

# Estimating the Causal Impact of Green Delivery Slot Labeling

An Applied Case Study from an E-commerce Platform in Singapore

Dung Nguyen

## 1. Motivation and Background Context

Digital platforms increasingly use choice architecture to steer users toward actions that support platform-level goals. In last-mile delivery, a common strategy is labeling certain delivery time slots as environmentally preferable in order to encourage customers to select options that allow deliveries to be batched more efficiently. If successful, such interventions can reduce both carbon emissions and logistics costs.

This project evaluates the causal impact of a green-labeling mechanism implemented by an online grocery platform in Singapore. At the time of ordering, customers are presented with a set of available delivery time slots, some of which are labeled “Less CO<sub>2</sub>” based on whether other orders are already scheduled to the same district within the upcoming delivery window. Customers observe the labels but are not informed about the underlying labeling rule.

The analysis focuses on two questions:

- **RQ1:** Does increased exposure to green-labeled delivery slots causally increase the likelihood that customers select a green option?
- **RQ2:** Does higher uptake of green-labeled slots translate into measurable improvements in delivery batching efficiency?

## 2. Estimation Challenges

Estimating the causal impact of the green-labeling mechanism is non-trivial for three reasons.

**First, the data are fully observational.** The available data cover a 15-day window during which the green-labeling feature was already active for all users. No randomized experiment or phased rollout information is available, ruling out standard experimental or time-based quasi-experimental designs. As a result, causal identification must rely on cross-sectional and within-district variation rather than pre-post comparisons.

**Second, green labeling is only partially observed.** Although the platform labels delivery slots in real time, it does not record whether a specific slot is labeled as “Less CO.” Instead, the data only reports the number of green-labeled slots and the total number of available slots at the time of ordering. Therefore, the analysis must focus on the intensity of exposure to green labeling rather than discrete slot-level choice. Moreover, exposure rates mechanically differ across districts with varying order density, further complicating naive comparisons.

**Third, exposure may exhibit limited dependence across orders.** By design, whether a slot is labeled green depends on existing orders scheduled to the same district and delivery window. In principle, earlier orders could therefore affect the labeling environment faced by later customers, raising concerns about dynamic feedback and interference. However, because labeling is driven by aggregate district-level conditions rather than individual choices, this dependence is likely small after conditioning on district and slot characteristics. The analysis explicitly evaluates this possibility through aggregation and robustness checks.

These challenges motivate a design that relies on within-district variation rather than time-series comparisons, models exposure intensity rather than slot-level labels, and validates individual-level results using district-level aggregation. District fixed effects and slot controls isolate exposure-driven variation, while aggregation mitigates residual dependence across orders generated by the platform’s labeling rule. As a result, estimated effects correspond to policy-relevant changes in green-slot exposure under realistic operational constraints.

### 3. Data and Institutional Setting

The data consist of order-level records from December 1 to December 15, 2023, covering approximately 15 delivery dates. For each order, the platform records:

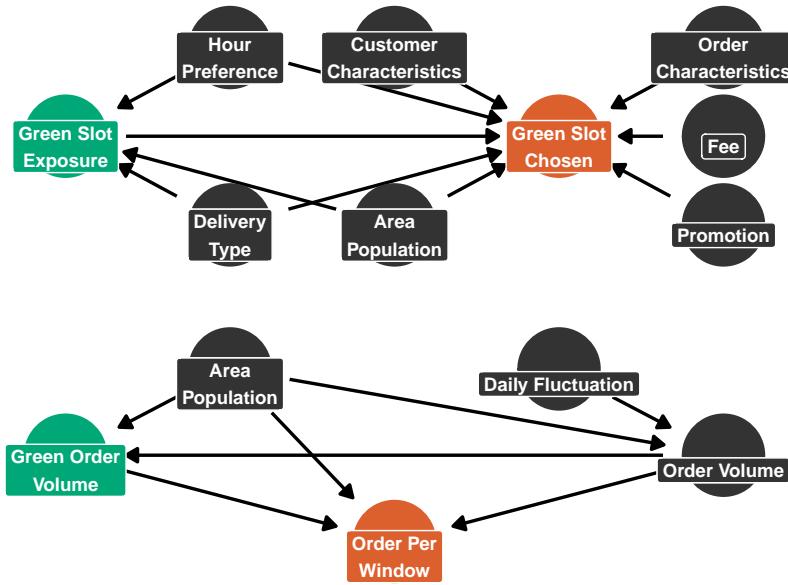
- the number of available delivery slots,
- the number of available green-labeled slots,
- the selected delivery slot,
- delivery district, slot type, slot hour, and other order characteristics.

