Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2019-7 | February 05, 2019 | http://www.economics-ejournal.org/economics/discussionpapers/2019-7

The only child, birth order and educational outcomes

Yehui Lao and Zhiqiang Dong

Abstract
The one-child policy was implemented in September 1980 and abolished in late 2015. With this change in the demographic policy, the fertility decision of families also changed. Such decisions can result in an increase in the number of siblings in a family. Individuals' educational outcomes may be affected by a change in their parents' fertility decision. The objective of this paper is to provide evidence of the difference of educational outcomes between the only child and the first born. The authors try to estimate the change of educational outcomes when the only child of a family turns to the first born of a family. Moreover, they estimate different channels to interpret these effects. They employ the dataset of China Education Panel data in this paper. In the part of mechanism check, the Sobel-Good test is used for checking the mediation effects of different channels. They found the only child has significant higher educational outcomes comparing to a child who has siblings. Furthermore, the middle child has the lowest educational outcomes of a family. The last born has higher educational outcomes than his or her siblings. To explain these effects, the authors use three channels to interpret: (1) money resource, (2) parenting time, and (3) closeness of parent-child relationships. The policy implication is to help the policymaker estimate and predict the impact of the new demographic policy.

JEL I20 J13
Keywords Onld child; birth orders; educational outcomes

Authors
Yehui Lao, South China Normal University, Guangzhou, China, yehui.lao@foxmail.com
Zhiqiang Dong, South China Normal University, Guangzhou, China

Citation Yehui Lao and Zhiqiang Dong (2019). The only child, birth order and educational outcomes. Economics Discussion Papers, No 2019-7, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2019-7

Received January 3, 2019 **Accepted as Economics Discussion Paper** January 28, 2019
Published February 5, 2019

1. Introduction

The one-child policy was implemented in September 1980 and abolished in late 2015. With this change in the demographic policy, the fertility decision of families also changed. Such decisions can result in an increase in the number of siblings in a family. Individuals' educational outcomes may be affected by a change in their parents' fertility decision. Current research indicates that a tradeoff exists between the quantity of children (Becker 1960), meanings that the quality of children decreases when the number of children in a family increases. Furthermore, the sensitivity of the "quantity-quality" effect may be heterogeneous for children in different birth orders. This study provides evidence of the "sibship effect" on individuals educational outcomes according to their birth order. In particular, we focus on the difference in educational outcomes between only children and firstborn children, which we define as the only-child effect. To achieve this aim, we use the China Education Panel Survey dataset combined with a treatment effect model. Two waves are employed in the investigation of this dataset: the academic year 2013-14 and that of 2014-15. Because the grade 7 students of the academic year 2013-14 are followed in the wave of the academic year 2014-15, we combine this sample as a pool to obtain cross-sectional data. Because schools adopt difference systems; for example, some schools adopt 150 scores as the full mark to assess students outcomes in math, Chinese and English, we translate all outcomes of students into a 100-mark system. Furthermore, fertility decisions are endogenous and affected by parents background, including political status, ethnicity, and their age at having their first child. Therefore, the treatment effect model is employed in this paper. To explain these effects, we use three channels for interpretation, namely (1) financial resources, (2) time spent parenting, and (3) closeness of parent-child relationships. We reveal that the only children are bestowed with significantly greater money resources, more parenting time, and closer parent-child relationships than a child who has siblings. Moreover, the last-born child has a closer relationship with their parents than their siblings do. The Sobel-Goodman test reveals that financial resources,

parenting time, and the closeness of parent-child relationships are mediators of these effects. An only child has superior academic attainment compared with a child with siblings, and educational outcomes among children differ according to the birth order. Difference in financial resources, parenting time, and the closeness of parent-child relationships can be used to interpret these effects.

2. Literature Review

Evidence in the field of sociology reveals that an individual's family structure during childhood affects their educational attainment, occupation, and economic status in adulthood. The number of siblings plays a crucial role in these factors in adulthood (Biblarz and Raftery 1993; Biblarz, Raftery, and Bucur 1997). Much of the literature (McLanahan 1985; Downey 1995; Sandefur and Wells 1999) pays particular attention to the negative effect of sibship on the educational outcomes of adults. Studies suggest that having more siblings dilutes a family's financial resources (Thomson, Hanson and McLanahan 1994; Downey 1995). For individuals, the marginal cost of siblings is a reduction in the number of schooling years they receive, and the marginal cost of each sibling is approximately one-fifth of that of schooling years (Featherman and Hauser 1978; Blake 1981,1989; Heer 1985; Powell and Steelman 1990). For families with more children, parents must allocate their limited material and nonmaterial resources (such as time and energy) to different children. Each child from a large family, compared with each child from families with fewer children, must obtain diluted material and nonmaterial resources so that the number of siblings has a negative effect on resources allocated to them, regardless of education level, occupation, or even intelligence.

Children of different birth orders have different degrees of sensitivity to this effect. Children born earlier than their siblings are generally allocated lower financial resources than siblings born later. This is because they are born in the early stages of their parents' careers, their parents have fewer resources. However, children born later, in the middle of their parents' careers, have access to

more resources than their older siblings due to an increase in family income and the accumulation of wealth (Parish and Willis 1993). If low-income families are subject to credit constraints, older children in the family are likely to drop out of school and begin working early in their lives to provide income and resources for their younger siblings. Lao, Dong and Yang (2018) found that financial resources have a negative effect on individuals' education. Young children born relatively late (when their mother is relatively older) may have a relatively low birth weight (Dammert 2010).

