

# The Quantity-Quality Tradeoff and the Formation of Cognitive and Non-cognitive Skills\*

Chinhui Juhn<sup>†</sup>                      Yona Rubinstein<sup>‡</sup>  
University of Houston              London School of Economics

C. Andrew Zuppann<sup>§</sup>  
University of Houston

October 31, 2012

## Abstract

We estimate the impact of increases in family size on childhood and adult outcomes using matched mother-child data from the National Longitudinal Survey of Youth. We use two approaches: using twin births as exogenous shocks to family size and utilizing the precise timing of family expansion to separate out family size increases from total family size effects. We find evidence that families face a substantial quantity-quality tradeoff: increases in family size decrease childhood cognitive abilities, decrease parental investment, decrease educational attainment, and decrease measures of adulthood non-cognitive abilities.

*JEL Classification: J13, J24*

---

\*Please do not circulate without authors' permission.

<sup>†</sup>University of Houston. E-mail: [cjuhn@uh.edu](mailto:cjuhn@uh.edu)

<sup>‡</sup>London School of Economics. E-mail: [y.rubinstein@lse.ac.uk](mailto:y.rubinstein@lse.ac.uk)

<sup>§</sup>University of Houston. E-mail: [cazuppann@uh.edu](mailto:cazuppann@uh.edu)

# 1 Introduction

Researchers have consistently found that larger families lead to poorer educational outcomes for children in cross sectional data (see (Blake 1989) and (Hanushek 1992) among others). This empirical regularity is most often attributed to a quantity-quality trade-off in parental investments as suggested by the well-known model of Becker (1960), Becker & Lewis (1973), and Becker & Tomes (1976). A key feature of the quantity-quality model is the interaction of child quality and child quantity in the budget constraint, whereby an exogenous increase in quality raises the shadow price of having more children, and reductions in the number of children reduces the marginal cost of investing in quality. Becker and various co-authors originally formulated this model to explain how rising incomes, by raising the demand for quality, could lead to declining fertility, even when children are not inferior goods. The negative reinforcing mechanism between quantity and quality also plays a central role in macro growth models with endogenous fertility where higher fertility leads to less human capital investment and lower levels of growth ((Becker & Barro 1988), (Becker, Murphy & Tamura 1990), (Moav 2005)).

Despite the pre-eminence of the quantity-quality model in many empirical and theoretical papers, establishing a causal relationship between family size and education has been surprisingly challenging. The key challenge is to correct for selection - that is, allow for the fact that parents who choose to have more children may be inherently different from those who choose to have fewer children. Rosenzweig & Wolpin (1980) first addressed this problem by using twin births to instrument for quantity. They found that exogenous shocks in family size do indeed lower average schooling of children in Chinese families. Similar results have also been found in more recent Chinese data in Rosenzweig & Zhang (2009). In the U.S. context, Caceres-Delpiano (2006) and Conley & Glauber (2006) have found, using twin

births and sibling composition to instrument for family size, that children in larger families are less likely to attend private school.

In contrast to these studies, several prominent papers have found little evidence of the quantity-quality tradeoff. Black, Devereux & Salvanes (2005) use large samples from Norway and find that the negative relationship between family size and education disappears once birth order controls are included or twin births are used to instrument for family size. Similarly, Angrist, Lavy & Schlosser (2010) use the Israeli Census and alternative instruments, twin births and sex composition of children, and find no relationship between family size and education.

In this paper, we revisit these issues using detailed data of matched mothers and children from the National Longitudinal Survey of Youth (NLSY). We feel our data provides several advantages over previous work in this area. First, most previous work have focused on adult outcomes such as education and earnings. However, adult outcomes are likely to be a function of institutions as well as parental investments. For example, the lack of family size effects in Black et al. (2005) may reflect the existence of a strong public education system in Norway where at the margin, investments in child quality may not result in variations in educational outcomes. It is useful therefore to have more direct measures of children's skill levels as well as measures of parental investments. The NLSY has measures of both childhood outcomes such as cognitive and non-cognitive test-scores as well as the adult outcomes. The data also allows a more in-depth look at the channels by which parents are adjusting resources within the household, as the survey includes measures of the home environment and parental activities.

