Birth Spacing and Educational Outcomes

Elaine L. Hill
David J.G. Slusky

October 5, 2016

Abstract

Purpose: Virtually all parents want their children to succeed academically. How to achieve this goal, though, is far from clear. Specifically, the temporal spacing between adjacent births has been shown to affect educational outcomes. While many of these studies have produced substantial and statistically significant results, these results have been relatively narrow in their application due to data limitations.

Design/methodology/approach: Using Colorado birth certificates matched to schooling outcomes, we investigate the relationship between birth spacing and educational attainment. We instrument birth spacing with a previous pregnancy that did not result in a live birth.

Findings: We find no overall effect of spacing on either the first or second children's grade 3-10 test scores. Stratifying by the sexes of the children, we find that when the first child is a boy and the second a girl, an extra year of spacing increases the first child's math, reading, and writing test scores by 0.07-0.08 SD, while there is no impact on the second child.

Originality/value: This is the first study to do such an analysis using matched large scale birth and elementary to high school administrative data, and to leverage a very large data set to stratify our results by the sexes of the children.

JEL codes: J13, I14, I20

Keywords: human capital, educational attainment, birth spacing, pregnancy loss, miscarriage

Affiliations

Hill (corresponding author): Department of Public Health Sciences, University of Rochester School of Medicine and Dentistry, 265 Crittenden Blvd, Box 420644, Rochester, NY 14642, elaine_hill@urmc.rochester.edu

Slusky: Department of Economics, University of Kansas, 1460 Jayhawk Blvd., 415 Snow Hall, Lawrence, KS, 66045, david.slusky@ku.edu

Introduction

Virtually all parents want their children to succeed academically. How to achieve this goal, though, is far from clear. Specifically, a child's environment in utero (Almond and Currie 2011), weight at birth (Figlio, Guryan, Karbownik, and Roth 2014), birth order (Hotz and Pantano 2013) and the temporal spacing between adjacent births (Buckles and Munnich 2012) all have been shown to affect educational outcomes. While many of these studies have produced substantial and statistically significant results, these results have been relatively narrow in their application due to data limitations.

In this paper, we surmount many of these limitations using a newly matched dataset: administrative records on Colorado birth certificates matched to schooling outcomes. As far as we know this is the first study to do such an analysis using matched large scale birth and elementary to high school administrative data.¹

Background

Birth spacing can be viewed as a proxy for the quantity of resources available for parental investment in children (Buckles and Munnich 2012). It can have long lasting effects both on a child's cognitive and overall human capital development (Black, Devereux, and Salvanes 2007; Buckles and Munnich 2012).

Our work also follows a long line of literature stemming from Rosenzweig (1986) and Rosenzweig and Wolpin (1988), both of which examine the relationship between optimal child spacing and investment in the human capital of those children. Whereas these studies use a very small sample size (in the hundreds), and the Buckles and Munnich (2012) work uses survey data with a medium sample size (in the thousands), here we are able to use administrative data with

¹Figlio et al. 2014 only go through 8th grade.

hundreds of thousands of observations. This will allow us to be more restrictive in our overall sample to increase the robustness of our empirical setup, and also delve into stratifications of the data.

Previous research has found that larger spacing between births has a positive effect on K-12 educational attainment, with an extra year increasing test scores by 0.17 standard deviations for the older siblings but no impact on the younger ones. (Buckles and Munnich 2012; Black, Devereux, and Salvanes 2005). Newer research has suggested that a larger space between births increases the probability of the mother re-entering the labor market between births and results in increased labor income and long-run wages for mothers with larger spaces between their children's births (Karimi 2014). Additionally, research suggests that short birth intervals may cause adverse birth outcomes for younger siblings due to maternal nutritional depletion (Conde-Agudelo, Rosas-Bermúdez, and Kafury-Goeta 2006).

Given this literature, we ask two previously unanswered questions: 1) Does the effect of birth spacing on educational attainment persist through high school, using a previous pregnancy that did not result in a live birth as an instrument? and 2) Does this relationships change depending on the sexes of the children?

Conceptual Framework

Consider parents faced with the choice of how many resources to invest in the human capital development of their offspring, given a fixed endowment of time. Having a second child can be seen as a particular investment decision, i.e., increasing the time spent investing in the human capital of the second child from zero to a positive number.

This situation can be parameterized as follows. Each child i in a family ($i = 1$ or 2) has a human capital stock H_t^i at time t . This stock is an aggregation of a function of all of the investments, I , previously made by the child's parents:

$$H_t^i = \sum_{\tau=0}^{t-1} \delta(\tau) F(I_\tau^i; H_\tau^i) \quad (1)$$

F is the production function of human capital, where $F' > 0$ and $F'' < 0$, and δ is some discount function. H_τ^i appears in the production function since the marginal product of investment in human capital is also affected by the stock of human capital at the time of investment.

The parents' utility is a function of the human capital stocks of their children:

$$U_t = G(H_t^1, H_t^2) \quad (2)$$

where $G_1 > 0$, $G_2 > 0$, $G_{11} < 0$, and $G_{22} < 0$.

Finally, the parents are subject to a budget constraint, such that

$$T_t = H_t^1 + H_t^2 \quad (3)$$

where T is the time endowment of the parents.

