Local Average Treatment Effect

The problem is inspired from Angrist and Evans, 1998, "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size", *American Economic Review,* 1998, 88(3):450-477.

The dataset consists of variables for 927,267 households in the US in the 1980 with at least two children. We restrict the sample to married mothers of 21-35 years old because their oldest child is likely to be less than 18 and still at home. We restrict also somewhat the range of ages for the father and end up with 557,482 households.
See in appendix, for a description of variables

1) What are the economic channels through which fertility and household labor supply are simultaneously determined and probably depend on the same unobservable?

An increase in fertility causes an increase in the amount of time devoted to the care of children for both women and men, but probably more for women (at least in Western societies). This increases the opportunity costs for time and thus decreases labor market participation, hours of work and earnings.
Conversely, if labor market involvement is important, then the number of desired children might decrease because its cost is increasing. Observables affecting behavior are education, family background and the like. However, tastes for work and tastes for large families which are unobserved are likely to be the main drivers of labor supply and fertility behavior.

2) In a standard OLS setting, regress labor supply variables (participation, earnings, hours of work) of married women and married men on fertility and other covariates like the ages of the mother and father, their education and the father's income. What is the treatment? Explain why these estimates are biased and try to predict the direction of the bias.

One example for treatment is the number of children. If there is a negative association between tastes for work and for children then the estimates are biased downwards, i.e., the effect is biased in the direction of a larger magnitude. If there is a positive association this is the reverse. In the remainder, we will use the father's income as an exogenous regressor. In the case of the United States in the 80s, this might be the case. We can argue that only the mother's supply of work is marginal to total household labor decision

3) The treatment we consider is the birth of a third child and its effect on labor market variables for both wife and husband. Explain carefully what the treatment effect is. The first instrument we are going to use is the sex composition of the first two children.
State cases in which, conditional on other covariates, the sex composition of the first two children might affect labor supply variables. Are these stories credible or second order?

Having the first two children with the same sex increases the probability of having a third one because there might be a "cultural" preference for diversity, parents might chase a boy after two girls, or a girl after two boys.

What we will argue that the sex composition of the first two children affects labor supply variables only by increasing the probability of having a third child (after controlling for co-variates). This is the condition for it to be a valid instrument. However, this might not be the case.
If having the same gender negatively impacts the number of rooms (boys and girls are separated), it might increase the incentive to supply more labor to be able to afford a bigger house. In this case $Y(0)$ is negatively affected and the standard orthogonality assumption need not to be verified.

4) Explain why the instrument might affect the treatment probability and what estimator we can estimate.

The treatment parameter is the effect of having a third child in the population of households with more than 2 children and who had a third child because the first two had the same sex and who would not have had their third child if the sex composition was varied. This is a LATE parameter and the interpretation above assumes the absence of defiers, i.e., those who would have decided to stop having their third child because of a homogenous sex composition. Furthermore, it is a weighted effect between households with 3, 4 or more children.

5) Explain why we restrict the sample to families of two children or more. How does this sample selection affect the parameters to estimate? Comment.

If we do not restrict the sample to two children or more, the households with less than 2 children would necessarily be in the control group and this would generate a sample selection bias. First, it would make the parameter of interest different from the one we are interested in (and its interpretation would be quite difficult to understand).Second, the differences in labor outcome variables are very different between households with different number of children.
By restricting the sample, we select a specific population which treatment effect is clearer to understand. Also, the sample presents similar characteristics, which gives the idea that the treatment effect would be similar between them. In this sense, the treatment parameter is restrictive and does not refer to the whole population.

6) We start by assessing the instrument's strength by estimating the propensity score as a function of age of mother and father, their education and father's income. What is the treatment variable? Use linear models, Probit and Logit for the propensity score using the sex composition of the first two children.

The treatment variable is to have three or more kids. The treatment parameter mixes effects on labor supply for the third, fourth, etc. kid. All models give that sex composition is a strong predictor of the birth of a third child (the p-values of *instrument* are $<$ 0,001). The t-statistic of *instrument* is high in all these cases, so we conclude that the instrument is not weak.

7) Predict the estimated propensity score and compare its distribution function in the treatment and control groups [*Hint*: use the command *kdensity* to graphically plot the density of the propensity scores in the two cases]. Comment. Is the common support assumption reasonably established?

The propensity score is slightly higher in the treatment than in the control group. The propensity score's range is 0.06 to 0.94 and the common support assumption is satisfied.