Our paper contributes to the quantity-quality tradeoff literature in two ways. First, we document the extent of the quantity-quality trade-off in the U.S. context using the established

empirical strategies and examining a wider array of available outcomes. Second, we exploit the panel structure of the NLSY to introduce a novel empirical approach to estimating the quantity-quality tradeoff. Since we know both the precise timing of expansions to family size as well as the final eventual family size once a mother ends her child-bearing, we can combine these two measures to estimate the relative importance of current versus final family size on child outcomes. This offers a valuable robustness check to the usual strategy of using twin births to instrument for family size.

Both empirical approaches find evidence of a strong tradeoff between quantity and quality within households. We find that children in larger families have significantly lower childhood abilities as well as lowered parental investment. The quantity-quality tradeoff persists into young adulthood, with children in larger families obtaining less education, having lower labor market attachment, and being more likely to experience a teenage pregnancy.

We argue that differences between these findings and previous research may arise from differences in institutions across countries. In particular, the return to parental investment may differ due to differences in the labor market or returns to skill. Differences in governmental support for additional children may also play a role in affecting the marginal cost of investing in children as family sizes increase.

The paper proceeds as follows. Section 2 describes the construction of the matched mother-child dataset. Section 3 discusses our empirical methodology. Section 4 presents our estimates and section 5 concludes.

## 2 Data

The National Longitudinal Survey of Youth provides a unique opportunity to evaluate the quantity-quality tradeoff as it contains information on childhood development as well as adult education and labor market outcomes. We match mothers from the 1979 survey with all their children from the Children and Young Adult survey. Children were surveyed biannually from 1986 to 2010. By matching children to their mothers and siblings we can identify not only identify the number of siblings and twin siblings that a child has but the precise timing of when family size expands.

Twins are not directly identified in either survey. However there are data on the month and birth of each child. We identify twins (and triplets and quadruplets) as two (or more) siblings who have the same mother and share the same year and month of birth. Out of 11,476 children respondents living in 4,510 households we identify 117 pairs of twins. Crucial to our IV identification strategy, there are 201 children who share households with twins of which 142 were born prior to the birth of twins.

The matched NLSY mother-child data contains detailed information about childhood cognitive and non-cognitive abilities as well as longer term outcomes. Children aged 4 to 14 are given Peabody Individual Achievement Tests (PIATs) that measure cognitive skills in mathematics, reading recognition, and reading comprehension. To measure non-cognitive abilities, the survey calculates a Behavioral Problem Index (BPI) and the subindices which measure particular problems including antisocial behaviors, anxiety, dependence, headstrongness, hyperactivity, and social problems. To measure parental investment, the NLSY asks questions to construct a HOME (Home Observation Measurement of the Environment-Short Form) score, "a unique observational measure of the quality of the cognitive stimulation and emotional support provided by a child's family". Examples of these questions include

how many books a child has, how often parents read to the child, and whether parents assist with homework. HOME scores have been shown to be a significant determinant in a child's development.

An advantage of these childhood measures is that we have repeated measurements over a child's life from ages 4 to 14. This brings two notable benefits. First, it allows us to perform placebo tests of how family size influences child quality *prior* to the increase in family size. Second, by combining precise timing of shocks to family size with test scores at various ages we can measure how the quantity-quality tradeoff may differ at different stages of childhood development.

The NLSY continues to follow children into their adult life allowing us to estimate how the quantity-quality tradeoff may have long term effects. The survey measures years of education, labor force participation, and wage/salary information. Additionally, there are measures of adult non-cognitive traits such as self-esteem scores and locus of self-control metrics.

Table 1 presents summary statistics of our matched mother-child sample. Almost 4% of individuals live in a household with twins. Blacks and hispanics are over-represented in our sample due to NLSY79 oversampling. Only 61% of the survey were at least 22 years old at the time of the 2010 survey. Roughly a quarter of the sample is unemployed in 2010 and almost a quarter of them report having been convicted of a crime.

Since many children in our sample are not yet adults by the last survey wave measuring adult outcomes such as educational attainment or wages is problematic. To minimize the measurement error of these outcomes we drop observations of adult outcomes for respondents who are younger than 22 at the time of the 2010 survey.<sup>1</sup> Given the IV strategy relies on

---

<sup>1</sup>The outcomes that we drop for respondents who are younger than 22 are: total years of education, high school completion, wages, hours worked, criminal convictions, and teen pregnancy.

identification through the older children in the household these dropped variables do not have a large impact on our estimates.

