

Partial Identification

Head Start is a major federally funded preschool program in the U.S. It is targeted at children from low-income parents and provides these children with schooling, health, nutrition, and social welfare services. In this assignment we will use the National Longitudinal Study of Youth 1979 (NLSY79) to investigate the average causal effect of Head Start participation on the probability of completing high school. The NLSY79 is a sample of 14 to 22-year-olds living in the U.S. in 1979 who were interviewed annually up to 1994 and every other year after. The data set `headstart.dta` (<https://www.dropbox.com/s/60aqtqt6i4pctst/headstart.dta?dl=0>) is a subset of the NLSY79 data set. It includes information on Head Start participation, high school completion (y) and parent's level of education for 4876 individuals who were born between 1960 and 1964.

1. Estimate the effect of participating in Head Start on the probability of obtaining a high school degree by OLS. Explain whether we can interpret this OLS estimate as an estimate of the average causal effect of Head Start on the probability of obtaining a high school degree.
2. Compute the share of children that obtained a high school degree for:
 - (a) the Head Start participants
 - (b) the nonparticipants in addition compute the share of Head Start participants and the share of nonparticipants in the data set.
3. Compute the worst case bounds around the following two mean potential outcomes:
 - (a) the mean potential probability of obtaining a high school degree with no Head Start as potential treatment $E[y(0)]$,
 - (b) the mean potential probability of obtaining a high school degree with Head Start as potential treatment $E[y(1)]$.
4. Compute the Worst case lower and upper bounds on the average causal effect of Head Start on the probability of obtaining a high school degree. Interpret the estimated lower and upper bounds.
5. Head Start was targeted at children from disadvantaged backgrounds and children from families with an income below the poverty line were eligible to participate in the program. Suppose we would impose a monotone treatment selection assumption. Write down the MTS assumption and explain the MTS assumption in this context, what would we assume?

6. Compute MTS lower and upper bounds on
 - (a) the mean potential probability of completing high school with no Head Start as potential treatment $E[y(0)]$.
 - (b) The mean potential probability of completing high school with Head Start as potential treatment $E[y(1)]$
 - (c) the average causal effect of Head Start on the probability of completing high school.
7. The variable `edu_parent` contains the maximum level of education obtained by the parents. Compute MTS lower and upper bounds on $E[y(0)]$ and $E[y(1)]$ separately for each value of parents level of education
8. Write down the MIV assumption that we are imposing when we would use parent's level of education as a (positive) monotone instrument and explain this MIV-assumption in words.
9. Use the results obtained in question 7) to compute MTS-MIV upper and lower bounds on
 - (a) the mean potential probability of completing high school with no Head Start as potential treatment $E[y(0)]$.
 - (b) the mean potential probability of completing high school with Head Start as potential treatment $E[y(1)]$.
10. Compute MTS-MIV lower and upper bounds on the average causal effect of Head Start on the probability of obtaining a high school degree. Interpret the estimated lower and upper bounds.