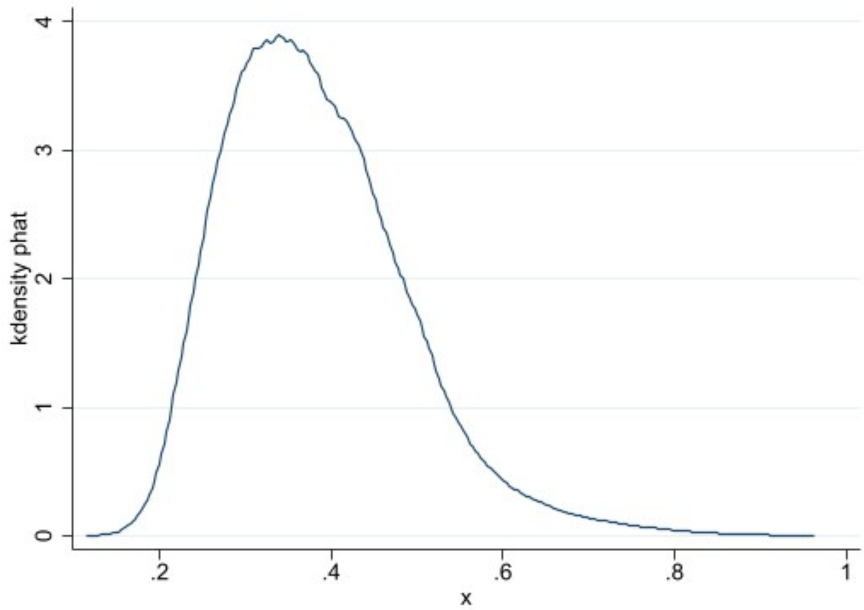

Figure 1: Propensity score

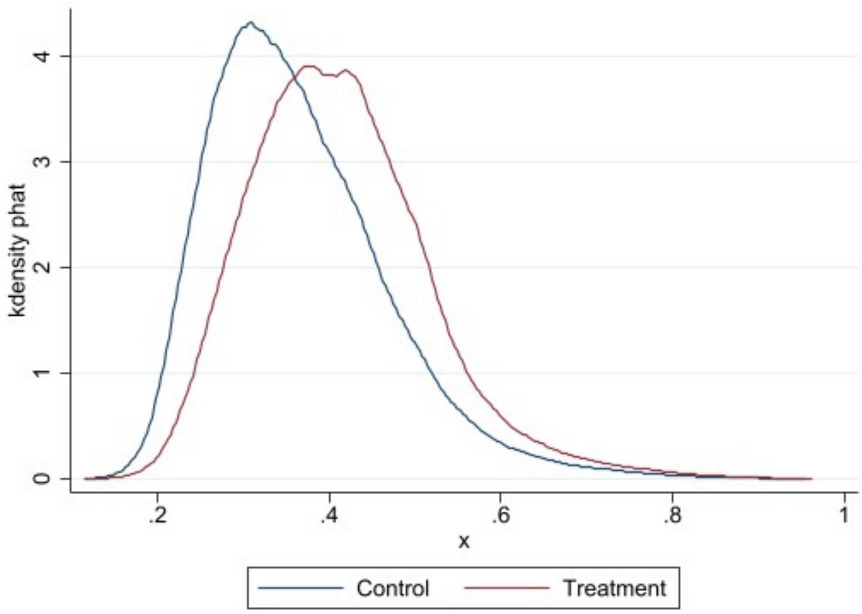

Figure 2: Propensity score by treatment

8) Estimate by 2SLS the treatment effect without covariates and compare to the OLS estimates. Compare to your priors of Question 2. What is the identified treatment effect?

For *incomem* (mother's total income) the 2SLS is -766$; OLS with no covariates is -990$. For *hourym* (number of hours worked in a year): -187 hours the 2SLS and -149 hours the OLS. For *weeksm* (number of weeks worked): -4.9 weeks for the 2SLS and -4.1 weeks the OLS.

Table 1: 2SLS vs OLS

	Income		Yearly hours worked		Weeks worked	
	2SLS	OLS	2SLS	OLS	2SLS	OLS
kidcount >= 3	-766.9**	-990.5***	-187.9***	-149.5***	-4.976***	-4.088***
	(-3.18)	(-73.09)	(-4.80)	(-67.83)	(-4.95)	(-72.26)
Constant	3823.2***	3908.9***	807.7***	793.0***	23.25***	22.91***
	(41.31)	(465.82)	(53.67)	(581.14)	(60.19)	(654.04)
N	655169	655169	655169	655169	655169	655169
R^2	0.008	0.008	0.007	0.007	0.008	0.008

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The estimated effect is the LATE.

9) One possible estimate of the treatment effect is the ratio of the coefficient on the instrument in the reduced form model to the coefficient on the instrument in the first stage regression. Relate it to the Wald/IV estimator, calculate and report that ratio.