Table 2 shows the distribution of households of various family sizes as well as the timing of twin births in our sample. Using twins as an IV for family size shock is identified off individuals who have twins as younger siblings. For example, there are 35 families where twins are the second and third children born in the family, implying that there are 35 older children in these families who are in our sample. Likewise, there are 19 families where twins are third and fourth children born implying there are 38 older children added to our sample from these families. Altogether, the table indicates that there are 142 of these individuals in our data. It is common in this literature to attempt to adjust for parity of total births in a household with the hope of ensuring that the twin birth does represent an unplanned shock to family size. Table 2 suggests that the NLSY sample may be too small to perform adequate parity adjustments and still obtain precise estimates. As such, our baseline estimates will include the full sample without parity controls although some robustness checks are performed that suggest that parity concerns do not appear to be significant in our data.

## 3 Empirical Approach

### 3.1 Twin births as a shock to family size

Our instrumental variable strategy directly follows Black et al. (2005) in estimating the effects of family size on outcomes as described by the following instrumental variable equations:

$$Y = \beta_0 + \beta_1 FAMSIZ E + X\beta_2 + \varepsilon \tag{1}$$

$$FAMSIZ E = \alpha_0 + \alpha_1 TWIN + X\alpha_2 + \nu \tag{2}$$

where  $Y$  is an outcome of interest, e.g., years of education or test scores,  $FAMSIZ E$  is the number of children in the family, and  $X$  is vector of family characteristics.  $TWIN$  is a dummy variable indicating whether the family has any twins. We restrict our attention to children who were born prior to the twin birth. As a conceptual example of the source of identification, we are comparing outcomes of firstborn children across families where the second birth was either a singleton or twins. Families with twin births have had a plausibly exogenous increase in family size. Differences in outcomes between firstborns in households with twins versus households without twins can thus be interpreted as a causal estimate of how family size affects outcomes.

Having panel data on childhood outcomes allows us to go further than just comparing the cross-sectional differences between families with and without twins. We can also differentially look at the effects of larger family sizes on children *before* the increase in family size. Looking at the effect of twin births on outcomes prior to the birth is a test of the exogeneity of the twin instrument. If the presence of twins in the household affects earlier siblings' outcomes through the quantity-quality tradeoff then we should not expect it to influence outcomes



before the quantity expanded. To test for the presence of pre-trend effects of family size we estimate the IV regression specified in (1) and (2) but subsample to children whose parents will have siblings but before those siblings' births.

### 3.2 Current vs future family size

Another approach to evaluate the effects of family size is to estimate the relative importance of a household's family size at a given moment in time compared to the household's eventual family size after the mother has ended childbearing. This is a direct test of whether the family effect or the family size effect matters more in child development. If endogenous selection of fertility confounds estimates of the quantity-quality tradeoff then we would expect final family size is a substantial force in explaining the variation of child outcomes. If parents do face an important tradeoff in allocating resources among their children then current family size should be important, even after controlling for the eventual size of the family.

As an example of this logic, if selection on fertility is a concern children in families that will *eventually* be large but are currently small should have the same outcomes as children in families that are currently large. However, if parents do face constraints in allocating resources across children then a child in a currently small family should have better outcomes than a child in a currently large family, controlling for eventual family size.

To formalize this test of the important of current family size versus eventual family size, we estimate regressions of the following form:

$$Y_t = \beta_0 + \beta_1 FAMSIZ E_t + \beta_2 FAMSIZ E_T + \varepsilon \tag{3}$$

where the subscripts denote the year of observation and  $FAMSIZ E_T$  is the final family size

after the mother is finished child-bearing. The quantity-quality tradeoff hypothesizes that the current family size should matter more than the future family size as it is the present resource scarcity within the household that drives parental decisions. By comparing  $\beta_1$  and  $\beta_2$  from this regression we can test this hypothesis.

Having data that combines the timing of births in a household along with panel data on childhood outcomes is crucial in estimating equation (3). By observing children before and after the arrival of their siblings as well as knowing the eventual size of the family allows us to identify the separate channels of current and final family size.

### **3.3 On NLSY households with twins**

One potential concern with using the NLSY sample in our empirical IV strategy is the relatively small number of households with twins in the sample. A mere 142 children in the matched dataset are born prior to the arrival of twins. Although concerns about sample size are usually sufficiently addressed in calculating standard errors, it is worth investigating these twin households in more detail to ensure that there are not substantial observable differences between households with and without twins.

