

The code is for replicating the main results of the paper. Note that there are separate Excel sheets for the grade data and for the survey data. A short note about variable names:

Gender = 0 male, 1 female.

Bet = grades for courses $i = 1, 2, \dots, N$.

Year = course fixed effects (the cohort fixed effects are called Cohort in the variable list). Please refer to the do-file for additional details.

Superficial = Desirability. It is defined as the average of the standardized beauty and standardized income. Note that it is restandardized in line 38 of the do-file to ensure that the variance is equal to unity. The latter procedure is also described in the paper.

MinusMean = The unstandardized Manski (1993) leave-out-one mean with respect to income.

MinusMeanBeauty = Same as above, but for beauty.

SSI = The Mehic-Olofsson (2021) socioeconomic status index. This is used as a control variable in the survey data.

For data not originating from the survey, the data sources are either the grade administration system at the university (grade data), the Swedish Tax Agency (income data), and the beauty jury data (the beauty data), respectively. For additional details, please see Appendix C of the paper.