The IV estimator is defined as:.

$$\beta_{IV} = \frac{\text{Cov}(Y, Z)}{\text{Cov}(D, Z)}$$

In the case of no covariates, the numerator (same argument for the denominator):

$$\begin{aligned}
\text{Cov}(Y, Z) &= \mathbb{E}[ZY] - \mathbb{E}[Y]\mathbb{E}[Z] = \mathbb{E}[ZY] - p\mathbb{E}[Y] \\
&= \mathbb{E}[ZY] - p(p\mathbb{E}[Y|Z = 1] + (1-p)\mathbb{E}[Y|Z = 0]) \\
&= p\mathbb{E}[1 \cdot Y|Z = 1] + \cancel{(1-p)\mathbb{E}[0 \cdot Y|Z = 0]} - p(p\mathbb{E}[Y|Z = 1] + (1-p)\mathbb{E}[Y|Z = 0]) \\
&= p(1-p)E(Y \mid Z = 1) - E(Y \mid Z = 0)
\end{aligned} \tag{1}$$

Therefore:

$$\begin{aligned}
\beta_{IV} = \frac{\text{Cov}(Y,Z)}{\text{Cov}(D,Z)} &= \frac{\cancel{p(1-p)}E(Y|Z=1) - E(Y|Z=0)}{\cancel{p(1-p)}E(D|Z=1) - E(D|Z=0)} \\
&= \frac{E(Y|Z=1) - E(Y|Z=0)}{E(D|Z=1) - E(D|Z=0)} = \beta_{\text{Wald}}
\end{aligned} \tag{2}$$

Then β_{IV} is the ratio of the coefficient on the instrument in the reduced form model to the coefficient on the instrument in the first stage regression. We apply equation 2 to obtain that the Wald estimator is -766.881.

10) Estimate by 2SLS models with different covariates and different instruments. For instance, use the sex of first kid interacted with sex composition.

Table 2: 2SLS vs OLS

	Income		Yearly hours worked		Weeks worked	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
kidcount>=3	-938.1***	-706.4**	-158.4***	-193.2***	-4.321***	-4.759***
	(-66.04)	(-3.05)	(-67.70)	(-5.06)	(-71.01)	(-4.79)
gradem	387.3***	392.6***	34.45***	33.65***	1.034***	1.024***
	(111.83)	(61.86)	(60.39)	(32.19)	(69.68)	(37.66)
agem	118.0***	114.8***	20.16***	20.64***	0.614***	0.620***
	(56.75)	(30.09)	(58.87)	(32.85)	(68.91)	(37.93)
aged	6.496***	6.064***	0.325	0.390	-0.0166*	-0.0158*
	(3.80)	(3.44)	(1.16)	(1.34)	(-2.27)	(-2.09)
graded	-116.5***	-115.0***	-21.45***	-21.66***	-0.428***	-0.431***
	(-41.73)	(-36.75)	(-46.66)	(-42.02)	(-35.81)	(-32.14)
income1d	-0.0296***	-0.0295***	-0.00957***	-0.00958***	-0.000219***	-0.000219***
	(-50.29)	(-49.83)	(-98.75)	(-98.19)	(-86.74)	(-86.21)
= sex 2 first children	13.47		-2.025		-0.0255	
	(1.00)		(-0.91)		(-0.44)	
_cons	-3757.2***	-3820.9***	80.29***	89.87***	-1.891***	-1.770***
	(-69.78)	(-43.16)	(9.05)	(6.16)	(-8.20)	(-4.67)
N	561459	561459	561459	561459	561459	561459
R^2	0.052	0.052	0.043	0.042	0.046	0.045

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We also add an additional instrument: the first two children are both female (so we differentiate between the two cases M+M+3rd child and F+F+3rd child). Regressing the treatment on the instruments and exogenous regressors, we find that both instruments are strong predictors but are not highly correlated with the treatment.

Table 3: 2SLS vs OLS

	Income		Yearly hours worked		Weeks worked	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
kidscount>=3	-938.7***	-560.3*	-158.5***	-161.3***	-4.324***	-3.910***
	(-66.08)	(-2.47)	(-67.74)	(-4.32)	(-71.06)	(-4.03)
gradem	387.3***	396.0***	34.45***	34.38***	1.034***	1.043***
	(111.83)	(63.37)	(60.39)	(33.43)	(69.68)	(39.00)
agem	118.0***	112.8***	20.17***	20.20***	0.614***	0.608***
	(56.76)	(30.02)	(58.88)	(32.68)	(68.92)	(37.81)
aged	6.496***	5.793**	0.325	0.331	-0.0166*	-0.0174*
	(3.80)	(3.29)	(1.16)	(1.14)	(-2.27)	(-2.30)
graded	-116.5***	-114.1***	-21.45***	-21.47***	-0.428***	-0.426***
	(-41.73)	(-36.62)	(-46.66)	(-41.84)	(-35.82)	(-31.90)
income1d	-0.0296***	-0.0295***	-0.00957***	-0.00957***	-0.000219***	-0.000219***
	(-50.29)	(-49.76)	(-98.75)	(-98.14)	(-86.74)	(-86.14)
= sex 2 first children	-15.85		-7.714**		-0.181**	
	(-0.98)		(-2.89)		(-2.60)	
two girls	61.99**		12.03***		0.329***	
	(3.26)		(3.84)		(4.04)	
_cons	-3757.8***	-3865.5***	80.19***	80.14***	-1.894***	-2.029***
	(-69.79)	(-44.28)	(9.04)	(5.58)	(-8.21)	(-5.43)
N	561459	561459	561459	561459	561459	561459
R^2	0.052	0.051	0.043	0.043	0.046	0.045

