

Problem Set I
Labor Economics

You may work in groups, but you must do the coding and write-up on your own. Include your code with your writeup. Do not submit STATA output in raw form.

1. This problem set examines the labor supply effects among single mothers from having their youngest child in junior kindergarten. Most of the problems involve replicating results in:

Gelbach, J., "Public Schooling for Young Children and Maternal Labor Supply," American Economic Review, 92, March 2002, 307-322.

The dataset is all single mothers from the 1980 U.S. Census with at least one child aged 5 on April 1st, 1980. The following variables are included:

Variable definitions

youngest: age of youngest child

hours: hours worked last week for mother (in 1980)

public: child aged 5 enrolled in public school

quarter: 0 if born 75, 1st quarter, 1 if born 74, 2nd quarter, 2 if born 74, 3rd quarter, 3 if born 74, 4th quarter

num05: number of children = 0 – 5

num612: number of children 6 – 12

num1317: number of children 13 – 17

othrlt18: number of household members <18

othrge18: number of household members >=18

state: state of residence (don't worry about which state each one stands for)

age: age of mother

grade: highest grade attained for mother

hours79: usual hours worked per week in 1979

salary: wage and salary income in 1979

work79: employment in 1979

lfp: employed last week

white: white

centcity: living in central city

Part I.

For Part I consider the sample of only single mothers with the youngest child either born in the second quarter of 74, or the first quarter of 1975. Most states allow children to begin junior kindergarten in the Fall (September) if they attain the age of 5 by December 31st that year. Thus, those children born in the second quarter of 74 were eligible to enter school in the Fall of 79 and be in school on April 1, 1980, the reference date for the 1980 census, whereas those born in the first quarter of 75 were generally not eligible to enter school until the Fall of 80, and so would likely not be in school on April 1, 1980.

Consider the model:

$$(1) \quad H_{it} = \beta_0 + \beta_1 \text{daycare}_{it} + e_{it}$$

where H_{it} is hours worked for individual i in period t , and daycare_{it} is an indicator variable for whether the youngest child went to daycare (say, for example, at age 4) or not.

Suppose we had data for daycare_{it} and hours worked when the child was 4 and when they were 10. Discuss the expected signs of β_1 , and whether the estimated responses are transitory, permanent, or anticipated.

Discuss the potential for omitted variable bias.

Now consider the model:

$$(2) \quad H_{it} = \delta_0 + \delta_1 \text{inschool}_{it} + e_{it}$$

where inschool_{it} is an indicator variable for whether the youngest child went to school at age 5.

What is the interpretation of δ_1 and how similar or different is it to β_1 ? What might be the omitted variables bias from estimating (2) using OLS?

Is quarter of birth for 5 year olds in April 80 a plausible instrument?

Let $\text{bornearly}_{it} = 1$ if the 5 year old in April 1, 1980 was born in the second quarter of 74, and 0 if born in the first quarter of 75 (you may have to generate this variable).

Compute the Wald Estimator for the effect of inschool_{it} and show that it is the same estimate for δ_1 using $H_{it} = \delta_0 + \delta_1 \text{inschool}_{it} + e_{it}$ and bornearly_{it} as an instrument.

Who are those that might not respond to the instrument? What is LATE here?

Discuss why estimates instead with the sample of single mothers with a child aged 5, and at least one other child younger than 5 may differ?

Part II.

Use the full dataset of single mothers with the youngest child age 5, and use all quarters of birth as separate instruments.

Produce a table showing the means, OLS and IV estimates for the dependent variables hours worked in 1980 (like before), current employment, wage and salary income in 1979, receipt of public assistance in 1979. Why would using outcomes from 1979 underestimate the estimated effects?

Add controls to your estimates and replicate Columns 2 and 3 of Gelbach's table 7. Does adding the controls make much difference? In one paragraph, summarize the results and draw conclusions.