
Social information	0.484	0.488	0.486	0.0995	0.0333	-0.407	-0.0552
	(0.657)	(0.642)	(0.562)	(0.731)	(0.227)	(1.880)	(0.388)
Payments	-0.703	-0.699	-0.697	-0.0281	0.0219	-0.766	-0.0885
	(0.687)	(0.672)	(0.596)	(0.760)	(0.234)	(1.929)	(0.407)
IdleGuard	1.023*	1.060*	0.980*	0.848	0.397**	2.049	-0.264
	(0.592)	(0.578)	(0.514)	(0.599)	(0.200)	(1.497)	(0.419)
Control post-period mean	27.560	26.620	20.818	32.576	10.231	80.680	10.414
Observations	701441	701441	701441	701441	701441	701441	701441
Vehicles	6428	6428	6428	6428	6428	6428	6428
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint p-value	0.248	0.208	0.160	0.567	0.262	0.502	0.920

Notes: The unit of observation is the vehicle-day. Social information compares *Private & Social Info* with *Private Info*; payments compare *Private & Social Info + Pay* with *Private & Social Info*; and *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. Outcomes correspond to the behavioural margins in Appendix Table D.2, with each winsorised at its 99.9th percentile in the full observed vehicle-day distribution. The estimation sample, periods, outcome construction, fixed effects, and clustered standard errors follow Table D.9. The joint p-value tests whether all three displayed incremental effects equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

D.5 Additional Dynamics and Persistence

Figure D.6 extends the treatment-dynamics analysis in the main paper to include the twelve pre-treatment study weeks. Table D.11 then reports incremental treatment effects during the main intervention and the two post-treatment periods, while Figure D.7 complements these pooled estimates with week-specific treatment effects after the main intervention ended.

Figure D.6: Treatment dynamics before and during the main intervention



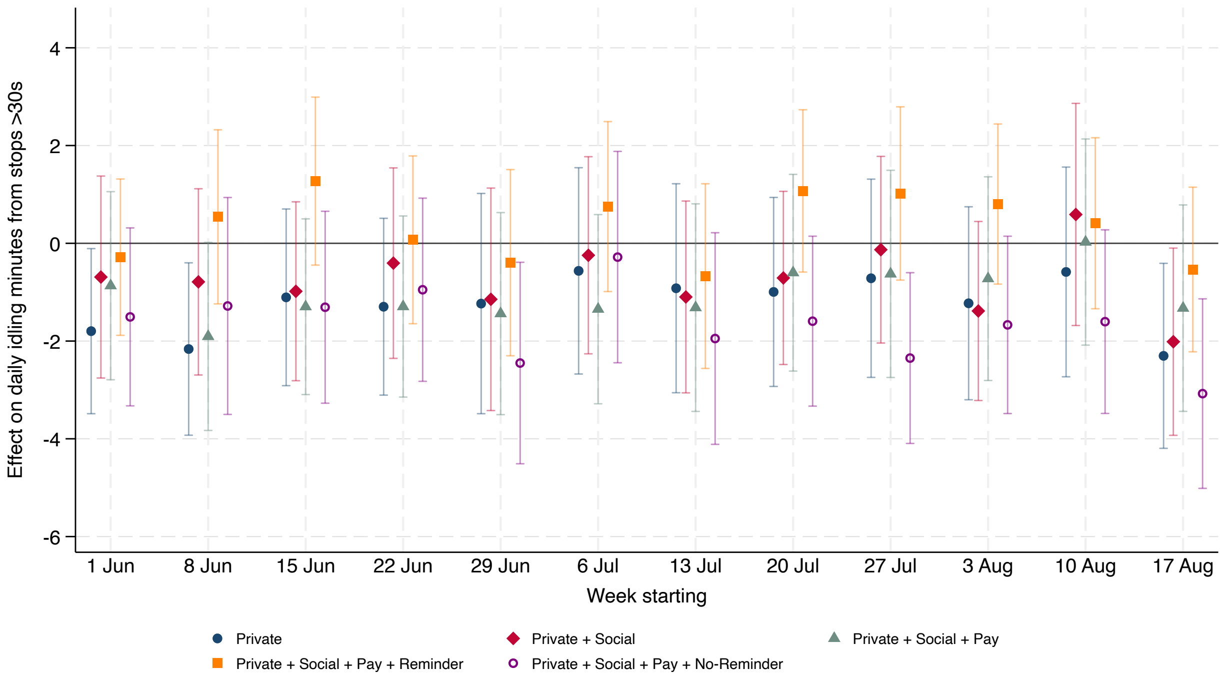
Notes: The figure reports study-week-specific treatment effects on daily idling minutes from stops longer than 30 seconds relative to the control arm. Points report treatment-arm-by-study-week coefficients, and capped lines report 95% confidence intervals. Estimates are normalised relative to the average treatment-control difference across the twelve pre-treatment study weeks, so the displayed pre-treatment coefficients average to zero within each treatment arm. Event time zero denotes the first treatment week, beginning 6 April 2026; the dashed vertical line marks treatment launch. Study weeks index the twelve pre-treatment analysis weeks observed between 17 November 2025 and 4 April 2026 and the eight main-treatment weeks; excluded calendar periods are omitted from the study-week sequence. Estimates are horizontally offset within each study week for readability. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. A joint test that treatment-control differences are constant across the twelve pre-treatment study weeks yields $p = 0.618$.

Table D.11: Incremental treatment effects during and after the main intervention

	Idling minutes from stops >30s		
	Treatment weeks	First 2 post weeks	Remaining post weeks
Social information	0.541	1.269	0.357
	(0.616)	(0.908)	(0.803)
Payments	-0.580	-0.579	-0.235
	(0.629)	(0.988)	(0.795)
IdleGuard	0.948*	1.541*	2.106***
	(0.531)	(0.869)	(0.682)
Control period mean	20.972	22.032	22.070
Observations	1,076,021	1,076,021	1,076,021
Vehicles	6,429	6,429	6,429
Vehicle FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes
Joint p-value	0.229	0.160	0.021

Notes: The unit of observation is the vehicle-day, and the outcome is daily idling minutes from stops longer than 30 seconds. Social information compares *Private & Social Info* with *Private Info*; payments compare *Private & Social Info + Pay* with *Private & Social Info*; and *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. The pooled pre-treatment period comprises study weeks 1–12. The main intervention comprises weeks 13–20, followed by weeks 21–22 and 23–32. The social-information comparison pools Early-Off and Late-On *Private Info* vehicles in both post-treatment periods. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. The joint p-value tests whether all three displayed effects in the corresponding period equal zero. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Figure D.7: Post-treatment dynamics in idling minutes from stops longer than 30 seconds



Notes: The figure reports week-specific treatment effects on daily idling minutes from stops longer than 30 seconds relative to the control arm. Points report treatment-arm-by-week coefficients, and capped lines report 95% confidence intervals. Study weeks 1–12 form the pooled reference period. The displayed weeks begin on 1, 8, 15, 22, and 29 June; 6, 13, 20, and 27 July; and 3, 10, and 17 August 2026. Estimates are horizontally offset within each week for readability. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. The first two *Private Info* estimates pool Early-Off and Late-On vehicles.

At the end of the eight-week main intervention, I randomised half of the *Private Info* group to continue receiving the same messages for two additional weeks. Table D.12 compares these Late-On vehicles with the Early-Off vehicles whose messages stopped after 30 May 2026. A negative coefficient would indicate that continuing the messages reduced idling relative to stopping them. The estimate is close to zero but imprecise, providing no detectable evidence that two additional weeks of messages reduced idling while leaving economically meaningful effects possible.

Table D.12: Effect of continuing Private Info messages after the main intervention

	Idling duration (minutes)		
	>30s	>10s	Beyond first 10s
Late-On × continuation window	0.048	-0.222	-0.019
	(0.881)	(1.036)	(0.880)
Early-Off continuation-period mean	22.765	28.614	23.078
Late-On continuation-period mean	21.408	26.870	21.666
Observations	20,222	20,222	20,222
Vehicles	867	867	867
Early-Off vehicles	428	428	428
Late-On vehicles	439	439	439
Vehicle FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day, and the sample is restricted to *Private Info* vehicles. Columns 1–2 report daily idling minutes from spells longer than 30 and 10 seconds, respectively; Column 3 excludes the first 10 seconds of each spell longer than 10 seconds. Early-Off vehicles stopped receiving messages after 30 May 2026, while Late-On vehicles continued through 13 June. The sample covers study weeks 19–22: weeks 19–20 precede the continuation split, and weeks 21–22 form the continuation period. Each coefficient reports the change for Late-On relative to Early-Off vehicles during the continuation period. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.13 provides a complementary stacked switch-off test. It compares each *Private Info* switch-off cohort with the control arm around the date when its messages stopped. A positive coefficient would indicate that idling rose after switch-off. The estimate is close to zero but imprecise, providing no detectable evidence of rebound while leaving economically meaningful decay possible.

