Power Calculation for Social Preferences of Different Income Groups

General

The power calculation is based on our earlier study, which included 282 spectators. There were 39 spectators with average income, 147 spectators with below-average income, and 66 spectators with above-average income (30 participants didn't provide information regarding their income). In the calculations below, we used only the participants with above- or below-average income, as the differences between these groups are the current study's focus.

Our primary interest is to examine differences in the redistribution choices of above-average-income and below-average-income individuals. The findings in our earlier study suggested that the two income groups differ in their average redistribution amounts within and between treatments. Participants with below-average income redistribute more in the Taste and Merit treatments and less in the Luck treatment compared to above-average-income participants. Between treatments, below-average-income participants treat the Taste and Merit treatments differently, i.e., they redistribute different amounts (on average) in those treatments but redistribute relatively similar amounts in the Taste and Luck treatments. On the other hand, those with above-average income treat the Taste and Merit treatments similarly but differ in their average redistribution amounts between the Taste and Luck treatments. Therefore, we calculate the required sample size to detect two differences: (1) The difference between the two income groups in their redistribution choices (reflected through the level of implemented inequality) within each treatment (three comparisons), and (2) The difference in the redistribution choices of the two income groups between treatments (three comparisons).

The sample-size calculations below are based on the effect sizes found in the previous study, $\alpha = 0.05$ (including a Holm-Bonferroni correction for multiple hypotheses) and power of 0.8 or higher. We target a sample size that will satisfy the above requirements for all six hypotheses.

Calculation

We conducted a data simulation power analysis to test the power to reject the above six hypotheses. [1]

Based on the results of the previous study, we created six groups by combining all possibilities of income levels (2 levels) with treatments (3 treatments). We have information regarding the average implemented inequality and its standard deviation for each of these groups. For each group, we randomly drew (assuming that the data is distributed normally) N observations. The number of observations, N, was taken from 5 to 400 with jumps of 5.

For each $N \in \{5,10,15,,,395,400\}$ we made 1000 iterations of a sample + regression for the simulated data. The dependent variable was the implemented inequality, and the explanatory variables were six dummy variables of each treatment interacted with above-average or below-average income.

For each regression, we ran a Wald-test-based comparison (using the R function linearHypothesis) with a Holm-Bonferroni correction to examine each hypothesis. For each N and each hypothesis, we examined the percentage of regressions for which this difference was statistically significant. This

[1] This kind of simulation-based calculation is described in detail in Arnold et. al. (2011) and several other papers.

percentage reflects the power for that specific N. The results are presented in Figure 1. We find that in order to reject the hypothesis “no difference between above-average and below-average income groups” with a power of 0.8, in the Luck treatment, we need 95 participants in each income group in the Luck treatment. For the Taste and Merit treatments, we need 45 and 130 participants in each income group, respectively. To detect “no difference between above-average and below-average income groups between the Luck and Taste treatments, we require 35 participants in each income group in each of these treatments. Similarly, we need 65 participants in each income group for the Luck and Merit treatments, and 190 participants for the Merit and Taste treatments. We thus plan on having 190 participants in each income group in each treatment, for a total of 190*6=1140 participants.

<u>Detailed Calculation</u>

For each $N \in \{5,10,15,,,395,400\}$, following the simulation of the data, we ran 1000 regressions. Each regression used the simulated level of implemented inequality as the dependent variable. Denoting the treatments by their names and Above_Income/Below_Income as the dummy variables that receive one if the participant has above-average income/below-average income, the independent variables were: Luck*Above_Income, Merit*Above_Income, Taste*Above_Income, Luck*Below_Income, Merit*Below_Income, Taste*Below_Income. (We used no constant in this specification).

For the within-treatment hypotheses, we test whether the coefficients of the Above_Income and Below_Income variables are significantly different from each other in each treatment.

For the between-treatment hypotheses, we compare the relevant coefficients for the difference between the two income groups’ implemented inequality levels between treatments. For example, to compare Taste and Luck, we test whether –

(Luck*Above_Income -Taste*Above_Income) - (Luck*Below_Income – Taste*Below_Income)

is significantly different from 0. We make a similar calculation for the two other comparisons – Merit vs. Taste and Merit vs. Luck.

