

Social Norms, Incentives, and Congestion at Electric Vehicle Charging Stations

Pre-Analysis Plan

Nicolas Astier* Jan Domagalski[†]

30 September 2026

Abstract

The increasing adoption of electric vehicles is creating congestion at public charging stations, where available spots are more and more often all occupied at the same time. Previous research has shown that drivers are strongly responsive to financial incentives such as peak vs off-peak electricity prices. However, electricity prices are not well suited to managing congestion because electricity consumption falls to zero once a vehicle's battery is full. An overstay fee addresses this problem by charging, when the station is busy, a per-minute fee to cars that stay at the station despite being almost fully charged. This project studies whether the informational content of the notification sent shortly before the overstay fee is triggered affects drivers' behaviour. We compare a control message with three messages appealing to prosocial behaviour, social norms, and information about decreasing charging speed.

Keywords: electric vehicles, charging, nudges, congestion

JEL codes: D12, Q41, R41

*Paris School of Economics and École des Ponts, France. Email: nicolas.astier[at]psemail.eu.

[†]Paris School of Economics, France. Email: jan.domagalski[at]psemail.eu.

1 Background

In the context of electric vehicle (EV) charging at a public charging station, an overstay fee is a per-minute fee that applies when two conditions are satisfied:

1. all charging spots at the station are occupied; and
2. the vehicle is close to being fully charged (e.g. the battery is charged above 80%).

The purpose of the fee is to encourage drivers whose vehicles are nearly fully charged to leave the station when other drivers want to access to the charging infrastructure.

Quasi-experimental evidence from the introduction of overstay fees shows that drivers respond to the financial incentive. However, less is known about the motivations underlying this response. This project examines whether the content of how the information sent to drivers shapes their response.

2 Research question

The main research question is:

How does the response to an overstay fee vary with the informational content of the notification sent to drivers?

3 Setting of the study

We partner with a charging-point operator that operates hundred of charging stations in Europe, mostly in France, and has implemented an overstay fee at most of these stations. The experiment is scheduled to run for six weeks, from 25 September 2026 to 6 November 2026.

4 Treatments and treatment assignment

The study population consists of eligible overstay-fee notifications sent to users over the period of the experiment. These notifications are sent 5 minutes before a 0.4€/min overstay fee starts binding.

The analysis sample is restricted to sessions initiated through the mobile application by non-commercial users.

Treatment is assigned at the charging-session level. Each eligible session is assigned with equal probability to one of four notification messages:

1. **Control:** “5 min avant les frais” (“5 minutes before fees”);
2. **Prosocial:** “Laissez la place aux autres” (“Leave the spot for others”);
3. **Social norm:** “Stop à 80%, c’est la norme” (“Stop at 80%, it’s a norm”);
4. **Information:** “Sauvez temps et argent ! La charge ralentit après 80%” (“Save time and money. Charging slows down after 80%”).

The treatment is assigned automatically when the notification is triggered. Assignment is therefore determined before the outcomes are realised.

5 Outcomes of interest

5.1 Primary outcomes

The two primary outcomes are:

1. **Compliance within the grace period:** an indicator equal to one if the charging session ends within five minutes of the notification and zero otherwise.
2. **Time to end of session:** the number of minutes between the notification and the end of the charging session.

The five-minute threshold corresponds to the grace period before the overstay fee applies.

5.2 Secondary outcomes

The main secondary outcomes are:

1. **Acceptability:** the driver’s rating of the session on a five-star scale;
2. **Revenue:** the total payment associated with the charging session;
3. **Retention:** the time until the driver’s next charging session in the operator’s network;
4. **Efficiency:** the average charging speed during the session;
5. **Final state of charge:** the battery state of charge at the end of the session.

6 Data collection

The variables used in the analysis will be collected by the charging-point operator. The raw data will not be shared outside the company because of data-protection restrictions. The researchers will access the data at the operator’s premises.

The analysis dataset will contain one observation per charging session receiving an eligible notification. Notification records will be linked to charging sessions using the charge identifier contained in the notification event. Sessions will then be linked to their charging measurements and financial records.

7 Empirical analysis

7.1 Randomization and balance checks

Randomization takes place at the charging-session level. Every eligible session is assigned to one of the four messages with equal probability.

