Senior Project
Department of Economics

"Birth Order and Education: How Does IQ Play a Part?"

Olivia Smith
May 2011

Advisors: *Francesco Renna*

This paper examines the effect that birth order has on education while accounting for children's genetic endowments. It has been theorized that past studies looking at educational attainment overstated the effect of birth order because they did not account for students' natural ability in their empirical models. Using the Wisconsin Longitudinal Study, this study tests the hypothesis that birth order has no effect on class rank when accounting for IQ. The findings suggest that this is the case. While a one point increase in IQ raises a student's class rank by 1.12 percentage points, birth order has no significant effect. A further analysis of IQ suggests that this variable is accounting for environmental factors also included in birth order, negating the empirical birth order effect.

Table of Contents

According to the United States Census Bureau, a college master's degree is worth $1.3 million more in lifetime earnings than a high school diploma. An educated workforce is essential for the growth and development of the country and for the human capital development of individuals. This leads researchers to question what makes one person likely to receive more education than another. In the past, economists had examined differences between families to explain some of the disparities in educational attainment. It was not until Gary Becker developed his demand curve for children in the 1970s that economists began to look within families to determine what these factors are. Specifically, economists began to look at birth order and family size as an explanation for some of the differences in educational attainment.

Because studies in this area are relatively new, there is still much disagreement on the theoretical impact that birth order has on educational attainment. Economists have argued that this effect is positive (later-born children receive more education than their siblings), U-shaped (first and last-born children receive more education than middle-borns), and negative (first-born children receive the most education). The theoretical model behind the birth order effect posits that the effect found (positive, negative, or U-shaped) is based on the inputs that parents put into the child and the child's genetic endowments. While all of the literature would agree with this, no study thus far has accounted for children's genetic endowments in their empirical model. This study will examine the effect that birth order has on education while taking into account a measure of children's genetic endowments.

This study will also add to the previous research by adjusting the way educational attainment is measured in the empirical model. In the past, educational attainment has been measured in terms of overall years of education received or attendance. While these are both easily quantifiable and very relevant in societies where school attendance is not mandatory, they may not be as pertinent to this study. It is important to know how much education students receive, but in terms of human capital development, the quality of this education may be more relevant. For this reason, this study will use class rank as a measure of academic success. This is assuming that a higher class rank means that a student is gaining more benefit out of the same years of education than students with lower class ranks.

In the following section I will discuss the theoretical framework that underlies most studies on the birth order effect. I will then discuss the different arguments for positive and negative birth order effects as presented in the literature. The empirical model will be developed in the following section, and I will discuss the data used in our study, which was obtained from the Wisconsin Longitudinal Study. After reporting the findings, I hope to determine if the previous models suffered from omitted variable bias by not accounting for intelligence and give a final report on the true birth order effect.