From this model, the following optimal investment strategy emerges: (1) Invest everything in the first child in the first period. (2) Choose to have a second kid whenever the marginal utility of putting some of the investment into the second child exceeds the marginal cost of no longer being able to invest it in the first child. (3) Split the investment between the two children to maximize the next period's utility. This will start out with the majority of the investment in the second child (due to the greater initial marginal utility) but will eventually converge to a split such that the marginal utilities are equal.

What now happens if this family is unable to have their children optimally spaced and instead has a longer space (e.g., due to a previous pregnancy that did not result in a live birth)? In the short term, the first child is better off as he or she receives the entire endowment as an

investment. In the medium term, though, if the marginal product of investment declines as the stock of human capital rises, the first child may become worse off. This is because the ratio of the marginal utility of investing in that child versus in the first is large and so the first child will receive a lower investment in the first years of the second child's life. The second child, on the other hand, may receive greater overall investment due to the larger investment in the early years from this high marginal product. This effect, though, may be very small, as parents may ultimately attempt to equalize the stocks of the two children. It also depends on the particular functional forms for the model above.

This illustrative model shows the challenge in pinning down a one-sided hypothesis for the empirical analysis below. Having an unexpected extra year could result in a persistent human capital rise for the first child. Or it could result in a persistent human capital loss as even more resources are diverted to the new child. The second child likewise could see better or worse outcomes. It is also possible that any suboptimal birth spacing (e.g., one affected by a previous pregnancy that did not result in a live birth) results in worse overall outcomes for both children as it diverges from the utility maximizing path. Finally, it is possible that the parents will adjust the investments in to the two children such that the human capital of each eventually reaches the previous optimum, albeit under a different path, and so there is no overall resulting change. The analysis below will provide quantitative results for these hypotheses. At the moment though, is it reasonable from the model to test both as two-sided.

Methodology

Our estimations below use the following empirical structure

$$Education_{igmt} = \alpha + Birthspacing_{imt} + \rho_t + \mathbf{X}'_{imt}\boldsymbol{\beta} + \varepsilon_{igmt} \quad (4)$$

The unit of observation is individual i , born in time t to mother m and tested in grade g . *Education* refers to the individual's standardized test scores in the matched education data in a particular grade. *Birthspacing* is a measure of parental investment in the individual (i.e., birth spacing). ρ are year-of-birth and month-of-birth fixed effects. X are demographic controls for the mother at the time of the birth, including maternal age, education, race, ethnicity, marital status, the sex of the child, and the ultimate number of children born to that mother. Finally, robust standard errors are clustered at the child level.

There is a substantial concern that birth spacing is endogenous to social economic and demographic factors, which in turn are correlated with educational outcomes. Following the literature, we use the presence of a previous pregnancy that did not result in a live birth (i.e., a pregnancy loss at any gestational age, as defined on birth certificate reporting) as an instrument for birth spacing (as in Buckles and Munnich 2012). On the birth certificate, each woman is asked to report number of live births still living, number of live births now dead, and number of pregnancy losses. They also report the dates of the last live birth and the last pregnancy loss date. For each unique mother in our data (using a unique mother id), we calculate the spacing between each reported birth in the vital records. We then form our instrument defined as a binary variable for whether the mother experienced a pregnancy loss (using last pregnancy loss date) in between each sibling pair.

We restrict our sample in two ways to improve the exogeneity of this instrument. First, we exclude all mothers who experienced a pregnancy loss before the first birth, as these women may have known that they had a greater propensity to miscarry and expected one between their first and second children. Secondly, we only study first and second births and not higher order births, as again any pregnancy loss between the first and second child would have caused the

mother's expectation of one between the second and third to increase. Additionally, we exclude mothers with more than three children as they are less likely to be comparable with our main sample. Finally, we further restrict the sample to records that link to the public school test score data as this linkage is crucial for our analysis.

Pregnancy loss is only an appropriate instrument in the case where mothers are planning to have another child and would carry that baby to term in the absence of a pregnancy loss (Fletcher and Wolfe 2009). We believe that this is not a concern because we are limiting our analyses to women who already have a child and in the future will have additional children.

We are also potentially concerned that a sizable portion of the pregnancy losses reported on the birth certificates are elective abortions, as birth certificates at this time did not distinguish among pregnancy losses that were spontaneous, induced out of a medical necessity, or elective.² This concern is founded in part by the fact that the total women in our data who reported their last pregnancy loss in a particular year (e.g., 2006) is 11,284, whereas the number of abortions in that year in Colorado is only roughly comparable at 10,636 (CDC 2008).

However, we believe that these numbers are not actually comparing the same thing (and so the apparent comparability is not a concern), as the number of the pregnancy losses reported in our data undercounts the total number of pregnancy losses for two reasons: (1) We only see mothers who had another child (otherwise there is no second child's birth certificate reporting the last pregnancy loss date as between the two births) and (2) Only the last pregnancy loss is reported and so if there were multiple we'd only count them as one. Furthermore, some induced abortions are for fetal abnormalities, which are effectively exogenous. Finally, approximately

² See questions 12 and 13 in the "Facility Worksheet for the Live Birth Certificate" (<https://www.cdc.gov/nchs/data/dvs/facwksBF04.pdf>) and in the "Guide to Completing the Facility Worksheets for the Certificate of Live Birth and Report of Fetal Death" (<http://www.cdc.gov/nchs/data/dvs/GuidetoCompleteFacilityWks.pdf>), both of which were current at the time of the births of individuals studied in this paper.