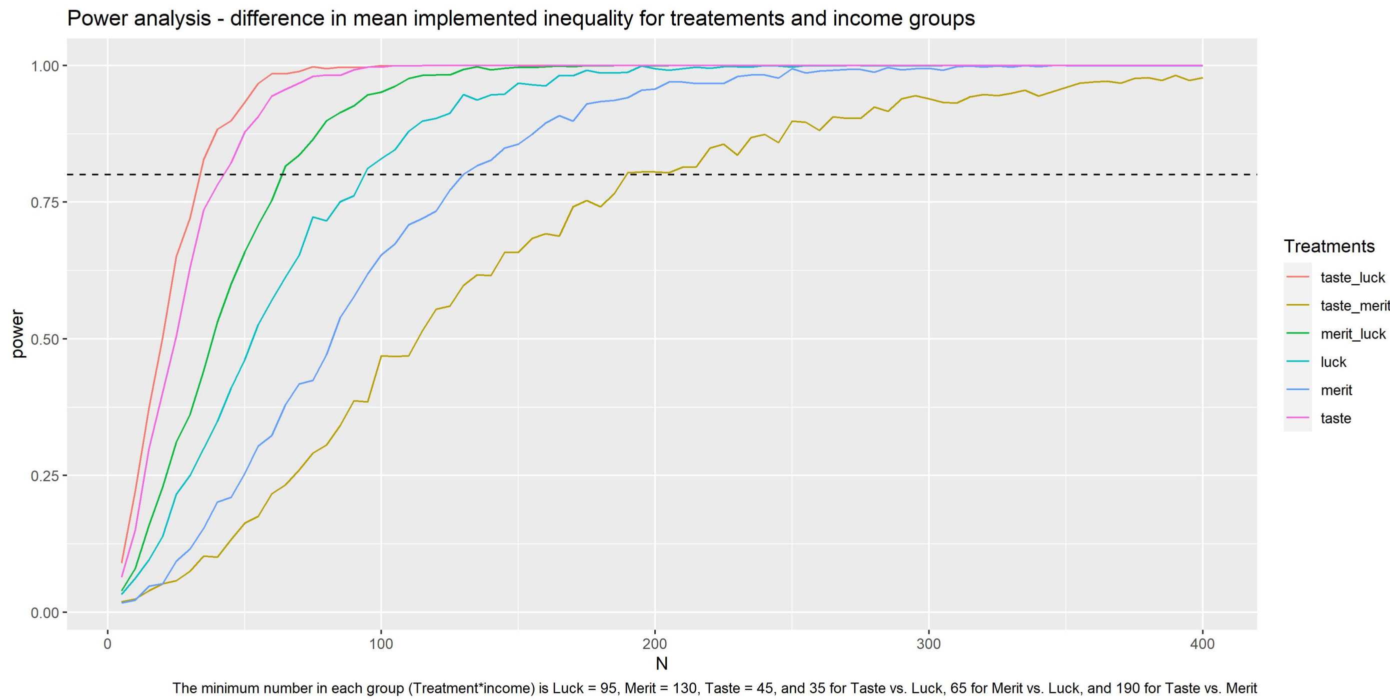

Figure 1: Power Analysis

Notes: The values on the vertical axis represent the share of regressions with a statistically significant result. The red line, for example, depicts, for each N, the share of regressions for which a significant difference arises in how the two income groups differ between the taste and the luck treatments.

Notes:

1. Panel members who participated in the previous study will not participate in this study as spectators (they may participate in other roles).

2. When we reach the last stage of the experiment in which spectators are recruited, we will ask the panel to reach out only to panelists with above-average or below-average income. In other words, unlike the first study, we should not have income information missing for any of our participants.

References

Arnold, Benjamin F., Daniel R. Hogan, John M. Colford, and Alan E. Hubbard. "Simulation methods to estimate design power: an overview for applied research." *BMC medical research methodology* 11, no. 1 (2011): 1-10.