Table 3 shows demographic information for mothers in our sample split by the presence of twins in the household. There is not a significant difference in mothers that have twins and those that do not. If anything, the mothers without twins are slightly less able in both cognitive and non-cognitive measures, with lower AFQT scores and lower self-esteem. These numbers are reassuring that the presence of twins within the household provide a plausible shock to family size and can overcome the problems of negative selection of family size and child outcomes.

## 4 Results

### 4.1 IV estimates using twin births

Table 4 presents estimates of the impact of family size on adult outcomes for the children of the NLSY. Columns 1 and 2 are OLS estimates with and without birth order controls and column 3 uses the presence of a younger twin birth as an instrument for family size. Column 4 shows the first stage coefficient for the IV estimates of the full sample.

Looking first at years of educational attainment, our OLS estimates are consistent with previous findings in the literature. We find a significantly negative relationship between larger families and years of education but this relationship declines substantially if controls for birth order are included. Our estimates using twin births as an instrument for family size are significantly different from previous finding. Instead of small, possibly zero magnitudes, our IV estimates are negative and larger than OLS estimates. Older children in families with twins have significantly reduced eventual education compared to children in families without twins. We estimate the causal impact of an additional sibling as reducing years of education by roughly a third of a year.

A similar pattern is found for other adult outcomes: a significant negative relationship in simple OLS specifications that declines substantially upon the inclusion of birth order effects but IV estimates indicate a large and significantly negative impact. The IV estimates indicate that larger families not only have lower educational attainment, but lower earnings, lower labor market participation, and increased likelihood of criminal behavior and teenage pregnancies.

Given the strength of these negative relationships, it is an advantage of the NLSY that we can investigate the potentials channels by which the quantity-quality tradeoff are occurring

in these families. Table 5 presents estimates of how family size impacts outcomes during childhood. We find evidence of family size having a substantial and negative effect on childhood cognitive scores. IV estimates suggest that an additional sibling can lower both math and reading scores by approximately a tenth of a standard deviation.

We also find a significant and negative impact of family size on a children's HOME score - the NLSY's measure of parental investment in a child. This finding speaks directly to the trade-offs in a quantity-quality model of child-rearing. Parents in larger families reduce their per-child investment.

As a test of the exogeneity of the instrument we run a series of placebo tests that estimate the effect of family size increases on outcomes prior to the births. Table 6 shows the results of these tests. Children with younger twin siblings do not have a significantly different birthweight than those with younger singleton siblings. Further, IV estimates fail to find an impact of family size shocks on child outcomes prior to the shock. We fail to find evidence that twin births are endogenously related to the previous children or household.

## 4.2 Current vs future family size

Our second empirical approach uses the panel structure of the NLSY sample combined with the timing of increases in family size to simultaneously estimate the effect of current versus final size on outcomes during childhood. Results are presented in Table 7.

Current family size is significantly negatively related with all four measures of cognitive ability while eventual family size appears to have close to no effect in all of our measures of childhood ability. Coefficients on the current family size variable are all negative, significant, and similar to baseline OLS estimates of the effect of family size on outcomes. Having one more child in the household lowers achievement by roughly one tenth of a standard

deviation. Importantly, this effect is *conditional* on final family size, which has a much smaller and statistically insignificant relation to ability measures.

We interpret these estimates as strong evidence for a quantity-quality tradeoff within the household. Children in families that will eventually be large but are currently small have significantly better outcomes than those children already living in large families.

### 4.3 Robustness

One notable difference in our approach from previous studies is that our sample is unadjusted for parity in family sizes across twin households and non-twin households. Tables 8 and 9 show estimates that control for parity by restricting our sample to families that had at most 3 births. Point estimates are not substantially changed by making parity adjustments although there is a loss of precision. Due to the relatively small number of households in our sample, we prefer the more precise estimates that do not control for parity.

Another concern with our estimates is that we have not sufficiently controlled for possible shocks at the household level. By clustering at the level of individual children, it is possible that we have overestimated the precision of our point estimates. We address this concern in two ways: first by clustering at the household level and second by restricting our sample to only firstborn children. Tables 10 and 11 present estimates with errors clustered by mother. Results are unchanged for adult outcomes although the precision of our estimates do worsen. Tables 12 and 13 show estimates when we restrict the sample to only firstborn children. Point estimates are similar to estimates from the full sample but there is a marked increase in standard errors for all estimates.