6% of all births in our data have a pregnancy loss between births, whereas Buckles and Munnich (2012) find a rate of 5.8% miscarriages or stillbirths between births, suggesting that only approximately 3% of the pregnancy losses are induced abortions.

There is one additional threat to the exclusion restriction which is that even spontaneous terminations are not exogenous and are instead correlated with unobserved heterogeneity across mothers. We believe this is not a major concern for three reasons. First, we are following a well-established practice in the literature of using pregnancy loss as an instrument for birth timing, not only in the work by Buckles and Munnich (2012) mentioned above but also Hotz, Mullin, and Sanders (1997), Hotz, McElroy, and Sanders (2005), and Miller (2011). Second, as described above, we are excluding women who had a pregnancy loss of any time before the first birth, which should eliminate many of the women for whom this is a concern. Finally, we control for numerous observable behavioral differences across mothers in the first stage, and still find that the instrument is extremely strong, suggesting that it not proxying for unobserved heterogeneity. Given all of this and our sample restrictions above, we believe that pregnancy losses between first and second births is an exogenous instrument for birth spacing for mothers who did not have a pregnancy loss before the first birth.

Data

The first set of data that we utilize is the birth certificate records from the Colorado Department of Public Health and the Environment (CDPHE) for the years 1990 through 2005. These natality records contain detailed information on every birth in the state including health at birth and background information on the mother and father which includes race, education, and

marital status as well as prenatal care and whether the mother smoked or drank alcohol during her pregnancy. These data also contain a unique id for each mother so we can study siblings.

Figure 1 shows the distribution of birth spacing between first and second children.

Figure 1: Distribution of birth spacing

As expected, birth spacing shows a skewed distribution, left censored at a little less than a year with a long right tail. Given that almost no pairs of adjacent children are more than ten years apart, we drop sibling pairs with spacing above ten years. (We wouldn't have much education data for those two children anyway in that case.) There is also some evidence of heaping at whole number of years. Unlike in other work where intentional heaping on a particular side of an econometrically meaningful discontinuity is a concern (e.g., Barreca, Guldi, Lindo, and Waddell 2011) we are not using any sharp cutoffs around birth spacing and so think any resulting measurement error should be classical.

Figure 2 shows the distribution of time between the last pregnancy loss and the second child's birth.

Figure 2: Distribution of time between pregnancy loss and second child's birth

Like Figure 1, this data is censored at little under a year, is skewed, and has a long right tail. Unlike above though, we see minimal heaping in the time between the last pregnancy loss and the second child's birth. We also see that the majority of the last pregnancy losses occur within two years of second child's birth. To maximize the strength of our pregnancy loss instrument, in our primary specification we will only set the instrument value to 1 if the last pregnancy loss occur within the past two years. This is because we believe that miscarriages that occurred more

than two years before the birth of the second child had a substantially reduced if any impact on the space between the two children.³

We link the birth certificate records to the Colorado Department of Education individual level education statistics data for 1997-2013. These data contain state standardized test scores for, math, reading, writing, and science. The state annually tests the students in the public schools for all four subjects starting in 3rd grade. Science is tested less frequently. For the following analysis, we focus on reading, math and writing scores, which we are able to measure from 3rd through 10th⁴ grade for cohorts born in the early 90s. For more recent cohorts (early 2000s), we can only observe scores for primary school.⁵

Results

Table 1 shows summary statistics for our matched data set of child-test score data.

³ In the appendix we repeat our results counting all last pregnancy losses between the first and second birth for the pregnancy loss instrument. The results are comparable.

⁴ The exams stop after 10th grade.

⁵ Appendix Figure A1 shows the distributions of standardized values for math, reading, and writing test scores. They are reasonably continuous distribution of test scores with minimal censoring. While some of them have moderate skew, we don't expect this will affect our results as we aren't assuming a normal distribution of any of the outcome variables.

Table 1: Summary Statistics

	First Child	Second Child	Difference
Mother Characteristics:			
Teen	0.15	0.02	-0.13***
Mom age 19-25	0.32	0.27	-0.05***
Mom age 25-34	0.48	0.57	0.09***
Hispanic	0.19	0.17	-0.02***
Less than high school	0.15	0.09	-0.06***
High school	0.28	0.30	0.01***
Some College	0.17	0.20	0.03***
Black	0.032	0.031	0.001***
Asian	0.03	0.02	0.00***
Other Race	0.012	0.009	0.003***
Mom Married	0.71	0.86	0.14***
Pregnancy Characteristics:			
Mother smoked	0.089	0.091	0.002***
Weight gain (lbs)	33.16	30.13	-3.03***
	(11.70)	(11.12)	(11.57)
Child Characteristics			
Female	0.49	0.50	0.01***
Birth Year	1996.8	1998	1.2***
	(3.28)	(2.78)	
N	390,607	249,714	
Sibling-pairs			
Pregnancy loss between	0.138		
Pregnancy loss between within past 2 years	0.086		
Mean Years between	3.30		
Median Years between	2.84		
Fraction less than 2 years apart	0.220		

As can be seen from the table, the second child is less likely to be born while the mother is a teen, more likely to be born within a marriage, and of mothers with more education. These differences are likely driven by our inclusion of sibling pairs with up to 10 years between their births. We also have more observations for the first child as those children have more years of test score data in our sample than the younger children.