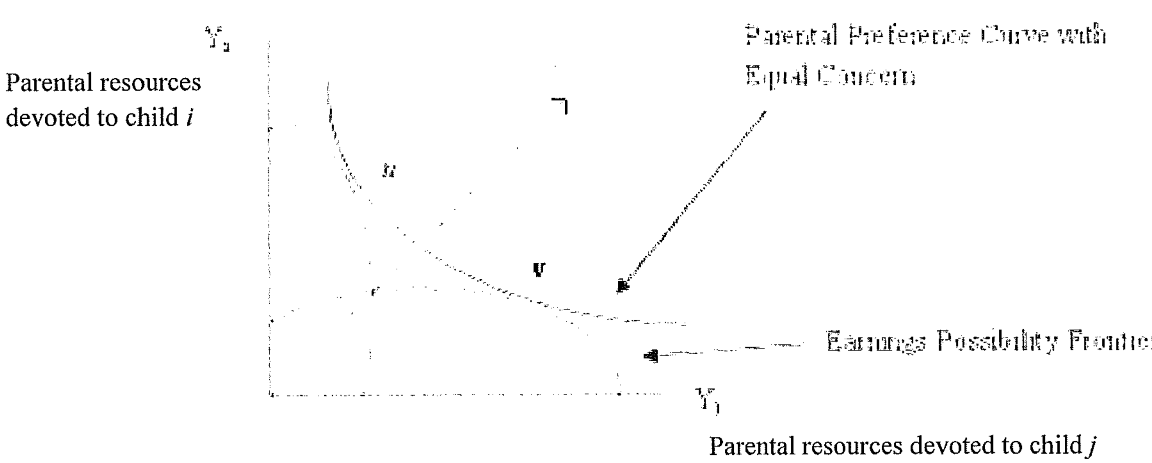
Theoretical Model

Multiple models have been developed to explain the way resources are divided among children, but the most widely accepted is Behrman's intrahousehold allocation model (Behrman 1986). Behrman's model is based on a utility function that parents maximize for their "consumption" of children. Parents maximize their utility function by taking into account the earnings capacity of each child, which is a function of their genetic endowments and human-capital investments. Genetic endowments and human capital investments are subject to diminishing returns, which is illustrated in the concave curve of the earnings possibility frontier in Figure 1. Parents are also faced with monetary, time, and budget constraints that determine their financial investment in children.

Parents are then left with the decision of how to allocate their resources across their children in order to maximize their utility function based on the expected future earnings of each child. The equilibrium of these functions will also be determined by parents' level of concern for their children. Therefore, the birth order effect could be positive, negative, or U-shaped depending on their parental preference curve. Because of this, child *j,* who is favored in Figure 1, could be either the first or last born, depending on the situational birth order effect. Arguments for the effect to be positive or negative will be presented in the following section.

If parents have equal concern for their children, the parental preference curve will be symmetric around a 45-degree ray from the origin, as in Figure 1. Other levels of concern will shift the preference curve and shift the equilibrium for the distribution of resources among children, but the focus of this study will be on the equal concern scenario as this is the most relevant.

Figure 1



Evidence from Literature

As stated earlier, given the intrahousehold allocation model, the birth order effect could be positive, negative, or U-shaped (a combination of the two). Booth and Joo (2009) outlined five reasons why this effect could be positive, meaning that younger children receive more education than older children. Because parents are younger when they have their first children, they will likely have fewer resources than they will later in life. This growth of income over the life cycle is expected to benefit later-born children. Parents also gain more child-raising experience over time, to the benefit of younger children. Cultural factors may also play a part. If the first-born child receives the biggest inheritance, parents may compensate for this by investing more heavily in their younger children's educations. Also, older siblings may leave school early to help provide for the family, putting them at a disadvantage when it comes to receiving education. Finally, younger children may gain time inputs from both their parents and their siblings, which could aid in their development (Booth 2009).

Ejrnaes and Portner (2004) outline reasons for negative birth order effects. Most importantly, parents are faced with time and financial constraints over their life, making it impossible to distribute resources equally among children. Because of this, it is often the first-born that benefits the most from their parents' resources. Also, later-born children will be born into households with lower average IQ's. That is, as more children are added, the average education and IQ of the family will decrease, putting later-born children at a disadvantage to first-borns. Biology may also play a part in creating birth order effects as older mothers have children with higher birth weights. Children with lower birth weights have been found to be more successful and have fewer defects than higher birth weight children, favoring first-borns. Finally, Booth and Joo (2009) also give a reason for negative birth order effects, this being that older children may be more responsible and therefore be higher achievers, leading them to gain more education (Booth 2009).

Methodology

In order to determine the effect that birth order has on education, this study develops an empirical model based on the one used by Booth and Joo (2009). While their data necessitated an ordered probit regression, this study is better suited to a simple ordinary least squares regression because the dependent variable is qualitative. The model is specified as follows:

$$C = \beta X + \alpha N + \gamma_1 D_1 + \gamma_2 D_2 + \omega IQ + \varepsilon$$

Where C = class rank

N = family size

D_1 = first born children dummy

D_2 = last born children dummy

I = IQ measure

X = vector of family demographics

A description of the included variables with their expected signs and descriptive statistics is provided in Table 1.