Table D.13: Private Info decay after message switch-off

	Idling duration (minutes)		
	>30s	>10s	Beyond first 10s
Private Info × Post-switch-off	0.164	0.318	0.191
	(0.525)	(0.604)	(0.523)
Control post-switch-off mean	21.219	26.665	21.505
Observations	59,916	59,916	59,916
Vehicles	1,727	1,727	1,727
Early-Off vehicles	429	429	429
Late-On vehicles	437	437	437
Vehicle-by-stack FE	Yes	Yes	Yes
Calendar-date-by-stack FE	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. Columns 1–2 report daily idling minutes from spells longer than 30 and 10 seconds, respectively; Column 3 excludes the first 10 seconds of each spell longer than 10 seconds. Stack 1 compares Early-Off *Private Info* vehicles with the control arm over study weeks 19–22, with messages stopping after 30 May 2026. Stack 2 compares Late-On vehicles with the control arm over weeks 21–24, with messages stopping after 13 June. Each coefficient reports the post-switch-off change for *Private Info* relative to the control arm within the corresponding stack. Regressions absorb vehicle-by-stack and calendar-date-by-stack fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. In Stack 1, control-arm safety messages also ended after 30 May; interpretation assumes that stopping those messages did not affect idling. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.14 reports the Early-Off *Private Info* effect during the active campaign and after switch-off. The pooled twelve-week post-switch-off estimate provides the persistence input used in the cost–benefit calculation.

Table D.14: Private Info effects during and after an eight-week campaign

	Idling duration (minutes)		
	>30s	>10s	Beyond first 10s
Active campaign, weeks 13–20	-2.394***	-2.480***	-2.332***
	(0.852)	(0.926)	(0.843)
First two weeks after switch-off, weeks 21–22	-2.638***	-2.488**	-2.467***
	(0.915)	(1.041)	(0.903)
Remaining ten weeks after switch-off, weeks 23–32	-1.926*	-2.131*	-1.873*
	(1.084)	(1.171)	(1.061)
All twelve weeks after switch-off, weeks 21–32	-2.050**	-2.193**	-1.976**
	(0.989)	(1.069)	(0.967)
Control mean, active campaign	20.972	26.770	21.339
Control mean, first two weeks after	22.032	27.654	22.298
Control mean, remaining ten weeks	22.070	27.571	22.460
Control mean, all twelve weeks after	22.063	27.586	22.432
Observations	251,288	251,288	251,288
Vehicles	1,512	1,512	1,512
Early-Off vehicles	503	503	503
Vehicle FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. The sample comprises Early-Off *Private Info* and control-arm vehicles. Columns 1–2 report daily idling minutes from spells longer than 30 and 10 seconds, respectively; Column 3 excludes the first 10 seconds of each spell longer than 10 seconds. The active-campaign period comprises study weeks 13–20, followed by weeks 21–22 and 23–32 after switch-off. The pooled post-switch-off estimate combines weeks 21–32 in a separate specification. Each coefficient reports the Early-Off *Private Info* effect relative to the control arm, using weeks 1–12 as the pooled reference period. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. The pooled twelve-week estimate provides the persistence input for the cost–benefit calculation. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

D.6 Fuel-Expenditure Effects and Robustness

Appendix Tables D.15 and D.16 report fuel-expenditure effects for the estimands that do not exhibit differential valid-outcome availability in Appendix Tables C.7 and C.8. Table D.15 reports the primary estimates, while Table D.16 assesses robustness to winsorising the outcome at the 99.9th percentile.

Table D.15: Effects on daily fuel expenditure

	Daily fuel expenditure (Rs.)
Panel A. Treatment effects	
Private Info	-872.946**
	(383.668)
Private & Social Info	-781.747*
	(407.416)
Control mean	3730.432
Vehicle-weeks	2,474
Vehicles	880
Joint p-value	0.067
Panel B. Incremental effects	
Social information	91.199
	(312.378)
IdleGuard	3.946
	(247.572)
No-Reminder mean	3073.208
IdleGuard-comparison vehicle-weeks	2,750
IdleGuard-comparison vehicles	825
Survey-week FE	Yes
Assigned-day FE	Yes
Randomisation-stratum FE	Yes

Notes: The unit of observation is the vehicle-week. The outcome is cleaned fuel expenditure, in Pakistani rupees, for the randomly assigned survey day. Panel A reports effects for *Private Info* and *Private & Social Info* relative to the control arm. Panel B reports the incremental effects of social information and *IdleGuard*. Social information compares *Private & Social Info* with *Private Info*; *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. Payment-containing treatment effects relative to the control arm and the incremental effect of payments are omitted because assignment changes valid-outcome availability for those comparisons. Regressions include survey-week, assigned-day, and original randomisation-stratum fixed effects. Standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table D.16: Effects on daily fuel expenditure with a p99.9-winsorised outcome

	Daily fuel expenditure (Rs.)
Panel A. Treatment effects	
Private Info	-707.158**
	(320.061)
Private & Social Info	-616.264*
	(348.422)
Control mean	3559.409
Vehicle-weeks	2,474
Vehicles	880
Joint p-value	0.072
Panel B. Incremental effects	
Social information	90.894
	(311.868)
IdleGuard	24.699
	(240.483)
No-Reminder mean	3045.635
IdleGuard-comparison vehicle-weeks	2,750
IdleGuard-comparison vehicles	825
P99.9 cap (Rs.)	50,000
Survey-week FE	Yes
Assigned-day FE	Yes
Randomisation-stratum FE	Yes

Notes: The outcome is cleaned fuel expenditure for the randomly assigned survey day, winsorised at the 99.9th percentile of all non-missing observations in the full six-arm fuel-survey panel. Values above the reported cap are set equal to the cap; no observations are dropped. The retained estimands, samples, regression specifications, fixed effects, and clustered standard errors follow Appendix Table D.15. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

E Additional Mechanism and Heterogeneity Results

E.1 Belief Correction and Decision Rules

Table E.1 summarises baseline beliefs about the minimum stop duration for which shutting off the engine saves fuel or reduces environmental harm. Tables E.3 and E.4 then examine whether treatment responsiveness varies with drivers' stop-duration decision rules and baseline fuel-saving threshold beliefs.

Table E.1: Baseline beliefs about when turning off the engine saves fuel and reduces environmental harm

Belief domain	N	Threshold (minutes)		Share of respondents (%)	
		Mean	Median	$\leq$ 10 seconds	$>$ 30 seconds
Fuel savings	6,466	6.60	5.00	3.1	93.6
Environmental harm	6,452	5.01	2.00	8.4	84.6

Notes: The unit of observation is the vehicle. The sample comprises the 6,596 vehicles in the final six-arm randomised sample, excluding the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. The table reports baseline beliefs about the minimum stop duration for which turning off the engine saves fuel or reduces environmental harm. Thresholds are reported in minutes. The final two columns report the shares stating a threshold of 10 seconds or less and greater than 30 seconds.