Table 2 shows our first stage estimates that regress three different birth spacing variables on pregnancy loss.

Table 2. First Stage: Birth Spacing and Pregnancy Loss

	(1) Space (yr)	(2) Space (yr)	(3) ln(space)	(4) ln(space)	(5) gap < 2 yr	(6) gap < 2 yr
Pregnancy loss in past 2 years	0.598*** (0.0233)	0.535*** (0.0218)	0.221*** (0.00605)	0.198*** (0.00568)	-0.193*** (0.00422)	-0.174*** (0.00418)
Demographic controls	N	Y	N	Y	N	Y
Birth Year & Month Fixed Effects	N	Y	N	Y	N	Y
N	55,342	55,342	55,342	55,342	55,342	55,342
R-squared	0.011	0.137	0.017	0.141	0.016	0.088
F stat	659.2	552.5	1335	649.9	2101	1173

Notes: Demographic controls (at the time of the second child's birth) include dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, the mother's pregnancy weight gain, and whether she smoked while pregnant. *** p<0.01, ** p<0.05, * p<0.1

We find that a pregnancy loss in the two years before the birth of the second child is a strongly relevant instrument for birth spacing, with the presence of one increasing the spacing by 0.54 years (F stat = 553), or an increase of about 18% percent using the log of spacing as the dependent variable (F stat=650). Pregnancy loss also decreases the likelihood of having a gap

between children of shorter than 2 years by 17 percentage points (F stat= 1173). The inclusion of our covariates attenuates the effect of the instrument and decreases the size of the F stat slightly, but we still consider this a strong first stage. Convinced of pregnancy loss's relevance as an instrument, we estimate the effect of birth spacing on standardized test scores for reading, writing and math for grades 3-10.

Our main tables present results for two groups: 1) “first children” defined as the first child in two and three child families (birth spacing defined as the space before), and 2) “second children” defined as the second child in two and three child families (birth spacing defined as the space after). Table 3 shows the impact of birth spacing on test scores for first children. Here birth spacing refers to the space between this child and the next child, and the pregnancy loss is during that interval. To put the estimates in context, Dahl and Lochner (2012) find that increasing household income by \$1000 raises test scores by 0.06 SD.

Table 3: Birth Spacing (Instrumented by Pregnancy Loss) and Test Scores: First Children

	(1) OLS Math	(2) IV Math	(3) OLS Reading	(4) IV Reading	(5) OLS Writing	(6) IV Writing
Birth spacing	-0.0120*** (0.00157)	0.00942 (0.0158)	-0.00190 (0.00153)	0.0240 (0.0150)	-0.00751*** (0.00145)	0.0203 (0.0152)
N	371,673	371,673	387,054	387,054	387,040	387,040
R-squared	0.232	0.230	0.226	0.224	0.240	0.238

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** p<0.01, ** p<0.05, * p<0.1

While a naïve OLS specification finds a statistically significant negative effect of birth spacing on math and writing scores (i.e., the more space between the first and second child the lower the first child's test scores), the IV specifications find no statistically significant results.

Table 4 shows the same analysis for the second children. Here birth spacing is defined before this child, and the pregnancy loss occurs then.

Table 4: Birth Spacing (Instrumented by Pregnancy Loss) and Test Scores: Second Children

	(1) OLS Math	(2) IV Math	(3) OLS Reading	(4) IV Reading	(5) OLS Writing	(6) IV Writing
Birth spacing	-0.0381*** (0.00238)	0.0171 (0.0228)	-0.0231*** (0.00233)	0.00466 (0.0221)	-0.0244*** (0.00204)	0.000911 (0.0207)
N	240,922	240,922	247,453	247,453	247,433	247,433
R-squared	0.228	0.222	0.210	0.208	0.234	0.233

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Analogous with above, a naïve OLS specification finds statistically negative significant results of increased birth spacing on test scores, but using pregnancy loss as an instrument for birth spacing makes those estimates statistically insignificant.

We now estimate our education IV models separately for grades less than 8th grade and greater than 8th grade. Table 5 below reports a subset of these models for the first child.

Table 5: Birth Spacing (Pregnancy Loss IV) and Test Scores: First Children, By Division

	(1) GRADE ≤8 Math	(2) GRADE >8 Math	(3) GRADE ≤8 Reading	(4) GRADE >8 Reading	(5) GRADE ≤8 Writing	(6) GRADE >8 Writing
Birth spacing	0.0133 (0.0162)	-0.0132 (0.0207)	0.0274* (0.0155)	0.00195 (0.0181)	0.0222 (0.0143)	-0.000813 (0.0270)
N	308,491	63,182	323,951	63,103	323,907	63,133
R-squared	0.210	0.247	0.204	0.234	0.216	0.264

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

These results are consistent with our main results in table 3 above, finding no statistically significant IV relationship between birth spacing and test scores.

Table 6 repeats the results of Table 4 above, again by less than 8th grade and greater than 8th grade.

