Lucy Page and Lisa Ho
July 2022
Analysis Plan: On-Platform Carbon Labeling Experiment

1 Introduction

This experiment will be run on the online platform of a market-leading meal kit company, and it will test whether implementing carbon footprint labels in this commercial, real-stakes environment effectively shifts customers to more sustainable meal choices. The experiment will include four experimental arms—one control group and three different labeling schemes—to be randomized at the individual level. We will then observe the impacts of the labels on customers' meal choices and retention, as well as on the company's per-customer cost, revenue, and profit.

Following this on-platform experiment, we will run an off-platform follow-up experiment testing the mechanisms behind any effects, as well as the impacts of the carbon labels (and related carbon-footprint campaigns) on customers' engagement with political climate advocacy. We will describe this follow-up experiment in a separate pre-analysis plan.

2 Sample population

The sample was allocated from the partner company's standard pool of customers between mid-April and mid-May 2022. In total, the sample includes over 200,000 customers that visited the company's website during the recruitment period.

3 Carbon labeling intervention

3.1 Labeling schemes

The experimental sample is randomized across four experimental arms: a control arm and three treatment arms with different labeling schemes. Our partner company developed each of these labeling schemes by first using the Agribalyse dataset to approximate the carbon footprints of each meal on a given week's menu. They then split the list of meals for each week into three tiers based on their estimated footprints: "Climate Superstars," which have an estimated footprint below 2 kg CO_2, "Good," which have estimated footprints between 2 kg CO_2 and 7 kg CO_2, and "Fair," which have estimated footprints above 7 kg CO_2.

Customers select meals off of a main menu page showing all of the meals available in a week. From this page, they can then click on menu cards for each meal with more detail on its ingredients, nutritional content, etc. Our partner company routinely labels meals with a variety of indicators for being low-calorie, vegetarian, spicy, etc., including small versions of each label on the main menu page and more detailed versions of each label on the menu cards. Likewise, each of our labeling schemes will include a set of small carbon labels portrayed on the main menu and somewhat larger labels displayed on the menu cards. Under each of the labeling schemes, the menu cards will also include descriptions of the labels and what each category means. The labeling schemes are as follows:

- Control: No carbon labels are included on the menu, though meals have standard labels for low-calorie, spicy, etc.
- Letter labels: Meals are labeled as A, B, or C if they are in the "Climate Superstar," "Good," or "Fair" footprint categories, respectively.

Figure 1: On main menu

Figure 2: On menu cards

- Abstract labels: Meals will be labeled with a globe icon and the text of their category name.

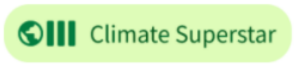
Figure 3: On main menu

Figure 4: On menu cards

- "Climate Superstar" labels: Meals in the "Climate Superstar" category are labeled with a globe icon and their category name.

Figure 5: On main menu

Figure 6: On menu cards



3.2 Treatment timeline

Treatment labels were first released on the company's online menu on May 18, during the company's Week 20, meaning that meal choices for Weeks 21 onwards could have been affected by the labels. However, our partner company always posts menus six weeks in advance of delivery, and customers can select their meal choices at any point in that window. Thus, many customers may have made meal selections for Weeks 21 through 25 without being exposed to the carbon footprint labels. We will estimate treatment effects separately for each week (though these results may be in an appendix), but we may decide to define the "experimental period" as Weeks 23 and onwards, for example.

4 Research Questions

Our primary research questions for the on-platform experiment are as follows. Note that our follow-up survey experiment will include more detailed research questions as well.

1. How much do people change their meal choices in response to climate labels?
2. Do the climate labels change customers' retention, decision to purchase meal boxes, or referrals to friends?
3. Do the climate labels affect costs, revenue, or profit per customer?
4. How do these impacts vary with customer characteristics, like gender, environmentalism, political affiliation, or taste in different foods?

5 Outcomes for analysis

5.1 Primary outcomes

- **Customers' meal choices:** We will observe customers' meal choices for the full experimental period and for several weeks before and after the experimental period. We will create several outcome variables based on customers' meal choices:

  - Estimated carbon footprint of meals that customers choose, including total carbon footprint across their three meal choices
  - Number of meals that customers purchase in the "Climate Superstar" category or in the "Fair" category and whether they purchase a meal in a given category;
  - Number of meals that customers choose containing a given protein source (e.g. poultry, fish, pork, beef, vegetarian) and whether customers choose a meal with a given protein source;
  - Whether customers switch the protein in a given meal to a lower- or higher-carbon protein.
    * Our partner company only recently introduced the option for customers to change the protein of a meal, so we may eliminate this outcome if the data suggests that customers do not use this option.

