The More The Better? The Effect of Private Educational Expenditures on Academic Performance: Evidence from Exogenous Variation in Birth Order

Changhui Kang
Department of Economics
National University of Singapore

Preliminary Version: September 2005

Abstract

The question of the causal relationship between educational expenditures and student outcomes continues to attract attention. Empirically establishing the causality, however, is proven difficult due to the endogeneity of educational investment decisions. To shed light on the causality from educational investments to outcomes, this paper examines the effect on standardized test scores of private tutoring which is widely employed by South Korean parents in order to supplement public school education. To deal with the endogeneity of expenditures on tutoring, this paper exploits a student's birth order as a source of identification. Using IV methods, the paper shows that a 10 percent increase in private educational expenditure leads to a 0.67 percentile improvement in test score. The estimated effect of private educational expenditures is comparable to the effect of public school expenditures on earnings estimated by previous studies.

JEL *Classification* : I20, C30

Keywords : Education, Private Tutoring, Test Scores, South Korea

* Address: Department of Economics, National University of Singapore, 1 Arts Link, Singapore 117570, Singapore; E-mail: ecskch@nus.edu.sg, Phone: +65-6874-6830, Fax: +65-6775-2646.

The More The Better? The Effect of Private Educational Expenditures on Academic Performance: Evidence from Exogenous Variation in Birth Order

Abstract

The question of the causal relationship between educational expenditures and student outcomes continues to attract attention. Empirically establishing the causality, however, is proven difficult due to the endogeneity of educational investment decisions. To shed light on the causality from educational investments to outcomes, this paper examines the effect on standardized test scores of private tutoring which is widely employed by South Korean parents in order to supplement public school education. To deal with the endogeneity of expenditures on tutoring, this paper exploits a student's birth order as a source of identification. Using IV methods, the paper shows that a 10 percent increase in private educational expenditure leads to a 0.67 percentile improvement in test score. The estimated effect of private educational expenditures is comparable to the effect of public school expenditures on earnings estimated by previous studies.

JEL *Classification* : I20, C30
Keywords : Education, Private Tutoring, Test Scores, South Korea

1 Introduction

The question of the causal relationship between educational expenditures and student outcomes continues to attract the attention of parents, educational policymakers and researchers alike. Despite decades of intensive study, however, general consensus of the existence (not to mention magnitudes) of the effects of monetary educational investments on student achievement and academic performance has yet to be reached.

On the one hand, papers that summarize the debate on the effectiveness of public school expenditures often advocate conflicting views. For example, while Greenwald, Hedges, and Laine (1996), Card and Krueger (1996) and Krueger (2003) are in favor of the effectiveness of public school expenditures, Hanushek (1986, 1997, 2003) and Betts (1996) cast doubt on the conclusion of these researchers and suggest several factors that can explain discrepancies in conclusion (e.g Betts 1996).

Studies on the effects of private schools (e.g. Catholic schools) do not shed much light on the effectiveness of educational expenditures for student outcomes as well. Evans and Schwab (1995) show that attending a Catholic high school substantially raises the probability of finishing high school or entering a four-year college. Neal (1997) present the effectiveness of Catholic secondary schooling among urban minorities. In contrast, Goldhaber (1996), Figlio and Stone (1999) and Altonji, Elder, and Taber (2005) examine differences in test scores between public and private schools, and find no significant gaps between two types of schools. Importantly, most of the studies on private and Catholic schooling suffer from reliable exogenous variation for identifying the causal effect on a variety of students outcomes. Altonji, Elder, and Taber (2002) argue that two frequently used instrumental variables—religious affiliation of the parents and geographical proximity of Catholic schools—are not a useful source of identification of Catholic school effect.

A major reason why researchers often disagree about the effectiveness of expenditures on public school resources and private school attendance is that empirically establishing the causal relationship between *any* educational expenditures and student outcomes is difficult due to several empirical challenges. For example, a spurious negative association between public school resources and student outcomes may be observed, because weak students can be assigned to remedial classes with higher resources per students, small students and better teachers. When children from wealthier family often attend public or private schools with better school resources, omitting family background variables may produce biased results in an estimation of the effects of educational spending. In addition, if parents spend educational funds in a compensating manner within family—invest more for low-performing children—, we observe a spurious negative correlation between the educational expenditures and student performance. If they spend them in a reinforcing manner—invest more for high-performing siblings—, on the other hand, we observe the reverse.

These obstacles can largely be circumvented by either an experiment in which students are randomly assigned to different resource environment or a reliable exogenous variation of educational investment to identify the causal effect of the expenditures. In the arena of public school expenditures, there exists a large-scale randomization experiment of class size reduction in the U.S. (Tennessee's Student/Teacher Achievement Ratio (STAR) Project). Yet, researchers fail to reach agreement as to how to interpret results from this experiment. Emphasizing the credibility of random assignment implemented in the Project STAR, Krueger (1999) shows substantial advantages in student performance that give rise to students in small classes. In contrast, casting doubt on the validity of randomness that arises from student mobility and class reassignment during the experiment, Hanushek (1999) raises questions on the credibility of the estimated effects of small classes. As for the merits of attending a private school, a most recent view is that frequently used instrumental variables do not seem to be a good source of identifying the causal effect, hence the estimated effects revealed so far are doubtful (Altonji, Elder, and Taber 2002).

In this paper we attempt to shed light on the effectiveness of educational expenditures on students' outcomes by examining the effect of *private* educational investments on standardized test scores. Specifically, we examine the effectiveness on academic performance of the expenditures on private tutoring that is widely employed by South Korean parents in order to supplement public school education. In South Korea, secondary school students are not essentially free to choose their middle and high schools in the school district. Since 1969 student allocation to private as well as public schools has largely been under the government's strict control, especially in urban regions. Under this system (labeled 'Leveling Policy'), students are basically assigned to secondary schools within their residential school district by either a pure lottery or an application-accompanied-by-lottery system under the supervision of the local Department of Education office. Moreover, within schools ability grouping is rarely implemented due to the government's egalitarian policy on secondary education and parents objections. Curricula are also controlled by the Department of Education for the most part. In response to such a rigid public education system, parents in South Korea spend a great deal of money in private tutoring for their children. According to a statistic, South Korean parents spend money on private tutoring 85 percent as much as on public schooling in 1998 (Korean Educational Development Institute, *Survey on Educational Expenditures*, 1998).

Given this situation in South Korea, many including parents as well as educational policy makers are concerned about the effectiveness of private tutoring expenditures on student academic performance. From a broader perspective, an examination on the effect of private tutoring serves to illuminate the debates on the effectiveness of public and/or private educational inputs on students' outcomes.

Given that educational expenditures on a student are not exogenously and randomly determined, there is no doubt that expenditures on private tutoring are correlated with a student's personal, family and academic characteristics. In the absence of a random experiment on private tutoring, a causal estimation therefore calls for an extra variable for a student which strongly affects her parents' decision to invest in her education, but is independent of her various characteristics (academic capability among others) when the amount of educational expenditures is controlled for. For such a variable this paper employs a student's birth order in the family. A large body of literature theoretically and empirically documents that parents favor a certain-parity child (e.g. first-born or last-born) in education. And, as long as a student's birth order is exogenously determined by how many older siblings were born before her, it is unlikely to affect her academic capability.

Using a student's birth order as an IV, this paper shows that a 10 percent increase in private educational expenditure leads to a 0.67 percentile improvement in test score. Evaluated at the mean value, this amount of effect is equivalent to a 1.4 percent increase in test score. Our estimated effect of private educational expenditures on test scores is fairly comparable to the effect of public school expenditures on earnings estimated by previous studies (e.g. Card and Krueger 1996). From an analysis of private educational expenditures, it is revealed that holding other factors constant, birth order significantly influences the amount of educational investment received by a student. Birth order, however, is not significantly associated with a student's academic capability.