The analysis sample consists of sessions that were initiated through the mobile application by non-commercial users and received an overstay-fee notification during the study period. Sessions with more than one associated notification record will be dropped.

We assess balance across treatment groups using characteristics determined before the notification is sent:

- hour of the day at which the notification was sent, measured in local CE(S)T time;
- day of the week on which the notification was sent;
- week of the study period;
- state of charge at the time of the notification, measured as the most recent battery reading at or before the notification;
- state of charge at the beginning of the charging session;
- elapsed charging time before the notification;
- highway versus non-highway station, if a reliable station classification is available;
- charging station;
- session initiation method, such as automatic initiation upon plugging in or manual initiation through the application;

- number of previous charging sessions associated with the same user before the focal session.

For continuous variables, we report means, standard deviations, differences relative to the control group, and standardized differences. For binary and categorical variables, we report treatment-group proportions and differences relative to the control group. We also report joint tests of equality across the four treatment groups.

Balance checks are descriptive and are not used to select covariates after observing the results. Missing values are reported for each balance variable. Variables measured after the notification, including the overstay fee, session duration after the notification, final state of charge, revenue, and user ratings, are not included in the balance checks.

7.2 Main specification

For each treatment group T , the main specification is:

$$Y_s = \alpha + \beta_T \mathbf{1}[s \in T] + \varepsilon_s, \quad (1)$$

where Y_s is one of the primary or secondary outcomes for session s , and $\mathbf{1}[s \in T]$ is an indicator for assignment to treatment group T .

The control group is the reference category. The coefficients β_T measure the difference between each treatment group and the control group.

Because treatment is assigned at the session level but users may contribute multiple sessions, standard errors will be clustered at the user level.

The primary family consists of three treatment-versus-control comparisons for each of two primary outcomes. Our main specification will therefore use a Bonferroni adjustment across six tests (3 treatments and 2 main outcomes).

7.3 Specification with covariates

We will estimate a second specification that adds the prespecified pre-treatment covariates listed in the balance-check section. These covariates will be included regardless of whether the balance tests reject equality across treatment groups.

The unadjusted specification in Equation (1) is the primary specification. The covariate-adjusted estimates are reported as a precision-enhanced sensitivity analysis.

7.4 Heterogeneity analyses

State-of-charge when the overstay is triggered

Triggering the overstay fee requires two conditions:

1. the station must be full; and
2. the vehicle must be charged above 80%.

The first condition applies to all drivers at the station, whereas the second is vehicle-specific, and depends on the progress of the charging session. Drivers may therefore receive the notification at different states of charge.

We will test whether the state of charge at the time of the notification affects the response to treatment. For each treatment group T , we estimate:

$$Y_s = \alpha + \beta_{1,T}\mathbf{1}[s \in T] + \beta_2 SOC_s^0 + \beta_{3,T} (SOC_s^0 \times \mathbf{1}[s \in T]) + \varepsilon_s, \quad (2)$$

where SOC_s^0 is the state of charge at the time of the notification.

We also estimate a specification in which SOC_s^0 is an indicator equal to one when the state of charge is strictly above 80% and zero when it is exactly 80%.

8 Power calculations

We conduct an *ex ante* power analysis for the two primary outcomes. The experiment has four equally sized arms. The primary analysis compares each of the three treatment arms with the control arm for each outcome, resulting in six primary hypothesis tests.

The calculations below thus use a Bonferroni-adjusted significance level of:

$$\alpha = \frac{0.05}{6} = 0.00833. \quad (3)$$

8.1 Expected sample size

The study is planned to run for six weeks, corresponding to 42 days. We consider three notification-volume scenarios (based on data for France only, as a conservative assessment):

1. a **lower planning scenario** of 100 notifications per day;
2. a **conservative observed scenario** of 350 notifications per day;
3. an **observed-rate scenario** of 450 notifications per day.

The lower planning scenario corresponds to the operator’s original planning estimate. The observed-rate during the seven days, 22-28 September 2026 is actually higher, approximately 465 notifications per day. The conservative observed scenario is in between these two cases.