Table 6: Birth Spacing (Pregnancy Loss IV) and Test Scores: 2nd Children, By Division

	(1) GRADE ≤8 Math	(2) GRADE >8 Math	(3) GRADE ≤8 Reading	(4) GRADE >8 Reading	(5) GRADE ≤8 Writing	(6) GRADE >8 Writing
Birth spacing	0.0175 (0.0225)	0.0182 (0.0418)	0.00908 (0.0221)	-0.0268 (0.0362)	0.00720 (0.0193)	-0.0460 (0.0508)
N	212,086	28,836	218,613	28,840	218,583	28,850
R-squared	0.202	0.219	0.190	0.209	0.211	0.239

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Here again using the IV specification we find no statistically significant impact of birth spacing on test scores.

Finally, Tables 7 and 8 re-estimate our model for subsets of the data based on the sex of the siblings. Specifically, we re-estimate for the subsets that are the same sex (boy-boy and girl-girl) or are of different sexes (boy-girl and girl-boy).

Table 7: Birth Spacing (Instrumented by Pregnancy Loss) and Test Scores: First Children, By Sex

Panel A: Math								
	(1) BoyBoy OLS	(2) BoyBoy IV	(3) GirlGirl OLS	(4) GirlGirl IV	(5) BoyGirl OLS	(6) BoyGirl IV	(7) GirlBoy OLS	(8) GirlBoy IV
Birth spacing	-0.0143*** (0.00332)	0.0104 (0.0354)	-0.0164*** (0.00311)	-0.0150 (0.0290)	-0.00878*** (0.00315)	0.0779** (0.0355)	-0.00915*** (0.00293)	-0.0312 (0.0276)
N	95,291	95,291	87,023	87,023	94,330	94,330	94,993	94,993
R-squared	0.232	0.230	0.238	0.238	0.229	0.206	0.234	0.232
Panel B: Reading								
	(9) BoyBoy OLS	(10) BoyBoy IV	(11) GirlGirl OLS	(12) GirlGirl IV	(13) BoyGirl OLS	(14) BoyGirl IV	(15) GirlBoy OLS	(16) GirlBoy IV
Birth spacing	-0.00293 (0.00335)	0.0344 (0.0345)	-0.00574* (0.00298)	0.0142 (0.0270)	0.00117 (0.00308)	0.0696** (0.0338)	-0.000187 (0.00280)	-0.0154 (0.0254)
N	99,257	99,257	90,583	90,583	98,281	98,281	98,897	98,897
R-squared	0.206	0.202	0.229	0.227	0.209	0.195	0.227	0.226

Panel C: Writing

	(17) BoyBoy OLS	(18) BoyBoy IV	(19) GirlGirl OLS	(20) GirlGirl IV	(21) BoyGirl OLS	(22) BoyGirl IV	(23) GirlBoy OLS	(24) GirlBoy IV
Birth spacing	-0.00863*** (0.00297)	0.0422 (0.0334)	-0.0118*** (0.00295)	-0.00546 (0.0290)	-0.00448 (0.00286)	0.0714** (0.0332)	-0.00548* (0.00283)	-0.0203 (0.0273)
N	99,247	99,247	90,566	90,566	98,291	98,291	98,900	98,900
R-squared	0.207	0.199	0.229	0.229	0.209	0.190	0.226	0.225

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8, Panel A shows the results on math scores. The IV results are statistically insignificant as above, with the exception of the boy-girl results on math scores in column (6). Here we see a positive impact of birth spacing on test scores, with an extra year of spacing between a boy and girl raising the boy's test scores by 0.08 SD. The p-value here is 2.8% and so given that we are only testing four hypotheses (i.e., boy-boy, girl-girl, boy-girl, and girl-girl) this result is likely not a Type I and more likely a real result.⁶

Panel B shows analogous results for reading scores. Here, we see a very similar effect in column (14), where an extra year of spacing between a boy and girl raises his test scores by 0.07 SD. Panel C in column (22) also has an analogous result for writing scores, where the effect is again approximately 0.07 SD. Given that this effect is consistent across all three test scores, we are reasonably confident that additional birth spacing has a particular positive effect on the older boy when his sibling is a girl.

Table 8 repeats this stratification for the second child.

⁶ This is even the case when also counting the four hypothesis tested for the second child.

Table 8: Birth Spacing (Instrumented by Pregnancy Loss) and Test Scores: Second Children, By Sex

Panel A: Math								
	(1) BoyBoy OLS	(2) BoyBoy IV	(3) GirlGirl OLS	(4) GirlGirl IV	(5) BoyGirl OLS	(6) BoyGirl IV	(7) GirlBoy OLS	(8) GirlBoy IV
Birth spacing	-0.0434*** (0.00490)	-0.00439 (0.0420)	-0.0354*** (0.00477)	-0.0209 (0.0455)	-0.0329*** (0.00441)	0.0765 (0.0465)	-0.0417*** (0.00499)	0.0178 (0.0490)
N	62,188	62,188	55,875	55,875	63,209	63,209	59,601	59,601
R-squared	0.235	0.233	0.237	0.236	0.232	0.207	0.214	0.208
Panel B: Reading								
	(9) BoyBoy OLS	(10) BoyBoy IV	(11) GirlGirl OLS	(12) GirlGirl IV	(13) BoyGirl OLS	(14) BoyGirl IV	(15) GirlBoy OLS	(16) GirlBoy IV
Birth spacing	-0.0278*** (0.00489)	-0.0135 (0.0431)	-0.0207*** (0.00450)	-0.0152 (0.0417)	-0.0196*** (0.00424)	0.0445 (0.0449)	-0.0253*** (0.00497)	0.00394 (0.0472)
N	63,811	63,811	57,410	57,410	64,967	64,967	61,216	61,216
R-squared	0.197	0.197	0.218	0.218	0.207	0.199	0.188	0.186