From an evolutionary perspective, both theoretical and empirical studies have shown that parents do not express their feelings toward and invest in their children equally (Daly and Wilson, 1987). Parental favoritism may be correlated with a child's birth order (Kiracofe and Kiracofe, 1990; Kowal and Kramer, 1997). Although parents may attempt to invest in their children equally, the fact that investment in children is heterogeneous due to parental favoritism may affect the perceptions of favoritism (Hertwig, Davis and Sulloway, 2002). These parental perceptions are considered to be catalysts for different processes related to personality development among siblings, affecting their approach to dealing with family, friends, partners, and colleagues (Salmon and Schumann, 2011). Numerous studies have consistently shown that the firstborn child has distinct advantages in terms of education and income in developed countries with different cultures (Black et al. 2005; Kantarevic and Mechoulan 2006; Booth and Kee 2009; de Haan 2010). In developing countries, the evidence is more complicated; some studies have shown that the firstborn child is actually less popular than subsequent children (Ejrnaes and Portner 2004; Edmonds 2006; Emerson and Souza 2008). Regarding the influence of the sibling structure on academic achievement, scholars posit that the effect of birth order on cognitive achievement is mainly influenced by the family's intellectual environment and access to intellectual resources (Zajonc and Markus 1975). As the size of a family increases, its intellectual environment declines. Therefore, a firstborn child is born into a superior intellectual environment than its siblings. However, this effect depends on time gaps between siblings. A large time gap is good

for younger siblings but not for older siblings. By contrast, a small time gap is good for older siblings but not for younger siblings. Older children can provide their young siblings with intellectual resources, which is conducive to the intellectual development of older siblings. Because last-born and only children do not have the opportunity to be a "tutor" to the younger siblings, their academic performance is inferior to those with younger siblings.

The resource dilution hypothesis (Downey 2001) posits that parental resources (such as money and personal concern) are limited and diluted as the number of siblings increases. According to this hypothesis, parents can fully devote themselves to only children or firstborn children. However, the arrival of newborns causes parents to reallocate their resources. Downey (2001) believes that the second child in the family has the most severe effect on resource dilution. Most studies that have investigated the educational effects of the number of siblings in a family indicate that only children have the same academic performance as children in two-child families, or that their academic performance is slightly poor in terms of test scores and years of schooling (Blake 1989). In addition, this hypothesis suggests that the relative abundance of parental resources affects a person's educational attainment. Therefore, only children and firstborn children (who are bestowed with all parental resources a few years before the birth of siblings) academically outperform children born later in families. Downey (2001) argues that different types of parental resources are crucial in the different stages of their child's life. For example, children require the concern of their parents in childhood, savings for college tuition fees while in high school, and their parents' heritage in adulthood. In addition, parental resources may only be part of parents' total family resources; parents may invest resources in activities that are not targeted at children (such as participation in bowling leagues and expenditure related to book clubs or golf courses). This means that the proportion of child's resources from parental investment in relation to household resources is not fixed. Some resources (such as books) can be shared, and there is little or no dilution effect of resources. However, other resources (such as savings for college in the future) can not be shared. Therefore, Powell and

Steelman (1989) believe that certain family resources are more sensitive to the number of children in a family than others. Parental resources are classified as base and surplus parental resources from attributes. Surplus parental resources are not essential for the survival of children; instead, they aimed at improving childrens long-term human capital by, for example, reading with children face-to-face, hiring math tutors, buying computers, providing special learning spaces, and saving money for their college education. By contrast, base parental resources are those that meet a child's general survival needs, such as the provision of adequate food, clothing, and shelter. The sibship effect has different degrees of sensitivity to these two resources. Although few parents question whether their children require basic resources, most attempt to determine the optimal allocation of surplus resources, in part because they are expensive and optional (Downey 2001). A specific threshold can be observed for the size of the child. Before this threshold is reached, parents do not consider the surplus needs of their children and are concerned more with their basic needs (Downey 1995).

3. Variable and Data Description

The data must have two features. First, a background to observations, such as the size of the family, the gender of family members, and parents' backgrounds (e.g., socioeconomic), must be included. Second, education background, such as test scores for each subject, must be included.

For the aforementioned reasons, this paper uses data from the China Education Panel Survey (CEPS). The data were collected by the National Survey Research Center at the Renmin University of China through administering questionnaires to students, parents, homeroom teachers, main subject teachers (but not homeroom teachers), and school administrators. This is a school-based, nationally representative, longitudinal survey of over 20,000 seventh and ninth graders in 438 classrooms of 112 schools in 28 county-level units in mainland China. This survey concerned the 2013-2014 academic year. The contents of the

CEPS include: basic personal and family information, mobility and migration status, personal experiences, cognitive ability, non-cognitive ability, relationship with parents, in-school performance, extracurricular activities, relationship with teachers and peers , family member information, living environment information, health status, and family spending on education. The CEPS also collects students' test scores in each subject, such as Chinese, mathematics, and English. In the 2013-2014 school year, the first round of surveys was conducted, and in the 2014-2015 school year, the original seventh graders were followed; most of the sample students were tracked successfully.

We merged two waves of data (2013-2014 and 2014-2015 academic year). Because the seventh graders were tracked in two waves, this part of the sample was retained. Students with more than six siblings account for approximately 1% of the entire sample; these students were excluded to remove extreme values. Therefore, the number of students in the sample is 8931. Descriptive statistics for variables are shown in Table 1.