Table E.2: *Private Info* effects by baseline beliefs

	Idling minutes from stops $>$ 30s		
	Warm-up engine life	Fuel-saving threshold	Threshold + confidence
Less mistaken	-1.660	-1.204	-1.741**
	(1.084)	(0.803)	(0.694)
More mistaken	-1.934***	-2.310***	-2.333**
	(0.714)	(0.791)	(1.117)
Difference: more minus less	-0.275	-1.106	-0.592
	(1.099)	(0.898)	(1.124)
Control post-period mean, less mistaken	19.600	20.858	21.067
Control post-period mean, more mistaken	21.479	21.097	20.472
Observations	213670	216447	216447
Vehicles	1969	1989	1989
Vehicle FE	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day, and the outcome is daily idling minutes from stops longer than 30 seconds. Each column reports the effect of *Private Info* relative to the control arm separately for drivers with less- and more-mistaken baseline beliefs. In Column 1, more mistaken denotes believing that warming up the engine increases engine life; the comparison group comprises drivers who reported that warming up decreases engine life or has no effect, with unsure responses excluded. In Column 2, more mistaken denotes reporting a fuel-saving shutoff threshold of at least 300 seconds; Column 3 additionally requires being confident or very confident in that belief. Difference rows test whether the *Private Info* effect differs between more- and less-mistaken drivers. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.3: Treatment effects by stop-duration decision rule

	Idling minutes from stops >30s
Panel A. Stop duration not very important	
Private Info	-1.979***
	(0.702)
Social information	0.450
	(0.694)
Payments	-0.958
	(0.716)
IdleGuard	0.931
	(0.601)
Panel B. Stop duration very important	
Private Info	-1.181
	(1.007)
Social information	0.656
	(1.259)
Payments	0.991
	(1.292)
IdleGuard	1.253
	(1.140)
Panel C. Very important minus not very important	
Private Info	0.798
	(1.026)
Social information	0.206
	(1.438)
Payments	1.949
	(1.477)
IdleGuard	0.322
	(1.289)
Control post-period mean, not very important	20.868
Control post-period mean, very important	21.420
Observations	701441
Vehicles	6428
Vehicle FE	Yes
Calendar-date FE	Yes

Notes: The unit of observation is the vehicle-day, and the outcome is daily idling minutes from stops longer than 30 seconds. The table reports the *Private Info* effect relative to the control arm and the incremental effects of social information, payments, and *IdleGuard*. The heterogeneity indicator equals one if the respondent reports that expected stop duration is very important for deciding whether to keep the engine on or turn it off. The comparison group comprises respondents with valid answers who report that duration is important, not important, or not at all important. Difference rows test whether each effect varies with the decision rule. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Respondents with missing or invalid baseline responses are excluded. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.4: Treatment effects by long fuel-saving threshold belief

	Idling minutes from stops >30s
Panel A. Fuel-saving threshold below 300 seconds	
Private Info	-1.207
	(0.803)
Social information	-0.856
	(0.949)
Payments	-0.625
	(1.087)
IdleGuard	2.285***
	(0.746)
Panel B. Fuel-saving threshold at least 300 seconds	
Private Info	-2.313***
	(0.791)
Social information	1.661**
	(0.798)
Payments	-0.623
	(0.758)
IdleGuard	-0.0170
	(0.761)
Panel C. Long threshold minus short threshold	
Private Info	-1.106
	(0.898)
Social information	2.517**
	(1.240)
Payments	0.00139
	(1.325)
IdleGuard	-2.302**
	(1.066)
Control post-period mean, threshold below 300s	20.858
Control post-period mean, threshold at least 300s	21.097
Observations	687912
Vehicles	6303
Vehicle FE	Yes
Calendar-date FE	Yes

Notes: The unit of observation is the vehicle-day, and the outcome is daily idling minutes from stops longer than 30 seconds. The table reports the *Private Info* effect relative to the control arm and the incremental effects of social information, payments, and *IdleGuard*. A long fuel-saving threshold is defined as reporting that the engine must be turned off for at least 300 seconds to save fuel. The comparison group comprises respondents with valid thresholds below 300 seconds. Difference rows test whether each effect varies with the baseline threshold belief. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

E.2 Social-Information Mechanisms

Table E.5 examines whether social-information effects vary with baseline environmental-harm misperceptions. Table E.6 then decomposes the incremental effect of social information across warming, traffic, and waiting idling.

Table E.5: Social-information effects by environmental-harm misperception

	Idling minutes from stops >30s	
	Threshold at least 300s	Threshold at least 300s + confidence
Panel A. Less mistaken		
Private Info	-1.841**	-1.737**
	(0.805)	(0.689)
Private & Social Info	-2.453***	-1.555**
	(0.774)	(0.667)
Social information	-0.613	0.182
	(0.884)	(0.674)
Panel B. More mistaken		
Private Info	-1.826**	-2.562**
	(0.780)	(1.229)
Private & Social Info	0.103	0.268
	(0.731)	(0.950)
Social information	1.930**	2.830**
	(0.823)	(1.395)
Panel C. Difference: more minus less		
Private Info	0.014	-0.826
	(0.888)	(1.233)
Private & Social Info	2.557***	1.823*
	(0.817)	(0.938)
Social information	2.542**	2.648*
	(1.207)	(1.549)
Control post-period mean, less mistaken	20.244	20.747
Control post-period mean, more mistaken	22.048	23.363
Observations	323763	323763
Vehicles	2964	2964
Vehicle FE	Yes	Yes
Calendar-date FE	Yes	Yes

Notes: The unit of observation is the vehicle-day, and the outcome is daily idling minutes from stops longer than 30 seconds. Column 1 defines more mistaken as reporting an environmental-harm shutoff threshold of at least 300 seconds; Column 2 additionally requires being confident or very confident in that belief. Panels A and B report effects for less- and more-mistaken drivers, respectively, and Panel C reports differences between them. Within each panel, rows report the effects of *Private Info* and *Private & Social Info* relative to the control arm and the incremental effect of social information. Regressions absorb vehicle and calendar-date fixed effects. Post covers 6 April through 30 May 2026. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.6: Social-information effects by environmental-harm threshold belief and idling context

	Total	Warming	Traffic	Waiting
Social information: threshold below 300s	-0.613	0.141	-0.280	-0.474
	(0.884)	(0.415)	(0.407)	(0.425)
Social information: threshold at least 300s	1.931**	0.664**	0.796**	0.471
	(0.822)	(0.288)	(0.331)	(0.453)
At least 300s minus below 300s	2.544**	0.523	1.076**	0.945
	(1.207)	(0.505)	(0.525)	(0.621)
Control post-period mean, threshold below 300s	20.244	4.674	8.686	6.884
Control post-period mean, threshold at least 300s	22.048	4.256	9.351	8.442
Observations	685895	685895	685895	685895
Vehicles	6286	6286	6286	6286
Vehicle FE	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. Outcomes are daily idling minutes from stops longer than 30 seconds, reported in total and decomposed into warming, traffic, and waiting idling. The table reports the incremental effect of social information, comparing *Private & Social Info* with *Private Info*, separately by whether the respondent reports an environmental-harm shutoff threshold of at least 300 seconds. The comparison group comprises respondents with valid thresholds below 300 seconds. Difference rows test whether the social-information effect varies with the threshold belief. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.7: Effects of social information on endline environmental beliefs

	Endline environmental beliefs	
	Believes warm-up harms environment	Environmental-harm shutoff threshold $\leq$ 30 seconds
Social information	-0.031	-0.010
	(0.024)	(0.019)
Private Info mean	0.451	0.169
Observations	1570	1552
Baseline outcome control	Yes	Yes

Notes: The unit of observation is the vehicle. The sample comprises vehicles assigned to *Private Info* or *Private & Social Info* with a usable endline outcome and corresponding baseline belief. Column 1 equals one if the respondent believes that warming up the engine harms the environment; Column 2 equals one if the respondent reports an environmental-harm shutoff threshold of 30 seconds or less. Each column reports the incremental effect of social information, comparing *Private & Social Info* with *Private Info*, and adjusts for the corresponding baseline belief. Endline availability does not differ significantly between the two groups for either outcome. Robust standard errors are shown in parentheses. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.8: Endline diagnostics for social-information mechanisms

	Correct recall	Low action relevance	Too much information	More likely to follow	Chooses messages
Social information	-0.010	-0.018	0.009	0.047	0.048
	(0.007)	(0.032)	(0.016)	(0.038)	(0.031)
Private Info mean	0.028	0.255	0.043	0.531	0.626
Observations	1584	740	740	696	951

Notes: The unit of observation is the vehicle. Each column reports the incremental effect of social information, comparing *Private & Social Info* with *Private Info*, on the indicated endline outcome. Correct 10-second recall equals one if the respondent correctly recalls the private-cost shutoff threshold; respondents who report not remembering or not knowing the threshold are coded as zero. Low action relevance equals one if the respondent describes the messages as not requiring personal action, containing too much information, being too general to act on, or as telling them what to do. More likely to follow and chooses messages over Rs. 250 are indicators for the corresponding survey responses. Columns 2–4 are observed among respondents routed to the message-reader follow-up questions; Column 5 is observed among respondents eligible for the incentivised continued-message choice. Robust standard errors are shown in parentheses. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.9: Social-information effects on message credibility by baseline environmental prior

	Low credibility
Panel A. Baseline environmental-harm threshold below 300 seconds	
Social information	-0.046***
	(0.016)
Panel B. Baseline environmental-harm threshold at least 300 seconds	
Social information	-0.001
	(0.017)
Panel C. At least 300 seconds minus below 300 seconds	
Social information	0.045*
	(0.024)
Private Info mean, threshold below 300s	0.056
Private Info mean, threshold at least 300s	0.022
Observations	736

Notes: The unit of observation is the vehicle. Low credibility equals one if the respondent reports that the treatment messages were not very believable or not believable at all. The outcome is observed among respondents who reported reading all, most, or some of the treatment messages. Panels A and B report the incremental effect of social information, comparing *Private & Social Info* with *Private Info*, for drivers whose baseline environmental-harm shutoff threshold was below and at least 300 seconds, respectively. Panel C reports the corresponding at-least-300-seconds-minus-below-300-seconds difference in the social-information effect. Robust standard errors are shown in parentheses. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

E.3 Payment Mechanisms

Table E.10 assesses whether pooling the two paid arms without an active reminder improves precision for estimating the incremental effect of payments.