- **Customer retention:** We will observe whether customers buy a meal box in a given week, and if so, how many meals they buy.
  - Note that if we find significant effects of the labels on customer retention, that could induce bias in our estimates for the treatment effects of the labels on meal choices, which we can only observe among customers buying meal boxes in a given week. In that case, we will use techniques to account for differential attrition when bounding our treatment effect estimates.

5.2 Secondary outcomes

- **Company costs and profit per customer**: We will observe our partner company's costs, revenue, and profits for each customer each week before, during, and after the experimental period.
  - We will only test for the impacts of the carbon labeling treatments on costs and profit per customer if initial cross-sectional analysis suggests that there are significant differences in costs and profit margins for meals with different protein composition.

- **Customer referrals to friends:** As a measure of customers' approval of our partner company, we will test the impacts of the labels on customers' referrals to friends.

- **Meal kit donations:** Customers always have the option to donate their meal kit to households experiencing food insecurity via a program run by our partner company. We will observe whether customers donate their meal box in a given week.
  * Note that we may not include this outcome if we find that it is extremely rare for customers to donate their mealboxes.

6 Estimating treatment effects

We will estimate our basic treatment effects using the following simple regression models:

$$y_i = \alpha + \sum_{j=1}^{3} \beta_j T_{ji} + X_i'\gamma + \epsilon_i \tag{1}$$

where $\{T_{1i}, T_{2i}, T_{3i}\}$ are dummy variables for the three label treatment arms, y_{it} is our outcome of interest, and X_i is a vector of participants' baseline characteristics.

- In our baseline models, we will control for lags of our outcome variables, the number of meals each customer buys, and the meal plan preference in which each customer is enrolled (e.g. vegetarian, chef's choice), which determines the default meals they are offered and the order in which meals are presented on the menu.

- In other models, we will include controls for baseline characteristics that may include customers' gender, state, political affiliation, food preferences, education levels, and environmentalism.
- Note that we are not yet in a position to lay out precisely what these control variables will be. We will be merging in data on customer characteristics collected from Experian, but we do not yet know what these data will include.

6.1 Restricting analysis to active meal choosers

The meal company with which we're partnering selects default meal choices for each customer based on their meal plan, and customers then have the option to switch these default meals to another meal choice. Since we would only expect the climate labels to affect meal choices among those actively engage with the menu, we will restrict in our preferred regressions to customers that alter their meal choices away from the defaults.

6.2 Heterogeneity by respondent characteristics

Using the Experian data that we will collect, we will look at heterogeneity of our treatment effects by certain baseline characteristics. Our regressions would look like the following:

$$y_i = \alpha + \sum_{j=1}^{3} \beta_j T_{ji} + X_i'\gamma + \sum_{j=1}^{3} \delta_j T_{ji} x_i \ + \epsilon_i \tag{2}$$

where x_i is a baseline characteristic for individual i. Note that the main effect of x_i would be included in the vector X_i. If the Experian data includes these characteristics, we may examine heterogeneity by political affiliation, gender, environmentalism, and educational attainment. We may also test for heterogeneity by customers' historic meal choices on the platform, e.g. whether they always choose beef or only choose beef occasionally.

6.3 Baseline survey for heterogeneous effects

To examine heterogeneous treatment effects, we have also collected a baseline sample from among the customers in our experimental sample. We sampled this baseline sample before treatment labels were released on our partner company's website, eliciting participants' beliefs about climate change, the extent to which they believe they are environmentally conscious, altruistic, health-conscious, etc., their political affiliation, and their baseline knowledge about the carbon footprints of different foods.

Our partner company recruited this sample by emailing all participants allocated to the experiment informing them of the opportunity to complete the short survey in exchange for being entered into a lottery to win a $100 or $200 gift card. (We changed the gift card amount in a second round of emails and data collection.) In total, we recruited only about 7,000 completed responses, about 3% of all customers ultimately enrolled in the experiment.

Given this very low recruitment rate, we will not put much stock in heterogeneous treatment effects using data collected via this baseline sample. In particular, we may not present results using this baseline survey if Experian data or administrative data suggests that they are substantially different from the full experimental sample or if baseline customer characteristics are not balanced across treatment arms within this baseline sample.