I will be using this model to test the hypothesis that birth order has no effect on education after accounting for IQ. There are theoretical and empirical reasons why birth order should have no effect. In their theoretical analysis, Ejrnaes and Portner (2004) argue that parents observe the genetic endowments of their last child and then decide whether or not to have another. If the child has lower than expected genetic endowments, the parents will try to have one more. In this way, the last child will have the highest level of natural ability. Because of this, last born children should be highly correlated with IQ, which leads to the omitted variable bias when IQ is not included in the model.

This is assuming that IQ is a measure of natural ability. While it does include genetic endowments to an extent, there are also nature versus nurture concerns regarding IQ. It has been found that acquired skills and natural abilities that are included in IQ cannot be differentiated (Heckman 2007). While IQ is a flawed measure of strictly natural intelligence, it is easily quantifiable, and for this reason I will continue to include it in the model, while being aware of its limitations.

From an empirical standpoint, one can examine the differences in average IQ between first-, last-, and middle-borns to see potential causes of omitted variable bias. I used a t-test to compare group means. Unfortunately, I had three groups to compare, as opposed to the usual two. Because of this, I was only able to examine the differences in means between first born children and all others, and last born children and all others. When examining the means, there is evidence of correlation between first-borns and IQ, potentially leading to omitted variable bias. The mean class rank of first-borns is 2.59 percentage points higher than other children, with a t-value of -6.65 (significant at the 1% confidence level).

The mean IQ of last born children is not found to be statistically significant, meaning that their IQ is no different from other children. The birth order effect cannot be seen in this case because last-borns are being compared to both first-born and middle-born children simultaneously. Because of this, the high average IQ for first-borns and low average IQ for middle-borns cancel each other out. This leaves the last-born dummy variable to have no difference in IQ from other children. Because of the significantly higher IQ of first born children, however, there is cause to suspect that this model may suffer from omitted variable bias if IQ is not included.

It is also important to note that this regression omits family sizes smaller than three. This is to avoid any problems that may arise with only children. Only children may have a stronger effect that would be included in the dummy for first-borns, inflating its importance. Also, eliminating families of two avoids the problem that twins present—they are both first and last born at the same time. Families of two children would also give a full-rank model, having no reference group with which to compare first and last-borns. For these reasons, the sample has been adjusted to include only families of three or more.

Data

All of the data was obtained from the Wisconsin Longitudinal Study (WLS), which includes information on 10,317 seniors who were graduating from Wisconsin high schools in 1957. This study began in 1957 and followed participants throughout their life, conducting surveys again in 1964, 1975, and 1993. The WLS is particularly helpful because of the wide array of variables they have available for download. There is an extensive set of information on each individual, including their birth order, class rank, years of education over their lifetime, and other family characteristics. Most importantly, however, the WLS contains a measurement of IQ, which is normally nearly impossible to obtain.

Because the data is strictly based on Wisconsin high schools in 1957, it is fairly limiting. It is not clear if data from 1957 is as relevant today, but it is the only study to include all of the data needed for this research question. Also, because the data set is based strictly on Wisconsin schools, these findings are relevant for at most the United States, and are certainly not indicative of birth order effects in developing countries. Some of these limits have turned out to be strengths, however. For example, less than 2% of the sample was non-white. While this would be a problem for most studies, it was helpful for this study because I am including IQ. Because the sample includes very few minorities, I did not need to account for the cultural bias of IQ (Zax 2002).

The IQ score reported by the WLS is from the Henmon-Nelson Test of Mental Ability, which respondents took in their junior year. Other variables included in the empirical model are as follows: percentile class rank of the student; dummy variables for first and last born based on the respondent's birth order; total number of children in the respondent's family; a dummy variable equal to one if the respondent is a female; the respondent's age at the time of the study; total years of education of both the respondent's mother and father; the log (to normalize the data) of the respondent's family income in 1957; a dummy variable equal to one if the respondent's mother was working during the student's senior year; and the number of students in the respondent's high school.