The rest of the paper is organized as follows. Section 2 presents various theories and empirical evidence on birth order and educational investment and outcomes. Data and the empirical strategy are discussed in Sections 3 and 4, respectively. Empirical results are shown in Section 5. Section 6 concludes the paper.

2 Birth Order and Educational Investment

2.1 Various Theories

The phenomenon of parental favor for a certain-parity child in educational investments has been documented in a number of studies. First, studies suggest priorities in education for first-born and last-born children. The resource dilution model in sociology argues that parents are faced with time or financial constraints over the life cycle. This makes it impossible to equalize resources over children. Generally, if there is a constraint, first-born and last-born children will benefit because they spend more quality time and resources in smaller families than middle-born children (Birdsall 1991).

Second, other scenarios are possible when the availability of a family's financial and time resources is important to a child. Parents may favor later-borns over earlier-borns in education if their earnings increase over their life cycle (Parish and Willis 1993). And later-born children may benefit from more parental attention, if earliest-born children have moved out of the family home. Provided that parents gain more experience in child care and education with each child, they can invest more efficiently in later-born children. It is also possible for disadvantaged families that a scarce financial resource could cause older children to enter the labor market earlier, thereby increasing the available resources for younger children.

Third, parental preferences may vary by birth order. If parents seek their security in old age, they may favor the oldest sons and devote greater resources to them, as they become economically independent first.[1] If parents are focused on their career when younger, however, they may be less willing to spend time with their children than when they are older and more established professionally.

Besides the birth order effects on educational investments in children, there are also plenty of reasons that a different-parity child may have different academic capability and intelligence that are either endowed or developed over time.

A genetic factor may play a role in forming a child's innate capability. Children of higher birth-order (i.e. younger siblings) naturally have older mothers, and older mothers tend to have children of lower birth weight. This tend to give the oldest children an advantage (Behrman 1988). On the other hand, there are studies which find that both younger and older mothers give birth to children that are less healthy in many dimensions (Strobino, Ensminger, Kim, and Nanya 1995; Royer 2004) . If this is the case, a bad future outcome can take place to both oldest and youngest children.

The intrahousehold allocation of resources in early childhood may also contribute to later differences in a child's capability among siblings, since early nutritional and health status is known to affect children's educational outcomes such as IQ. Especially in developing countries, inequalities are found in the allocation of nutrients and health care among siblings by birth

[1] A cultural factor can explain parental favor for the oldest child, especially oldest sons (Horton 1988). In some societies like South Korea they play important roles in funeral rites and ancestor-worship rituals.

order (Behrman 1988; Horton 1988).

When the intellectual environment is an important determinant of children's development, older siblings might be an asset for younger children, which would confer an advantage on later-borns. As long as children's learning is interactive, however, older siblings can also benefit from teaching their younger siblings (Blake 1989). In addition, the 'confluence theory' (Zajonc 1976) suggests that the average intellectual environment deteriorates when a higher proportion of household members are young children. This causes birth order effects on intelligence that are favorable for younger children and unfavorable for older children.

Independent of the genetic and resource allocation factors, the optimal fertility-stopping rule can give rise to the birth order effect. When parents employ the optimal childbearing rule which states that they make fertility decisions based on the quality of the prior-born child, the youngest children tend to have an extreme quality—either worst or best (Ejrnaes and Portner 2004). If parents stop having children when they have a child with lower than expected genetic endowments, then the last-born child has the lowest quality. On the contrary, if parents finish childbearing when they give birth to a child with a high genetic endowments, then youngest child has the best quality.

2.2 Empirical Evidence

Empirical analyses of the birth-order effects often confound educational investments for a child with the endowed quality of a child.[2] Typically, studies on birth order effects examine educational attainments (such as completed years of education, college attendance, high school graduation, private school attendance and test scores), and sometimes labor market outcomes (such as a full-time employment and earnings), not separating the amount of investment on a child by parents from her own capability. Therefore, empirical evidence on the effects of birth order on parents' educational investment and a child's own quality is limited, because both are not explicitly distinguished. Nevertheless, a trend (albeit arguable) reported by studies is

[2] Empirical analyses of the birth-order effects are also frequently complicated by the existence of the independent sibling-size effects. While theoretical discussions on birth order effects are close to none, economic theories on the sibship size effects have a long tradition and extensive empirical supports in the name of a quantity-quality trade-off (Becker and Tomes 1976). However, recent evidence suggests that the effects of sibship size on a child's educational attainment may proxy for those of birth order. Black, Devereux, and Salvanes (2005) convincingly show that a negative correlation between sibship size and children's education becomes negligible when the birth order is added as an explanatory variable along with the sibling size. See also Conley and Glauber (2005) for a causal analysis of the effects of sibship size and birth order.

that high birth order (e.g. being younger) exerts significantly negative impacts on educational attainment, and little or negligible impacts on a child's performance in schools and the labor market.

Behrman and Taubman (1986) find empirical evidence that lower birth orders or first-borns are favored in schooling, while there exists no statistically significant birth-order effects on earnings for young U.S. adults. Hanushek (1992) finds that while there is no birth order effects among small families with less than 5 siblings, in large families test scores and education are higher among last-born and first-born children than among middle-borns. Black, Devereux, and Salvanes (2005) document that for every sibship size up to 10 first-born children have significantly higher years of completed education than later-born children. They also report that first-born children have greater earnings, higher probability of full time employment and lower probability of teen childbearing than later-borns. Interestingly, they suggest that much of the birth order effects on earnings is likely to work through education.

On the other hand, Olneck and Bills (1979) and Blake (1981) find that there exist a negligible influence of birth order on individuals' educational attainment and test scores. Hauser and Sewell (1985) show no significant or systematic effects of birth order on schooling outcomes (high school graduation, postsecondary educational attainments of those graduates, or educational attainments). Kessler (1991) presents that birth order fails to significantly influence the level or growth rate of wages. Retherford and Sewell (1991) find no support for the Confluence theory (Zajonc 1976) that attempts to explain the negative effect of birth order on intelligence as well as some peculiar effects of birth order on the intelligence of last-born children. Given these preceding mixed and limited results of birth order on educational attainment and other outcome measures, the validity of a child's birth order as an IV for private educational expenditures is a matter to be resolved empirically in the context of our main analysis.

3 Data

3.1 Description of the Main Sample

This study employs the data from the Korean Education and Employment Panel (KEEP). KEEP is a longitudinal study that is conducted since the year 2004 by the Korea Research Institute for Vocational Education and Training (KRIVET)—a government-funded research institute.

The basic structure of KEEP is designed to follow the National Education Longitudinal Studies (NELS) in the U.S. The beginning cohorts of KEEP consist of 6,000 students from three different populations: 2,000 students each from middle schools (grade 9), academic high schools and vocational high schools (both grade 12, the final year of secondary education). Students of each group are sampled by the stratification method to reflect the national population of the group. More specifically, for each group 100 schools are selected in consideration of the regional distribution of schools and students. For each school 4 different classes are randomly chosen, and for each class 5 students are sampled at random. The sampled students are administered a wide variety of students' personal, family and school-related questionnaires. In addition, students' homeroom teachers, school principals and parents are separately surveyed to collect a wide range of background information on the sampled student.

One of the important features of the KEEP data is that the survey collects detailed information on a student's private tutoring experience and expenditures, and the sibling composition.[3] It enables us to construct main explanatory variables and instrumental variables of this study. Also unique in the KEEP data set is the availability of the College Scholastic Ability Test (CSAT) scores for high school graduates. CSAT is the national college-entrance examination of Korea which is annually administered under the supervision of the Department of Education and whose scores are importantly used by colleges and universities to determine the admission of the applicants. Using the resident registration number of a student, the KEEP data are merged with the administrative data base of the 2004 CSAT scores for the test writers. As a measure of a student's academic performance, we employ the CSAT percentile scores of the following three subjects: the Korean language, mathematics, and English.[4] The percentile score of each individual subject ranges from 0 (lowest score) to 100 (highest score). Although vocational high school graduates are eligible for CSAT, the majority of the CAST writers are academic high school graduates. Therefore, we restrict our analysis to the academic high-school sample of

[3]The KEEP survey asks the monthly average amount of expenditures on private tutoring during the last six months before grade 12—roughly nine to fourteen months prior to the test whose scores are used as our main measure of student academic performance.