Matching tests indicate that 99% of distinct notification identifiers can be linked to a charging session. Among linked sessions, 96% have a valid outcome observation. Applying these rates to the three volume scenarios gives the following expected sample sizes.

Table 1: Expected sample size under alternative notification-volume scenarios

Scenario	Notifications per day	Total notifications	Usable sessions	Usable sessions per arm
Lower planning scenario	100	4,200	4,001	1,000
Conservative observed scenario	350	14,700	14,004	3,501
Observed-rate scenario	450	18,900	18,006	4,501

The observed daily volume during the seven days before the finalization of this pre-analysis plan was 464.9 notifications per day, with a standard deviation of 73.7.

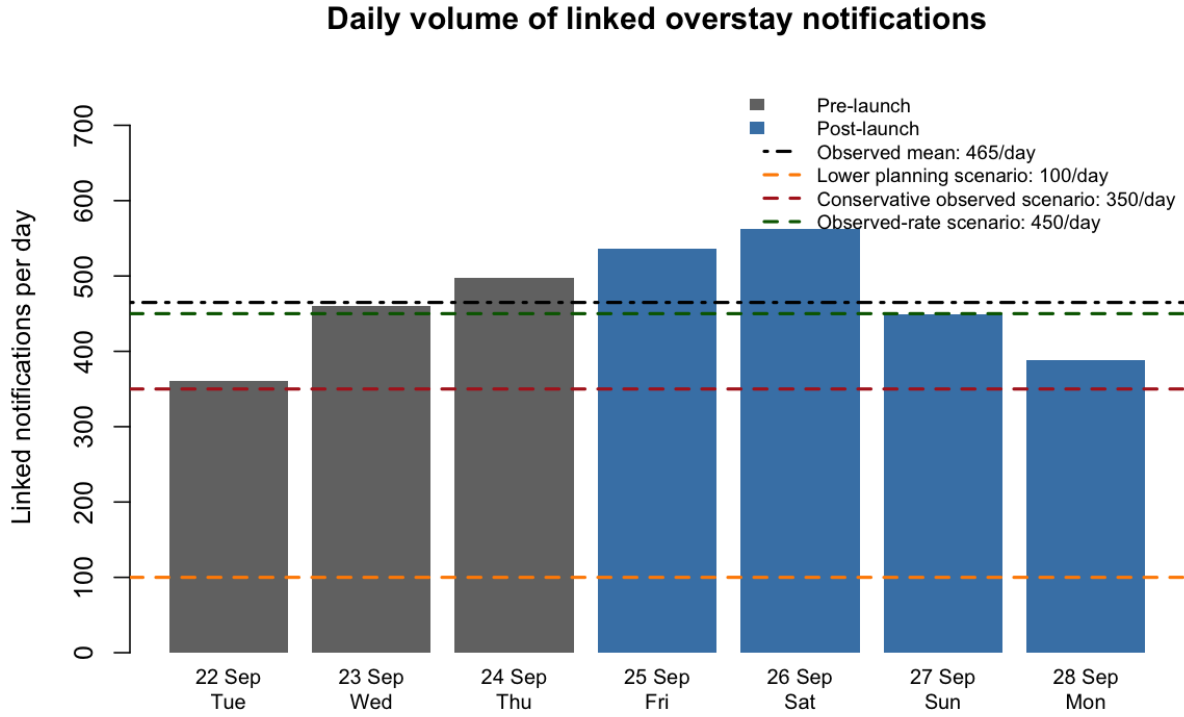


Figure 1: Observed daily notification volume and volume scenarios

8.2 Baseline values of the main outcomes

The baseline outcome parameters are estimated using the pre-randomization period from 1 September through 24 September 2026. During this period, the control message was used before the four message conditions were randomly assigned at noon on 25 September.

The continuous primary outcome is the number of minutes between the notification and the end of the charging session. The baseline sample contains 1,427 valid observations. The mean duration is 7.85 minutes, the median is 4.52 minutes, the 90th percentile is 20.70 minutes, and the 99th percentile is 48.99 minutes.

The binary primary outcome equals one when the session ends within five minutes of the notification. Its baseline rate is 54.0%.