A key limitation is that slot-level labels are not observed. Instead, exposure to green labeling is measured as the share of available slots that are labeled green at the time of ordering. This motivates focusing on exposure intensity rather than discrete slot choice.

### 4. Causal Framework and Identification Strategy

**Key challenge:** Green labeling is not randomly assigned. Whether a slot is labeled green depends on existing orders scheduled to the same district and delivery window.

The graph below summarizes the causal structure guiding the analysis. Answering RQ1 and RQ2 requires estimating the impact of green nodes on orange nodes, highlighted in the graph:



**Key identifying assumption:** Conditional on district and slot characteristics, short-run fluctuations in green-slot exposure are driven by platform state rather than by customers' unobserved preferences or strategic behavior.

This assumption is plausible because:

- customers do not control whether a slot is labeled green,
- labeling rules are mechanical and district-specific,
- customer and order characteristics influence *choice*, but not labeling itself.

Dynamic feedback is possible in principle (earlier orders affect later exposure), but diagnostic checks (see Section 7) show smooth exposure dynamics with no sharp short-run feedback. To further mitigate this concern, results are validated using district-level aggregation.

While the RQ1 focuses on individual choice behavior, the second research question evaluates whether increased green-slot uptake translates into operational delivery clustering. Identification for RQ2 relies on a district-by-date panel specification that exploits within-district variation in the share of green orders across delivery days. By aggregating orders to the district level, this analysis intentionally absorbs individual-level preference heterogeneity and instead treats green uptake as an equilibrium outcome of platform demand conditions.

Importantly, the feedback mechanism that complicates individual-level inference strengthens identification at the aggregated level: because green labeling is triggered by aggregate demand within a district, changes in green-order share reflect shifts in delivery density that are relevant for routing efficiency. Controlling for total order volume and district fixed effects isolates the extent to which green uptake increases average orders per delivery window, rather than merely capturing days with higher overall demand. Under the assumption that, conditional on volume, short-run fluctuations in green-order share are not driven by unobserved district-specific shocks to routing efficiency, the estimated coefficient captures the causal effect of green uptake on delivery clustering.

## 5. Empirical Models

### 5.1. Customer Choice Response (RQ1)

To estimate how exposure to green-labeled slots affects customer choice, I estimate the following logit model for each observation at order  $i$ :

$$\text{logit}(P(\text{GreenChoice}_i = 1)) = \beta \text{GreenExposure}_i + \gamma \text{SlotControls}_i + \delta_d$$

where

- **GreenExposure** is the share of available slots labeled green,
- **SlotControls** include slot type and slot hour,
- $\delta_d$  are district fixed effects.

This specification identifies the effect from within-district variation across comparable delivery options, isolating the behavioral response to green labeling.

### 5.2. Delivery Clustering Efficiency (RQ2)

To assess operational impact, orders are aggregated to the **district**  $\times$  **delivery-date** level (for each delivery type). Here, each delivery type is conservatively assumed to operate separately from the others, although they may have overlapping delivery windows, hence estimating 3 separate models for 3 delivery types. The outcome is the average number of orders per delivery slot, a proxy for clustering efficiency.

$$\log(\text{AvgOrdersperSlot}_{d,t}) = \alpha \text{GreenOrderShare}_{d,t} + \lambda \text{TotalVolume}_{d,t} + \theta_d + \epsilon_{d,t}$$

where  $\theta_d$  are district fixed effects, and  $\epsilon_{d,t}$  captures idiosyncratic variation.