Table E.10: Pooling diagnostic for payment effects across behavioural margins

	Idling duration (minutes)			Idle spells (count)		Vehicle use	
	Total	>10s	>30s	>10s	>30s	Moving minutes	Trips
Payments	-0.564	-0.562	-0.557	0.030	0.064	-0.762	-0.228
	(0.712)	(0.698)	(0.629)	(0.768)	(0.244)	(1.940)	(0.476)
Pooled paid no-reminder arms	-1.981***	-1.941***	-1.799***	-0.986*	-0.434**	-2.506	-0.465
	(0.647)	(0.628)	(0.571)	(0.595)	(0.192)	(1.537)	(0.452)
Incremental effect of pooled payments	-0.458	-0.474	-0.483	0.101	0.045	-0.798	0.087
	(0.595)	(0.583)	(0.519)	(0.656)	(0.216)	(1.707)	(0.338)
p-value, paid-arm effects equal	0.765	0.800	0.814	0.837	0.865	0.965	0.196
Control post-period mean	27.721	26.770	20.972	32.590	10.238	80.756	10.567
Private & Social Info post-period mean	28.742	27.808	21.845	33.730	10.918	84.082	9.442
Pooled paid no-reminder arms post-period mean	27.054	26.128	20.434	32.039	10.169	80.059	9.913
Observations	434,844	434,844	434,844	434,844	434,844	434,844	434,844
Vehicles	3,966	3,966	3,966	3,966	3,966	3,966	3,966
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. The table assesses whether pooling the two paid arms without an active reminder improves precision for the incremental effect of payments. Payments compare *Private & Social Info + Pay* with *Private & Social Info*; pooled payments compare the pooled *Private & Social Info + Pay* and *Private & Social Info + Pay + No-Reminder* arms with *Private & Social Info*. The reminder arm is excluded because activating *IdleGuard* changes the attention and self-monitoring environment. The reported p-value tests whether the effects of the two paid arms are equal. Regressions absorb vehicle and calendar-date fixed effects. Post equals one from 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

E.4 *IdleGuard* Beliefs and Baseline Behaviour

Tables E.11–E.13 examine whether the incremental effect of *IdleGuard* varies with baseline beliefs and warming behaviour and how any differences are distributed across warming, traffic, and waiting idling.

Table E.11: *IdleGuard* effects by strong fuel-saving threshold misperception and idling context

	Total	Warming	Traffic	Waiting
Incremental effect of IdleGuard: no strong misperception	2.154***	1.005***	0.430	0.719*
	(0.737)	(0.369)	(0.290)	(0.412)
Incremental effect of IdleGuard: strong misperception	0.0547	-0.0611	0.278	-0.162
	(0.771)	(0.335)	(0.309)	(0.374)
Difference: strong minus no strong misperception	-2.100**	-1.066**	-0.152	-0.882
	(1.066)	(0.499)	(0.424)	(0.556)
Control post-period mean: no strong misperception	20.872	4.755	8.521	7.597
Control post-period mean: strong misperception	21.091	4.296	9.292	7.502
Observations	687912	687912	687912	687912
Vehicles	6303	6303	6303	6303
Vehicle FE	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. Outcomes are daily idling minutes from stops longer than 30 seconds, reported in total and decomposed into warming, traffic, and waiting idling. The table reports the incremental effect of *IdleGuard*, comparing *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm, separately by strong fuel-saving threshold misperception. Strong fuel-saving threshold misperception is defined as reporting that the engine must be turned off for at least 300 seconds to save fuel and being confident or very confident in that answer. The comparison group comprises respondents with valid fuel-saving threshold and confidence responses who do not satisfy this definition. Difference rows test whether the *IdleGuard* effect varies with strong fuel-saving threshold misperception. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.12: *IdleGuard* effects by warm-up engine-life belief and idling context

	Idling minutes from stops >30s			
	Total	Warming	Traffic	Waiting
Incremental effect of IdleGuard: no warm-up belief	0.549	1.068	-0.134	-0.384
	(1.460)	(0.826)	(0.493)	(0.747)
Incremental effect of IdleGuard: warm-up belief	0.893	0.205	0.392*	0.296
	(0.548)	(0.219)	(0.235)	(0.281)
Difference: belief minus no belief	0.344	-0.863	0.526	0.681
	(1.559)	(0.855)	(0.546)	(0.797)
No-reminder post-period mean: no belief	20.934	4.797	8.829	7.308
No-reminder post-period mean: belief	19.640	4.276	8.625	6.740
Observations	247734	247734	247734	247734
Vehicles	2139	2139	2139	2139
Vehicle FE	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. Outcomes are daily idling minutes from stops longer than 30 seconds, reported in total and decomposed into warming, traffic, and waiting idling. The table reports the incremental effect of *IdleGuard*, comparing *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm, separately by whether the respondent reported at baseline that warming up the engine before driving increases engine life. The comparison group comprises respondents who reported that warming up decreases engine life or has no effect; unsure responses are excluded. Difference rows test whether the *IdleGuard* effect varies with the warm-up engine-life belief. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.13: *IdleGuard* effects by baseline warming intensity and idling context

	Idling minutes from stops >30s			
	Total	Warming	Traffic	Waiting
Incremental effect of IdleGuard: low baseline warming	-0.193	0.142	-0.109	-0.226
	(0.513)	(0.242)	(0.221)	(0.254)
Incremental effect of IdleGuard: high baseline warming	1.999**	0.747*	0.698**	0.554
	(0.915)	(0.402)	(0.351)	(0.472)
Difference: high minus low baseline warming	2.191**	0.605	0.807*	0.780
	(1.049)	(0.470)	(0.414)	(0.536)
No-reminder post-period mean: low baseline warming	12.736	2.068	6.496	4.172
No-reminder post-period mean: high baseline warming	27.551	6.838	10.843	9.870
Observations	254782	254782	254782	254782
Vehicles	2200	2200	2200	2200
Vehicle FE	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. Outcomes are daily idling minutes from stops longer than 30 seconds, reported in total and decomposed into warming, traffic, and waiting idling. The table reports the incremental effect of *IdleGuard*, comparing *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm, separately by whether the vehicle had above-median baseline warming idling. Baseline warming idling is the vehicle-level average of warming idling minutes from stops longer than 30 seconds over the two clean immediate pre-main-treatment GPS weeks. Low baseline warming denotes vehicles at or below the median, and high baseline warming denotes vehicles above the median. Difference rows test whether the *IdleGuard* effect varies with baseline warming intensity. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

E.5 *IdleGuard* Threshold-Crossing and Long-Spell Decompositions

Tables E.14–E.18 examine where the incremental *IdleGuard* effect arises and whether the threshold-crossing pattern persists after treatment ends. They assess whether more spells cross the 30-second reminder threshold, how the resulting duration effect is distributed across portions and terminal durations of spells, and whether the increase in very long spells reflects entry into the long-spell category or continued idling beyond 120 seconds.