For the reasons mentioned earlier, I expect first and last born to have either positive or negative coefficients, but to be statistically insignificant. IQ should have a significant positive effect on class rank because it is controlling for some of the genetic and environmental factors that impact one's ability to succeed academically. Family size is expected to be negative because as the number of children increases, the amount of time and money parents are able to devote to each child's education decreases. Parents' income should have a positive effect on class rank because increasing income increases the amount of resources available to help children succeed. A dummy variable for mothers who were working during the child's senior year should have a negative coefficient, as it is a proxy for parents' time constraints. Females have been shown to have higher scholastic achievement, and therefore the coefficient for the female dummy variable should be positive.

The effect of age is ambiguous. While it may be assumed that all students in the same grade are in the same age group, there is actually some variation in students' ages (they vary from 17-20 years old). Younger students may be either more immature (if their parents sent them early) or higher achievers (if they were able to skip grades). Similarly, older students may be under-achievers (if they were held back) or more mature (if their parents sent them to school late). The number of students in the respondent's high school may also have a positive or negative effect on a student's class rank. Small schools may allow the student to receive more personal attention and increase their ability to succeed, but they may also hamper the student by not being able to take advantage of economies of scale as well as larger schools are able to. In this case, larger high schools may have better access to tools such as labs and libraries.

Father and mother's education should have a positive effect because it is a proxy for parents' value of education. More highly educated parents will be more likely to place a high value on education and therefore more likely to support their children academically. The value that parents place on education is probably one of the environmental characteristics that is included in IQ. Because of this, it is likely that parent's education is also suffering from omitted variable bias because it is correlated with IQ, which can be seen in Table 2. For these reasons, while I expect mother and father's education to have a positive coefficient, I expect it to be statistically insignificant when IQ is included in the empirical model.

Results

The results of the regression are presented in Table 3. To replicate the previous studies, I ran the model twice—once without IQ and once with it. Because the model may suffer from heteroskedasticity—the error term is in unlikely to have constant variance—I use the White test to obtain heteroskedasticity consistent estimates of the errors and t-values. When IQ is not included in the model, the first and last born dummies are both statistically significant at the 10% level. The birth order effect that we obtain is U-shaped: first-borns are ranked 1.78 percentage points higher than middle-children, and last-borns are ranked 2.11 percentage points higher than middle-children.

For the most part, the other variables are significant and have the signs that were expected. Family size, age, working mothers, and high school size all decrease students' class rank, other things constant. Unexpectedly, the log of family income was also found to have a negative effect on class rank, but this may be because the model is misspecified. Education of respondents' fathers and mothers was found to be positive as expected, as well as the dummy for females. This had the greatest effect on class rank in our model: other things equal, females are ranked 14.9 percentage points higher than males. Because this variable is so much larger than the rest, it may be best to estimate the model separately for males and females. It is also important to note that the adjusted R-squared is very low—the model only explains 13.3 percent of the variation in class rank. This is not surprising considering the large size of the data sample, however.

When IQ is included in the model, the results change drastically. The dummy variables for first and last-borns become insignificant. This would imply that the previous model suffered from omitted variable bias—the birth order dummies were picking up some of the effect that IQ has on class rank, causing them to appear statistically significant. Also, the adjusted R-squared increases greatly—from 0.134 to 0.413. This means that adding IQ to the model explains an additional 27.9 percent of the variation in class rank, more than doubling the explanatory power of the previous model. It is obvious from these effects alone that it is necessary to include a measure of intelligence in this model to receive unbiased results.

All of the other variables had signs that were as expected, and most variables were found to be significant. As before, family size, age, working mothers, and high school size have negative effects on class rank. In this model, the log of family income was significant and positive, and while education of respondents' parents was positive, it was insignificant. This was as expected because the effect that parents' education has on class rank is now measured as a part of IQ; parents' education was suffering from omitted variable bias. Also as in the last model, females were found to have much higher class ranks than males—everything else constant, females are ranked 15.0 percentage points higher than males.