Panel C: Writing

	(17) BoyBoy OLS	(18) BoyBoy IV	(19) GirlGirl OLS	(20) GirlGirl IV	(21) BoyGirl OLS	(22) BoyGirl IV	(23) GirlBoy OLS	(24) GirlBoy IV
birth_space	-0.0261*** (0.00408)	-0.0239 (0.0370)	-0.0232*** (0.00422)	-0.0200 (0.0422)	-0.0219*** (0.00386)	0.0642 (0.0433)	-0.0266*** (0.00417)	-0.0145 (0.0443)
N	63,837	63,837	57,385	57,385	64,974	64,974	61,188	61,188
R-squared	0.212	0.212	0.224	0.224	0.218	0.203	0.196	0.196

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** p<0.01, ** p<0.05, * p<0.1

Unlike above, none of the IV coefficients in any of the panels of Table 9 are statistically significant.

Robustness Checks

The Appendix contains the result of two robustness checks. First, we apply the extra specifications for the birth spacing variable from Table 2 (the log of spacing and a dummy for spacing less than two years) to the test scores data from Tables 3 and 4. These results are in Appendix Tables A1 and A2 and are consistent with our main results.

Secondly, we repeat the primary analysis in Tables 2-4 using all pregnancy losses between the first and second births as pregnancy losses as opposed to only pregnancy losses that occurred in the two years before the birth of the second child. Appendix Tables A3-A5 contains these results. While our instrument is weaker than before (which is as expected), the IV results are consistent.

Discussion

Our results for the impact of birth spacing on test scores are different in sign from some of the point estimates in Buckles and Munnich (2012), though none of their positive estimates are statistically significant at even the 10% level (in OLS or IV specifications). This is partly to be expected, as our sample size is two orders of magnitude larger than theirs and so our estimates can be substantially more precise. Using a naïve OLS analysis, we see a negative impact of birth spacing on test scores for both first and second children, consistent with the one effect that Buckles and Munnich find that is statically significant at the 5% level (a negative impact of spacing on younger child's math scores). This effect, however, may be due to a selection effect

resulting from families whose children are spaced closer together being systematically different from those whose children are farther apart.

The IV results should correct for any selection effect, as we have limited the sample to the cases where we are more confident that pregnancy loss is exogenous. In doing so, all the main pooled coefficients lose all statistical significance. This suggests that either there is no effect of birth spacing on test scores for either the first or second child, or that the parents compensate for any differences by altering their human capital investments. Another possibility is that while birth spacing has a negative impact on test scores, having a miscarriage has a positive impact, increasing the investment of the parents in their children as the parents have a greater understanding of the challenges in having children. These two effects would therefore cancel each other out, producing the null results seen in our IV analysis above.

We do, however, find statistically significant IV results for one other outcome: the positive impact of spacing on the first child's test scores (across all three measures) when the first child is a boy and the second child is a girl. The stratified analysis by the sex of the children is a new result, and one that is only possible using the larger Colorado matched administrative dataset.

One possible mechanism for this finding is son preference (e.g., Lundberg and Rose 2003; Dahl and Moretti 2008; Abrevaya 2009; Ebenstein and Leung 2010). When a family is exogenously given more time to invest in the first kid due to a pregnancy loss, they do so, since the investment has nowhere else to go. If the second kid is a boy (or both are girls), then they compensate, and we observe no overall effect. But if the second is a girl and the first a boy, then the families appear to not compensate and let the asymmetry in investment persist, hence the result seen above.

Another possible explanation is the psychological effect of having boy-girl siblings that are unexpectedly farther apart in age. Here, the mechanism would not be differential parental investment but more the positive impact of having a younger sister that is farther apart in age and therefore less of a rival as opposed to a younger sister who is closer in age. Perhaps the firstborn son has more time to choose and establish his role in the family and is less threatened by a younger sister who is farther apart in age (Sulloway 1999). Future research could differentiate between these mechanisms using more direct data on pre-schooling parental investment in siblings.

A third explanation is the “confluence theorem” (Zajonc and Markus 1975; Zajonc, Markus, and Markus 1979) which posits that the intellectual development of a child is driven by the average intellect of the other members of the family present. A firstborn child with a longer space before the second child would have more time with an average intellect that is only determined by the parents and not diluted by the presence of a younger sibling. Our results suggest that if this is the mechanism then it has a stronger effect when the first child is a boy and the second child is a girl.