Table E.14: *IdleGuard* effects on spells crossing the 30-second threshold

	Daily number of idling spells > 30s			
	All	Traffic-like	Waiting-like	Warming-like
Reminder arm × Post	-0.066	0.087	-0.061	-0.092*
	(0.206)	(0.155)	(0.039)	(0.049)
No-Reminder arm × Post	-0.453**	-0.252	-0.098**	-0.104**
	(0.222)	(0.163)	(0.046)	(0.050)
IdleGuard	0.387*	0.338**	0.037	0.012
	(0.207)	(0.152)	(0.043)	(0.048)
No-Reminder post-period mean	9.893	6.347	1.719	1.826
Observations	372,907	372,907	372,907	372,907
Vehicles	3,412	3,412	3,412	3,412
Vehicle FE	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day, and the outcomes are daily numbers of idling spells longer than 30 seconds. Columns report all idling and its traffic, waiting, and warming components. The first two rows report effects for *Private & Social Info + Pay + Reminder* and *Private & Social Info + Pay + No-Reminder* relative to the control arm; the third reports the incremental effect of *IdleGuard*. Positive coefficients indicate that more spells cross the 30-second threshold. Regressions absorb vehicle and calendar-date fixed effects. Post covers 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.15: Persistence of *IdleGuard* effects on spells crossing the 30-second threshold

	Daily number of idling spells >30s			
	All	Traffic-like	Waiting-like	Warming-like
Prior Reminder arm × Post	0.303	0.278	0.026	-0.002
	(0.269)	(0.172)	(0.050)	(0.103)
Prior No-Reminder arm × Post	-0.582**	-0.360**	-0.113*	-0.109*
	(0.256)	(0.179)	(0.058)	(0.061)
Prior *IdleGuard* assignment	0.885***	0.639***	0.139**	0.107
	(0.288)	(0.181)	(0.058)	(0.107)
No-Reminder post-period mean	9.813	6.214	1.741	1.857
Observations	430,812	430,812	430,812	430,812
Vehicles	3,412	3,412	3,412	3,412
Vehicle FE	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day, and the outcomes are daily numbers of idling spells longer than 30 seconds. Columns report all idling and its traffic, waiting, and warming components. The first two rows report post-treatment effects for vehicles previously assigned to *Private & Social Info + Pay + Reminder* and *Private & Social Info + Pay + No-Reminder* relative to the control arm; the third reports the corresponding prior-*IdleGuard*-assignment difference, defined as *Reminder* minus *No-Reminder*. Positive coefficients in the third row indicate more threshold-crossing spells among vehicles previously assigned to *IdleGuard*. Regressions absorb vehicle and calendar-date fixed effects. The post-treatment period comprises study weeks 21–32, from 1 June through 22 August 2026, after all treatment delivery and the *IdleGuard* reminder had ended; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

The threshold-crossing difference persists after the reminder is removed. Over the twelve post-treatment weeks, vehicles previously assigned to the no-reminder arm continue to have fewer spells longer than 30 seconds than the control arm, while the corresponding estimate for vehicles previously assigned to *IdleGuard* is close to zero. The resulting prior-*IdleGuard*-assignment difference is 0.9 additional threshold-crossing spells per vehicle-day and is statistically significant at the 1% level. Most of this difference arises in traffic-like idling. Because the reminder was no longer active during this period, the persistence of the gap is consistent with more durable preventive self-monitoring among drivers who had not been able to rely on the external cue.

Table E.16: Decomposing the incremental *IdleGuard* effect on idling duration

	Idling minutes from stops > 30s by duration layer							Daily spell count
	Full duration	First 30s	Beyond 30s	30–35s	35–60s	60–120s	>120s	>30s
Reminder arm × Post	-0.730	-0.033	-0.697	-0.013	-0.073	-0.099	-0.512	-0.066
	(0.581)	(0.103)	(0.512)	(0.017)	(0.063)	(0.082)	(0.408)	(0.206)
No-Reminder arm × Post	-1.726***	-0.227**	-1.500***	-0.038**	-0.154**	-0.236***	-1.071**	-0.453**
	(0.631)	(0.111)	(0.551)	(0.018)	(0.069)	(0.090)	(0.432)	(0.222)
IdleGuard	0.996*	0.194*	0.803*	0.025	0.082	0.137	0.559	0.387*
	(0.532)	(0.104)	(0.457)	(0.017)	(0.064)	(0.084)	(0.346)	(0.207)
Share of *IdleGuard* duration effect (%)	100.0	19.4	80.6	2.5	8.2	13.7	56.1	
No-Reminder post-period mean	19.874	4.946	14.928	0.802	2.952	3.497	7.677	9.893
Observations	372,907	372,907	372,907	372,907	372,907	372,907	372,907	372,907
Vehicles	3,412	3,412	3,412	3,412	3,412	3,412	3,412	3,412
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. Columns 1–7 decompose daily idling minutes from spells longer than 30 seconds into mutually exhaustive duration components; Column 8 reports the daily number of such spells. Full duration equals the first 30 seconds of each qualifying spell plus all idling beyond 30 seconds, which is further divided into the 30–35-, 35–60-, 60–120-, and beyond-120-second layers. The first two rows report effects for *Private & Social Info + Pay + Reminder* and *Private & Social Info + Pay + No-Reminder* relative to the control arm; the third reports the incremental effect of *IdleGuard*. The share row reports each duration component's contribution to the full incremental *IdleGuard* effect. Regressions absorb vehicle and calendar-date fixed effects. Post covers 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.17: *IdleGuard* effects by terminal spell duration

	Idling minutes from stops > 30s by total spell duration				Daily spell count		
	Full duration	30–60s	60–120s	>120s	>30s	>60s	>120s
Reminder arm × Post	-0.730	-0.033	-0.072	-0.624	-0.066	-0.085	-0.056
	(0.581)	(0.077)	(0.092)	(0.502)	(0.206)	(0.114)	(0.060)
No-Reminder arm × Post	-1.726***	-0.114	-0.176*	-1.436***	-0.453**	-0.304**	-0.182***
	(0.631)	(0.080)	(0.100)	(0.537)	(0.222)	(0.125)	(0.065)
IdleGuard	0.996*	0.081	0.104	0.811*	0.387*	0.219*	0.126**
	(0.532)	(0.075)	(0.095)	(0.442)	(0.207)	(0.117)	(0.060)
Share of *IdleGuard* duration effect (%)	100.0	8.1	10.5	81.4			
No-Reminder post-period mean	19.874	3.553	4.116	12.205	9.893	5.148	2.264
Observations	372,907	372,907	372,907	372,907	372,907	372,907	372,907
Vehicles	3,412	3,412	3,412	3,412	3,412	3,412	3,412
Vehicle FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. Columns 1–4 report daily idling minutes from spells longer than 30 seconds, decomposed into mutually exclusive bins defined by total spell duration; Columns 5–7 report cumulative daily spell counts crossing the 30-, 60-, and 120-second thresholds. The first two rows report effects for *Private & Social Info + Pay + Reminder* and *Private & Social Info + Pay + No-Reminder* relative to the control arm; the third reports the incremental effect of *IdleGuard*. The share row reports each terminal-duration bin's contribution to the full incremental *IdleGuard* effect on idling duration. Regressions absorb vehicle and calendar-date fixed effects. Post covers 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days. The sample excludes the 25 vehicles removed before the second-stage *Reminder* versus *No-Reminder* randomisation. Standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table E.18: Entry versus continuation within long *IdleGuard* spells

	Full >120s min	First 120s of >120s spells	Excess beyond 120s	Count >120s
IdleGuard vs Placebo	-0.522	-0.049	-0.474	-0.024
	(0.448)	(0.121)	(0.359)	(0.060)
No IdleGuard vs Placebo	-1.360***	-0.325**	-1.035***	-0.163**
	(0.455)	(0.137)	(0.341)	(0.068)
IdleGuard added	0.837*	0.276**	0.561	0.138**
	(0.497)	(0.136)	(0.390)	(0.068)
Share of >120s duration gap (%)	100.0	33.0	67.0	
NoIG post mean	12.205	4.528	7.677	2.264
Observations	179,604	179,604	179,604	179,604
Vehicles	3,133	3,133	3,133	3,133
Vehicle FE	Yes	Yes	Yes	Yes
Calendar-date FE	Yes	Yes	Yes	Yes

Notes: The unit of observation is the vehicle-day. The sample comprises the control arm, *Private & Social Info + Pay + Reminder*, and the matched *Private & Social Info + Pay + No-Reminder* arm. The table decomposes daily idling minutes from spells longer than 120 seconds into an entry component and a continuation component. The entry component equals the first 120 seconds of each qualifying spell, while the continuation component equals idling beyond 120 seconds in those spells. The count column reports the daily number of spells longer than 120 seconds. The incremental *IdleGuard* row compares the reminder arm with the matched no-reminder arm. The share row reports each component's contribution to the incremental *IdleGuard* effect on minutes from spells longer than 120 seconds. Regressions absorb vehicle and calendar-date fixed effects. Post covers 6 April through 30 May 2026; the pooled pre-treatment period comprises 12 study weeks observed between 17 November 2025 and 4 April 2026. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

E.6 *IdleGuard* Endline Mechanism Evidence

Tables E.19 and E.20 provide complementary endline evidence on the mechanisms underlying the *IdleGuard* effect. The first reports drivers' experiences with the reminder, while the second tests whether *IdleGuard* changed beliefs about required engine warm-up.