## 5 Conclusion

We find a significant tradeoff between quantity and quality of children in for NLSY79 mothers and their children. Children in larger families have worse outcomes both during childhood and into young adulthood. They have lower test scores, decreased parental investment, are more likely to get pregnant as a teenager, obtain fewer years of education, have lower labor force attachment, and have lower wages.

These results differ substantially from previous research so it is worthwhile to consider why our estimates are so different. One potential explanation for our results in contrast to (Black et al. 2005) may be institutional differences between Norway and the U.S. In particular, at the margin, parental investments may matter more in the U.S. where a substantial fraction of young men and women, particularly from lower income backgrounds, are at risk of not finishing high school. A recent paper by (Black, Devereux & Salvanes 2010) offers some support for the idea that the particular country and the particular cohort examined matters. In contrast to their earlier which examined an older cohort, their recent paper finds a negative impact of family size on IQ among younger birth cohorts in Norway. Likewise, Li, Zhang & Zhu (2008) also find that the negative family size effect on schooling is particularly strong in rural areas of China where the public education system is less well developed. Finally, child allowances and subsidies may have mitigated the impact of family size in the Israeli data examined by Angrist et al. (2010) although the authors minimize the importance of these subsidies.

## References

- Angrist, J. D., Lavy, V. & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children, *Journal of Labor Economics* **28**: 773–824.
- Becker, G. S. (1960). An economic analysis of fertility, *Demographic and Economic Change in Developed Countries* .
- Becker, G. S. & Barro, R. J. (1988). A reformulation of the economic theory of fertility, *The Quarterly Journal of Economics* **103**: 1–25.
- Becker, G. S. & Lewis, H. G. (1973). On the interaction between the quantity and quality of children, *Journal of Political Economy* **81**: S279–288.
- Becker, G. S., Murphy, K. M. & Tamura, R. (1990). Human capital, fertility, and economic growth, *Journal of Political Economy* **98**: S12–37.
- Becker, G. S. & Tomes, N. (1976). Child endowments and the quantity and quality of children, *Journal of Political Economy* **84**: S143–162.
- Black, S. E., Devereux, P. J. & Salvanes, K. G. (2005). The more the merrier? the effects of family size and birth order on children’s education, *The Quarterly Journal of Economics* **120**: 669–700.
- Black, S. E., Devereux, P. J. & Salvanes, K. G. (2010). Small family, smart family? family size and the iq scores of young men, *Journal of Human Resources* **35**: 33–58.
- Blake, J. (1989). *Family Size and Achievement*, University of California Press.
- Caceres-Delpiano, J. (2006). The impacts of family size on investment in child quality, *Journal of Human Resources* **41**: 738–754.

- Conley, D. & Glauber, R. (2006). Parental educational investment and children's academic risk: Estimates of the impact of sibship size and birth order from exogenous variation in fertility, *Journal of Human Resources* **41**: 722–737.
- Hanushek, E. (1992). The trade-off between child quantity and quality, *Journal of Political Economy* **100**: 84–117.
- Li, H., Zhang, J. & Zhu, Y. (2008). Cheap children and the persistence of poverty, *Demography* **45**: 223–243.
- Moav, O. (2005). Cheap children and the persistence of poverty, *The Economic Journal* **115**: 88–110.
- Rosenzweig, M. R. & Wolpin, K. I. (1980). Testing the quantity-quality fertility model: The use of twins as a natural experiment, *Econometrica* **48**: 227–240.
- Rosenzweig, M. R. & Zhang, J. (2009). Do population control policies induce more human capital investment? twins, birth weight and china's one-child policy, *Review of Economic Studies* **76**: 1149–1174.