[4]Guidelines of CSAT stipulate that students are free to choose individual subjects for their examination. Yet a majority of students choose either the Korean language, mathematics, English or all, because they are required by many universities for application. Out of a national total of 574,218 CSAT writers in 2004, 98.9 percent selects the Korean language, 87.8 percent mathematics, 99.3 percent English, 34 percent science and 59.1 percent social studies (Korea Institute of Curriculum and Evaluation (KICE), Media Briefing in December, 2004). In the KEEP data set, 99.3 percent of a total of 1,733 CSAT writers in 2004 choose the Korean language. The corresponding figures are 90.2 and 99.9 percent for math and English, respectively.

2,000 students.

For a refined analysis, the original academic high-school sample has been further narrowed down. First, those students whose guardian, they report, is not one of the parents are excluded from the sample. Patterns of private educational investments and academic performance among these students may be far from typical due to the absence of a parent. A total of 85 students are removed from the original sample as their guardian is either a grandparent, sibling, or close relatives. Students from single-parent families, however, are retained.

Second, students are excluded from the main sample if they either attend a special high school for music, fine arts and athletics, take private tutoring to major in these subjects for higher education in universities, or both. Tutoring costs among them are generally much greater than costs of a normal tutoring of academic subjects. And these students are likely to be poor performers in a general subject test like CSAT.[5] The number of such students are 168 in total. Considering the fact that more than one restrictions may be applied to a single student, the preceding two restrictions leave a total of 1,752 students for futher analysis. Descriptive statistics of the main sample and their difference between first-born students and later-borns are documented in Table 1.

3.2 Descriptive Statistics

Before discussing the descriptive statistics, it is necessary to explain how our measure of student performance are constructed from the raw CSAT percentile scores. Not every student of academic high schools takes CSAT upon graduation. In the KEEP data set, there exist students who do not apply for the 2004 CSAT. Some students are absent for a few or all subjects of the examination upon application. Thus we consider a student to be a non-writer of the 2004 CSAT, and treat her CSAT scores as unknown and missing, when she either do not apply for the 2004 CSAT, or are absent (upon application) for the tests of all three main subjects (the Korean language, math, and English). If a student misses the test of only one or two subjects, on the other hand, her CSAT score for the missed subject is set as zero, not missing. Those

[5] According to the KEEP data, students who attend special high schools for music, fine arts and athletics or take private tutoring of these fields expend a monthly average of 513,000 Korean Won—roughly equivalent to US$430. Other students, on the other hand, spent nearly a half as much on private tutoring—a monthly average of 283,000 Korean Won. In addition, the average CSAT scores are substantially different between these two groups of students. The average percentile scores of the arts and atheletic students are 31.3, and those of non-specialist group are 48.3. The difference is significantly far from zero.

who are, according to the preceding criteria, considered to not take the 2004 CSAT account for a total of 235 students—13.4 percent of students in the main sample. Test-writing behaviors do not seem to vary between first-born and later-born students. Difference in the proportion of CSAT absentees are negligible in Table 1.

It should be noted that a student's test-writing behavior may be closely related with academic performance prior to CSAT. It may also have some bearings with the amount of private tutoring expenditures. Given these possibilities, we later attempt to take those non-writers of CSAT into account in our investigation. Up to such a point, our main sample does not include those who miss all the three tests and have their test scores missing.

Among test-writers, the mean percentile score of the Korean language is 49.3 and the mean scores of math and English are 48.7 and 48.5, respectively. While mean math scores are close between the two groups, mean scores of the Korean language and English among first-borns significantly exceed those of later-born students. The mean of the percentile scores averaged over the three subjects is also significantly greater for first-borns than for later-born students. Yet it is not clear whether these differences between the two groups are endowed by birth order or created by variations in educational investments.

As for the amount of spending on private tutoring, first-born students receive more intensive investments by their parents than later-born counterparts. While the overall average monthly spending on private tutoring is about W285,400—roughly equivalent to US$239, the average spending for first-born students (W324,000) is 35 percent greater than that for later-borns (W239,800). This amount of gap is significantly different from zero. The proportion of those who have received private tutoring for any subject—those with positive monthly spending—is also far higher among first-born students (83.1 percent) than among later-borns (71.4 percent).

When tutoring experiences are focused on the three main subjects, math is the subject for which students most likely take private tutoring (51.8 percent). It is followed by English (41.0 percent) and the Korean language (30.1 percent) in frequency. Weekly tutoring hours also are longest for math (2.5 hours), and English (1.7 hours) and the Korean language (1.3 hours) follow it in terms of duration. Whether in terms of frequency or duration, the preceding statistics suggest that first-born students receive a significantly greater amount of private tutoring than later-borns.

Using the questionnaires for grade-12 homeroom teachers, we create a measure of a student's pre-tutoring performance. This measure is important to our analysis, since it reveals pre-existing differences in student quality before the treatment (i.e private tutoring). The teachers are asked to report a student's approximate rank from 0 (lowest quality) to 100 (highest quality) within a school or a classroom during the second semester of grade 11.[6][7] According to this measure, students are on average rated to have a medium pre-treatment quality (46.1 percentile). Later-borns have a pre-treatment quality slightly higher than first-born students. The difference, however, is indistinguishable from zero. Weekly hours of self-study excluding private tutoring hours are also similar between first-borns (11.6 hours) and later-born students (11.1 hours). The preceding statistics indicate that there seem to exist no substantial differences in pre-treatment quality and self-investment in study between first-born and later-born students. Parents, on the other hand, invest more intensively in first-borns than in later-born children.

A student's age, sex, presence of both parents and school characteristics do not vary substantially between first-born and later-born students. Parents' average age is lower, and their average education level and family monthly income is higher among first-borns than among later-born students. This reflects a tendency that better-educated (hence higher-income) people get married, give birth to a first child later in life and have fewer offsprings (Rosenzweig 1986). In our main sample, first-born students occupy 54.2 percent, middle students 7 percent, and last-born students 38.8 percent. The mean number of siblings in a family is 2.2 and the maximum is 6.

4 Empirical Framework

For our empirical analysis we consider a value-added model of educational production function given by:

$$Y_i = \beta_0 + T_i\beta_1 + \tilde{Y}_i\beta_2 + X_i\beta_3 + u_i \tag{1}$$

[6]Because of the Korean government's traditional leveling policy in secondary education, ability mixing is widely applied in Korean high schools. As a result, a student's rank will not vary substantially, whether an entire school or a single classroom is employed as a reference group for ranking students. We do not attempt to convert the ranks reported by grade-12 homeroom teachers across the school and class levels, since there is no sufficient information to objectively support this conversion in the KEEP data set.

[7]As the KEEP survey collects information on private tutoring during the second semester of grade 11 (as mentioned in the footnote (3), an assumption required in our analysis is that it takes some time for private tutoring to take effects in a student's performance. Otherwise, our measure of a student's pre-treatment quality is error-ridden due to the fact that it contains part of short-term effects of private tutoring.

where Y_i is the percentile test score of student i; T_i is the monthly spending on tutoring (in natural log) for i[8]; $\tilde{Y}_i$ is i's pre-tutoring performance in grade 11[9]; X_i is the vector of i's personal and family backgrounds as well as school characteristics; and u_i is the random error term.