Table E.19: Endline reports of experience with *IdleGuard*

	Relied on beep	Delayed correction	More comfortable
Share	0.314	0.596	0.421
Observations	309	329	321

Notes: The unit of observation is the vehicle. The table reports descriptive shares among drivers assigned to the *IdleGuard* arm who provide a substantive response to the corresponding endline question; responses indicating uncertainty, not noticing or remembering the beep, or an incomplete interview are excluded. Column 1 equals one if the respondent reports relying more on the beep, rather than continuing to monitor oneself or relying on both equally. Column 2 equals one if the respondent reports not usually turning the engine off or starting to drive promptly after the beep. Column 3 equals one if the respondent reports becoming more comfortable leaving the engine running because the beep would remind them; among substantive responses to this question, 30% report no change and 28% report becoming less comfortable.

Table E.20: Effects of *IdleGuard* on endline warm-up beliefs

	Any warm-up	At least 30 seconds	At least 1 minute
IdleGuard	-0.017	0.004	0.016
	(0.019)	(0.025)	(0.026)
No-Reminder mean	0.864	0.709	0.362
Observations	1413	1413	1413

Notes: The unit of observation is the vehicle. The sample comprises vehicles assigned to the *Reminder* or matched *No-Reminder* arm with a usable endline outcome. Columns 1–3 equal one if the respondent believes, respectively, that any engine warm-up is needed, that at least 30 seconds is needed, and that at least one minute is needed. Each column reports the randomised effect of *IdleGuard*, comparing the *Reminder* arm with the matched *No-Reminder* arm. Endline completion and outcome availability do not differ significantly between the two groups. Robust standard errors are shown in parentheses. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

E.7 Pro-Social Preference Heterogeneity

Table E.21 examines whether treatment responsiveness varies with revealed pro-social preference, measured using the respondent's incentivised baseline donation to WWF-Pakistan.

Table E.21: Treatment effects by revealed pro-social preference

	Idling minutes from stops >30s
Panel A. WWF donation = Rs. 50	
Private Info	-0.941
	(0.972)
Social information	0.262
	(1.109)
Payments	-1.030
	(0.963)
IdleGuard	1.112
	(0.793)
Panel B. WWF donation = Rs. 100	
Private Info	-2.004***
	(0.682)
Social information	0.501
	(0.652)
Payments	-0.550
	(0.684)
IdleGuard	0.991*
	(0.587)
Panel C. Rs. 100 minus Rs. 50	
Private Info	-1.063
	(0.874)
Social information	0.239
	(1.144)
Payments	0.480
	(0.992)
IdleGuard	-0.121
	(0.831)
Control post-period mean	20.985
Observations	688864
Vehicles	6313

Notes: The unit of observation is the vehicle-day, and the outcome is daily idling minutes from stops longer than 30 seconds. The table reports the *Private Info* effect relative to the control arm and the incremental effects of social information, payments, and *IdleGuard*. The heterogeneity variable is the respondent's baseline donation to WWF-Pakistan, scaled from zero to one. Panels A and B report implied effects evaluated at donations of Rs. 50 and Rs. 100 out of the Rs. 100 endowment, respectively; Panel C reports the implied change in each effect when moving from Rs. 50 to Rs. 100. Social information compares *Private & Social Info* with *Private Info*; payments compare *Private & Social Info + Pay* with *Private & Social Info*; and *IdleGuard* compares *Private & Social Info + Pay + Reminder* with the matched *Private & Social Info + Pay + No-Reminder* arm. Respondents with missing or invalid donation amounts are excluded. Regressions absorb vehicle and calendar-date fixed effects. Outcomes are computed over observed GPS vehicle-days, and standard errors are clustered by vehicle. Significance levels are denoted by $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

F Cost–Benefit Analysis

This appendix documents the construction of the cost–benefit estimates reported in Section 7 of the main paper. The exercise evaluates the *Private Info* campaign because it isolates the low-cost informational intervention that generated the clearest behavioural response without requiring social-information, payment, or reminder components. Benefits include avoided private fuel expenditure, avoided global climate damages, and avoided local $PM_{2.5}$-related health damages. Reduced engine wear and other possible maintenance benefits are not monetised.

F.1 Treatment-Effect Inputs

The calculation uses the treatment effect on the full duration of idling spells longer than 10 seconds. During the eight-week treatment period, the estimated reduction in the *Private Info* arm, expressed as a positive magnitude, is

$$\widehat{\beta}^{PI}_{\text{active}} = 1.952 \text{ minutes per observed vehicle-day.} \tag{F.1}$$

The active-period calculation covers eight treatment weeks and six study days per week. To express benefits per treated vehicle, I apply the estimated effect per observed vehicle-day to the 48 possible study days during this period, treating temporary tracker non-reporting as missing measurement and assuming that the estimated effect applies similarly on unobserved study days. Appendix Table C.6 finds no evidence of differential post-launch observability across treatment arms.

To incorporate the twelve directly observed post-switch-off weeks, I use the *Early-Off* subgroup of the *Private Info* arm. These vehicles stopped receiving messages after the eight-week campaign. Appendix Table D.14 estimates an active-period effect of 2.480 minutes and a pooled twelve-week post-switch-off effect of 2.193 minutes for this subgroup. The estimated post-switch-off effect is therefore

$$\frac{\widehat{\beta}^{EO}_{\text{post}}}{\widehat{\beta}^{EO}_{\text{active}}} = \frac{2.193}{2.480} = 0.884. \tag{F.2}$$

of the corresponding active-period effect. I apply this persistence ratio to the more precisely estimated full-arm effect in Equation F.1. The resulting full-arm-equivalent post-switch-off reduction, expressed as a positive magnitude, is

$$\widehat{\beta}^{PI}_{\text{post}} = 1.952 \times \frac{2.193}{2.480} = 1.726 \text{ minutes per observed vehicle-day.} \tag{F.3}$$

This scaling preserves the post-to-active relationship estimated within the randomly assigned *Early-Off* subgroup while anchoring the treatment-period calculation to the full *Private Info* arm.

F.2 Physical-Benefit and Monetisation Inputs

For each evaluation horizon h, let M_h denote the cumulative reduction in idling minutes per treated vehicle. Fuel savings and avoided emissions are calculated as

$$F_h = M_h \times 0.0246, \tag{F.4}$$

$$CO_{2,h} = M_h \times 0.06133, \tag{F.5}$$

$$PM_{2.5,h} = M_h \times 0.02282, \tag{F.6}$$

where F_h is measured in litres, $CO_{2,h}$ in kilograms, and $PM_{2.5,h}$ in grams.

Table F.1: Inputs to the cost–benefit calculation

Input	Value	Unit	Use in calculation
Idling fuel consumption	1.476	litres/hour	Conversion from avoided idling minutes to fuel saved
Petrol price	310.71	PKR/litre	Valuation of avoided private fuel expenditure
Idling CO_2 emissions	3.68	kg/hour	Conversion from avoided idling minutes to CO_2 avoided
Social cost of carbon	229.31	2025Q4 USD/tonne	Valuation of avoided global climate damages
Exchange rate	278.12	PKR/USD	Conversion of climate and health values into PKR
Idling $PM_{2.5}$ emissions	0.02282	g/minute	Conversion from avoided idling minutes to $PM_{2.5}$ avoided
Local $PM_{2.5}$ health damage	137.0	PKR/g	Valuation of avoided local health damages
Programme cost	375.68	PKR/vehicle	Eight-week app, email, and WhatsApp delivery cost

Notes: The idling fuel-consumption rate follows the engineering input used to construct the treatment messages (Gaines and Rask, 2012). The petrol price is taken from Pakistan State Oil's price archive (Pakistan State Oil, 2026). The CO_2 emissions rate combines the fuel-use rate with the average carbon content of petrol and diesel reported by the U.S. Environmental Protection Agency (U.S. Environmental Protection Agency, 2005). The social cost of carbon begins from the USD 185-per-tonne estimate in Rennert et al. (2022) and is updated from 2020 to 2025Q4 dollars using the U.S. GDP implicit price deflator (U.S. Bureau of Economic Analysis, 2026). The exchange rate is the USD-to-PKR rate observed on 14 July 2026 (XE, 2026). The derivation of the $PM_{2.5}$ emissions and health-damage inputs is described in Appendix F.3. Programme cost includes the marginal app, email, and WhatsApp delivery costs of the eight-week campaign.

