Analysis plan for "Nudging in complex environments"

Alexander K. Koch[a,b], Dan Mønster[a] & Julia Nafziger[a,c,*]

[a]Department of Economics and Business Economics, Aarhus University,

Fuglesangs Alle 4, 8210 Aarhus, Denmark

[b]Center for Hybrid Intelligence, Aarhus University, Denmark

[c]Centre for Economic Policy Research (CEPR), UK

For our analysis, we rely on the regression analysis of covariance (ANCOVA) because it is more efficient than the alternative of a difference-in-differences estimator (McKenzie, 2012; Burlig et al., 2020; Ek, 2020). We estimate the following equation via a least squares regression:

$$Y_{i,t} = \alpha_t + \beta\,\mathbf{X}_i + \gamma\,T_{i,t} + \theta\,\bar{Y}_{i,Pre} + \epsilon_{i,t}, \tag{1}$$

where i refers to the individual and the post-treatment date t to the recipe number, $Y_{i,t}$ is the outcome variable of individual i at date t, α_t are recipe fixed effects (in place of a constant), $\mathbf{X}_i$ is a set of control variables, $T_{i,t} \in \{0,1\}$ is the treatment status of i at date t, and $\bar{Y}_{i,Pre}$ is the mean for individual i over the pre$-$treatment measures in game modules 1 and 2 (seven recipes) for individual i. ANCOVA is a convex combination of an ex-post means comparison and a difference-in-differences comparison between treatment and control, as (ignoring for simplicity the covariate vector $\mathbf{X}_i$) the ANCOVA estimator of the treatment effect is given by

$$\hat{\gamma} = \left(\bar{Y}_{Post}^{T} - \bar{Y}_{Post}^{C}\right) - \hat{\theta}\left(\bar{Y}_{Pre}^{T} - \bar{Y}_{Pre}^{C}\right), \tag{2}$$

where superscript T and C refer to treatment and control group, respectively. Note that in Hypotheses 2 and 5, *Reminder* serves as baseline C category for treatment *ManyReminders*.

We report three specifications: (i) without covariate vector $\mathbf{X}_i$, (ii) with the set of basic controls (cf. Table 1), (iii) with the set of extended controls (cf. Table 2). Further, for robustness, we add specifications where we address potential confounds.

Table 1: Control variables

Variable	Description	Values
Scorediff12	Difference in the average score per recipe in game modules 2 and 1	$1, \ldots, 5^a$
Enjoyment	The game is fun	$1, \ldots, 5^a$
SpeedPref	Important to be quick to complete a recipe in the game	$1, \ldots, 5^a$
HygienePref	Importance of preparing food under hygienic circumstances	$1, \ldots, 5^b$
PriorKnowledge	Food safety related knowledge based on Koch et al. (2021)	continuous
PriorBehaviors	Food safety related behaviors based on Koch et al. (2021)	continuous
Age	Age in years	$20, \ldots, 60$
Gender	base category: male, dummy for female, dummy for non-binary/third gender	
SingleHousehold	Dummy=1 if the participant lives in a single-person household	0,1
CurrentSituation	University/college is baseline, dummies for high-school, vocational training, employed in public sector, self-employed, unemployed, other category	
FreqMeat	How often the participant prepares a warm lunch or dinner with meat (including poultry) on average	
FreqComputerGames	Frequency of playing computer games	$0, \ldots, 7^c$
WorkedFoodSector	Dummy for whether the participant has ever worked in the food industry or in gastronomy/food service, coded 1 if yes and 0 if no	0,1
HealthSector	Dummy for whether the participant has ever worked as a health professional (health worker, nurse, doctor, physician, nutritionist, …); include health students and health vocational training, coded 1 if yes and 0 if no	0,1
HadFoodPoison	Dummy for whether the participant has ever had food poisoning, coded 1 if yes and 0 if no/don't know	0,1
SharePreprepared	Share of pre-prepared meals that just need to be heated	$0, \ldots, 5^d$
Risk tolerance	Question of Dohmen et al. (2011)	$1, \ldots, 11$

Notes: ^{a} 1=strongly disagree, 2=somewhat disagree, 3=neither agree nor disagree, 4=somewhat agree, 5=strongly agree ^{b} 1=very unimportant, 2=slightly unimportant, 3=neutral, 4=slightly important, 5=very important ^{c} 1=never, 2=0 - 1 hour a month, 3=1 - 3 hours a month, 4=1 - 2 hours a week, 5=4 - 6 hours a week, 6=1 - 2 hours a day, 7=More than 2 hours a day ^{d} 1=none (0%), 2=1 in 4 (25%), 3=half (50%), 4=3 in 4 (75%), 5=all (100%)

Table 2: Extended set of control variables

Variable	Description	Values
InfoSeek	Information seeking about food safety in week prior to experiment	$1,\ldots,5^a$
Children	Dummy=1 if living with children	0,1
Stressed	How often the participant felt stressed when cooking because of time pressure	$1,\ldots,5^a$
ConcernedFoodPois	Food-related risk tolerance	$0,\ldots,11^b$
KitchenCleanPref	Importance of not messing up the kitchen when cooking	$1,\ldots,5^c$
AvoidWastePref	Importance of avoiding waste	$1,\ldots,5^c$
HygieneMyth	Agreement with statement "Too much hygiene in the kitchen is the cause of allergies and prevents building up a good immune system"	$1,\ldots,5^d$

Notes: [a] 1=never, 2=once, 3=twice, 4=3-4 times, 5=5 times or more [b] Concerned about getting sick from food poisoning? 1=not at all concerned … 11=very concerned [c] 1=very unimportant, 2=slightly unimportant, 3=neutral, 4=slightly important, 5=very important [d] 1=strongly disagree, 2=somewhat disagree, 3=neither agree nor disagree, 4=somewhat agree, 5=strongly agree

References

BURLIG, F., L. PREONAS, AND M. WOERMAN (2020): "Panel data and experimental design," *Journal of Development Economics*, 144, 102458.

DOHMEN, T., A. FALK, D. HUFFMAN, U. SUNDE, J. SCHUPP, AND G. G. WAGNER (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, 9, 522–550.

EK, C. (2020): "ANCOVA power calculation in the presence of serial correlation and time shocks: A comment on Burlig et al.(2020)," *Working Paper*.

KOCH, A. K., D. MØNSTER, J. NAFZIGER, AND N. VEFLEN (2021): "Fostering safe food handling: Causal evidence on game- and video-based online interventions," .

MCKENZIE, D. (2012): "Beyond baseline and follow-up: The case for more T in experiments," *Journal of Development Economics*, 99, 210–221.