Mat, chn, and eng are the original scores of mathematics, Chinese, and English, respectively, which are all translated into a 100-mark system. Schools have different marking systems for each subject, with full marks for the respective subjects being 100, 120, 130, and 150 respectively. Only the data of the 2014-2015 academic year provides the full marks for each subject. Because a school generally does not change the marking system it has adopted, (for example, if a school adopts the 130-mark system, then this system will be adopted for all grades in the school for a long period), full marks in the 2014-2015 academic year were matched to the exam scores of individuals in the 2013-2014 academic year. Exam scores in mathematics, Chinese, and English in the two waves of the survey, which are translated into the 100-mark system by dividing them by full marks of the subject, are used. Table 1 shows that the average scores of students in these three subjects are 65.47, 68.43, and 67.18, respectively, which are approximately at the pass level. The standard deviations are 24.62, 14.37, and 23.07, respectively. Of the subjects, variances in Chinese test scores are the smallest, whereas variances in mathematics scores are the largest. This may be

Table 1: Variable and Data Description

Variable	Observation	Mean	Std	Minimum	Maximum
mat	17340	65.391	24.694	0	100
chn	17334	68.282	14.521	0	98.33334
eng	17340	66.943	23.150	0	100
expense	16038	1030.396	3530.714	0	98618.34
concern_par	17237	2.496	0.549	1	3
self_museum	17958	2.364	1.347	1	6
self_show	18041	2.345	1.448	1	6
relation_fa1	18123	2.566	0.570	1	3
relation_mo1	18117	2.723	0.496	1	3
sibr1	17340	0.256	0.436	0	1
sibr2	17340	0.032	0.177	0	1
sibr3	17340	0.227	0.419	0	1
only_child	17340	0.456	0.498	0	1
steco_5c	17267	2.881	0.604	1	5
birth_age_fa	15585	27.014	5.056	14	65
birth_age_mo	15567	28.835	5.340	14	70
ethnicity_fa	16716	1.405	1.576	1	8
political_fa	16506	2.703	0.705	1	3
ethnicity_mo	16622	1.416	1.578	1	8
political_mo	16448	2.858	0.504	1	3
stsex	17340	0.518	0.500	0	1
stprhedu	17340	4.633	2.029	1	9
birth_year	17808	2000.464	0.701	1996	2002
ethnicity	17251	1.458	1.658	1	8
hukou_place	17474	1.610	0.762	1	4
clsids	17340	228.491	126.901	1	436
schids	17340	59.248	32.686	1	112
time	17298	2013.500	0.500	2013	2014

because the relationship between mathematics and cognitive ability is relatively large, whereas Chinese is a common language.

The expense refers to an individual's expenses for extracurricular activities per semester. Because the survey year spans the 2013-2014 and 2014-2015 academic years, the real expenditure is based on the year 2014, which is identified based on whether the individual was surveyed in the spring semester of 2013-14 or the fall semester of the 2014-15. If an individual's survey time is the 2013-14 fall semester or the 2014-15 spring semester, the expenditure on extracurricular activities is used as the 2013 price and the 2015 price, respectively; thus, the 2014 consumer price index (CPI) and 2015 CPI are used to obtain the actual value of the 2014 base year.[1] The variable concern_par is "how strict your parents are with your homework and exams." It is a dummy variable, with answers being "not strict", "average" and "very strict." The variables self_museum and self_show are respectively "the frequency of visiting museums, zoos, science museums, etc. with the parents" and "the frequency of watching shows with parents", both of which are dummy variables, with answers being "never", "once a year", "every six months", "once a month", "once a week" and "more than once a week".

The variables relation_fa1 and relation_mo1 are relationship with father and relationship with mother, respectively, which measure a child's closeness to their parents. These are dummy variables, with answers being "not close", "average" and "close" respectively.

According to the mean values of only_child, sibr2, and sibr3, only children are the largest proportion of those in the study sample, with middle children comprising the smallest proportion.

The reproductive age, the ethnicity, and political status of parents are instrumental variables related to their fertility decision (i.e., whether the study participant is an only child). The reproductive ages of fathers and mother-

[1] In 2014, the CPI was 1.5%, and the 2015 CPI was 1.6% (source: China Statistics Bureau www.stats.gov.cn).

s range from 14 to 65 years and from 14 to 70 years, respectively. Samples that parents' reproductive age is less than 14 are excluded. The variables "father's ethnicity" and "mother's ethnicity" relate to Han, Mongolia, Manchu, Hui, Tibetan, Zhuang and other ethnic groups. The dummy variables "father's political status" and "mother's political status" relate to Community Party of China, democratic parties and the general public. The variable "parents' highest education level" relate to the following responses: "illiterate", "primary school", "junior high school", "secondary school/technical school", "vocational high school", "high school", "university college", "university undergraduate" and "graduate and above". The year of birth of the participants ranges from 1996 to 2002. The ethnicities of individuals relate to Han, Mongolia, Manchu, Hui, Tibetan, Zhuang, and other ethnic groups. The hukou status at birth includes agricultural hukou, non-agricultural hukou, resident hukou, and others. We use a treatment effect model with a series of personal background factors related to individuals' fathers and mothers as instrumental variables.

4. Empirical Analysis

Columns (1), (4), and (7) in Table 2 report the results using the ordinary least squares (OLS) approach. Considering that a class is taught by the same teacher and that the teaching concept, learning progress, and class climate have the same effect on all individuals in the class, the maximum likelihood estimation (MLE) is adjusted by the class ID clustering standard error. The results of (1) and (7) indicate that the test scores of an only child are higher than that of a child with siblings, with mathematics scores being 2.94 points higher and English scores being 4.09 points higher. The results of (4) show the Chinese scores are not significantly different.

Because of parental self-selection in fertility decision-making, parents with poor cognitive ability are more likely to prefer to have more children, whereas parents with higher cognitive ability are more likely to have fewer children (Grotevant, Scarr, Weinberg 1977). Therefore, the variable only_child is en-

