1 Methodology

The study was designed as a cluster crossover quasi-experiment. Section 1.1 summarizes relevant information on the setting where the study was implemented, Section 1.2 provides the details of the study design and Section 1.3 illustrates the statistical models adopted.

1.1 Setting

The study was conducted at the canteen of the University of Gastronomic Sciences of Pollenzo (UNISG), a private institution located in northern Italy and attended by Italian and international students. The canteen is presented by the university as a distinctive project called *Academic Tables*, that *"combines education, sustainability, and local products at a fair price"*.

1.1.1 Booking methods

At the canteen, each item was to be ordered singularly, either:

1. *pre-ordering*: on the pre-ordering website, where every Friday the menu of the following week was uploaded by the canteen staff. To pre-order on the website, students must place their order at least 24 hours in advance.

2. *last-minute*: on site, where touch-screen totems installed at the entrance of the canteen allowed patrons to print the receipts of their pre-orders as well as to order food for same-day consumption;

1.2 Study design

The experiment involved the naming of three vegetable plates served in a university canteen during lunch on weekdays. These three vegetable plates - a green salad, a plate of cooked vegetable and a pulses salad - were prepared every day with organic and sustainable ingredients and with little-to-no-variation in the recipes across days. The intervention consisted in altering the naming of the vegetable plates by including key words that emphasized the organic and/or sustainable origin of the ingredients. Specifically, during the 8-weeks in which the experiment was in place, the three vegetable plates alternatively received one of four labelings: *standard*, *organic*, *sustainable*, *organic sustainable*.

1.2.1 Treatment assignment

Due to the way that the experimental labels were assigned to vegetable plates, this study is a *cluster crossover quasi-experiment*: a longitudinal study where clusters (e.g., a canteen) receive a sequence of treatments. Every week, the label *standard* was assigned to vegetable items on Monday and Friday and the experimental labels *organic*, *sustainable* and *organic & sustainable* were alternated between Tuesday, Wednesday and Thursday in a fashion defined prior to the study to ensure variation across weeks. The main implications of this study design are as follows: (i) we observe the same level of intervention multiple times; (ii) in each time period all customers are assigned to the same intervention; (iii) the same customer can be observed under different interventions. These implications are visually summarized in Figure 1.

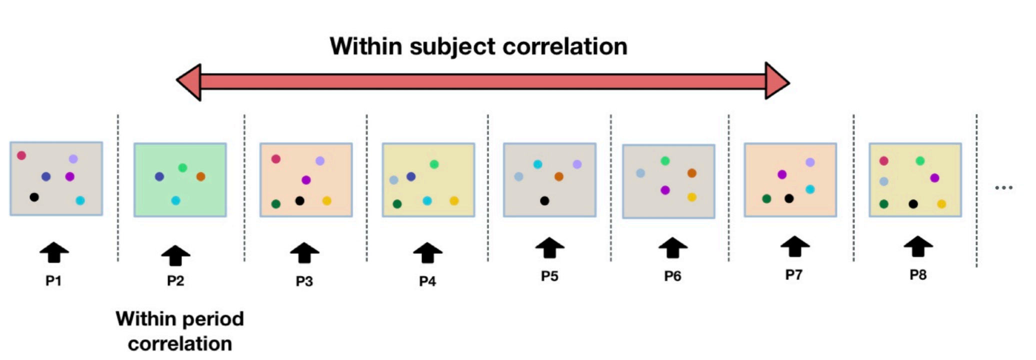

Figure 1: A diagram that summarizes the study design. P1, P2, ... represent the periods of the experiment, which correspond to the menu days. The rectangles represent the canteen on different days. The color of the rectangle represents the label assigned to the vegetable items on a given day (gray is for the standard label). The circles in the rectangles represent the students on a given day. The color of a circle identifies a specific student; a repeated color indicates that the same student attended the canteen on multiple days. This diagram is an adaptation of Figure 1 in Parienti and Kuss (2007).

1.3 Statistical methodology

In light of the characteristics of the study design and considering that prior information was available through the pilot data, the statistical modeling approach chosen to address both primary and secondary aims of the study was a hierarchical Bayesian model. This approach corresponded to a multilevel analysis that accounted for the correlation between repeated observations on the same subject and that extracted relevant prior information by using the pilot data to inform the choice of prior distributions for the model parameters of the main analyses.

