Nudging plant-based protein products in a real-life online supermarket: a randomized controlled trial

Analysis plan

Nina van der Vliet & Josine M. Stuber, et al

Research questions

We will primarily evaluate whether the implementation of nudges will result in a between group mean-difference in the number of plant-based protein purchases in an online supermarket. As secondary outcome, we will evaluate between groups the mean-differences in the number of purchased plant-based meat replacements, animal-based meat products, plant-based dairy purchases, and animal-based dairy purchases. As a tertiary outcome we will investigate the mean-differences in retailer revenue as a relevant business related outcome. In addition, we will explore if potential effects are sustained after the intervention period has ceased and investigate potential interactions with the primary outcome for neighborhood socio-economic positon (SEP), age and sex.

Population

All online supermarket customers during the study period.

Outcome measure (dependent variables)

1. The number of plant-based protein products purchased.
2. The number of purchases from (a) meat, (b) plant-based meat substitutes, (c) dairy products, and (d) plant-based dairy products.
3. Total revenue in Euros.

Independent variable

Categorical variable (two trial arms)

1. Arm 1: Regular online supermarket (control).
2. Arm 2: Online supermarket with the addition of nudges on plant based protein products.

Confounders

Randomization results in unobserved confounders being randomly and equally distributed across arms.

Effect-modifiers

We will test for effect modification by sex and neighborhood deprivation levels in the primary outcome model. All results will be stratified in the case of a significant ($p < 0.05$) interaction term. If both variables turn out to be effect modifiers, the variable with the smallest p-value for the interaction term will be used as stratification variable and the other variable will be included as covariate in the models.

Survey data collected among sub-sample

All customers shopping during the duration of the experiment will be invited to participate voluntarily in a survey after checkout. Data to be collected via the survey:

- Frequency of meat consumption;
- Eating pattern (meat eater, flexitarian, vegetarian, vegan and asking "how many days a week do you eat meat during dinner?" "how many days a week do you eat a plant-based alternative to meat during dinner?");
- Meat attachment (measured with a shortened version/one or two subscales of the Meat Attachment Scale (7-point Likert scale, including items such as "Eating meat is one of the pleasant things in life");
- Perception of norms (7 point-Likert scale, including items on descriptive and injunctive norms regarding plant-based alternatives to meat such as: "How likely is it that other customers like you would choose a plant-based alternative in the webshop?";
- Efforts and salience of plant-based alternatives to meat (on a 7 point Likert scale, including items such as "Plant based products stood out");
- Customer satisfaction and experience during shopping in the online supermarket;
- Highest attained educational level.

Statistical analysis

Descriptive statistics will include the proportion of females (n(%)), the proportions per age category (n(%)), the proportion of shoppers from deprived areas (n(%)), and the total amount of items purchased per shopper per food group (e.g., median(IQR)). Detailed customer characteristics derived from the sub-sample who complete an additional questionnaire will be included as explorative population descriptives and opinions on nudging of plant-based products.

Shoppers who did not purchase any animal or plant-based protein products will be excluded from the analyses of the primary and secondary outcomes, along with the upper 1% of shoppers in terms of number

of products purchased to avoid over- or underestimation of nudging effects due to extreme cases. All shoppers will be included for the analysis of differences in revenue.

We expect a high number of true zero values in the purchasing data, especially in the plant-based food groups, since not all customers will purchase products from all protein groups. Furthermore, we expect highly right skewed data destitution as most customers will buy one or two products from a specific group, whereas a smaller share will purchase products in bulk. We therefore expect that the assumption of the Poisson distribution (i.e., equal variance to the mean) will be violated. To analyze this type of data, we consider a generalized linear mixed model with a Conway–Maxwell Poisson distribution as the most suitable approach. It is a flexible generalization of the Poisson distribution that is suitable for modeling of overdispersed as well as underdispersed count data. We will therefore apply GGLMs with a Conway–Maxwell Poisson distribution using a random intercept at the customer level while adding the store as fixed effect for our primary and secondary outcomes, to estimate the incidence rate ratios (IRRs) and 95% CI's for products purchased in the nudging group compared to the control group. IRRs in which the 95% CIs do not include 1 will be considered statistically significant. Modeled with the *glmmTMB* R package by Magnusson et al. (2020), the Conway–Maxwell Poisson distribution is interpretable as a log-linear model. The IRR therefore reflects the percentage difference in purchases in the nudging arm compared with the control arm. If the data does fit the assumptions of a Poisson distribution, we will use the analyses above, but without the Conway-Maxwell component. Modelled with (glmer R package).

We expect the continuous tertiary outcome, total Euros spent, to be approximately normally distributed. Therefore, a linear mixed model with a random intercept at the customer level while adding store as a fixed effect will be used to analyze between group differences in total revenue.

To explore the combined effects of nudges with price promotions we will use explorative analyses where we only select the intervention cases and compare purchases in the weeks with regular price promotions on meat substitutes and/or plant based dairy products with the weeks without within regular price promotions (3 versus 3 weeks), Moreover, we will explore the number of clicks, add-to-carts, and conversion ratio on plant-based protein products following the same analytical approach as the primary outcome. Last, for a subsample of the population we will additionally explore moderating psychological determinants of purchasing behavior based on the survey data.