Estimating this model by OLS may fail to produce a causal consistent estimate for β_1, because of the endogeneity problem (i.e. $E(T_i u_i) \neq 0$). If educationally-motivated families spend more on private tutoring and these families are not appropriately measured, the estimate for β_1 is likely to be biased upwards. If parents determine the expenditure on private tutoring according to the pre-tutoring performance of the child, then the estimate can also be biased. For example, if parents tend to spend more on lower performers and less on higher performers, the estimated is biased downwards. When parents spend less on lower performers and more on higher performers, the estimated is biased upwards.

In order to deal with such an endogeneity problem, we estimate (1) by 2SLS. The first-stage model for expenditures on private tutoring is specified as follows:

$$T_i = \gamma_0 + BO_i\gamma_1 + \tilde{Y}_i\gamma_2 + X_i\gamma_3 + \epsilon_i \tag{2}$$

We use two sets of variables for i's birth order BO_i. The first set is a dummy variable that takes 1 if i is the first-born child in a family and 0 otherwise. The second set consists of a dummy taking 1 if i is a middle child and 0 otherwise, and another dummy taking 1 if i is a last-born child and 0 otherwise. The second set of BO_i is designed to formally test the validity—IVs having a significant influence on T_i—and relevance—IVs not having an association with u_i—of birth order as an IV by means of the over-identification test. In the second set of birth order, the reference group is a group of first-born students.

In a quality/quantity model of fertility, the number of siblings is determined simultaneously with the human-capital investment in the children (Becker and Tomes 1976). Consequently,

[8] In order to minimize outlier problems and make the interpretation convenient, the raw amounts of tutoring spending are converted into the natural log metric. And the log-transformed values are multiplied by 10 so as to see the changes in Y_i associated with 10 percent changes in tutoring spending. In order to deal with zero spending in the log transformation, a value of 10 is added to every student's raw value of tutoring spending. The value of 10 is used since it is the smallest unit of monetary amount reported in the survey (10,000 Korean Won) and it is about 3.5 percent of the mean expenditure on private tutoring. When a smaller value (e.g. 1) is added to every expenditure, the results are qualitatively similar.

[9] Five dummy variables are generated for pre-tutoring performance: first (lowest quality) represents pre-tutoring performance between 0 and 25th percentiles, second between 25th and 50th percentiles, third between 50th and 75th percentiles, and fourth (highest quality) between 75th and 100th percentiles. The final variable is an indicator for missing pre-tutoring performance. (About 27 percent of the main-sample students have their pre-tutoring performance missing.) The reference group is the lowest-quality group.

if included in (2), the number of siblings is not likely to be exogenous to ϵ_i. However, failure to control for it may confound birth order effects with sibship size effects in (2). Noting the possibility of bias in its estimates, we include the number of siblings in X_i of (2) as well as (1).

It is well known that when there is heterogeneity in effects, the treatment effect that is identified in our 2SLS analysis is local (LATE). Angrist, Imbens, and Rubin (1996) show that the treatment effects estimated by 2SLS are applicable only to the group or groups whose behavior is influenced by the instruments. In this sense we are estimating the effect of private tutoring for those group of students whose amounts of expenditures on private tutoring are likely to vary according to hypothetically assigned birth order—specifically those whose expenditures will be higher when they are first-born than when they are later-born (*compliers*). We do not identify the treatment effect for a potential group of students whose expenditures on tutoring are likely to be fixed whether or not they are first-born (*never-takers* or *always-takers*). In addition, we do not identify the treatment effect for those whose expenditures on tutoring will be always lower when they are first-born than when they are later-born (*defiers*). To shed light on heterogeneous treatment effects, below we stratify the entire sample by the level of pre-tutoring performance, sex, family income and self-rating on usefulness of private tutoring.

5 Estimation Results

5.1 Results with the First-born Indicator as the IV

5.1.1 Reduced-form Results by OLS

Basic OLS results of the effect of expenditures on private tutoring on student performance are presented in columns (1)-(3) of Table 2. As expected, higher pre-tutoring performances lead to higher current test scores. A single fatherhood (relative to the presence of both parents) and parents' average education level are significantly related to the higher test score of a student. In contrast, a single motherhood, the number of books at home, family income and parents' average age are not marginally associated with test scores. While a student's age and sex have a strong connection, being an only child and the number of siblings fail to have a significant bearing with test scores.

In column (3) being a first-born child is not strongly related with higher test scores, even if there is a positive connection. Although examining the significance of the effect of being first-born in an OLS framework is not a formal test for the validity of birth order as an IV—the consistency of the estimated effect of being first-born depends crucially upon exogeneity of expenditures on private tutoring—, this finding is suggestive of the possibility that birth order can be exogenous to u_i in (1). The results of a formal test (the over-identification test) will be presented below.

As regards the expenditures on tutoring, whether or not pre-tutoring performances are controlled for, the association between tutoring expenditures and percentile test scores (averaged over the three main subjects) is positive and similar in magnitude. Although it is statistically significantly different from none, the estimated association is quite small. A 10 percent greater monthly expenditures on private tutoring is related to nothing but a 0.1 percentile higher test score. In contrast, weekly hours of self-study, which do not include tutoring hours, are strongly associated with test scores. One-hour longer weekly hours of self-study are related to an approximate 0.35-0.52 percentile higher test score.[10] The one-hour increase in hours of self-study (equivalent to a 8.8 percent longer hour than the mean) is roughly 4 to 5 times more strongly associated with higher test scores than a 10 percent higher expenditure on tutoring.

As explained previously, the association between tutoring expenditures and test scores estimated by OLS may not be consistent and causal. Depending the correlation between T_i and u_i, the estimate may be biased upwards or downwards. We present causal estimates of β_1 when we discuss the 2SLS estimation results later.

5.1.2 First-stage Results of 2SLS

The results of the first-stage regression of tutoring expenditures are presented in columns (4) and (5) of Table 2. Whether or not pre-tutoring performances are controlled for, being a first-born child significantly and positively affects parents' expenditures on private tutoring. First-born students receive about 31 to 33 percent more expenditures on tutoring than later-born students. (Recall that a log of a monthly tutoring expenditure is multiplied by 10.) This implies potential validity of being first-born as an IV for expenditures on tutoring for a student. Provided that being first-born has no direct association with test scores, birth order can serve as a legitimate

[10]The effect of hours of self-study on the test score may not be causal, because both of them are likely to be simultaneously determined. For an expositional convenience, however, we often compare the effect of tutoring expenditures with that of hours of self-study.

instrument for spending on tutoring.

Tutoring expenditures show an inverted-U shape relationship with a student's pre-tutoring performance. While parents spend significantly more on students with medium-low prior performance (the 25th and 50th percentiles) than those with the lowest prior performance, the expenditure start to fall as the level of prior performance rises above the 50th percentile. If a linear term of prior performance is used rather than dummies, the relationship is significantly negative—negative 5.115 (s.e. 1.026). This inverse relationship between pre-tutoring performance and expenditures on tutoring suggests that the OLS estimates of β_1 in (1) are more likely to be biased downwards than upwards.

Other variables, which significantly affect expenditures on a student's private tutoring, include the number of siblings and the variables reflecting a family's economic strength such as family income and parents' average education level. The negative relationship between the sibship size and tutoring expenditures is consistent with a quality/quantity tradeoff in fertility. In contrast, hours of self-study, single-parenthood, the number of books at home, parents' age and a student's age, sex and only-child status do not have a strong association with the amount of tutoring expenditures.

