Truett Bloxsom (tsb962)

ECO 348K, Problem Set 4

Professor Vogle

TA: Bokyung Kim

1. Table 1 shows summary statistics



21% of the sample participated in Head Start. There is a large portion of blacks, 32%, in the sample compared to the 12.7% estimated population in the US. Comparing the racial make-up in Head Start, 51.8% where black, 29.4% where white, and 18.8% where Hispanic. Whites make up 48% of the sample, so there is a large portion of blacks in Head Start. Males consist of 51% of the sample. Tabulating father being present in the child’s life from ages 0 to 3, shows that as the father lived less time with the child, the probability of the child being in head start increased. Children who were the first born of a couple are less likely to be in Head Start.

2. Table 2 shows regression of child test scores (ages 5 to 6) on Head Start participation clustering by mother



Table 3 shows regression of child test scores (ages 5 to 6) on Head Start participation and family income clustering by mother



If we assume that Head Start is exogenous, which means that it is not affected by any other variables in the model, then the Head Start coefficient shows the casual effects on test scores. A child participating in Head Start is predicted to have a 5.53 point lower test score compared to a child not in Head Start. This coefficient is highly significant with a t-value of -4.59. But it is not reasonable to assume Head Start participation is exogenous since both test score and Head Start participation are correlated with income which is shown in Table 3. Since including family income in the regression reduces the significance of Head Start participation, it can be shown that Head Start is not exogenous.

3. Table 4 shows random effects regression of child test scores (ages 5 to 6) on Head Start participation clustering by mother



The coefficient using the random effects model is smaller and less significant that the OLS regression. The coefficient can be interpreted as a child who participated in Head Start is predicted to have a lower test score by 2.50 points compared to a child who did not participate. The comparison makes me less confident that the OLS estimate is efficient since the coefficients are different between the OLS and random effects. Since the coefficients are so different, there may be omitted variable bias at the family level since clustering OLS and random effect model accounts for within family variation. If there is omitted variable biased, our coefficient estimates will be inaccurate. This can be solved by using the fixed effects model! Yay!

4. Table 5 shows fixed effects regression of child test scores (ages 5 to 6) on Head Start participation clustering by mother



Table 6 shows fixed effects regression of child test scores (ages 5 to 6) on Head Start participation and other family variables grouped by mother



We can include any variables that vary within family groups like if the child was first born or income of the family when the child was young. Since we created and controlled for the complete set of family dummies, we cannot estimate coefficients that vary across families since it is collinear with the full vector of family level dummies. The results show that after controlling for within family control variables, that the effects of Head Start on test scores are highly significant. The fixed effect results from Table 5 show that a child participating in Head Start is predicted to have an increased test score of 7.4 compared to a child who did not participate. The fixed effect results are almost the opposite of the OLS results. The results are so different because of omitted variable biased or family level unobservable that are correlated with other independent variables in the OLS regression which the fixed effects model takes into account.

5. When we ran the fixed effects model, we assumed that there was no difference within family characteristics. So, we are assuming the covariance between Head Start and within family variation is equal to 0 and within family variation and group wide characteristics is equal to 0. We can test if there is correlation of Head Start and within family variation by running regular clustered regression of family level variables on Head Start and doing the same with fixed effects. If the regression coefficients are different between the OLS and fixed effects (shown in tables 7 and 8) then we can say that Head Start is correlated with the residual and OLS will be inefficient. It also means the Fixed effects model is accurate since it controls for within family characteristics and family unobservable.

Table 7 shows OLS regression of first born on Head Start clustering by mother



Table 8 shows fixed effects regression of first born on Head Start controlling by mother



Table 7 and 8 show that there is a difference between OLS and fixed effects model for first born child variable. This can also be done with other within family variation variables like sex, income, and birth weight which has similar differences in OLS and fixed effect model. This proves that we should be using the fixed effect model since Head Start must be correlated with within family variables.

6. Table 9 shows fixed effects regression of standardized child test scores (ages 5 to 6) on Head Start participation and other family variables clustering by mother



Table 10 shows fixed effects regression of standardized child test scores (ages 7 to 10) on Head Start participation and other family variables clustering by mother



Table 11 shows fixed effects regression of standardized child test scores (ages 11 to 14) on Head Start participation and other family variables clustering by mother



As shown from Tables 7-9, the t-stat and coefficient for Head Start decrease as time progresses. But Head Start participation is still significant at the 1% level at ages 11-14 test scores. These regressions show proof that the effects of Head Start fade out with age.

7. After running simple regression of Head Start on long-term outcomes, these are the results. Head Start had a significant negative effect on repeating a grade at the 10% level. Head Start had a significant negative effect on having a learning disability at the 5% level. Head Start had a significant positive effect of completing high school at the 1% level. Head Start had a significant positive effect of attending some college at the 1% level. Head Start had a significant negative association on having poor self-reported health at the 1% level.

8. After running regression of Head Start, dummy for sex, and interaction variable between Head Start and sex on long-term outcomes, these are the results. In all of the regression, sex was a highly significant variables with males having a higher probability of repeating a grade and having a learning disability. And males had a lower probability of completing high school, attending some college, and self-reported poor health.

After running regression of Head Start, dummy for sex, and dummy for race on long-term outcomes, these are the results. For both blacks and Hispanics, Head Start had an insignificant effect on repeating a grade. For blacks, Head Start had a significant negative effect on learning disability at the 5% level but was insignificant for Hispanics. For both blacks and Hispanics, Head Start had a significant positive effect on completing high school. For blacks, Head Start had a highly significant positive effect on attending some college but was insignificant for Hispanics. For both blacks and Hispanics, Head Start had an insignificant effect on self-reported poor health.

9. Table 12 shows fixed effect regression of child test scores on Head Start and other family variables that include dad being home and PPVT score.



We decided not to include dad being home or PPVT score in our earlier regressions since it reduced the sample size to below 600. But including these two variables makes the Head Start variable insignificant on test scores. Since we found that the fixed effect model was the most unbiased and efficient model for our data, this last regression reduces my confidence that Head Start has a significant effect on test scores. Running the same regression but replacing test score with long-term outcome variables like repeating a grade or completing high school has the same results that Head Start has an insignificant effect on them. We must highlight that there is association between missing data in dad being home and being in Head Start. These mixed results lead me to believe that I am not confident in advocating for either a cut or increasing the budget for the Head Start program.