Table 2: Baseline outcome parameters

Outcome	Observations	Mean	Standard deviation	Median	90th percentile	Baseline rate
Minutes until session end	1,427	7.85	10.13	4.52	20.70	–
Ends within five minutes	1,427	–	–	–	–	54.0%

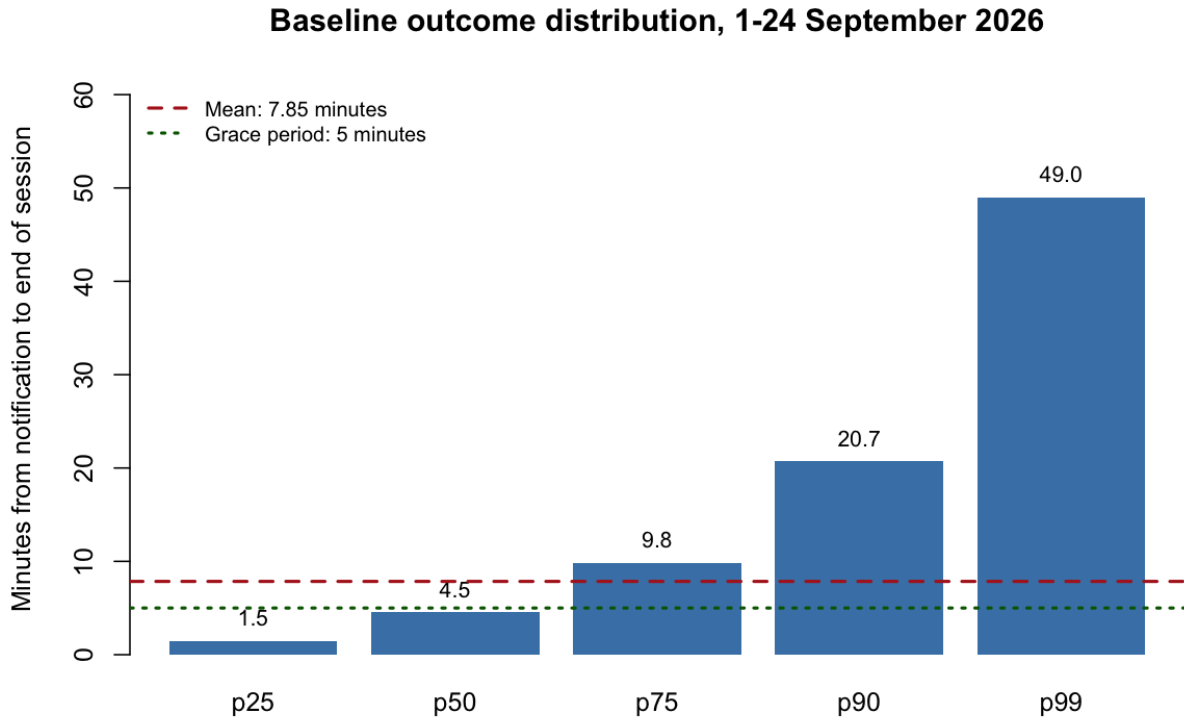


Figure 2: Distribution of the continuous outcome during the baseline period

8.3 Minimum detectable effects

For the continuous outcome, the minimum detectable effect is calculated using the standard two-sample comparison-of-means formula. The primary calculation uses the observed raw standard deviation of 10.13 minutes. A sensitivity calculation uses the winsorized standard deviation of 9.35 minutes to reduce the influence of the long right tail.

For the binary outcome, the minimum detectable effect is calculated using the baseline compliance rate of 54.0%. All calculations assume equal allocation across the four arms and 80%

power.

The table below reports the minimum detectable effects under the full-family significance level. The winsorized continuous-outcome column and the binary-outcome column include the provisional user-clustering adjustment.

Table 3: Minimum detectable effects (for a power of 80%) under alternative volume scenarios (target significance of $\alpha = 0.00833$)

Scenario	Continuous outcome raw standard deviation	Continuous outcome winsorized and clustered	Binary outcome clustered
Lower planning scenario	1.58 minutes	1.52 minutes	7.88 p.p.
Conservative observed scenario	0.84 minutes	0.81 minutes	4.24 p.p.
Observed-rate scenario	0.74 minutes	0.71 minutes	3.74 p.p.