Table 2: The Only-child Effect on the Individual's Educational Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ols1	mle1	ts1	ols2	mle2	ts2	ols3	mle3	ts3
Dep. Var	mat	mat	mat	chn	chn	chn	eng	eng	eng
main									
only_child	2.332***	11.19***	11.14***	0.152	5.220***	5.200***	3.466***	11.66***	11.58***
	(0.416)	(2.572)	(1.654)	(0.228)	(1.541)	(0.910)	(0.369)	(2.303)	(1.457)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
only_child									
birth_age_fa		-0.049***	-0.048***		-0.0470***	-0.048***		-0.049***	-0.048***
		(0.006)	(0.004)		(0.006)	(0.004)		(0.006)	(0.004)
birth_age_mo		0.021***	0.022***		0.019**	0.022***		0.022***	0.022***
		(0.007)	(0.004)		(0.008)	(0.004)		(0.007)	(0.004)
ethnicity_fa		-0.108***	-0.110***		-0.105***	-0.109***		-0.110***	-0.110***
		(0.019)	(0.012)		(0.019)	(0.012)		(0.019)	(0.012)
political_fa		-0.318***	-0.319***		-0.318***	-0.319***		-0.321***	-0.319***
		(0.028)	(0.017)		(0.027)	(0.017)		(0.027)	(0.017)
ethnicity_mo		-0.052***	-0.049***		-0.056***	-0.050***		-0.050***	-0.049***
		(0.015)	(0.012)		(0.016)	(0.012)		(0.015)	(0.012)
political_mo		-0.202***	-0.216***		-0.202***	-0.216***		-0.196***	-0.216***
		(0.039)	(0.024)		(0.039)	(0.024)		(0.039)	(0.024)
_cons		2.274***	2.269***		2.296***	2.268***		2.254***	2.269***
		(0.178)	(0.097)		(0.176)	(0.097)		(0.179)	(0.097)
$H_0 : \rho = 0$		13.89***			13.38***			15.25***	
P value		0.0002			0.0003			0.0001	
Year fixed effect	Y	Y	Y	Y	Y	Y	Y	Y	Y
School fixed effect	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	16096	13972	13972	16092	13969	13969	16096	13972	13972
adj. R2	0.104			0.169			0.207		

Note: $* p < 0.1$, $** p < 0.05$, $*** p < 0.01$. Robust standard errors are shown in parentheses. (1), (4) and (7) use personal ID clustering standard errors, and (2), (5) and (8) use class ID clustering standard errors. Controls includes economic status of a family, parents' highest education level, birth year, ethnicity, hukou status at birth.

dogenous.

The estimation of the treatment effect model involves the maximum likelihood estimation (MLE) and the two-step method. The results for each subject according to these two methods are shown in (2), (3), (5), (6), (8) and (9). MLE estimation is more efficient than the two-step method because the two-step approach causes efficiency losses due to errors of the first step affecting estimations in the second step. Therefore, our analysis is mainly based on MLE results, whereas the results of the two-step method are used to determine the robustness of the model. The results of ρ given in (2), (5), and (8) are all significant at the 1% level, indicating that the dummy variable "whether the individual is an only child" is endogenous. This means that the results of the treatment effect model are reliable. We use the probit model to estimate factors affecting fertility decision. The parameters of each instrumental variable are significant at the 1% level, indicating that each factor has a strong predictive ability to determine the fertility decision of the parents of only child.

The results of the only_child variable with MLE for mathematics, Chinese and English are 11.19, 5.22 and 11.66, respectively, which are all significant at the 1% level. This means that the mathematics, Chinese, and English scores of only children are respectively 11.19, 5.22, and 11.66 points higher than those of children with siblings after all the individual's characteristics are controlled. The MLE results are larger than the OLS results, which means that the OLS method underestimates the difference in academic performance between only children and those with siblings due to the problem of sample self-selection. Because individuals characteristics were controlled, the significant result is caused by the number of siblings an individual has.

Table 3 presents the results of children of different genders obtained using the treatment effect model estimation. MLE is a robust estimation method that adjusts standard errors through class clustering. The results shown in columns (1), (2), and (3) indicate that for females, only children scored higher than those with siblings with mathematics scores being 11.50 points higher, Chinese scores being 4.64 points higher, and the English scores being 10.09

Table 3: The Only-child Effect on the Individual's Educational Outcomes (Subsample of Genders)

	(1)	(2)	(3)	(4)	(5)	(6)
	mle1	mle2	mle3	mle4	mle5	mle6
Subsample:		female			male	
Dep. Variable:	mat	chn	eng	mat	chn	eng
main						
only_child	11.50***	4.636***	10.09***	10.18***	5.482**	12.32***
	(2.775)	(1.614)	(2.423)	(3.951)	(2.225)	(3.764)
Controls	Y	Y	Y	Y	Y	Y
only_child						
birth_age_fa	-0.043***	-0.040***	-0.041***	-0.055***	-0.054***	-0.056***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
birth_age_mo	0.024**	0.021**	0.023**	0.019**	0.017**	0.021**
	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.008)
ethnicity_fa	-0.077***	-0.075***	-0.078***	-0.139***	-0.135***	-0.140***
	(0.026)	(0.027)	(0.027)	(0.025)	(0.026)	(0.025)
political_fa	-0.360***	-0.361***	-0.366***	-0.273***	-0.273***	-0.275***
	(0.036)	(0.036)	(0.036)	(0.035)	(0.035)	(0.034)
ethnicity_mo	-0.085***	-0.086***	-0.084***	-0.025	-0.032	-0.024
	(0.024)	(0.025)	(0.024)	(0.023)	(0.025)	(0.022)
political_mo	-0.240***	-0.243***	-0.233***	-0.171***	-0.167***	-0.169***
	(0.055)	(0.054)	(0.056)	(0.049)	(0.047)	(0.048)
_cons	2.163***	2.194***	2.154***	2.406***	2.414***	2.380***
	(0.233)	(0.232)	(0.234)	(0.233)	(0.229)	(0.236)
$H_0 : \rho = 0$	11.47***	7.61***	8.33***	4.60**	7.44***	6.00**
P value	0.0007	0.0058	0.0039	0.0321	0.0064	0.0143
Year fixed effect	Y	Y	Y	Y	Y	Y
School fixed effect	Y	Y	Y	Y	Y	Y
N	6966	6966	6966	7006	7003	7006

Note: $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$. Class ID clustering standard errors are shown in parentheses. Controls includes economic status of a family, parents' highest education level, birth year, ethnicity, hukou status at birth.

points higher. The results in columns (4), (5), and (6) demonstrate that male only children outperformed males with siblings, with mathematics scores being 10.18 points higher, Chinese scores being 5.48 higher, and English scores being 12.32 points higher. All results are significant, meaning that the only-child effect is significant and robust regardless of gender. The only-child effect in girls affects mathematics scores to a greater extent than does this effect males, whereas the opposite is observed for Chinese and English scores.