Conclusion

In the past, theoretical models have discussed the importance of genetic endowments in determining educational outcomes among children. While it had been discussed in theory, no study had been able to test this empirically. It appears that once these genetic endowments are accounted for, any birth order effect that may have been present disappears. Previous studies suffered from omitted variable bias by not including a measure of children's natural ability, and this lead to the overstatement of birth order as an explanatory variable in education.

That being said, IQ is probably not the best measure of children's genetic endowments. Even this is a hard variable to find, however, so it would be difficult for future studies to find a more accurate, but accessible measure. IQ itself is suffering from omitted variable bias. As Heckman has stated, "measured abilities are the outcome of environmental influences…and also have genetic components" (Heckman 2007). Because the two, genetics and environment, are so intertwined, there is no way to separate their effects, leading IQ to account for more than just genetic endowments.

This means that IQ is overstated in the model. Because birth order is itself a measure of environmental influences, its effect is a part of IQ. Birth order controls for some of the nurture effects of IQ, and because of this, the two likely suffer from reverse causality. The effort that parents put into their children is reflected in the children's IQ, but that effort may also be determined by birth order. The relationship here is impossible to separate, and because of this, IQ must be included in the model. Not only does it include the nurture effects of birth order, but it also accounts for natural ability that has too great of an effect to omit entirely by only including birth order. One can therefore conclude that differences in education within households are not strictly related to birth order, but to the combination of genetic endowments and environmental influences.

An interesting thing to note from these results is the predictive ability of the model. The model presented here explains only approximately 40% of the variation in class rank. The variables included in the model are also only things that are "uncontrollable" by the child and education policy makers. For example, these are factors related to what family the child happens to be born into, the order that they are born in, their gender, etc.—all things that the child presumably has no control over. Future studies should attempt to determine what factors explain the other 60% of variation in class rank. It will likely include variables that cannot easily be measured, such as children's determination, willingness to work, etc. There may, however, be other factors that are influenced by education policy makers. Better understanding of these effects can eventually lead to more effective recommendations for education policy.

Future studies should also attempt to repeat this process and determine the true effect of birth order including more variables to measure intelligence and family background. In this study, it would have been beneficial to estimate the results using a high school fixed effects model to control for the variability in class rank between schools. Unfortunately, this data is not publicly released. Also, because females have such higher class ranks than males, it may be beneficial to run a Chow test to determine if the model would be better-specified by not pooling the data. Future studies would benefit by taking these things into consideration. Overall, however, I am confident in my results that birth order has no effect on educational attainment, specifically in terms of academic success, when genetic endowments are included along with environmental influences.

Bibliography

Behrman, Jere and Paul Taubman. "Birth Order, Schooling, and Earnings." *Journal of Labor Economics* 4.3 (1986): S121-S145. *JSTORE*. Web.

Booth, Alison L., and Kee Hiau Joo. "Birth order matters: the effect of family size and birth order on educational attainment." *Journal of Population Economics* 22.2 (2009): 367-397. *Academic Search Complete*. EBSCO. Web.

Dayioglu, Meltem, Murat G. Kirdar, and Aysit Tansel. "Impact of Sibship Size, Birth Order and Sex Composition on School Enrolment in Urban Turkey." *Oxford Bulletin of Economics and Statistics* 71.3 (2009): 399-426. *EconLit.* Web.

de Haan, Monique. "Birth order, family size and educational attainment." *Economics of Education Review* 29.4 (2010): 576-588. *Academic Search Complete*. EBSCO. Web. 13 Jan. 2011.

Ejrnaes, Mette, and Claus C. Portner. "Birth Order and the Intrahousehold Allocation of Time and Education." *Review of Economics and Statistics* 86.4 (2004): 1008-19. *EconLit.* Web.