5.1.3 Second-stage Results of 2SLS

The causal estimates of the effect of tutoring expenditures on test scores are shown in columns (6) and (7) of Table 2. When pre-tutoring performances are not controlled for, a 10 percent increase in the expenditure on private tutoring enhances a student's performance by 0.83 percentile points. Under the value-added specification, the estimate indicates a 0.67 percentile improvement in test score due to a 10 percent increase in expenditure. Evaluated at the mean percentile score, the estimates of β_1 imply a 1.4-1.7 percent increase in test score due to a 10 percent increase in expenditure on private tutoring. At its face value, a 10 percent increase in tutoring expenditure is roughly equivalent to a 1.6-1.9 more weekly hours of self-study in enhancing a student's test score.

Our 2SLS estimates are 6 to 7 times greater than the OLS estimates. This implies that the OLS estimates are severely biased downwards. As discussed previously, such a bias arises probably because parents of low-performing students tend to spend more on private tutoring, other things (especially family income) being equal.

Although evidence suggests a downward bias in the OLS estimates, the magnitude of improvement in test score estimated by 2SLS does not seem to be substantial. As far as one may argue the reverse, it would be interesting to compare our estimates of the effect of private educational expenditures with corresponding estimates of previous studies, precision of our estimates set apart. Unfortunately, existing literature on the effect of private educational expenditures on test scores is quite scarce. Thus, we rely on estimates of the effect of public school expenditures on student outcomes in order to gain some perspective of our estimates.

In the analysis of a randomization experiment on class size (Project STAR), Krueger (1999, Table VII) presents that a one student decrease in class size in grades K to 4 leads to a 0.67-0.88 percentile increase in test score. Evaluated at the mean values of 21 students per class and 51 percentile test score (Appendix Table), these estimates imply a 2.8-3.6 percent improvement in test score due to a 10 percent decrease in class size and the same amount of increase in per-pupil expenditure.[11] Compared with Krueger's (1999) estimates, our estimated effect of expenditures on private tutoring amount to a half.[12]

In terms of earnings in the labor market, Card and Krueger (1996, p.37) summarize that a 10 percent increase in school spending leads to about a 1-2 percent increase in subsequent earnings. For example, they report that a reduced-form re-analysis of Card and Krueger (1992) presents a 1.1 percent increase in weekly earnings associated with a 10 percent reduction in the average pupil-teacher ratio. On the other hand, other researchers find slightly weaker effects on earnings. Betts (1995) suggests that a 10 percent reduction in the average teacher-pupil ratio leads to a 0.4 percent increase in earnings. Grogger (1996) shows a 10 percent increase in mean spending per student leads to a 0.7 percent increase in wages. Our estimated effect of expenditures on private tutoring seems to be fairly comparable to the estimated effects of school expenditures on earnings, although ours are on a slightly higher side.

Nonetheless, it is necessary to note that our 2SLS estimates are not precisely estimated and statistically indistinguishable from zero effect. Therefore, we cannot completely rule out a possibility that private expenditures on tutoring may have no causal influences on student academic performance.

[11] Krueger (2003, F55-F56) infers that one percent decrease in class size will be approximately converted into the same amount of percentage increase in annual per pupil cost.

[12] In contrast, Hanushek (1986, 1997) summarizes that public school expenditures produce little if any measurable improvement in students' academic performance in school. In addition, Hanushek (1999) raises questions on the credibility of the effects of small classes reported in Krueger (1999).

5.2 Results with the Middle and Last-born Indicators as the IV

A merit of having detailed information on siblings in the KEEP data is that we can examine the effect of being born at different parities on the expenditures on private tutoring. Although such an analysis might be of its own interest, we can use this merit in our estimation of the effect of tutoring on student performance by exploiting the fact that the first-born instrument can be converted into two instruments. In particular, two separate indicators, middle-born and last-born, can be generated as potential instruments.[13] These two new instruments lead us to an over-identified model in which it is possible to more formally test the validity of birth order as an IV.

The results of the first and second-stage regressions with the two instruments are shown Table 3. There is essentially little change in the first-stage estimates when the middle and last-born indicators are employed as IVs. As shown previously, tutoring expenditures show an inverted-U shape relationship with a student's pre-tutoring performance. The number of siblings, the family income and parents' average education level strongly affect the expenditures on private tutoring of a student. Hours of self-study, a single-parenthood, the number of books at home, parents' age and a student's age, sex and only-child status, however, do not have a strong association.

The first-stage results also show that a student's birth order significantly affects parents' educational expenditures on her. Relative to first-born students, last-born students receive significantly less expenditures on private tutoring. The amount of spending for last-borns is about 33-35 percent lower than for first-borns. Middle students also receive smaller amount of educational expenditures than first-borns. The amount of spending on tutoring for middle students is about 17-18 percent lower than for first-borns. The difference in spending between middle and first-born students, however, is statistically indistinguishable from zero. Nevertheless, the middle and last-born indicators are jointly significant. From column (2) the F-test statistic for the joint significance of these two variables is 7.65, and the p-value is less than 0.001. This implies that the two IVs are jointly strong enough to serve as valid IVs.[14]

[13]Theories suggest that middle-born children may receive either more or less investments in eduction than first-born and last-born. See Section 2.

[14]Bound, Jaeger, and Baker (1995) show that the F statistic on the excluded instruments in the first-stage estimation of IV contains valuable information about the magnitude of the finite-sample bias. They warn that F statistics close to 1 should be cause for concern.

When the first-born indicator is replaced with the middle and last-born indicators, the causal estimates of the expenditures on private tutoring on a student's performance slightly fall. A 10 percent increase in the expenditure on private tutoring enhances a student's average test score by 0.64-0.77 percentile points. Still, these amounts of the effect are 6-7 times greater than those estimated by OLS. The magnitude of improvement in test score estimated by 2SLS seems to be modest compared with other studies on the effectiveness of school resources. Again, since the 2SLS estimates are not precisely estimated, the true effect of expenditures on private tutoring is likely to none.

The over-identification test confirms that the middle and last-born indicators are valid instruments. Whether or not pre-tutoring performances are controlled for, the p-value of the test is far above 0.1 in columns (3) and (4) of Table 3. That is, the two variables of birth order are uncorrelated with the error term and correctly excluded from the main equation (1).

5.3 Effects of Tutoring Hours on Subject Test Scores

In the preceding sections we examine the effect of expenditures on private tutoring on a student's test score averaged over three subjects (math, the Korean language and English). In fact, the tutoring expenditures used in previous regressions are total spending on tutoring that is paid by parents for unspecified kinds and subjects of private tutoring for a month on average. Consequently, if a student receives private tutoring for other subjects than math, the Korean language and English, the expenditures on private tutoring are contaminated by measurement errors. If such errors are widespread, the effect of tutoring expenditures on performance estimated by OLS is biased toward zero effect, although the 2SLS estimates will be relatively less severely affected.

In this section, instead of using undifferentiated total expenditures on private tutoring, we employ weekly average tutoring hours and percentile test scores for each particular subject. The OLS and 2SLS estimation results for each subject are documented in Table 4. The value-added specifications with controls for pre-tutoring performance are employed for OLS and 2SLS.

There are five numbered columns in the table. The OLS estimates are reported in column (5) as a basis. And we implement two versions of 2SLS. The first version uses the first-born indicator as an IV. Its first-stage and second-stage results are presented in columns (1) and (3), respectively. The second version employs the middle and last-born indicators as IVs. Columns

(2) and (4) show the first-stage and second-stage results, respectively. F statistic on the excluded instruments in the first-stage estimation and p-value of the over-identification test are also reported in columns (2) and (4), respectively.

As shown previously, the OLS estimates indicate quite a small association between tutoring hours and test scores for all three subjects. A one-hour longer tutoring is related to only a 0.01-0.2 percentile higher subject test score. When we examine the causal effect of tutoring for each subject by 2SLS, however, some different trends emerge for different subjects.