Table 1: Children of NLSY79: Summary Statistics

	Mean	Std. Dev.	# of Obs.
Age in 2010	23.8	6.41	10501
Older than 22 in 2010	0.61	0.49	10501
Female	0.49	0.50	10501
Black	0.28	0.45	10501
Hispanic	0.19	0.39	10501
Birth weight of child (oz)	116.1	22.8	9473
Family size	2.91	1.40	10501
Twins in family	0.037	0.19	10501
Yrs of education — older than 22	12.7	2.32	4522
HS grad — older than 22	0.75	0.43	4522
2009 Wage + salary income	21138.7	21162.0	2964
Log 2009 wage + salary income	9.79	1.04	2419
Worked 35+ hours at 1st job, 2010	0.73	0.44	2903
Child was convicted of a crime	0.24	0.43	4507
Child had a teenage pregnancy indicator	0.074	0.26	6379
Math test (percentile)	51.5	28.0	31554
Reading comprehension test (percentile)	52.8	27.7	26806
Reading recognition test (percentile)	58.1	28.4	31444
Picture vocabulary test (percentile)	35.9	30.3	18725
HOME Inventory score (percentile)	46.6	29.4	46360
Behavioral Problems Index (percentile)	58.9	28.1	35697
Rosenberg self-esteem index (percentile)	49.9	28.2	12589
Pearlin locus of control index (percentile)	50.0	28.2	12589

Data from Children of the NLSY79, 1986-2010. Adult outcomes in the middle box are conditional on being at least 22 years old in the 2010 survey wave.

Table 2: Distribution of family size and twin births

	Family size # of HHs	Birth order of twins within family # of HHs
1	1069	41
2	1814	34
3	1033	17
4	394	9
5	125	5
6	44	2
7	18	
8	6	
9	4	
10	1	
11	2	
Total	4510	108

Data from Children of the NLSY79, 1986-2010.

Table 3: Maternal characteristics by twin and non-twin households

	HH w/o twins	HH w/ twins	Difference	p-value
Mother's AFQT pctile	40.2	44.0	-3.794	0.161
Mother's self-esteem index	22.1	22.6	-0.556	0.168
Mother's locus-of-control index	8.82	8.74	0.0760	0.746
Age of mother at birth of child	22.9	23.3	-0.438	0.412
Black	0.26	0.26	0.00245	0.954
Hispanic	0.17	0.17	0	1

Data from Children of the NLSY79 sample, 1986-2010. The last column reports p-values of an unpaired t-test for difference in means between households with and without twins.

Table 4: Long-term outcomes

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	First Stage
Twin coefficient				0.712*** (0.0664)
Years of education	-0.144*** (0.0320)	-0.0430 (0.0360)	-0.359** (0.178)	
Graduated HS	-0.0361*** (0.00443)	-0.0186** (0.00599)	-0.0891** (0.0287)	
Earnings	-1080.2** (364.5)	-856.6* (466.0)	-6071.9*** (1362.7)	
Log earnings	-0.0502** (0.0189)	-0.0487** (0.0220)	-0.192** (0.0757)	
Labor market participation	-0.0219* (0.0112)	-0.0266* (0.0138)	-0.142* (0.0732)	
Works full-time	-0.0167** (0.00732)	-0.0107 (0.00852)	-0.155*** (0.0436)	
Criminal conviction	0.0287*** (0.00497)	0.0212*** (0.00594)	0.0396 (0.0280)	
Teen pregnancy	0.0138*** (0.00280)	0.0104** (0.00326)	0.0729*** (0.0206)	
Birth order controls	No	Yes	Yes	

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 5: Childhood outcomes

	(1)	(2)	(3)
	OLS	OLS	IV
Math	-1.758*** (0.171)	-1.003*** (0.211)	-2.533** (1.118)
Reading Comprehension	-2.111*** (0.174)	-1.053*** (0.220)	-1.837* (1.114)
Reading Recognition	-2.455*** (0.198)	-1.414*** (0.250)	-2.263* (1.302)
Picture vocabulary	-2.630*** (0.172)	-1.468*** (0.221)	-2.012** (1.014)
HOME Index	-2.722*** (0.153)	-1.218*** (0.208)	-2.749** (1.032)
Locus of control	-1.009*** (0.231)	-0.852** (0.275)	-1.947 (1.341)
Self-esteem	-1.596*** (0.237)	-1.400*** (0.284)	-0.757 (1.444)
Behavioral Problems	-0.261 (0.196)	-0.471* (0.248)	0.795 (1.218)
Birth order controls	No	Yes	Yes

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 6: Outcomes prior to twin birth

	(1)
	IV
Birthweight (oz)	-0.419 (1.510)
Reading Comprehension	0.842 (1.847)
Reading Comprehension	-0.348 (1.948)
Reading Recognition	-1.241 (2.007)
Picture vocabulary	-1.571 (1.644)
HOME Index	-1.511 (1.827)

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010, subsampled to years prior to the birth of younger siblings.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 7: Current versus final family size