Table F.2: Implied physical benefits per treated vehicle

Evaluation horizon	Fuel saved (litres)	CO_2 avoided (kg)	$PM_{2.5}$ avoided (g)
8-week treatment period	2.30	5.75	2.14
20-week observed period	5.36	13.37	4.97
1 year: faster decay	7.15	17.81	6.63
1 year: slower decay	9.44	23.53	8.76

Notes: The table reports physical benefits per average treated *Private Info* vehicle. Fuel savings use an idling consumption rate of 0.0246 litres per minute. Avoided emissions use rates of 0.06133 kg of CO_2 and 0.02282 g of $PM_{2.5}$ per idling minute. The eight-week and 20-week calculations use observed experimental effects. In the faster-decay scenario, the treatment effect declines linearly to zero 26 weeks after switch-off. In the slower-decay scenario, it declines linearly to zero by campaign week 52.

Private fuel benefits are

$$B_h^F = F_h \times 310.71. \tag{F.7}$$

Global climate benefits are

$$B_h^{CO_2} = \frac{CO_{2,h}}{1000} \times 229.31 \times 278.12. \tag{F.8}$$

Local air-pollution health benefits are

$$B_h^{PM} = PM_{2.5,h} \times 137.0. \tag{F.9}$$

Total gross benefits, net benefits, and the benefit–cost ratio are therefore

$$B_h = B_h^F + B_h^{CO_2} + B_h^{PM}, \tag{F.10}$$

$$NB_h = B_h - C, \tag{F.11}$$

$$BCR_h = \frac{B_h}{C}, \tag{F.12}$$

where $C = 375.68$ PKR is the eight-week programme cost per treated vehicle.

F.3 Local $PM_{2.5}$ Health Benefits

The local-health calculation first converts avoided idling fuel into avoided primary $PM_{2.5}$ emissions. The calculation uses a simple average of petrol and diesel densities. The petrol-density input is taken from laboratory analysis of commercial motor gasoline in Pakistan (Yasin et al., 2008), while the diesel-density input is taken from laboratory analysis of Pakistan State Oil high-speed diesel (Ashraf et al., 2016). Their midpoint is 0.7795 kilograms per litre. I combine this density with the fleet-average emissions factor of 1.19 grams of $PM_{2.5}$ per kilogram of fuel estimated from near-road vehicle measurements in Mumbai (Raparthi et al., 2023). Together with the idling fuel-use rate, these inputs imply

$$0.0246 \times 0.7795 \times 1.19 = 0.02282 \text{ grams of } PM_{2.5} \text{ per idling minute.} \tag{F.13}$$

I map avoided emissions into ambient exposure using the transport contribution to annual $PM_{2.5}$ emissions in Punjab's emissions inventory (FAO, 2018). The exposure calculation uses the combined population of Punjab, Sindh, and the Islamabad Capital Territory from the 2023 Population and Housing Census (PBS, 2023). I then translate the implied change in annual average exposure into life-expectancy gains using the Air Quality Life Index concentration–life-expectancy relationship (Greenstone et al., 2025).

Mortality-risk reductions are valued using the OECD value of a statistical life for lower- and middle-income countries (OECD, 2025). Because the Air Quality Life Index relationship represents the long-run effect of sustained exposure, I conservatively prorate the estimated life-expectancy gain linearly over an assumed 68-year life expectancy. This procedure implies a central local health value of approximately

$$137.0 \text{ PKR per gram of avoided } PM_{2.5}. \tag{F.14}$$

This valuation is illustrative. It relies on an emissions-to-exposure mapping rather than a local atmospheric-dispersion model and assumes that the health consequences of a temporary exposure reduction can be approximated by linear prorating of the long-run concentration–life-expectancy relationship.

F.4 Evaluation Horizons

I report four evaluation horizons.

Eight-week treatment period. This benchmark counts benefits over the same eight weeks in which messages were delivered. It uses the full-arm effect in Equation F.1 over 48 possible study days and requires no assumption about persistence after treatment.

Twenty-week observed period. This horizon adds the twelve directly observed post-switch-off weeks. The first eight weeks use the full-arm active-period estimate, while the subsequent twelve weeks use the scaled post-switch-off estimate in Equation F.3. The programme cost remains fixed because no additional messages are delivered during these twelve weeks.

One year: faster decay. The faster-decay scenario begins from the observed post-switch-off effect and assumes that it declines linearly to zero 26 weeks after switch-off. Because treatment ends after campaign week 8, the effect reaches zero at campaign week 34. Benefits after that point are set to zero.

One year: slower decay. The slower-decay scenario begins from the same observed post-switch-off effect and assumes that it declines linearly to zero by campaign week 52. No benefits are counted beyond the one-year evaluation horizon.

The eight-week and 20-week estimates are based on experimentally observed effects. The two one-year estimates are extrapolation scenarios rather than estimated one-year treatment effects. Both scenarios impose eventual complete decay and assume neither continued message delivery nor permanent treatment persistence.

F.5 Illustrative National Scaling

Let N_{PV} denote Pakistan's reported stock of registered motor cars, jeeps, and station wagons. The Pakistan Economic Survey reports

$$N_{\text{PV}} = 4{,}889{,}104 \tag{F.15}$$

registered vehicles in this category in 2025 (Government of Pakistan, Ministry of Finance, 2026).

Let s denote the share of this stock reached by the campaign. The number of reached vehicles is

$$N_{\text{reached}} = N_{\text{PV}} s. \tag{F.16}$$

For evaluation horizon h, national gross benefits and programme costs are

$$B_h^N = B_h N_{\text{PV}} s, \tag{F.17}$$

$$C^N = C N_{\text{PV}} s. \tag{F.18}$$

National net benefits are

$$NB_h^N = B_h^N - C^N. \tag{F.19}$$

Because benefits and costs scale proportionally with programme reach, the national benefit–cost ratio is equal to the corresponding per-vehicle ratio:

$$BCR_h^N = \frac{B_h^N}{C^N} = \frac{B_h}{C}. \tag{F.20}$$

The preferred illustration uses $s = 0.25$, corresponding to

$$4{,}889{,}104 \times 0.25 = 1{,}222{,}276 \tag{F.21}$$

reached vehicles. This is intended as an illustrative programme scale rather than a forecast of realised national coverage.

Table F.3: Illustrative national physical benefits at 25% programme reach

Evaluation horizon	Fuel saved (million litres)	CO_2 avoided (tonnes)	$PM_{2.5}$ avoided (tonnes)
8-week treatment period	2.82	7,024	2.61
20-week observed period	6.55	16,341	6.08
1 year: faster decay	8.73	21,776	8.10
1 year: slower decay	11.54	28,763	10.70

Notes: The illustration applies the per-vehicle *Private Info* intent-to-treat effects to 1,222,276 vehicles, equal to 25% of Pakistan's registered motor cars, jeeps, and station wagons. The eight-week and 20-week calculations use observed experimental effects. The one-year estimates assume linear decay to zero either 26 weeks after switch-off or by campaign week 52.