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

11) When there is more than one instrument, test the overidentifying restrictions. What does this tell you about the parameters that you estimated.

To test for overidentifying restrictions we use the Sargan test (a specific case of the Hansen test), a valid test under conditional homoskedasticity.

The null hypothesis H_0 is the joint hypothesis of correct model specification and the orthogonality conditions, i.e., all instruments are uncorrelated with the error term. A rejection might either mean that the specification is wrong or that one or both instruments are invalid. The intuition is that, if both instruments are valid, using both or just one of them should only differ by sampling variation.

The test rejects the overidentifying restriction, telling us that the LATE parameters are not the same across instruments.

Furthermore, by splitting the sample with respect to the sex of the first child, we indeed find different effects. For boys, the effect is much larger. A possible story would be that having two boys mean more housework in the future than with two girls because of sex specialization into household chores, so the opportunity costs of time are different. However, the two LATE effects refer to two populations so it could also be composition effects.

Table 4: 2SLS vs OLS

	Income		Yearly hours worked		Weeks worked	
	Girl	Boy	Girl	Boy	Girl	Boy
kidcount>=3	-447.6	-995.9**	-107.4*	-296.8***	-2.596*	-7.350***
	(-1.56)	(-2.63)	(-2.28)	(-4.72)	(-2.12)	(-4.51)
gradem	398.2***	386.4***	35.46***	31.44***	1.068***	0.970***
	(48.00)	(39.08)	(26.08)	(19.18)	(30.14)	(22.80)
agem	112.8***	117.3***	19.67***	21.90***	0.596***	0.650***
	(22.83)	(19.55)	(24.28)	(21.99)	(28.25)	(25.17)
aged	6.016*	6.169*	0.450	0.362	-0.0121	-0.0186
	(2.39)	(2.49)	(1.09)	(0.88)	(-1.12)	(-1.75)
graded	-117.1***	-113.3***	-21.47***	-21.98***	-0.431***	-0.434***
	(-26.68)	(-25.11)	(-29.83)	(-29.37)	(-22.97)	(-22.36)
income1d	-0.0277***	-0.0312***	-0.00938***	-0.00978***	-0.000213***	-0.000224***
	(-32.71)	(-37.70)	(-67.45)	(-71.14)	(-58.77)	(-62.94)
_cons	-3914.2***	-3723.8***	56.95**	126.5***	-2.614***	-0.840
	(-33.23)	(-27.55)	(2.95)	(5.64)	(-5.20)	(-1.44)
N	273129	288330	273129	288330	273129	288330
R^2	0.050	0.052	0.042	0.037	0.044	0.041

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix

Characteristics of household

- kidcount: counts of kids in the household
- faminc: family income
- twin1st: first two born are twins
- triplet: first three born are triplet
- poverty: family in poverty status
- twins_2: second and third born are twins

Characteristics of mother

- agem: age in years of mother
- gradem: mother's highest grade
- fingradm: did mother finish the highest grade of instruction
- weeksm: mother's number of weeks worked per year
- hoursm: mother's number of hours worked per week
- yobm: mother's year of birth
- incomem: mother's full income

Characteristics of father

- aged: age in years of father
- graded: father's highest grade
- fingradd: did father finish the highest grade of instruction
- weeksd: father's number of week worked per year
- hoursd: father's number of hours worked per week
- incomed: father's full income

Characteristics of children

First born

- sexk: gender, 0 = male, 1 = female
- grade: highest grade
- agek: age in years
- yobk: year of birth
- ageqk: age in quarters

Second born

- sex2nd: gender, 0 = male, 1 = female
- grade2nd: highest grade
- ageq2nd: age in quarters

Third, fourth, fifth born

- sex3rd: third kid gender, 0 = male, 1 = female
- sex4th: fourth kid gender, 0 = male, 1 = female
- sex5th: fifth kid gender, 0 = male, 1 = female
- ageq3rd: third kid age in quarters
- ageq4th: fourth kid age in quarters
- ageq5th: fifth kid age in quarters