First, math tutoring has little effect on enhancing the math test score. The 2SLS estimates show a negative impact of math tutoring on the test score, although they are statistically indistinguishable from zero effect. Second, tutoring for English is found efficient on improving the English test score. The 2SLS estimates imply that a one-hour increase in weekly tutoring hours causes a 0.6 percentile point increase in the English test score. This amount of effect is fairly precisely estimated. To put the estimate in perspective, a one hour of weekly tutoring is equivalent to 1.4 hours of self-study for English. Third, tutoring for the Korean language also seems to be effective in improving the test score, while caution is needed in interpretation. The 2SLS estimates suggest a 1.1-1.2 percentile point increase in the test score due to a one-hour increase in weekly tutoring hours. A one-hour increase in weekly tutoring hours is comparable to 4.7-5.3 hours of self-study. These estimates, however, are imprecisely estimated. As a result, we cannot rule out a possibility that tutoring for the Korean language has no effect on the test score.

For all three subjects, a student's birth order serves as a valid IV. First-born students receive significantly greater investments in private tutoring than later-born students. When the middle and last-born indicators are used, F statistics on excluded IVs in the first-stage estimation and p-values of the over-identification test confirm the validity of these variables as IVs.

5.4 Heterogeneity in the Effect of Private Tutoring

As mentioned previously, when there is heterogeneity in effects, the treatment effect that is identified in our analysis is a local treatment effect. That is, we are basically estimating the effect of private tutoring for those group of students whose amounts of expenditures on private tutoring are likely to be affected by birth order. Although it is not possible to a priori distinguish between those group of students whose private tutoring is influenced by birth order and those not, it would be illuminating to see if there exists heterogeneity in effect along measured dimensions of student and family characteristics. In order to address heterogeneity in the treatment effect, below we stratify the main sample by four measured dimensions: the level of pre-tutoring performance, sex, family income and self-rating on the usefulness of private tutoring. In empirical analysis we employ the total monthly expenditures on private tutoring for unspecified subjects, and the percentile test scores averaged over math, the Korean language and English.

5.4.1 Level of Pre-tutoring Performance

According to the pre-tutoring performance reported by grade 12 homeroom teachers, we split the main sample into three sub-samples: bottom-third, middle-third and top-third samples. The OLS and 2SLS estimation results for each sub-sample are documented in Table 5. Here and in subsequent tables, we follow the format of Table 4. The OLS estimates are reported in column (5). The first and second-stage results of 2SLS which uses the first-born indicator as an IV are shown in columns (1) and (3), respectively. Those of 2SLS which employs the middle and last-born indicators as IVs are shown in columns (2) and (4).

For each level of pre-tutoring performance, the OLS estimates imply a small association between expenditures on tutoring and test scores. A 10 percent larger expenditure on tutoring is related to only a 0.2 percentile points higher test score. The size of the association seems to vary little across different levels of pre-tutoring performance.

For students whose pre-tutoring performances are in the bottom third (panel A), birth order significantly affects tutoring expenditures. First-born students receive about 51 percent more expenditures on tutoring than later-born students. From column (2), the amounts of tutoring spending for middle and last-born students are about 68 and 49 percent smaller than for first-born students. And the 2SLS estimates suggest a modest increase in test score due to the increase in expenditures on tutoring. A 10 percent more spending increases a student's test score by a 0.62-0.69 percentile points. These amounts are equivalent to 2.8-3.2 weekly hours of self-study. Yet we do not rule out a possibility that tutoring for low ability students has no effect on test outcome, since the 2SLS estimates are not precisely estimated.

For other students whose pre-tutoring performances are in the middle (panel B) and top thirds (panel C), birth order does not seem to significantly affect the tutoring expenditures.[15] Birth order is much less effective in explaining the variation of tutoring expenditures for students with middle-level pre-tutoring performances than for those with high-level pre-tutoring performance. For the high-level students, the significance levels of the first-born indicator in column (1)—t-value is 1.90—and the last-born indicator in column (3)—t-value is 1.94—are on the borderline.

Although the validity of IVs is more or less questionable, the 2SLS estimates suggest a modest increase in test score due to the increase in expenditures on tutoring. For the middle-third students, a 10 percent more spending on tutoring increases a student's test score by a 0.54-1.30 percentile points. For the top-third students, a 10 percent more spending enhances the test score by a 0.48-0.54 percentile points. Likewise, since the 2SLS estimates are not precisely estimated, it is possible that expenditures on tutoring has no effect on test outcome.

5.4.2 Sex

The OLS and 2SLS estimation results for each sex are documented in Table 6. For both male and female students, birth order significantly affects tutoring expenditures. First-born boys and girls receive about 31 and 28 percent more expenditures on tutoring than later-born siblings, respectively. The amounts of tutoring spending for middle and last-born students are smaller than for first-born students. The two indicators are jointly strong enough to serve as valid IVs.

The 2SLS estimates show a stronger positive effect of private tutoring for boys than for girls. While a 10 percent more spending on tutoring enhances a female student's test score only by 0.25-0.28 percentile points, the same amount of spending increases a male student's score by 0.69-0.91 percentile points. Again, the 2SLS estimates are imprecisely estimated. Hence we do not rule out a possibility that expenditures on private tutoring has no influences on test scores.

[15] It is not clear whether non-influence of birth order on tutoring expenditures for the middle-third and top-third students implies the 2SLS estimates in Tables 2 and 3 are applicable only to the group of bottom-third students. Since the first-stage regressions are estimated by OLS, the significance of the estimates not only depends on the true relationship between birth order and tutoring expenditures, but the sample size. The sample size becomes smaller when the entire sample is stratified according to pre-tutoring performance.

5.4.3 Family Income

According to the total family income reported in the KEEP data, we split the main sample into three sub-samples: low-income, mid-income and high-income family samples.[16] The low-income families are those whose total monthly income is less than 2 million Won (roughly equivalent to US$1,675)—the first (lowest) quartile of the sample distribution of family income. The mid-income families are those whose total monthly income is between 2 million Won and 4 million Won (US$3,350)—the third quartile of the sample distribution of family income. The high-income families are those whose monthly income is greater than or equal to 4 million Won. The OLS and 2SLS estimation results for each family income group are presented in Table 7.

Birth order significantly affects tutoring expenditures only for students from mid-income families. For this group, first-born students receive about 48 percent more expenditures on tutoring than later-born students. The amount of tutoring spending for middle and last-born students is 34 and 50 percent smaller than for first-born students, respectively. For students from low-income and high-income families, first-born children receive more expenditures on tutoring than later-born siblings. The estimates, however, are not significantly different from zero.

The estimates for mid-income students imply that a 10 percent increase in expenditures on tutoring leads to a 0.43 percentile point increase in test score. This size of improvement seems modest. As the estimates are imprecisely estimated, it is not ruled out that the effect of expenditures on private tutoring may be none. Although the validity of IVs is questionable, expenditures on tutoring for low and mid-income students seem to slightly increase test scores. A 10 percent more spending on tutoring enhances a low-income and high-income student's test score by a 0.56 and 0.62-0.78 percentile points, respectively. Again, these estimates are not precisely estimated. Overall, the causal effect of expenditures on private tutoring does not vary substantially according to family income of a student.

5.4.4 Self-rating on the Usefulness of Private Tutoring

Here the main sample is divided according to whether a student personally regard tutoring as useful. According to an opinion questionnaire in the KEEP data, about 61 percent of students in

[16]Total family income includes average monthly labor incomes of family members residing in the household and their financial incomes.

our main sample believe that private tutoring is either very useful or useful in a five-point scale. When the focus is narrowed on those who have experienced private tutoring, this proportion is about 66 percent. Given the opinion of students, it would be illuminating to examine whether the effect of private tutoring in fact is greater among those who rate it to be useful than among those not. The OLS and 2SLS estimation results for each group are presented in Table 8.

