

1. In this assignment you are going to estimate the effect of fertility on mother's labour supply. More specifically, you are going to estimate the effect of having more than two children on labour supply of mothers who have at least two children. The data set `laboursupply_long.dta` contains information about the first and second births of mothers, and their labour supply.

Use heteroskedasticity robust standard errors in this assignment.

The following Stata commands might be useful when making this assignment:

**ivregress 2sls** instrumental variable regression

**help ivregress postestimation** for information about tests after instrumental variable regression

- (a) Prepare your dataset for analysis where you have 1 observation per mother (use `-reshape-`). You will need to generate three dummy variables:  $more2kids_i$  which equals one if the mother has more than 2 children.  $samesex_i$  which equals one if the mother's first two children had the same gender, and  $twins2_i$  which equals one if the mother had twins on the second birth.
- (b) Estimate the effect of having more than two children on the average number of hours that a mother works per week, by estimating the following equation by OLS:

$$hoursm_i = \alpha + \beta \cdot more2kids_i + u_i$$

Interpret the sign, magnitude and statistical significance of the estimated coefficient on  $more2kids_i$ .

- (c) Explain why OLS might not give an unbiased and consistent estimate of  $\beta$ .
- (d) A researcher claims that you can obtain a consistent estimate of the effect of having more than two children on mother's labour supply by performing 2SLS. He suggests using the sex mix of the first two children ( $samesex$ ) as instrument for having more than two children. He claims that if the first two children are of the same sex, parents are significantly more likely to have another child, because of the widely observed preference of parents for a mixed sibling sex composition. Estimate the effect of the first two children being of the same sex on the probability that a mother has more than two children. Discuss whether the instrument relevance and instrument exogeneity assumptions hold.
- (e) Estimate the effect of having more than two children on the average number of hours that a mother works per week, using the variable  $samesex$  as instrument, test whether  $more2kids_i$  is an exogenous variable and interpret your findings.

- (f) A second researcher suggests using a different instrument. She suggests using twins at second birth as instrument for having more than two children. Estimate the effect of a twin at second birth ( $twins2$ ) on the probability that a mother has more than two children. Discuss whether the instrument relevance and instrument exogeneity assumptions hold.
- (g) Estimate the effect of having more than two children on the average number of hours that a mother works per week, using twins at second birth as instrument, test whether  $more2kids_i$  is an exogenous variable and interpret your findings.
- (h) Estimate the effect of having more than two children on the average number of hours that a mother works per week, using both twins at second birth and the sex composition of the first two children as instruments. Perform an over-identifying restrictions test and interpret your findings.