This model examines whether higher green uptake results in denser delivery clusters, controlling for daily demand.

## 6. Results

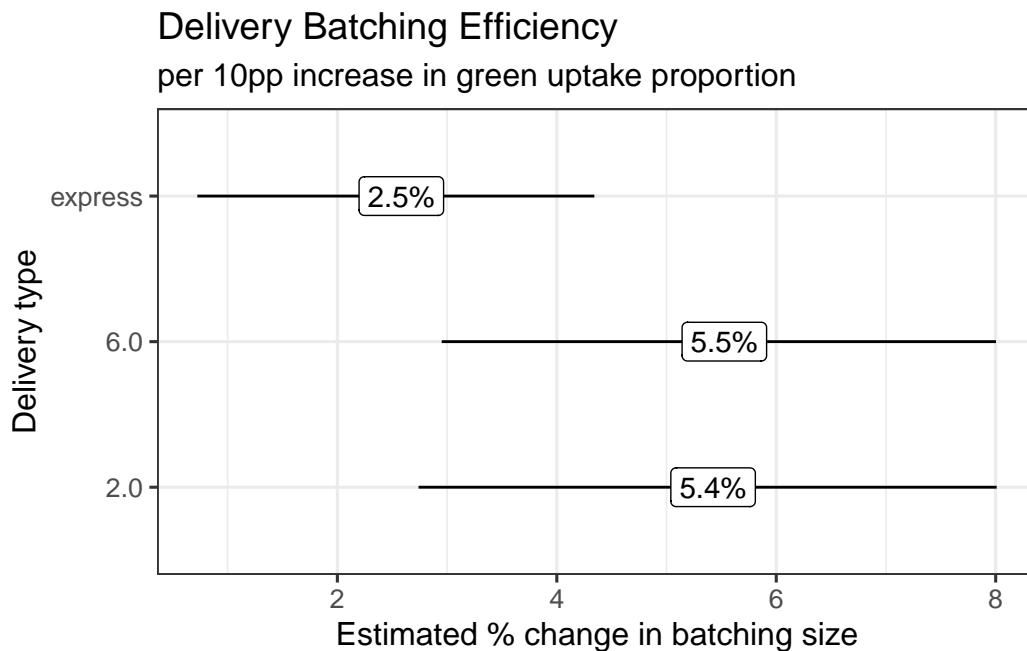
### 6.1. Effect of Green Exposure on Customer Choice

There are 15% of all orders where the chosen delivery time slots are green-labeled. On average, 8% of the time slots available to customers are green-labeled by the platform at the time of order. A 10 percentage-point increase in green-slot exposure increases the probability of selecting a green slot by approximately **9 percentage points**, holding district and slot characteristics constant.

This effect is stable across alternative specifications and aggregation levels.

### 6.2. Effect of Green Uptake on Delivery Clustering

The following graph shows the estimated effect of green uptake on delivery clustering efficiency.



Across delivery types, higher green uptake results in modest gains for express delivery and larger gains for 2-hour (saver) and 6-hour (standard) windows. In particular, a 10 percentage-point increase in the proportion of green orders results in a **2.5%**, **5.4%**, and **5.5%** increase

