Replication of "Children and Their Parent's Labor Supply: Evidence from Exogenous Variation in Family Size"

Matthew Borelli

INTRODUCTION

Angrist and Evan's project, "Children and Their Parent's Labor Supply: Evidence from Exogenous Variation in Family Size", is a study that estimates the causal effect of family size on several measures of labor supply using the combinations of the first two children's sexes as an instrumental variable. The labor supply measures are the labor force participation decision, number of weeks worked in a year, hours worked per week, and labor income. They note that many previous studies have reported findings that indicate a negative correlation between family size and female labor supply, but the interpretations of those correlations are unclear due to concerns of endogeneity in the fertility/labor supply decisions. Some former studies have claimed that there are "strong theoretical reasons" that point to family size and labor supply being jointly determined. There is a lack of clarity in the answer, but the general intuition is that family size and labor supply likely have a negative relationship.

For this project, Angrist and Evans utilized the 1970, 1980, and 1990 Census Public Micro Samples (PUMS). The 1970 sample is from the 1/100 state file, while the 1980 and 1990 samples are from the 5-percent samples. PUMS for these years include all relevant labor supply variables that the authors wanted to analyze. However, the samples only included the number of children as a measure of family size, so Angrist and Evans matched mothers to their children using subfamily identifiers in order to generate indicators for the variables of children. This is important as it ties closely to their choice of instrumental variable: sibling sex composition. Due to the aforementioned issues with endogeneity in previous studies on family size and labor supply, Angrist and Evans wanted to find an instrumental variable to examine the relationship between family size and labor supply due to exogenously having two children of the same sex. Their argument for instrument relevance uses the quantity / quality model of fertility, in which parents derive utility both from the number of children and qualities about those children. In those models, parents could gain utility from having children with different sexes, which means that women with two kids of the same sex would be more likely to have another kid. This is the exogenous change in family size that Angrist and Evans chose to exploit in order to find the true causal effect of fertility on labor supply.

The findings in this study show that family size and labor supply are negatively correlated. While this is not a different result than hundreds of other projects that came before on this topic, it is unique because of the plausibly exogenous variation in family size resulting from the instrumental variable of sibling sex composition. 2SLS and IV regression estimates confirm the OLS findings of negative relationships between family size and labor supply, but show that OLS tends to overestimate the causal effect of children. They did find some heterogeneity in these results: the causal effect is smaller or insignificant for college-educated women and women whose husbands have a high wage. This contradicts previous OLS results of the consequences of family size on labor supply being greater for more educated women.

SUMMARY STATISTICS

The first step in this replication is to analyze summary statistics for some of our variables. To note, the data given for this replication is a subsample of the data used in Angrist and Evan's paper. Our sample is women aged 21 - 35 with two or more children in the 1980 PUMS. Table 1 displays sample mean for number of children, the percent of women with 2 or more children, and the percent of women who worked in the previous year

Table 1. Fertility and Labor Supply Measures

Sample	1980 PUMS
Women aged 21-35 with 2 or more children	
Mean children ever born	2.508
Percent with more than 2 children	38.056
Percent worked last year	52.822
Observations	254,654

We see that women in this sample average approximately 2.5 children, with 38% of women having more than 2 children. 52% of the sample, approximately 134,513 women, worked in the previous year. Next, Table 2 displays more descriptive statistics for the sample.

Table 2. Descriptive Statistics, Women Aged 21-35 with 2 or More Children

Variable	1980 PUMS
Children ever born	2.51
	(0.77)
More than 2 children (=1 if mother had more than 2 children, =0 otherwise)	0.381
	(0.486)
Boy 1st(s_1)(=1 if first child was a boy)	0.514
	(0.500)
Boy 2nd(s_2)(=1 if second child was a boy)	0.513
	(0.500)
Two boys(=1 if first two children were boys)	0.267
	(0.442)
Two girls(=1 if first two children were girls)	0.239
	(0.427)
Same sex(=1 if first two children were the same sex)	0.506
	(0.500)
Twins-2 (=1 if second birth was a twin)	0.0083
	(0.0908)
Age	30.4
	(3.4)
Age at first birth(parent's age in years when first child was born)	20.8
	(2.9)
Worked for pay(=1 if worked for pay in year prior to census	0.528
	(0.499)
Weeks worked (weeks worked in year prior to census)	19.0
	(21.9)
Hours/week (average hours worked per week	16.7
	(018.3)
Labor income (labor earnings in year prior to census, in 1979 dollars)	6250
	(10210)

There are a few important points to note from Table 2. The probability of having two children of the same sex is 0.506, which means the probability of having two children with different sexes is 0.494. Having two boys is slightly more likely than having two girls, although it's not a statistically significant difference. For labor supply, we see that women work an average of 19 weeks a year for an average of 16.7 hours a week. The average labor income in 1979 dollars is $6250. It is interesting to note that if this study were replicated using more recent PUMS data, we would likely see increases in labor supply for women around the board, even though the potential relationship between fertility and family size might be the same.