Under the lower planning scenario of 100 notifications per day, six weeks provides approximately 1,000 usable observations per arm. The experiment can detect a continuous effect of approximately 1.5 minutes and a change in the binary compliance outcome of approximately 8 p.p..

Under the conservative observed scenario of 350 notifications per day, the corresponding minimum detectable effects are approximately 0.8 minutes and 4.2 p.p.. Under the observed-rate scenario of 450 notifications per day, they are approximately 0.7 minutes and 3.7 p.p..

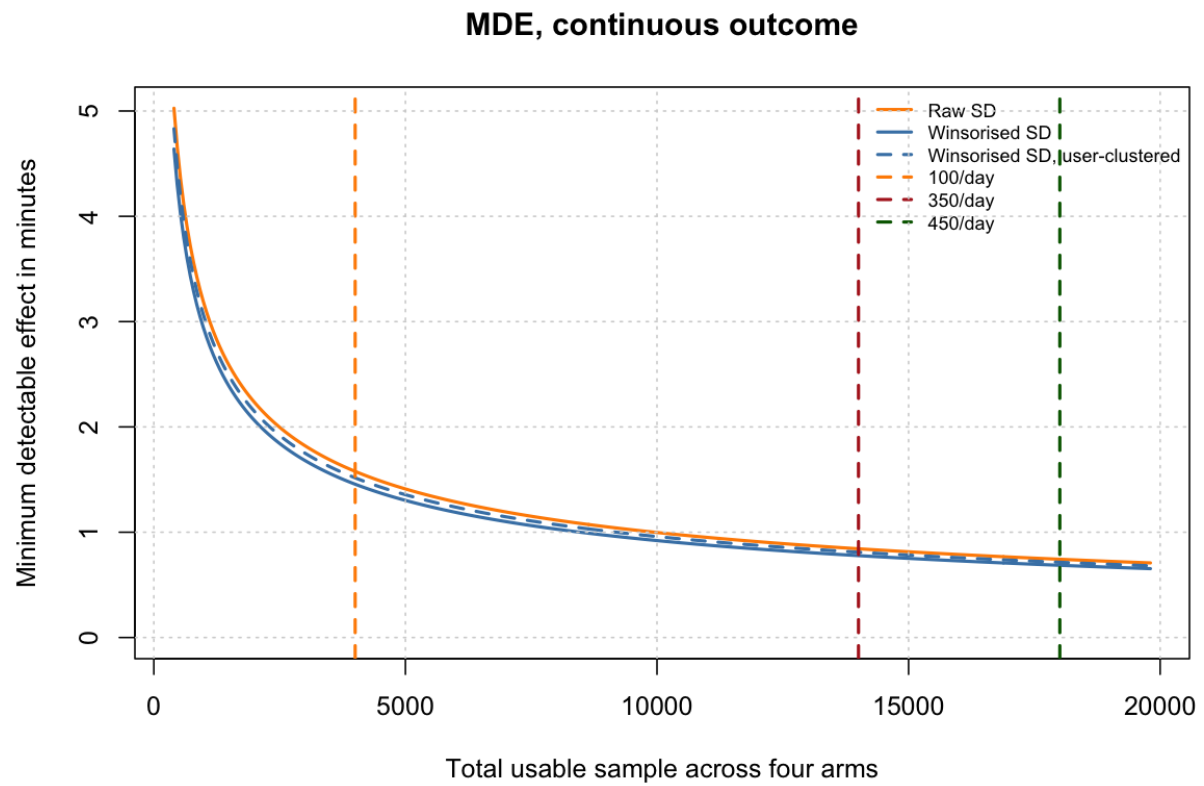


Figure 3: Minimum detectable effects for the continuous outcome

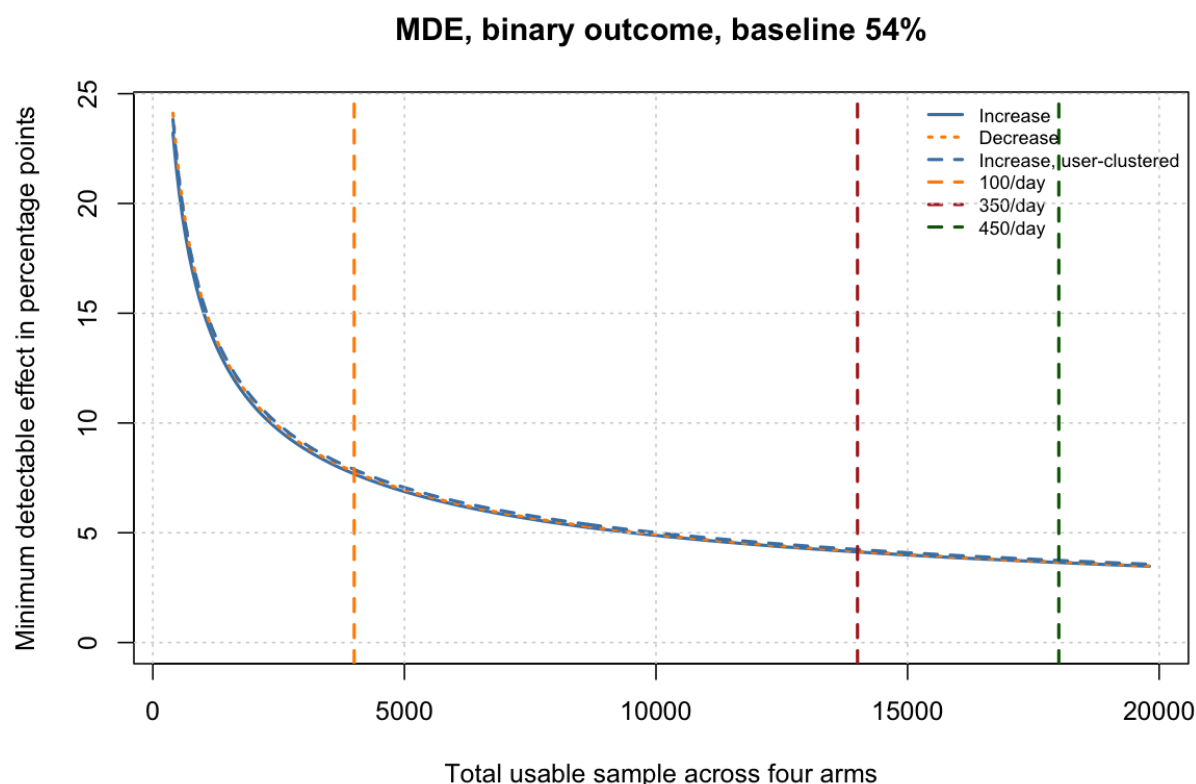


Figure 4: Minimum detectable effects for the binary outcome

8.4 Repeated observations and inference

Randomization takes place at the charging-session level, but the same user may appear in more than one session. The primary regressions will therefore report standard errors clustered at the user level.

The current estimates of within-user correlation are 0.424 for the continuous outcome and 0.269 for the binary outcome. These estimates imply modest increases in the minimum detectable effects.

8.5 Covariate adjustment

The prespecified baseline covariates explain approximately 2% of the variation in the continuous outcome in the current baseline extract. A sensitivity analysis shows that explaining 10% of the outcome variance would reduce the clustered continuous-outcome minimum detectable effect by approximately 5%, while explaining 20% would reduce it by approximately 11%.

Covariate adjustment is therefore expected to improve precision modestly. The unadjusted comparison remains the primary specification, while the covariate-adjusted specification is reported as

a precision-adjusted sensitivity analysis.

8.6 Conclusion

Our main power calculation uses the conservative observed scenario of 350 notifications per day. Under this scenario, six weeks yields approximately 3,501 usable observations per arm. At 80% power and the family-wise significance level, the study can detect an effect of approximately 0.8 minutes on the continuous outcome and approximately 4.2 p.p. on the binary compliance outcome.

This scenario is below the average notification volume observed during the initial complete days of the study and therefore provides a cautious, empirically grounded benchmark. If the notification rate remains closer to 450 per day, the corresponding minimum detectable effects fall to approximately 0.7 minutes and 3.7 p.p.

Even under the lower planning scenario of 100 notifications per day, the study would generate approximately 1,000 usable observations per arm. It would still detect effects of approximately 1.5 minutes and 8 p.p., respectively, although smaller effects would be difficult to identify under that scenario.