Heckman, James J. "The Economics, Technology, and Neuroscience of Human Capability Formation." *Proceedings of the National Academy of Sciences* 104.33 (2007): 13250-13255. Web.

Kantarevic, Jasmin, and Stephane Mechoulan. "Birth Order, Educational Attainment, and Earnings: An Investigation using the PSID." *Journal of Human Resources* 41.4 (2006): 755-77. *EconLit.* Web.

Kessler, Daniel. "Birth Order, Family Size, and Achievement: Family Structure and Wage Determination." *Journal of Labor Economics* 9.4 (1991): 413-426. *JSTORE.* Web.

Zax, Jeffrey S. and Daniel I. Rees. "IQ, Academic Performance, Environment, and Earnings." *The Review of Economics and Statistics* 84.4 (2002): 600-616. *EconLit.* Web.

Appendix

Table 1

Summary Statistics					
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Expected Sign**
Class rank	49.504	28.491	0	99.00	N/A
Family size	5.147	2.420	3.00	27.00	–
First born dummy	0.287	0.452	0	1	+/–
Last born dummy	0.195	0.396	0	1	+/–
Years of father's education	10.094	3.097	7.00	18.00	+
Years of mother's education	10.442	2.900	7.00	18.00	+
Log of family income (in $100s)	4.373	0.575	1.099	7.824	+
High school class size	161.962	130.544	5.00	482.00	–
IQ score	99.986	14.803	61.00	145.00	+
Dummy for working mothers	0.339	0.473	0	1	–
Dummy for females	0.527	0.499	0	1	+/–
Age of the student	18.166	0.510	17.00	20.00	–

Table 2

Pearson Correlation Coefficients						
Variable	IQ score	Father's education	Mother's education	First born	Last born	Family size
IQ score	1.000	0.2392 (<.0001)	0.2184 (<.0001)	0.07904 (<.0001)	-0.0081 (0.4968)	-0.1332 (<.0001)
Father's education	0.2392 (<.0001)	1.000	0.4858 (<.0001)	0.1287 (<.0001)	-0.0885 (<.0001)	-0.1855 (<.0001)
Mother's education	0.2184 (<.0001)	0.4858 (<.0001)	1.000	0.1089 (<.0001)	-0.0940 (<.0001)	-0.1733 (<.0001)
First born	0.0790 (<.0001)	0.1287 (<.0001)	0.1089 (<.0001)	1.000	-0.3125 (<.0001)	-0.1968 (<.0001)
Last born	-0.0081 (0.4968)	-0.0885 (<.0001)	-0.0940 (<.0001)	-0.3125 (<.0001)	1.000	-0.0796 (<.0001)
Family size	-0.1332 (<.0001)	-0.1855 (<.0001)	-0.1733 (<.0001)	-0.1968 (<.0001)	-0.0796 (<.0001)	1.000

Table 3

OLS Estimation, Heteroskedasticity Consistent		
	(1)	(2)
Dummy for first born	1.784 (1.66)*	0.0115 (0.01)
Dummy for last born	2.107 (1.72)*	0.877 (0.87)
Family size	-0.782 (-3.56)***	-0.413 (-2.26)**
Age	-8.978 (-9.37)***	-2.196 (-2.67)***
Years of father's education	0.998 (5.93)***	0.179 (1.24)
Years of mother's education	0.902 (4.97)***	0.176 (1.17)
Log of family income	-3.618 (-4.09)***	2.566 (-3.59)***
Dummy for working mothers	-1.742 (-1.82)*	-1.511 (-1.92)*
High school size	-0.0104 (-2.97)***	-0.0159 (-5.36)***
Dummy for females	14.895 (16.39)***	15.0058 (20.13)***
IQ		1.115 (44.83)***
Constant	209.833 (11.53)***	-16.123 (-0.99)
Observations	3,512	3,512
R^2	0.1362	0.4148
Adjusted R^2	0.1338	0.4129

Note. T-values in parenthesis. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level.