Conclusion

In this paper, we utilize a new data source – linked birth certificate records to schooling records in Colorado-- to examine the relationship between birth spacing and academic achievement. We use pregnancy losses that occur between live sibling births as an instrument for inter-pregnancy intervals as we are concerned that child spacing is endogenous and correlated with unobserved family characteristics that may influence observed health and educational outcomes. Our IV results suggest that while a one-year increase in spacing does not have a statistically overall effect on either the first or second child's test scores, for sibling pairs that are a boy and then a girl it increases the older child's grades 3-10 test scores by 0.07-0.08 SD. To our knowledge, our paper is the first to take advantage of a very large data set to stratify our results by the sexes of the children.

Acknowledgements: We are grateful to participants at the Human Capital and Health Behaviour Conference at Centre for Health Economics at the University of Gothenburg for their help and comments. We would like to thank Kirk Bol at Colorado Department of Health and the Environment for his assistance in linking and accessing our data. We also thank Kelly Thevenet-Morrison for excellent programming support. Special thanks to Sudhir Singh for all of his advice and support.

References

- Abrevaya, J. (2009). Are There Missing Girls in the United States? Evidence from Birth Data. *American Economic Journal: Applied Economics*, 1(2), 1–34
- Almond, D., & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, 25(3), 153.
- Barreca, A.I., Guldi M., Lindo J.M., & Waddell, G.R. (2011). Saving Babies? Revisiting The Effect Of Very Low Birth Weight Classification. *Quarterly Journal of Economics*, 126(4), 2117–2123.
- Black, S.E., Devereux, P.J., & Salvanes, K.G. (2005). The more the merrier? The effect of family size and birth order on children's education. *Quarterly Journal of Economics*, 120(2), 669–700.
- Black, S.E., Devereux, P.J., & Salvanes, K.G. (2007). From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. *Quarterly Journal of Economics*, 122(1), 409–439.
- Buckles, K.S. & Munnich, E.L. (2012). Birth Spacing and Sibling Outcomes. *Journal of Human Resources*, 47(3), 613–642.
- Centers for Disease Control and Prevention. (2008). Abortion Surveillance — United States, 2005. *Surveillance Summaries, Morbidity and Mortality Weekly Report*, 57(SS-13).
- Colorado Department of Public Health and Environment. Center for Health and Environmental Data, Health Statistics and Evaluation Branch. (2005). Colorado Vital Birth Statistics 1990-2005. Retrieved from <http://www.chd.dphe.state.co.us/>.
- Colorado Department of Education. (2013). Colorado Individual-Level Education Statistics 1997-2013. Retrieved from <https://www.cde.state.co.us/cdereval>.
- Conde-Agudelo, A., Rosas-Bermúdez, A., & Kafury-Goeta, A.C. (2006). Birth spacing and risk of adverse perinatal outcomes: a meta-analysis. *JAMA* 295(15), 1809–1823.
- Currie, J. (2011). Inequality at birth: some causes and consequences. *American Economic Review*, 101(3), 1–22.
- Currie, J., & Almond, D. (2011). Human capital development before age five. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, 4(B), (pp. 1315–1486), Amsterdam: North Holland.
- Dahl, G. & Moretti, E. (2008). The Demand for Sons. *Review of Economic Studies*, 75(4), 1085–1120.
- Dahl, G. & Lochner, L. (2012). The Impact of Family Income on Child Achievement:

Evidence from Changes in the Earned Income Tax Credit. *American Economic Review*, 102(5): 1927–56.

Ebenstein, A. & Leung, S. (2010). Son Preference and Access to Social Insurance: Evidence from China's Rural Pension Program. *Population and Development Review*, 36(1), 47–70.

Figlio, D., Guryan J., Karbownik, K., & Roth, J. (2014). The Effects of Poor Neonatal Health on Children's Cognitive Development. *American Economic Review*, 104(12), 3921–55.

Fletcher, J.M., & Wolfe, B.L. (2009). Education and labor market consequences of teenage childbearing evidence using the timing of pregnancy outcomes and community fixed effects. *Journal of Human Resources*, 44(2), 303–325.

Hotz, V.J, Mullin, C., & Sanders, S. (1997). Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analyzing the Effects of Teenage Childbearing. *Review of Economic Studies* 64(4), 575–603.

Hotz, V.J, McElroy, S.W., & Sanders, S.G. (2005). Teen Childbearing and Its Life Cycle Consequences. *Journal of Human Resources*, 45(3), 683–715.

Hotz, V.J. & Pantano, J. (2013). Strategic parenting, birth order, and school performance. *Journal of Population Economics*, 28(4), 1–26.

Karimi, A. (2014). The spacing of births and women's subsequent earnings: Evidence from a natural experiment *Working Paper, IFAU-Institute for Evaluation of Labour Market and Education Policy*, No. 2014:18

Lundberg, S. & Rose, E. (2003). Child Gender and the Transition to Marriage. *Demography*, 40(2), 333–349.

Miller, A.R. (2011). The Effects of Motherhood Timing on Career Path. *Journal of Population Economics* 24(3), 1071–1100.

Rosenzweig, M.R. (1986). Birth Spacing and Sibling Inequality: Asymmetric Information Within the Family. *International Economic Review* 27(1), 55–76.

Rosenzweig, M.R. & Wolpin, K.I. (1988). Heterogeneity, Intrafamily Distribution and Child Health. *Journal of Human Resources* 23(4), 437–61.