ESTIMATING CAUSAL EFFECTS OF HAVING MORE KIDS ON LABOR SUPPLY

The next step is to estimate the causal effect for the four outcome variables: probability of working, weeks worked, hours worked per week, and labor income per week. For each outcome variable, we run four regressions using different methods: Ordinary least squares (OLS) regression, covariate adjusted instrumental variable regression, manual two-stage least squares regression, and automated two-state least squares regression using the *IVregress* command in STATA.

Tables 3 and 4 list the estimated coefficients for the indicator variable set to 1 if the woman has more kids after the first two and 0 if not. These tables let us compare the estimated causal effects across each of the regression methods for each of the labor supply outcome variables. Each of these regressions includes controls for race, current age, and age when the subject had their first child.

Table 3. OLS, CIV, Manual 2SLS, and 2SLS regressions of More Kids on Working and Weeks Worked

Outcome Variable		**More Kids**		
	OLS	Covariate-Adjusted IV	Manual 2SLS	IVregress 2SLS
Probability of Working	-0.166***	-0.122***	-0.124***	-0.122***
	(0.002)		(0.028)	(0.028
Weeks Worked	-8.042956***	-5.472***	-5.579***	-5.472***
	(0.087)		(1.257)	(1.217)
First Stage Instrument				
Controls	x	x	x	x
F statistic for IV in first stage			1238.17	1238.17
N	254,654	254,654	254,654	254,654
Mean Dependent Variable	0.381	0.381	0.381	0.381
Std. Dev. Dependent Variable	0.486	0.486	0.486	0.486

Standard errors in parenthesis. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 4. OLS, CIV, Manual 2SLS, and 2SLS regressions of More Kids on Hours of Work per Week and Labor Income per Week

Outcome Variable		**More Kids**		
	OLS	Covariate-Adjusted IV	Manual 2SLS	IVregress 2SLS
Hours Worked per Week	-6.021***	-4.868***	-4.962***	-4.868***
	(0.074)		(1.021)	
Labor Income per Year	-3165.533***	-1340.519***	-1366.581**	-1340.52**
	(40.597)		(588.693)	(573.2723)
First Stage Instrument				
Controls	x	x	x	x
F statistic for IV in first stage			1238.17	1238.17
N	254,654	254,654	254,654	254,654
Mean Dependent Variable	0.381	0.381	0.381	0.381
Std. Dev. Dependent Variable	0.486	0.486	0.486	0.486

Standard errors in parenthesis. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Two stage least squares regression, both manual and automated, estimate a negative causal effect of family size on labor supply outcome variables. This means that increases in family size, exogenously changed by having two kids of the same sex, cause women to lower the likelihood that they work, the amount of time worked, average hours worked per week, and labor income per year.

OLS regression was biased and inconsistent because of unresolved endogeneity between the independent and dependent variables. As mentioned before, there is evidence that labor supply decisions effect family size, so estimating the causal effect of family size on labor supply is inconsistent because some part of change in family size is actually caused by labor supply. This also likely means that family size is correlated with the error terms, which biases the coefficients and generally overestimates causal effects. This means we can't access the true causal effect using OLS in this case.

IV solves the inconsistency problem found in OLS by utilizing exogenous changes in labor supply and family size caused by the sibling sex composition. These exogenous changes aren't directly related, thus we can estimate a true causal effect using IV. For this to be true, we have to fulfill two assumptions: Instrument relevance and instrument exogeneity. Instrument relevance dictates that the IV has to be correlated with the independent variable, otherwise there is no exogenous change to measure with. We want to see an F-statistic > 10 in order to reasonably claim instrument relevance. In this case the F-statistic = 1238.17 which is more than enough to claim instrument relevance. To fulfill the assumption instrument exogeneity, the IV cannot be related to unobservables that affect the dependent variable. We cannot test this assumption by definition, as we do not observe the unobservables. However, we know from the descriptive statistics in Table 2 that the sibling sex composition is approximately random (binary w/ $p = 0.5$). Therefore, it is reasonable to assume that our random IV is not related to unobservables that affect labor supply.

Lastly, we want to compare our regression methods. Our two stage least squares coefficient estimates have the same signs as our OLS coefficient estimates, but they are lower in magnitude. This falls in line with what we discussed earlier in saying that OLS tends to overestimate causal effects. There seems to be minute differences between manual two-stage and automated two stage, with automated having slightly lower magnitudes on coefficients as well as marginally lower standard errors. These are statistically significant differences, but do show that there are difference in the manual and automated processes of two-stage least squares. It's also important to note that the automated two-stage has exactly the same point estimates as covariate adjusted IV, as that is how STATA calculates two-stage least squares. Therefore, either method between the two will lead you to the same results.

CONCLUSION

In this replication, we find the same evidence of negative relationships between family size and labor supply that Angrist and Evans originally wrote about. This replication surely doesn't reach the same level of robustness as the original project, as we only have a subsample of Angrist and Evan's sample. However, there is still evidence of this negative causal effect of having a kid on a woman's labor supply levels. The only somewhat open question is whether the IV of sibling sex composition is truly exogenous, since we can't test this assumption. It seems reasonable to assume exogeneity since people cannot decide what sex they want their kid to be and it's approximately a 50/50 chance between having a girl or a boy. Overall, IV regression and two-stage least squares are effective ways of cutting through endogeneity between independent and dependent variables, allowing us to estimate causal effects more effectively where we were previously blocked off.