Whether or not students consider private tutoring to be useful for their performance, birth order significantly affects tutoring expenditures. First-born students receive about 26-30 percent more expenditures on private tutoring than later-born students. The amount of tutoring spending for middle and last-born students is smaller than for first-born students. The exception is the expenditures on tutoring for middle children in the group of students who rate private tutoring not useful. Their amount of expenditures on tutoring is not significantly different from those for first-born children. Yet F statistic in column (2) and p-value of the over-identification test in column (4) suggest that the middle and last-born indicators are jointly strong enough to serve as valid IVs.

Contrary to an expectation that the effect of private tutoring is stronger for those rating it useful than for those rating it not useful, the estimated causal effect is greater for the former group. While a 10 percent more spending on tutoring enhances test scores only by 0.33 percentile points for those students who rate private tutoring to be useful, the same amount of spending increases test scores by 1.1-1.3 percentile points for those rating it not useful. The 2SLS estimates, however, are imprecisely estimated. Hence it is not unlikely that expenditures on private tutoring has no influences on test scores for both groups of students.

In sum, the examination of heterogeneity in the treatment effect along measurable dimensions of student characteristics reveals no considerable differences in the effect of tutoring expenditures. While there exist differences in estimated effect between boy and girl students, and between students rating private tutoring useful and not useful among other dimensions, the size of the difference is within the range of statistical errors. Although the imprecision of 2SLS estimates limits an interpretation, the 2SLS estimates reported in Tables 2 and 3 do not seem to represent a very narrow group of students.

5.5 Robustness: Missing Test Scores Included

5.5.1 Decision to Take the Test

According to Table 1, about 13.4 percent of students in the main sample do not take CSAT and hence have no test scores. If test-writing decisions of students are systematically related with their either observable or unobservable characteristics, then our 2SLS estimates will be biased due to a sample-selection problem. To address such a possibility, we estimate a linear probability model for a student's decision to take CSAT, using the same specification as in (1) and (2) except that Y_i is replaced by an indicator for no test-writing. The OLS and 2SLS estimation results are shown in Table 9.

The OLS estimate in column (5) suggests that expenditures on private tutoring are positively associated with test-writing decisions. The size of the association, however, is not significantly different from zero. The 2SLS estimates in columns (3) and (4) also suggest no strong relationship between expenditures on private tutoring and test-writing behavior. The estimation results suggest that missing test scores of some students which may take place in association with expenditures on private tutoring is not likely to cause a serious bias in our 2SLS estimates.

On the other hand, a student's hours of self-study and pre-tutoring performance significantly affect the decision to take CSAT. Both of the variables are strongly and positively related with test-writing behavior. These two variables may not be a concern for the consistency of 2SLS estimates when they are uncorrelated with expenditures on private tutoring. However, it is hard to believe they are so. As long as students who have lower pre-tutoring performance and shorter hours of self-study, hence require greater expenditures on private tutoring by their parents systematically choose not to take the test, the strong correlation between a student's hours of self-study, pre-tutoring performance and test-writing decision may lead to a bias in our 2SLS estimates. We address such a potential bias next.

5.5.2 Considerations for Missing Test Scores

To deal with a potential sample selection due to systematic test-writing behaviors, we replace missing test scores with those that are randomly generated in reference to a student's pre-tutoring performance.[17] Specifically, we rely on the following method to fill missing test scores.

[17] The Heckman's method for correcting a sample-selection problem is not well applied in our context due to the absence of an appropriate variable that explains the selection process, while not affecting one's test score.

When a student's test score is missing while her pre-tutoring performance is available in the data, we assign a hypothetical test score (Y_i) which is randomly generated by the formula given by:

$$Y_i = \tilde{Y}_i + 5 \times \upsilon_i \quad (3)$$

where $\tilde{Y}_i$ is i's observed percentile value of the pre-tutoring performance in grade 11, and υ_i is a random number from $N(0, 1)$. In order to give a sufficient range of randomness in test score, a factor of 5 is multiplied to υ_i. Next, if both a student's test score and her pre-tutoring performance are not available in the data, we employ the following assignment process:

$$Y_i = \tau_i \quad (4)$$

where τ_i is a random number from a uniform distribution between 0 and 10. The reason for using the range between 0 and 10 is because Table 9 suggests that those not taking the test are likely to be low-performers than medium- or high-performers. By keeping simulated test scores at a low level, we attempt to take precautions against a potential downward bias of our estimates of the effect of private tutoring.[18] According to the preceding procedures, a total of 235 missing test score are substituted for. Their average value and standard deviation are 38.5 and 33.9, respectively.

The OLS and 2SLS estimation results with missing test scores plugged in are shown in Table 10. For convenience of comparison, we report the results for the total sample as well as the three sub-samples stratified by pre-tutoring performance.

Similar to our main results in Tables 2, 3 and 5, birth order serves as a legitimate IV for the total sample and the sub-sample of students with low pre-tutoring performances. For students with medium and high pre-tutoring performances, however, its role as a valid IV is more or less questionable. The first-born indicator in column (1) and the middle and last-born indicators in column (2) have a borderline significance.

If missing test scores are replaced by the preceding procedures, the causal effect of private tutoring expenditures on performance is more precisely estimated. For the whole sample, a 10 percent increase in expenditure on private tutoring enhances a student's performance by 0.95-

[18] Here, we use a factor of 5 in (3) and 10 in (4) for generating a hypothetical test score for a student whose test score is unknown. When we employ other values than 5 and 10, however, there are no substantial qualitative differences in the results. The results under different values are available upon request.

0.98 percentile points. These amounts of the effect are significantly different from zero effect. Yet the effect does not seem considerable. Evaluated at the mean percentile test score, a 10 percent increase in expenditure leads to a 2.0-2.1 percent increase in student performance. Although this is on the large side for the estimated effect of public school expenditures summarized by Card and Krueger (1996), it is modest compared with Krueger's (1999) estimated effect of class size reduction.[19]

If the whole sample is divided by the pre-tutoring performance of a students, the results are basically similar to those in Table 5. Although the 2SLS estimates are not precisely estimated, a 10 percent more spending on tutoring generates the improvement in student performance by 0.6, 0.5 and 1.6 percentile points for bottom-third, middle-third and top-third students, respectively. Again, these amount of the effects are relatively modest.

6 Concluding Remarks

In order to shed light on the debates on the effectiveness of educational investments on student outcomes, this paper examines the causal effect of private educational expenditures (expenditures on private tutoring in South Korea) on student standardized test scores. Given that educational expenditures on a student are not exogenously and randomly determined, the paper exploits that parents favor a certain-parity child (e.g. first-born) in education, while her academic capability can be little affected by birth order in the family.

From the analysis of expenditures on private tutoring, it is revealed that holding other factors constant, birth order significantly influences the amount of educational investment received by a student. First-born students are invested by parents significantly greater amount of money than later-born students. When birth order is sub-divided by first-born, middle and last-born indicators, last-born students receive significantly lower amount of expenditures on private tutoring than first-born students. In contrast, expenditures on private tutoring received by middle students do not seem to be significantly different from those for first-born students.

OLS estimates of the relationship between expenditures on private tutoring and student test

[19]Our estimated effect of expenditures on tutoring under simulation is dwarfed by a meta-analysis of Hedges, Laine, and Greenwald (1994) which yields an estimate that a 10 percent increase in public school expenditure produces an improvement in student performance of approximately 0.7 standard deviations. This amount is equivalent to 16.2 percentile point improvement in test score in our metric. The summary of Hedges, Laine, and Greenwald (1994) are, however, criticized by Hanushek (1997) for being biased in favor of large positive effects of school expenditures.

scores are likely to biased downwards due to a potential negative correlation between student pre-tutoring capability and expenditures on private tutoring. The negative bias seems to exist in OLS estimates. Although they are statistically significantly different from none, the OLS estimates suggest that a 10 percent greater monthly expenditure on private tutoring is related to nothing but a 0.1 percentile higher test score. The causal estimates of the relationship estimated by 2SLS confirm a downward bias in OLS estimates. The 2SLS estimates imply a 0.67 percentile improvement in test score due to a 10 percent increase in expenditure on private tutoring. Evaluated at the mean value, this amount of effect is modest and equivalent to a 1.4 percent increase in test score. Our estimated effect is fairly comparable to the effects of public school expenditures on earnings estimated by previous studies. Nonetheless, as our 2SLS estimates are not precisely estimated, a possibility cannot be completely ruled out that private expenditures on tutoring may have no causal influences on student academic performance.