		(1)
		OLS
Math	Current family size	-2.130*** (0.419)
	Final family size	0.0606 (0.393)
Reading Comprehension	Current family size	-2.242*** (0.473)
	Final family size	-0.100 (0.453)
Reading Recognition	Current family size	-2.112*** (0.477)
	Final family size	-0.540 (0.454)
Picture vocabulary	Current family size	-3.282*** (0.394)
	Final family size	0.142 (0.374)
HOME Index	Current family size	-2.472*** (0.323)
	Final family size	-1.111*** (0.311)
Locus of control	Current family size	0.0801 (1.590)
	Final family size	-1.184 (1.575)
Self-esteem	Current family size	-0.0960 (1.642)
	Final family size	-1.619 (1.631)
Behavioral Problems	Current family size	0.366 (0.459)
	Final family size	-0.492 (0.439)

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 8: No families with >3 births

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	First Stage
Twin coefficient				1.239*** (0.0771)
Years of education	-0.184** (0.0591)	-0.0395 (0.0678)	-0.287** (0.146)	
Graduated HS	-0.0373*** (0.00813)	-0.0171 (0.0112)	-0.0721** (0.0223)	
Earnings	-1395.6** (541.2)	-1314.6** (656.4)	-4808.7*** (969.5)	
Log earnings	0.000867 (0.0316)	-0.000442 (0.0350)	-0.169** (0.0641)	
Labor market participation	-0.0580** (0.0182)	-0.0587** (0.0242)	-0.105** (0.0502)	
Works full-time	-0.00923 (0.0153)	-0.00859 (0.0179)	-0.129*** (0.0358)	
Criminal conviction	0.0215** (0.00881)	0.0101 (0.0102)	0.0322 (0.0215)	
Teen pregnancy	0.0265*** (0.00580)	0.0213** (0.00656)	0.0522*** (0.0158)	

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010, subsampled to families with have 3 or fewer births.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 9: No families with >3 births

	(1)	(2)	(3)
	OLS	OLS	IV
Math	-2.277*** (0.305)	-1.226*** (0.369)	-2.103** (0.922)
Reading Comprehension	-2.346*** (0.314)	-0.802** (0.380)	-1.565* (0.923)
Reading Recognition	-2.785*** (0.349)	-1.226** (0.428)	-2.059* (1.061)
Picture vocabulary	-3.214*** (0.317)	-1.775*** (0.402)	-1.780** (0.847)
HOME Index	-1.792*** (0.306)	0.236 (0.364)	-2.642** (0.872)
Locus of control	-1.354** (0.430)	-1.091** (0.524)	-1.972* (1.061)
Self-esteem	-1.377** (0.451)	-1.137** (0.542)	-1.317 (1.132)
Behavioral Problems	-0.589* (0.338)	-0.878** (0.416)	0.584 (1.033)

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010, subsampled to families with have 3 or fewer births.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.



Table 10: Clustering at mother level

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	First Stage
Twin coefficient				0.712*** (0.103)
Years of education	-0.144*** (0.0367)	-0.0430 (0.0400)	-0.359* (0.196)	
Graduated HS	-0.0361*** (0.00555)	-0.0186** (0.00670)	-0.0891** (0.0306)	
Earnings	-1080.2** (392.1)	-856.6* (488.5)	-6071.9*** (1390.5)	
Log earnings	-0.0502** (0.0199)	-0.0487** (0.0229)	-0.192** (0.0764)	
Labor market participation	-0.0219* (0.0119)	-0.0266* (0.0141)	-0.142** (0.0715)	
Works full-time	-0.0167** (0.00745)	-0.0107 (0.00882)	-0.155*** (0.0435)	
Criminal conviction	0.0287*** (0.00537)	0.0212*** (0.00622)	0.0396 (0.0259)	
Teen pregnancy	0.0138*** (0.00319)	0.0104** (0.00339)	0.0729*** (0.0188)	
Birth order controls	No	Yes	Yes	

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the household level. Data from Children of the NLSY79 sample, 1986-2010.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 11: Clustering at mother level