The OLS results with controls are reported in columns (1), (3), and (5) of Table 4, and columns (2), (4), and (6) report the MLE results obtained using the treatment effect model. As shown in column (1), the mathematics scores of only children are 2.49 points higher than those of firstborns with siblings. The mathematics scores of middle children are 3.05 points lower than those of firstborns. However, their mathematics scores were not significantly different than those of last-born children. From the results in column (3), middle children scored 2.38 points lower than firstborns. From the results in column (5), the English scores of only children are 3.54 points higher than those of firstborns. The English scores of middle children are 5.22 points lower than those of firstborns. The English scores of last-born children are not significantly different than those of firstborns. All insignificant results mean no difference to the firstborn.

According to columns (2), (4), and (6) in Table 4, only children scored higher in all subjects than children with siblings (mathematics, Chinese, and English scores are 12.6, 5.81, and 12.84 points higher, respectively). Furthermore, the difference in performance between children with siblings and only children is heterogeneous and mainly dependent on birth order. As shown in columns (2), (4), and (6) of Table 4, of all the sampled child types, the difference between middle children and only children is the largest in all subjects; the mathematics, Chinese, and English scores of middle children are significantly lower (by 3.87, 2.10, and 4.87 points, respectively) than those of firstborns). This may confirm the existence of a "middle child syndrome" which is discussed in greater depth subsequently. The last-born child has the smallest gap in all subjects (mathematics, Chinese, and English scores are significantly higher by 2.08,

Table 4: The Only-child Effect and the Birth Order Effect on the Individual's Educational Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	ols1	mle1	ols2	mle2	ols3	mle3
Dep. Variable	mat	mat	chn	chn	eng	eng
main						
only_child	2.496**	12.60***	0.129	5.809***	3.544***	12.84***
	(1.029)	(2.749)	(0.558)	(1.582)	(0.835)	(2.475)
sibr2	-3.049*	-3.776**	-2.379**	-2.102*	-5.223***	-4.814***
	(1.588)	(1.744)	(0.990)	(1.143)	(1.283)	(1.362)
sibr3	0.832	2.080**	0.292	0.994**	0.944	1.950***
	(0.825)	(0.861)	(0.455)	(0.450)	(0.676)	(0.706)
Controls	Y	Y	Y	Y	Y	Y
only_child						
birth_age_fa		-0.049***		-0.047***		-0.049***
		(0.006)		(0.006)		(0.006)
birth_age_mo		0.021***		0.019**		0.022***
		(0.007)		(0.008)		(0.007)
ethnicity_fa		-0.108***		-0.105***		-0.110***
		(0.019)		(0.019)		(0.019)
political_fa		-0.316***		-0.317***		-0.320***
		(0.028)		(0.027)		(0.027)
ethnicity_mo		-0.052***		-0.057***		-0.050***
		(0.015)		(0.016)		(0.015)
political_mo		-0.200***		-0.200***		-0.194***
		(0.040)		(0.039)		(0.039)
_cons		2.281***		2.305***		2.260***
		(0.177)		(0.175)		(0.178)
$H_0 : \rho = 0$		15.18***		14.46***		16. 10***
P value		0.0001		0.0001		0.0001
Year fixed effect	Y	Y	Y	Y	Y	Y
School fixed effect	Y	Y	Y	Y	Y	Y
N	16096	13972	16092	13969	16096	13972
adj. R2	0.104		0.170		0.210	

Note: $* p < 0.1$, $** p < 0.05$, $*** p < 0.01$. Robust standard errors are shown in parentheses. (1), (3) and (5) use personal ID clustering standard errors, and (2), (4) and (6) use class ID clustering standard errors. Controls includes economic status of a family, parents' highest education level, birth year, ethnicity, hukou status at birth.

0.99, and 1.95 points, respectively, than those of firstborns). McGuirk and Pettijohn (2008) stated that last-born children and only children "who have never grown up" and their parents are more concerned about them. It can be assumed that parental concern and input leads to the smallest gap in the scores between only children and last-born children. This is discussed further in the following section.

5. Mechanism Check

5.1. Parental Material Reources for Different Children

The results discussed in the previous section show that the only-child effect and birth order effect significantly influence educational outcomes. The mechanism behind this effect is discussed in this section.

For results presented in Table 5, the expense for extracurricular activities per semester is employed as a proxy variable to estimate the resource allocation of parents to only children and children of different birth orders. To reduce the influence of bias caused by extreme values on the estimation, columns (1), (2), (3) and (4) in Table 5 exclude samples with a maximum of 1% of the dependent variable, column (5) eliminates samples with a maximum of 5% of the dependent variable and column (6) excludes samples with a maximum of 10% of the dependent variable. The results of column (1) estimates using OLS show that extracurricular activity expenses for only children per semester are 386.2 yuan (based on purchasing power in 2014) higher than that of children with siblings. Columns (2) and (3) respectively report the MLE and two-step estimation results, showing that the average expenses for the extracurricular activities of only children per semester are 751 yuan (based on purchasing power in 2014) at a significance level of 1% (MLE) and 890.9 yuan (two-step, based on purchasing power in 2014) more than that of children with siblings.

