ANALYSIS PLAN

Main model specification

In our main specification, we distinguish between German and French study groups according to the language of the study instructions, which also corresponds to the main schooling language of pupils.

The model takes the following form:

$$Y = f(DE, \boldsymbol{X}),$$

where Y is one of the key outcomes of interest, i.e., exponential discounting rate δ or present bias β explained in document "Outcomes explanation", DE is a dummy variable that takes value 1 if the study language is German and 0 if the study language is French, $\boldsymbol{X}$ is a vector of control variables. We estimate the effects of interest using OLS regression, i.e., we allow for linear relationship between the study language and the corresponding outcome. Causal inference is based on heteroscedasticity robust standard errors clustered on the class level.

We consider two model specifications depending on control variables $\boldsymbol{X}$ included into the regression. The first (simple) model controls for risk tolerance parameter; the second (extended) model conditions on risk tolerance parameter, a dummy variable for correctly answering all comprehension questions (relevant for the outcome), gender, age (given year and month of birth), migration background, family structure and material conditions, parents' background, schooling, values. This allows us to take into account potential differences in sociodemographic characteristics between German and French study groups.

Observations with missing values in control variables are excluded from the analysis.

Robustness checks

We investigate sensitivity of our main results to the definition of main language. To this end, we examine whether the estimates change if we exploit native language, language(s) spoken at home or with friends instead of the study language. To capture all language groups potentially present in our sample, we allow for several language dummies. For instance, if we observe French and German speakers, and bilingual individuals, we use one dummy variable for German speakers, another dummy for bilingual individuals, whereas French speakers remain the reference group in the model.

We re-estimate the main specification when dropping all observations with multiple switch points in the questions determining the outcome variables.

We exploit another definition of present bias as an outcome variable, i.e., we consider a dummy variable that takes value 1 if the elicited present bias parameter β is below 1 and 0 otherwise.

We use probit regression to estimate the language effect on binary outcomes or outcomes expressed in proportions.