Sample Data Perturbation and Randomization Methods

Data Perturbation

Researchers may complete a preregistration to request the sample dataset that contains perturbed response data.

For ordinal and ratio variables, each value was perturbed by adding noise from a normal distribution, with the standard deviation of that distribution relative to the size of each variable's value range. The standard deviation was set as 20% of the size of the value range for ordinal variables and 10% for ratio variables. Ordinal variables with country-specific values (e.g., Income) were only perturbed to values applicable to the country that the sample data was collected from (Sweden).

All nominal variables had a 60% chance to be perturbed from their original value into a randomly selected different category. Eligible categories were limited to those present in the original wave 1 data from Sweden.

For all perturbed variables, values that were recorded as "missing" (e.g., NA, (Saw, skipped), (Refused), (Don't Know)) were kept as-is and not perturbed. Additionally, variables recorded as strings (i.e., verbatim religious denomination) were not changed in any way.

Data Randomization

The sample dataset available to anyone without a study preregistration contains perturbed response data (see above for details on the procedure). Additionally, the data in each column has been randomized across rows such that the covariance between variables has been rendered meaningless.