Columns (4), (5), and (6), compared with column (2), add variables "whether the child is a middle child" (sibr2) and "whether the child is the last-born child" (sibr3). The variable "whether the child is the firstborn child" is an omitted

Table 5: The Only-child Effect and the Birth Order Effect on the Expense for Extracurricular Activities per Semester

	(1)	(2)	(3)	(4)	(5)	(6)
	ols	mle1	ts1	mle2	mle3	mle4
Dep. Variable	expense	expense	expense	expense	expense	expense
main						
only_child	386.2***	751.0***	890.9***	755.0***	524.8***	385.2***
	(29.61)	(88.34)	(119.7)	(92.17)	(57.00)	(41.39)
sibr2				-22.70	-42.23	-21.45
				(55.50)	(32.69)	(26.33)
sibr3				9.056	-11.16	11.10
				(31.29)	(18.76)	(15.00)
Controls	Y	Y	Y	Y	Y	Y
only_child						
birth_age_fa		-0.048***	-0.047***	-0.048***	-0.047***	-0.043***
		(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
birth_age_mo		0.022***	0.021***	0.022***	0.020***	0.017***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ethnicity_fa		-0.109***	-0.110***	-0.109***	-0.108***	-0.110***
		(0.013)	(0.012)	(0.013)	(0.013)	(0.013)
political_fa		-0.325***	-0.326***	-0.325***	-0.330***	-0.329***
		(0.017)	(0.017)	(0.017)	(0.018)	(0.018)
ethnicity_mo		-0.049***	-0.047***	-0.049***	-0.043***	-0.040***
		(0.012)	(0.012)	(0.012)	(0.012)	(0.013)
political_mo		-0.226***	-0.217***	-0.226***	-0.206***	-0.190***
		(0.026)	(0.025)	(0.026)	(0.027)	(0.028)
_cons		2.305***	2.304***	2.305***	2.271***	2.192***
		(0.106)	(0.101)	(0.106)	(0.109)	(0.111)
$H_0 : \rho = 0$		16.16***		16.12***	20.90***	17.08***
P value		0.0001		0.0001	0.000	0.000
Year fixed effect	Y	Y	Y	Y	Y	Y
School fixed effect	Y	Y	Y	Y	Y	Y
N	14944	13144	13144	13144	12460	12023
adj. R2	0.129					

$* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$. Personal ID clustering standard errors are shown in parentheses. In order to reduce the influence of extreme values on the estimation, (1), (2), (3) and (4) exclude the sample with the maximum value of the dependent variable at 1%; (5) exclude the sample with the maximum value of the dependent variable at 5%; (6) Exclude samples with a maximum of 10% of the dependent variable.Controls includes economic status of a family, parents' highest education level, birth year, ethnicity, hukou status at birth.

variable, meaning firstborn children are used as a reference group. All results in columns (4), (5), and (6) are positive at the significant level of 1%. These results reveal that only children are given more material resources that firstborns. However, no significant difference was observed in expenses for extracurricular activities per semester among firstborns, middle children and last-born children.

5.2. Parental Non-material Resources for Different Children

To examine the parental nonmaterial resources, the impacts of the only-child effect and birth order effect on parental concern are analyzed.

To reduce the influence of bias, we exclude children who live with either or neither of their parents. Columns (1) and (2) of Table 6 report parents' strictness toward individuals regarding assignments and exams, columns (3) and (4) report the frequency of visiting museums with parents, and columns (5) and (6) report the frequency of watching lives shows with parents. The correlations between treatment errors and outcome errors (atanh ρ) are significant in columns (2), (4), and (6), meaning that unobservable variables exist in fertility decision-making. Therefore, the variable only_child is endogenous; it is necessary to address this endogeneity by using a treatment effect model. Columns (1), (3), and (5) report the results of the ordered probit model, revealing that only children are more likely to obtain more concern from their parents than firstborns are. For example, the parents of only children may be stricter toward their children in terms of assignments and examinations than parents of firstborns, and they may spend more time with their children than parents of firstborns. Moreover, the last-born effect on parental concern and on the frequency of watching lives shows with parents is not significant, whereas the last-born effect on the frequency of visiting museums with parents is significant at the 1% significant level. Columns (2), (4), and (6) report the results of an ordered probit treatment effect model, in which the absolute value of the only_child's coefficient is larger than the coefficient of the OLS approach. Therefore, the results estimated using the two methods are consistent; only children are given access to more nonmaterial resources than other children.

Table 6: The Only-child Effect and the Birth Order Effect on the Parental Non-material Resource

	(1)	(2)	(3)	(4)	(5)	(6)
	oprobit	treatoprobit	oprobit	treatoprobit	oprobit	treatoprobit
Dep. Variable:	concern_par	concern_par	self_museum	self_museum	self_show	self_show
main						
only_child	0.0817**	0.323***	0.161***	0.548***	0.238***	0.856***
	(0.0330)	(0.109)	(0.0292)	(0.0965)	(0.0304)	(0.0946)
sibr2	-0.113*	-0.0899	-0.118*	-0.0646	-0.127*	-0.0571
	(0.0630)	(0.0736)	(0.0675)	(0.0690)	(0.0679)	(0.0685)
sibr3	-0.0115	0.00528	-0.128***	-0.103***	-0.0518	0.00336
	(0.0360)	(0.0384)	(0.0342)	(0.0348)	(0.0350)	(0.0345)
Controls	Y	Y	Y	Y	Y	Y
cut1						
_cons	182.9***	185.9***	226.0***	245.8***	148.7***	151.6***
	(38.72)	(68.82)	(37.11)	(0.0349)	(37.45)	(0.0412)
cut2						
_cons	184.9***	187.8***	226.7***	246.5***	149.2***	152.1***
	(38.72)	(68.82)	(37.11)	(0.0305)	(37.45)	(0.0351)
cut3						
_cons			227.5***	247.3***	149.9***	152.7***
			(37.11)	(0.0253)	(37.45)	(0.0279)
cut4						
_cons			228.2***	248.0***	150.6***	153.4***
			(37.11)	(0.0174)	(37.45)	(0.0178)
cut5						
_cons			228.5***	248.3	151.0***	153.8
			(37.11)	(.)	(37.45)	(.)
atanh_rho						
_cons		-0.144**		-0.244***		-0.413***
		(0.0664)		(0.0623)		(0.0698)
Year fixed effect	Y	Y	Y	Y	Y	Y
School fixed effect	Y	Y	Y	Y	Y	Y
N	9537	8571	9429	8477	9479	8519
pseudo R2	0.012		0.058		0.065	

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$. Personal ID clustering standard errors are shown in parentheses. In order to reduce the bias, we exclude the sample lives with either or neither of parents. Controls includes economic status of a family, parents' highest education level, birth year, ethnicity, hukou status at birth.