1.3.1 Primary analysis

The model chosen for the primary analysis accounted for two sources of correlation: (i) observations collected on the same subject; and (ii) observations collected in the same day (see Figure 1). The analysis population for the primary analysis was all available records of students in the data.

1.3.1.1 Model likelihood

Let $y_{it\ell} \in \{0, 1\}$ be a binary variable indicating if student i on a day t assigned to label ℓ ordered at least one vegetable plate. The observation likelihood for the model adopted is

$$Y_{it\ell} \mid p_{it\ell} \sim \mathsf{Bernoulli}\left(p_{it\ell} = \mathsf{logit}^{-1}(\beta_{i0} + \beta^*_\ell + \tau_t + \mathbf{x}_i^T \gamma + z_{it}^T \eta)\right) \tag{1}$$

where β_{i0} is the random effect of student i (between-periods correlation), β^*_ℓ is the label effect of label $\ell \in \{$*organic*, *local*, *organic & local*$\}$, τ_t is the random effect of menu date t (within-period correlation), $\gamma = (\gamma_{\text{male}}, \gamma_{\text{grad}}, \gamma_{\text{age}})^T$ is a vector of coefficients of subject-specific covariates $\mathbf{x}_i$ (indicator for male, indicator for graduate students and normalized age at the beginning of the study), $\eta = \eta_{\text{ LM}}$ is the coefficient of a subject-specific time-varying covariate z_{it} that indicates whether subject i's order on day t was pre-ordered or placed on the same day.

1.3.1.2 Model priors

The following describe the priors assigned to the parameters in Equation 1:

Student-level correlation

$$\beta_{i0} \sim \mathsf{Normal}\left(\mu_{i0}, \sigma^2_{i0}\right)$$

For all students that were observed in the pilot data, μ_{i0} and σ^2_{i0} were set as the estimated mean and standard deviation of the posterior distribution of β_{i0} obtained from the analysis of the pilot data. For the students that were not observed in the pilot data, μ_{i0} and σ^2_{i0} were set as the estimated mean and standard deviation of the distribution of μ_{i0} obtained by analyzing the pilot data.

Parameters of interest

$$\beta^*_\ell \mid \mu_\ell, \sigma^2_\ell \sim \mathsf{Normal}\left(\mu_\ell, \sigma^2_\ell\right)$$

The values of μ_ℓ and σ^2_ℓ were set, respectively, to 0 and 10 so that this prior would be broad and weakly informative.

Within-day correlation

$$\tau_t \sim \mathsf{Normal}\left(\mu_\tau, \sigma_\tau^2\right)$$

The values of μ_τ and σ_τ^2 were set at the estimated mean and standard deviation of the posterior distribution of μ_τ obtained from the analysis of the pilot data.

Precision variables

$$\gamma_k \overset{\text{ind}}{\sim} \mathsf{Normal}\left(\mu_k, \sigma_k^2\right), \qquad k \in \{\text{male, grad, age}\}$$

$$\eta_{\text{ LM}} \sim \mathsf{Normal}\left(\mu_{\text{ LM}}, \sigma_{\text{ LM}}\right)$$

Similarly as for β_{i0} and τ_t, μ_k, σ_k^2, $\mu_{\text{ LM}}$ and $\sigma_{\text{ LM}}$ were set as the estimated mean and standard deviation of the posterior distribution of the coefficients of male indicator, graduate indicator, standardized age and last-minute indicator obtained from the analysis of the pilot data.

1.3.2 Implementation

The Bayesian models were estimated using the Hamiltonian Monte Carlo (HMC) sampler, which is an efficient Markov Chain Monte Carlo (MCMC) method. The implementation of the samplers was performed using the R interface **rstan** of the Stan probabilistic programming language (Carpenter et al. 2017).

The GLME models were implemented using the R package **lme4** (Boeck et al. 2011)

References

Boeck, Paul De, Marjan Bakker, Robert Zwitser, Michel Nivard, Abe Hofman, Francis Tuerlinckx, and Ivailo Partchev. 2011. "The Estimation of Item Response Models with the Lmer Function from the Lme4 Package in R." *Journal of Statistical Software* 39 (March): 1–28. https://doi.org/10.18637/jss.v039.i12.

Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A. Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76: 1. https://doi.org/10.18637/jss.v076.i01.