Table F.4: Monetised benefits in the preferred national-scale application

Evaluation horizon	Gross benefit	Programme cost	Net benefit	BCR
8-week treatment period	1.68	0.46	1.22	3.66
20-week observed period	3.91	0.46	3.45	8.52
1 year: faster decay	5.21	0.46	4.75	11.35
1 year: slower decay	6.89	0.46	6.43	14.99

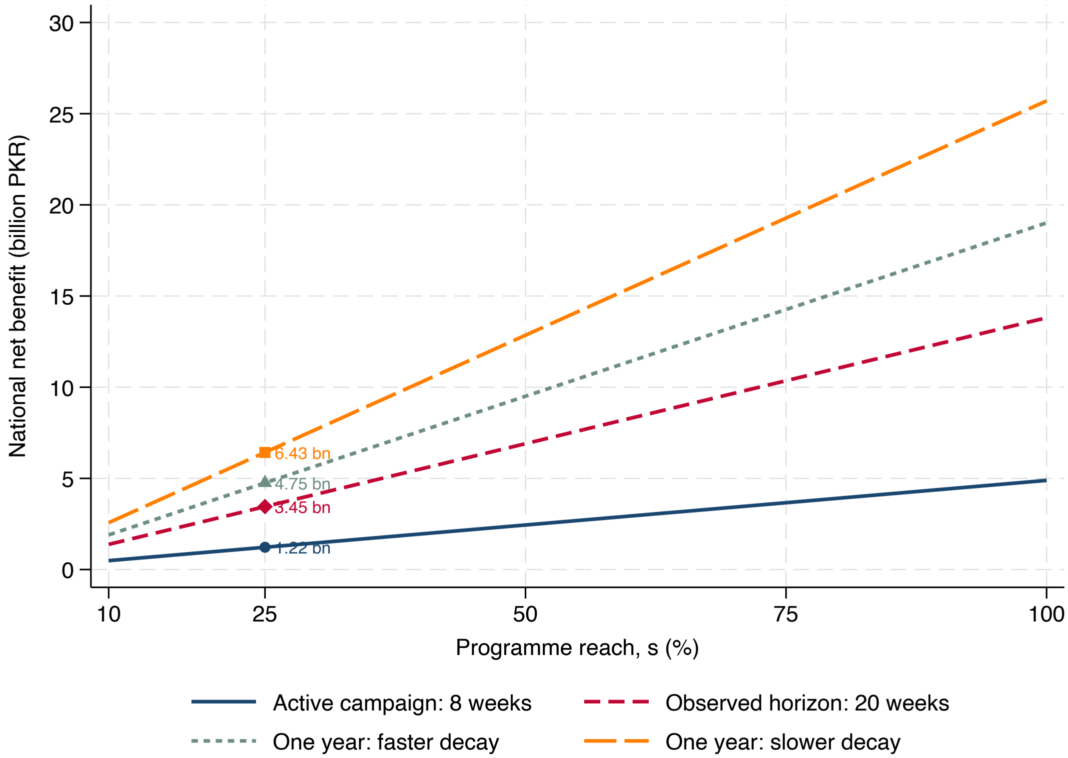
Notes: Monetary values are billions of PKR. The illustration assumes that the programme reaches 25% of Pakistan's registered motor cars, jeeps, and station wagons, corresponding to 1,222,276 vehicles. The experimental per-vehicle intent-to-treat benefit is applied directly to each reached vehicle. Benefits and programme costs scale proportionally with programme reach, so the national benefit–cost ratios equal the corresponding per-vehicle ratios. The eight-week and 20-week calculations use observed experimental effects. The one-year calculations apply the faster- and slower-decay scenarios described in Appendix F.4.

F.6 Interpretation and Limitations

The eight-week estimate is the most conservative benchmark because it counts benefits only while the campaign is actively delivered. The 20-week estimate additionally incorporates twelve weeks of directly observed persistence without adding programme costs. The one-year estimates show how returns vary under two transparent assumptions about decay beyond the observed period.

The exercise monetises private fuel savings, global climate benefits, and local $PM_{2.5}$-related health benefits. It excludes reduced engine wear and other possible maintenance gains, making the calculation conservative in the set of benefits included. Conversely, the national calculations assume that the experimental per-vehicle intent-to-treat effect applies directly to each vehicle reached. They

Figure F.1: Illustrative national net benefits under alternative programme reach



Notes: The figure reports illustrative national net benefits across four evaluation horizons and alternative programme-reach shares. Programme reach is defined as the share of Pakistan's registered motor cars, jeeps, and station wagons reached by the campaign. The per-vehicle *Private Info* intent-to-treat estimates are applied directly to each reached vehicle. Markers identify the preferred illustrative case of 25% reach. The eight-week and 20-week calculations use observed experimental effects. The one-year estimates assume linear decay to zero either 26 weeks after switch-off or by campaign week 52. The calculations are illustrative scaling exercises rather than forecasts of realised national programme coverage.

should therefore be interpreted as illustrative scaling exercises rather than forecasts of take-up, implementation fidelity, or realised national effects.

References

Ashraf, Talal, Shaheen Aziz, and Ahmed Nawaz, "Interpretation of Qualitative Analysis of Pakistan State Oil's High Speed Diesel," in "International Conference on Energy, Environment and Sustainable Development" November 2016.

FAO, "Remote Sensing for Spatio-Temporal Mapping of Smog in Punjab and Identification of the Underlying Causes Using GIS Techniques (R-SMOG)," Technical Report, Food and Agriculture Organization of the United Nations 2018.

Gaines, Linda and Eric Rask, "Which Is Greener: Idle, or Stop and Restart?," Argonne National Laboratory and Alternative Fuels Data Center fact sheet 2012. Accessed 28 June 2026.

Government of Pakistan, Ministry of Finance, "Pakistan Economic Survey 2025–26: Transport and Communications," Technical Report, Finance Division, Government of Pakistan, Islamabad 2026. Table 13.3, Number of Motor Vehicles Registered.

Greenstone, Michael, Tanushree Ganguly, Christa Hasenkopf, Nishka Sharma, and Hrishikesh Chandra Gautam, "Air Quality Life Index 2025 Annual Update," Technical Report, Energy Policy Institute at the University of Chicago 2025. Air Quality Life Index.

Kreindler, Gabriel, "Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications," *Econometrica*, 2024, *92* (4), 1233–1268.

OECD, "Mortality Risk Valuation in Policy Assessment: A Global Meta-Analysis of Value of Statistical Life Studies," Technical Report, OECD Publishing, Paris 2025.

Pakistan State Oil, "Petroleum Product Price Archive," https://psopk.com/fuel-prices/pol/archives?page=1 2026. Accessed 13 July 2026.

PBS, "Seventh Population and Housing Census 2023," https://www.pbs.gov.pk/content/final-results-census-2023 2023. Pakistan Bureau of Statistics. Accessed 29 June 2026.

Raparthi, Nagendra, Atique Barudgar, Mengyuan Chu, Zhi Ning, and Harish C. Phuleria, "Estimating Individual Vehicle Emission Factors from Near-Road Measurements in India," *Atmospheric Environment*, 2023, *308*, 119869.

Rennert, Kevin, Frank Errickson, Brian C. Prest, Lisa Rennels, Richard G. Newell, William Pizer, Cora Kingdon, Jordan Wingenroth, Roger Cooke, Bryan Parthum, David Smith, Kevin Cromar, Delavane Diaz, Frances C. Moore, Ulrich K. Müller, Richard J. Plevin, Adrian E. Raftery, Hana Ševčíková, Hannah Sheets, James H. Stock, Tammy Tan, Mark Watson, Tony E. Wong, and David Anthoff, "Comprehensive Evidence Implies a Higher Social Cost of CO2," *Nature*, 2022, *610*, 687–692.

U.S. Bureau of Economic Analysis, "National Income and Product Accounts: Implicit Price Deflators for Gross Domestic Product," https://apps.bea.gov/iTable/?reqid=19&step=2&isuri=1&categories=survey 2026. Quarterly series accessed 13 July 2026.

U.S. DOE, "Light-Duty Vehicle Idle Reduction Strategies," https://afdc.energy.gov/conserve/idle-reduction-light 2026. Alternative Fuels Data Center. Accessed 29 June 2026.

U.S. Environmental Protection Agency, "Average Carbon Dioxide Emissions Resulting from Gasoline and Diesel Fuel," https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1001YTF.TXT 2005. Office of Transportation and Air Quality.

U.S. EPA, "What You Can Do to Reduce Pollution from Vehicles and Engines," https://www.epa.gov/transportation-air-pollution-and-climate-change/what-you-can-do-reduce-pollution-vehicles-and 2025. Accessed 29 June 2026.

XE, "U.S. Dollar to Pakistani Rupee Exchange Rate," https://www.xe.com/currencyconverter/convert/?Amount=1&From=USD&To=PKR 2026. Exchange rate accessed 14 July 2026.

Yasin, Ghulam, Tariq Mahmood Ansari, Syed Muhammad Sibtain Raza Naqvi, and Farah Naz Talpur, "Analytical Studies on the Quality and Environmental Impact of Commercial Motor Gasoline Available in Multan Region of Pakistan," *Pakistan Journal of Analytical and Environmental Chemistry*, 2008, *9* (2), 84–91.