Del Boca, Flinn, and Wiswall (2013) state out that the time parents spend on their children is critical to their development, especially in terms of educational output (Blau and Currie, 2006; Knudsen, Heckman, Cameron, and Shonkoff, 2006). Therefore, this evidence provides support for the supposition that "parents' nonmaterial resources for different children are heterogeneous", which shows that the only-child effect on academic achievement is partly achieved through this channel.

5.3. Closeness of Parent-child Relationships for Different Children

Turning now to the closeness of parent-child relationships for different children.

Table 7 reports the results obtained using an ordered probit model. Columns (1) and (2) demonstrate the closeness of parent-child relationships between the respondent and their father, whereas columns (3) and (4) report the closeness of parent-child relationships between the respondent and their mother. The correlations between the treatment and outcome errors (atanh ρ) are significant in column (2) and (4), which means that unobservable variables are present in fertility decision-making. Therefore, the variable only_child is endogenous and it is necessary to address the endogeneity using a treatment effect model. Columns (1) and (3) report the results obtained using an ordered probit model. These results show that only children are more likely to have a closer parent-child relationship than firstborns are. Columns (2) and (4) report results obtained using an ordered probit treatment effect model. According to these results, the absolute value of the only children is larger at the 1% level and the coefficient of last-born children is significant at the 1% level. Therefore, results obtained by estimation using the two methods are consistent; that is, the difference parent-child relationships between parents and only children and that between parents and lastborn children is positive and significant. This means that both only children and last-born children have closer parent-child relationships than firstborns, whereas no difference is observed between middle children and firstborns in terms of parent-child relationships. These results support the findings of

Table 7: The Only-child Effect and the Birth Order Effect on the Closeness of Parent-child Relationships

	(1)	(2)	(3)	(4)
	oprobit	treatoprobit	oprobit	treatoprobit
Dep. Variable:	relation_fa1	relation_fa1	relation_mo1	relation_mo1
main				
only_child	0.215***	0.452***	0.279***	0.518***
	(0.0332)	(0.117)	(0.0369)	(0.124)
sibr2	-0.102	-0.0539	-0.0304	-0.0397
	(0.0643)	(0.0732)	(0.0713)	(0.0788)
sibr3	0.250***	0.294***	0.283***	0.340***
	(0.0364)	(0.0392)	(0.0407)	(0.0434)
Controls	Y	Y	Y	Y
cut1				
_cons	-82.27**	-91.24*	-10.51	1.377
	(39.79)	(51.27)	(44.87)	(47.17)
cut2				
_cons	-80.67**	-89.63*	-9.013	2.897
	(39.79)	(51.26)	(44.88)	(47.17)
atanh_rho				
_cons		-0.138*		-0.133*
		(0.0726)		(0.0768)
Year fixed effect	Y	Y	Y	Y
School fixed effect	Y	Y	Y	Y
N	9533	8566	9534	8568
pseudo R2	0.016		0.017	

$*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$. Personal ID clustering standard errors are shown in parentheses. In order to reduce the bias, we exclude the sample lives with either or neither of parents. Controls includes economic status of a family, parents' highest education level, birth year, ethnicity, hukou status at birth.

Del Boca, Flinn and Wiswall (2013). Therefore, the evidence supports the hypothesis that "the closeness of parent-child relationships to different children is heterogeneous", which shows that the only child effect and the last-born effect on academic achievement are partly achieved through this channel.

6. Conclusion

This paper examines the only-child effect and birth order effect on educational outcomes. The results show that the academic performance of only children is significantly better than that of children with siblings. Furthermore, in terms of birth order, middle children have the lowest academic performance, followed by firstborns. This is a result of differences in parental material and nonmaterial resources and the closeness of parent-child relationships. The conclusions presented in this paper can guide parenting decisions and human capital investment in children.

References

Biblarz, T.J. and A.E. Raftery. (1993). The Effects of Family Disruption on Social Mobility. American Sociological Review 58:97-109.

Biblarz, T.J, A.E. Raftery, and A. Bucur. (1997). Family Structure and Mobility. Social Forces 75:1319-41.

Blake, J. (1981).Family Size and the Quality of Children. Demography 18:421-42.

Blake, J. (1985). Number of siblings and educational mobility. American Sociological Review, 84-94.

Blake, J. (1989). Family size and achievement (Vol. 3). Univ of California Press.

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. The Quarterly Journal of Economics, 120(2), 669-700.

Blau, D., & Currie, J. (2006). Pre-school, day care, and after-school care: who's minding the kids?. Handbook of the Economics of Education, 2, 1163-1278.

Booth, A. L., & Kee, H. J. (2009). Birth order matters: the effect of family size and birth order on educational attainment. Journal of Population Economics, 22(2), 367-397.

Daly, M., & Wilson, M. (1988, January). The Darwinian psychology of discriminative parental solicitude. In Nebraska symposium on motivation (Vol. 35, pp. 91-144).

Dammert, A. C. (2010). Siblings, child labor, and schooling in Nicaragua and Guatemala. Journal of Population Economics, 23(1), 199-224.

Del Boca, D., Flinn, C., & Wiswall, M. (2013). Household choices and child development. Review of Economic Studies, 81(1), 137-185.

Downey, D.B. (1995). When Bigger Is Not Better: Family Size, Parental Resources, and Childrens Educational Performance. American Sociological Review 60:746-61.

Downey, D. B. (2001). Number of siblings and intellectual development: The resource dilution explanation. American Psychologist, 56(6-7), 497.