	(1)	(2)	(3)
	OLS	OLS	IV
Math	-1.758*** (0.231)	-1.003*** (0.249)	-2.533** (1.244)
Reading Comprehension	-2.111*** (0.226)	-1.053*** (0.255)	-1.837 (1.318)
Reading Recognition	-2.455*** (0.268)	-1.414*** (0.296)	-2.263 (1.481)
Picture vocabulary	-2.630*** (0.233)	-1.468*** (0.267)	-2.012* (1.205)
HOME Index	-2.722*** (0.254)	-1.218*** (0.301)	-2.749* (1.431)
Locus of control	-1.009*** (0.296)	-0.852** (0.322)	-1.947 (1.501)
Self-esteem	-1.596*** (0.285)	-1.400*** (0.316)	-0.757 (1.489)
Behavioral Problems	-0.261 (0.330)	-0.471 (0.347)	0.795 (1.471)
Birth order controls	No	Yes	Yes

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the household level. Data from Children of the NLSY79 sample, 1986-2010.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 12: Just first-borns

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	First Stage
Twin coefficient				1.014*** (0.120)
Years of education	-0.0477 (0.0488)	-0.0477 (0.0488)	-0.190 (0.281)	
Graduated HS	-0.0156** (0.00731)	-0.0179** (0.00805)	-0.120** (0.0440)	
Earnings	104.9 (624.5)	104.9 (624.5)	-7481.7** (3176.1)	
Log earnings	0.0109 (0.0290)	0.0109 (0.0290)	-0.125 (0.100)	
Labor market participation	-0.0280 (0.0200)	-0.0280 (0.0200)	-0.107 (0.129)	
Works full-time	0.000786 (0.0114)	0.000786 (0.0114)	-0.227** (0.0776)	
Criminal conviction	0.0223** (0.00816)	0.0223** (0.00816)	0.0267 (0.0424)	
Teen pregnancy	0.00866** (0.00404)	0.00866** (0.00404)	0.00877 (0.0231)	

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010, only first-borns.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 13: Just first-borns

	(1)	(2)	(3)
	OLS	OLS	IV
Math	-0.882** (0.295)	-0.882** (0.295)	-2.751 (1.853)
Reading Comprehension	-0.996** (0.311)	-0.996** (0.311)	-0.243 (1.954)
Reading Recognition	-1.437*** (0.355)	-1.437*** (0.355)	0.00867 (2.306)
Picture vocabulary	-1.356*** (0.333)	-1.356*** (0.333)	-2.167 (1.693)
HOME Index	-0.480 (0.309)	-0.480 (0.309)	-2.235 (1.654)
Locus of control	-0.740** (0.369)	-0.740** (0.369)	-3.292 (2.337)
Self-esteem	-1.237** (0.378)	-1.237** (0.378)	-1.034 (2.500)
Behavioral Problems	-0.631* (0.332)	-0.631* (0.332)	-0.687 (2.109)

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010, only first-borns.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 14: With and without birthweight controls

	(1)	(2)
	IV	IV
Years of education	-0.329*	-0.328*
	(0.168)	(0.169)
Graduated HS	-0.0915**	-0.0924**
	(0.0289)	(0.0290)
Earnings	-6049.4***	-6478.2***
	(1360.7)	(1444.3)
Log earnings	-0.193**	-0.199**
	(0.0756)	(0.0769)
Labor market participation	-0.144**	-0.151**
	(0.0734)	(0.0761)
Works full-time	-0.156***	-0.158***
	(0.0435)	(0.0433)
Criminal conviction	0.0407	0.0408
	(0.0281)	(0.0281)
Teen pregnancy	0.0724***	0.0720***
	(0.0206)	(0.0205)
Birthweight control	No	Yes

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.

Table 15: With and without birthweight controls

	(1)	(2)
	IV	IV
Math	-2.533** (1.118)	-2.580** (1.134)
Reading Comprehension	-1.837* (1.114)	-1.868* (1.112)
Reading Recognition	-2.263* (1.302)	-2.308* (1.301)
Picture vocabulary	-2.012** (1.014)	-2.037** (1.020)
HOME Index	-2.749** (1.032)	-2.769** (1.036)
Locus of control	-1.947 (1.341)	-1.981 (1.346)
Self-esteem	-0.757 (1.444)	-0.780 (1.445)
Behavioral Problems	0.795 (1.218)	0.810 (1.228)
Birthweight control	No	Yes

Each row presents estimates for a different outcome of interest. Robust standard errors are clustered at the individual level. Data from Children of the NLSY79 sample, 1986-2010.

\*: significant at 10% level. \*\*: significant at 5% level. \*\*\*: significant at 1% level.