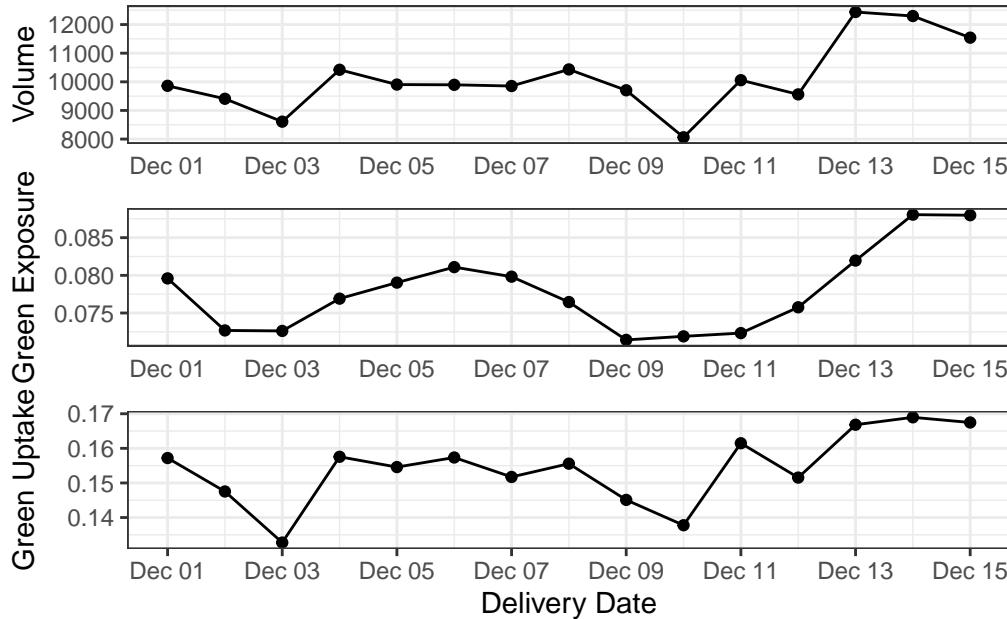
in the number of orders per delivery window for express, saver, and standard delivery, respectively.

These results suggest that the green-labeling mechanism not only changes customer behavior but also meaningfully improves delivery batching. The larger gains for wider delivery windows are consistent with greater flexibility for route consolidation.

## 7. Diagnostics and Robustness

The following plot shows delivery volume, green-slot exposure, and green uptake by delivery date. It is worth noting that there are weekend effects: delivery volume is lower on Saturdays and Sundays, reflecting operational scheduling. At the same time, exposure and uptake drop near the boundaries of the sampling window, likely due to mechanical truncation of the sampling window.

Neither pattern biases the estimated effects, which rely on within-district variation rather than time-series trends. Estimates are robust to excluding boundary dates and to alternative aggregation strategies.



This project focuses on point estimates that mainly answer stakeholders' questions. For another robustness check, I also validated the main results (in RQ1) using a Bayesian hierarchical binomial model to better account for uncertainty and partial pooling across districts. I found the results support the main findings.

## **8. Limitations and Interpretation**

This analysis relies on observational data collected over a short time window during which the green-labeling feature had already been fully deployed. As a result, causal identification depends on conditioning assumptions rather than experimental variation, and longer-term behavioral adaptation or learning effects cannot be evaluated. The platform does not record slot-level labeling status, which necessitates modeling exposure through the proportion of green-labeled slots rather than direct choice sets.

For the order-level analysis, identification assumes that, conditional on district and slot characteristics, short-run variation in green-slot exposure is not driven by unobserved customer preferences or strategic behavior. While the labeling mechanism introduces a mechanical dependence across orders, robustness checks using alternative specifications and district-level estimation suggest that such feedback effects do not materially bias the estimated behavioral response. At the district level, aggregation mitigates individual-level interference and allows green uptake to be interpreted as a system-level state variable relevant for delivery efficiency.

The findings are most applicable to high-density urban delivery settings where routing efficiency depends on geographic and temporal consolidation. Extrapolation to platforms with different fulfillment models, customer populations, or delivery constraints should be made cautiously. Finally, while green labeling is presented as a voluntary nudge, platforms should remain attentive to how such design choices may differentially affect users with tighter delivery constraints, ensuring that sustainability interventions do not inadvertently reduce access or convenience for certain customer groups.