Edmonds, E. V. (2006). Understanding sibling differences in child labor. Journal of Population Economics, 19(4), 795-821.

Ejrnæs, M., & Prtner, C. C. (2004). Birth order and the intrahousehold allocation of time and education. Review of Economics and Statistics, 86(4), 1008-1019.

Emerson, P. M., & Souza, A. P. (2008). Birth order, child labor, and school attendance in Brazil. World Development, 36(9), 1647-1664.

Featherman, D. L., Hauser, R. M. (1978). Opportunity and change. InStudies in Population. Academic Press. New York NY United States 1978..

Grotevant, H. D., Scarr, S., Weinberg, R. A. (1977). Intellectual development in family constellations with adopted and natural children: A test of the Zajonc and Markus model. Child Development, 1699-1703.

Heer, D. M. (1985). Effects of sibling number on child outcome. Annual

review of sociology, 27-47.

Hertwig, R., Davis, J. N., & Sulloway, F. J. (2002). Parental investment: how an equity motive can produce inequality. Psychological bulletin, 128(5), 728.

Kantarevic, J., & Mechoulan, S. (2006). Birth order, educational attainment, and earnings an investigation using the PSID. Journal of human resources, 41(4), 755-777.

Kiracofe, N. M., & Kiracofe, H. N. (1990). Child-perceived parental favoritism and birth order. Individual Psychology, 46(1), 74.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building Americas future workforce. Proceedings of the National Academy of Sciences, 103(27), 10155-10162.

Kowal, A., & Kramer, L. (1997). Children's understanding of parental differential treatment. Child Development, 68(1), 113-126.

Lao, Y., Dong, Z., & Yang, X. (2018, June). Scholarship, Admission and Application of a Postgraduate Program. In International Symposium on Distributed Computing and Artificial Intelligence (pp. 57-66). Springer, Cham.

McGuirk, E. M., Pettijohn, I. I., & Terry, F. (2008). Birth Order and Romantic Relationship Styles and Attitudes in College Students. North American Journal of Psychology, 10(1).

McLanahan, S.S. (1985). Family Structure and the Reproduction of Poverty. American Journal of Sociology 90:873-901.

Parish, W. L., Willis, R. J. (1993). Daughters, education, and family budgets Taiwan experiences. Journal of Human Resources, 863-898.

Powell, B., Steelman, L. C. (1990). Beyond sibship size: Sibling density, sex composition, and educational outcomes. Social Forces, 69(1), 181-206.

Salmon, C., & Schumann, K. (2011). The secret power of middle children: How middleborns can harness their unexpected and remarkable abilities. Penguin.

Sandefur, G. D., & Wells, T. (1999). Does family structure really influence

educational attainment?. Social Science Research, 28(4), 331-357.

Thomson, E., T.L. Hanson, and S.S. McLanahan. 1994. Family Structure and Child Well-being: Economic Resources vs. Parental Behaviors. Social Forces 73:221-42. De Haan, M. (2010). Birth order, family size and educational attainment. Economics of Education Review, 29(4), 576-588.

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

Dataset used in the article The Only Child, Birth Order and Educational Outcomes by Yehui Lao and Zhiqiang Dong

Description of dataset:

The empirical analysis of this article has been carried out using microdata from the China Education Panel Survey (CEPS).

The CEPS is a large-scale, nationally representative, longitudinal survey starting with two cohorts the 7th and 9th graders in the 2013-2014 academic year. Documenting educational processes and transitions by which students progress through various educational stages, the CEPS aims at explaining the linkages between individuals' educational outcomes and multiple contexts of families, school processes, communities and social structure, and further studying the effects of educational outcomes during people's life course.

The CEPS applies a stratified, multistage sampling design with probability proportional to size (PPS), randomly selecting a school-based, nationally representative sample of approximately 20,000 students in 438 classrooms of 112 schools in 28 county-level units in mainland China. The baseline survey of CEPS was completed in the 2013-2014 academic year, conducted by National Survey Research Center (NSRC) at Renmin University of China. The data are currently available for academic research. Follow-up surveys are annual as the sample adolescents matriculate throughout the junior-high stage and in the 1st, 3rd, 7th, 8th, 17th and 27th year after they graduate from junior-high. CEPS will last more than 30 years, during which a new cohort of 7th graders will be started in a 10-year interval.

The CEPS administers 5 different questionnaires to the sample students, parents, homeroom teachers, main subject teachers who are not the homeroom teacher, and school administrators.

The student questionnaire includes topics such as students demographic characteristics, mobility and migration status, childhood experience, health status, household structure, parent-child interactions, in-school performance, extra curricular activities, relationship with teachers and peers, social behavior develop-

ment, and expectations for the future.

Parent questionnaire consists of questions about parents' demographic characteristics and lifestyles, parent-child interactions, educational environment and investment for child, community environment, parent-teacher interactions, and parents' perceptions of school education and expectations for the future of the child.

The questionnaire for homeroom and main subject teachers involves questions concerning teachers' demographic characteristics, teaching experience, commentson student behaviors, parent-teacher interactions, comparison between local and non-local students, perceptions of education, and degree of stress and job satisfaction.

The questionnaire for school administrators asks about administrators demographic characteristics, perceptions of education, school's educational facilities, daily management, enrollment of students, statistics of the student body and staff body.

Copyright information:

China Education Panel Survey was designed by National Survey Research Center at Renmin University of China, cooperating with 19 local universities and institutions of China Social Survey Network (CSSN) system. NSRC and CSSN Co-PIs will continue their cooperation on CEPS in the coming years, initiating a new pattern of academic cooperation in social surveys in China. To achieve the permission of dataset, please contact:

Email: ceps@nsrcruc.org

Tel: +86 (10) 62510695

Mailing Address: China Education Panel Survey,

National Survey Research Center,

Renmin University of China, Haidian, Beijing, China, 100872