Sulloway, F.J. 1999. Birth Order. In Runco, M.A. & Prtizker, S.R. (Eds.) *Encyclopedia of Creativity*, 1, (pp. 189–202), San Diego, CA: Academic Press.

Zajonc, R.B. & Markus, G.B. (1975). Birth order and intellectual development. *Psychological Review* 82(1), 74–88.

Zajonc, R.B., Markus, H., & Markus, G.B. (1979). The birth order puzzle. *Journal of Personality and Social Psychology* 37(8), 1325–1341.

Appendix Figures and Tables:

Figure A1: Distribution of Standardized Values of Test Scores

Panel A: Math

Panel B: Reading

Panel C: Writing

Table A1: Birth Spacing (Instrumented by Pregnancy Loss) and Test Scores: First Children, Extra Measures

Panel A

	(1) OLS Math	(2) IV Math	(3) OLS Math	(4) IV Math	(5) OLS Reading	(6) IV Reading	(7) OLS Reading	(8) IV Reading
Ln (space)	-0.0336*** (0.00572)	0.0270 (0.0453)			0.00588 (0.00559)	0.0690 (0.0429)		
Gap < 2			-0.00466 (0.00699)	-0.0344 (0.0576)			-0.0386*** (0.00673)	-0.0882 (0.0548)
N	371,673	371,673	371,673	371,673	387,054	387,054	387,054	387,054
R-squared	0.231	0.230	0.231	0.231	0.226	0.225	0.226	0.226

Panel B:

	(9) OLS Writing	(10) IV Writing	(11) OLS Writing	(12) IV Writing
Ln (space)	-0.0155*** (0.00533)	0.0584 (0.0438)		
Gap < 2			-0.0193*** (0.00663)	-0.0747 (0.0559)
N	387,040	387,040	387,040	387,040
R-squared	0.240	0.239	0.240	0.239

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** p<0.01, ** p<0.05, * p<0.1

Table A2: Birth Spacing (Instrumented by Pregnancy Loss) and Test Scores: Second Children, Extra Measures

Panel A:

	(1) OLS Math	(2) IV Math	(3) OLS Math	(4) IV Math	(5) OLS Reading	(6) IV Reading	(7) OLS Reading	(8) IV Reading
Ln (space)	-0.103*** (0.00787)	0.0460 (0.0612)			-0.0568*** (0.00766)	0.0125 (0.0593)		
Gap < 2			0.0463*** (0.00810)	-0.0521 (0.0692)			0.0178** (0.00786)	-0.0140 (0.0665)
N	240,922	240,922	240,922	240,922	247,453	247,453	247,453	247,453
R-squared	0.227	0.223	0.225	0.224	0.209	0.208	0.209	0.208

Panel B:

	(9) OLS Writing	(10) IV Writing	(11) OLS Writing	(12) IV Writing
Ln (space)	-0.0663*** (0.00687)	0.00244 (0.0556)		
Gap < 2			0.0328*** (0.00741)	-0.00274 (0.0624)
N	247,433	247,433	247,433	247,433
R-squared	0.233	0.233	0.233	0.233

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** p<0.01, ** p<0.05, * p<0.1

Table A3: First Stage of Birth Spacing on Miscarriage Using All Pregnancy Losses Between Births

	(1) Gap < 2	(2) Gap < 2
Any pregnancy loss	-0.147*** (0.00429)	-0.137*** (0.00413)
Year and month of birth fixed effects	N	Y
Demographic controls	N	Y
N	55,342	55,342
R-squared	0.013	0.087

Notes: Demographic controls (at the time of the second child's birth) include dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, the mother's pregnancy weight gain, and whether she smoked while pregnant. *** p<0.01, ** p<0.05, * p<0.1

Table A4: Birth Spacing and Test Scores: First Children, OLS, and IV Using All Pregnancy Losses Between Births

	(1) OLS Math	(2) IV Math	(3) OLS Reading	(4) IV Reading	(5) OLS Writing	(6) IV Writing
Birth Spacing	-0.0120*** (0.00157)	-0.00296 (0.00967)	-0.00190 (0.00153)	0.0174* (0.00924)	-0.00751*** (0.00145)	0.00713 (0.00919)
N	371,673	371,673	387,054	387,054	387,040	387,040
R-squared	0.232	0.231	0.226	0.225	0.240	0.239

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** p<0.01, ** p<0.05, * p<0.1

Table A5: Birth Spacing and Test Scores: Second Children, OLS, and IV, Using All Pregnancy Losses Between Births

	(1) OLS Math	(2) IV Math	(3) OLS Reading	(4) IV Reading	(5) OLS Writing	(6) IV Writing
Birth Spacing	-0.0381*** (0.00238)	-0.0193 (0.0156)	-0.0231*** (0.00233)	-0.0134 (0.0154)	-0.0244*** (0.00204)	-0.0271* (0.0142)
N	240,922	240,922	247,453	247,453	247,433	247,433
R-squared	0.228	0.227	0.210	0.209	0.234	0.234

Notes: All regressions include fixed effects for the year and month of birth as well as demographic controls, including dummy variables for buckets of maternal age, education, race, ethnicity, marital status, the sex of the child, and ultimate number of children born to that mother. Robust standard errors clustered at the child level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$