By the local nature of 2SLS estimates, the effect of private tutoring identified by our 2SLS method is drawn from the group of students whose amounts of expenditures on private tutoring are likely to be affected by birth order. Generally, we cannot a priori tell which students belong to such a group. To illuminate heterogeneity in treatment effects, we stratify the entire sample by the level of pre-tutoring performance, sex, family income and self-rating on usefulness of private tutoring. Examination on the sub-samples reveals no considerable differences in the effect of tutoring expenditures. While there exist differences in estimated effect between boy and girl students, and between students rating private tutoring useful and not useful among other dimensions, the size of the differences is within the range of statistical error.

We also examine a potential sample selection that may arise from systematic test-writing behaviors of students. When missing test scores are replaced by simulated scores, the 2SLS estimates slightly rise above those obtained from the actual sample, and the size seems to be still modest. A 10 percent increase in expenditures on private tutoring enhances a student's performance by 0.95-0.98 percentile points. Again, since these estimates are imprecisely estimated, it is not unlikely that expenditures on private tutoring do not affect student academic performance.

Table 1: Descriptive Statistics of the Main Sample

	Total Sample			(1) First-borns		(2) Later-borns		Differences ((1)-(2))		
Variable	N	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.E.	T-value
Take test (No=1)	1752	0.134	0.341	0.136	0.343	0.132	0.339	0.004	0.016	0.24
Average score of three tests	1517	48.50	23.08	50.26	22.65	46.42	23.43	3.838	1.185	3.24
Test score of Korean	1504	49.28	26.14	51.46	25.80	46.73	26.33	4.731	1.348	3.51
Test score of math	1428	48.66	26.53	48.91	26.48	48.36	26.61	0.550	1.411	0.39
Test score of English	1516	48.52	26.77	50.91	26.48	45.73	26.85	5.181	1.373	3.77
Spending on Tutoring ($W1,000$)	1749	285.4	341.1	324.0	375.0	239.8	2.899	84.17	16.25	5.18
Any tutoring (Yes=1)	1749	0.778	0.416	0.831	0.375	0.714	0.452	0.117	0.020	5.90
Tutoring hours for Korean	1752	1.295	2.504	1.439	2.600	1.125	2.377	0.315	0.120	2.63
Tutoring for Korean (Yes=1)	1752	0.301	0.459	0.336	0.473	0.259	0.438	0.077	0.022	3.52
Tutoring hours for math	1752	2.464	3.136	2.739	3.248	2.139	2.967	0.599	0.150	4.00
Tutoring for math (Yes=1)	1752	0.518	0.500	0.574	0.495	0.452	0.498	0.122	0.024	5.14
Tutoring hours for English	1752	1.732	2.627	1.950	2.518	1.474	2.730	0.476	0.125	3.79
Tutoring for English (Yes=1)	1752	0.410	0.492	0.472	0.499	0.337	0.473	0.135	0.023	5.76
Prior performance	1285	46.13	26.68	45.10	26.17	47.36	27.25	2.262	1.494	1.51
Hours of self-study	1752	11.34	10.22	11.58	10.33	11.05	10.09	0.526	0.490	1.07
Age	1752	17.74	0.517	17.72	0.521	17.75	0.511	-0.028	0.025	-1.14
Male (Yes=1)	1752	0.580	0.494	0.577	0.494	0.584	0.493	-0.007	0.024	-0.28
Only child (Yes=1)	1752	0.075	0.263	0.138	0.345	0.000	0.000	0.138	0.012	11.33
Number of Siblings	1752	2.192	0.663	1.998	0.539	2.422	0.721	-0.424	0.030	-14.07
Parents' average age	1751	46.06	3.178	44.90	2.809	47.43	3.046	-2.522	0.140	-18.01
Parents' average education	1748	12.10	2.538	12.46	2.473	11.67	2.550	0.787	0.120	6.54
Both parents present (Yes=1)	1752	0.922	0.269	0.920	0.272	0.924	0.265	-0.004	0.013	-0.32
Single father (Yes=1)	1752	0.027	0.162	0.031	0.172	0.022	0.148	0.008	0.008	1.05
Single mother (Yes=1)	1752	0.051	0.221	0.050	0.217	0.054	0.225	-0.004	0.011	-0.38
Books at home	1752	184.1	211.3	195.5	216.8	170.6	203.9	-24.85	10.12	-2.46
Family income ($W1,000$)	1730	3,188	1,967	3,318	2,008	3,035	1,908	283.7	94.71	2.99
School characteristics										
Private school (Yes=1)	1752	0.568	0.495	0.576	0.494	0.559	0.497	0.017	0.024	0.73
Coed school (Yes=1)	1752	0.464	0.499	0.467	0.499	0.461	0.499	0.006	0.024	0.25
Boys-only school (Yes=1)	1752	0.307	0.461	0.312	0.464	0.301	0.459	0.011	0.022	0.48
Girls-only school (Yes=1)	1752	0.229	0.420	0.221	0.415	0.238	0.426	-0.017	0.020	-0.82
Rural area (Yes=1)	1752	0.193	0.395	0.177	0.382	0.213	0.410	-0.036	0.019	-1.90
Medium city (Yes=1)	1752	0.330	0.470	0.329	0.470	0.331	0.471	-0.002	0.023	-0.11
Metropolitan city (Yes=1)	1752	0.477	0.500	0.494	0.500	0.456	0.498	0.038	0.024	1.60
First-born child (Yes=1)	1752	0.542	0.498							
Middle child (Yes=1)	1752	0.070	0.255							
Last-born child (Yes=1)	1752	0.388	0.487							

Table 2: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Performance: First-born Child Indicator as an IV

	Reduced form models (OLS)					Structural models (2SLS)	
Dependent variable:	Average test score			Spending on tutoring (log)		Average test score	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tutoring spending	0.111 (0.041)**	0.101 (0.037)**	0.096 (0.037)**			0.829 (0.426)	0.665 (0.395)
First-born child			1.769 (1.128)	3.279 (0.823)**	3.105 (0.817)**		
Hours of self-study	0.546 (0.053)**	0.364 (0.049)**	0.363 (0.049)**	0.032 (0.035)	0.019 (0.035)	0.521 (0.061)**	0.352 (0.054)**
Prior quality (Q2)		13.363 (1.590)**	13.165 (1.600)**		2.966 (1.301)*		11.476 (2.208)**
Prior quality (Q3)		21.397 (1.596)**	21.202 (1.600)**		1.525 (1.279)		20.332 (1.869)**
Prior quality (Q4)		33.926 (1.680)**	33.806 (1.685)**		1.231 (1.318)		33.105 (1.923)**
Prior quality missing		18.312 (1.620)**	18.155 (1.625)**		-2.831 (1.251)*		19.769 (2.041)**
Single father	-5.804 (3.218)	-6.455 (2.424)**	-6.714 (2.425)**	-1.444 (2.214)	-1.850 (2.198)	-5.091 (3.789)	-5.660 (2.939)
Single mother	2.704 (2.662)	2.033 (2.421)	2.166 (2.424)	-2.915 (1.728)	-2.799 (1.712)	5.010 (3.109)	3.761 (2.742)
Books at home	0.004 (0.003)	0.004 (0.003)	0.004 (0.003)	0.001 (0.002)	0.001 (0.002)	0.003 (0.003)	0.003 (0.003)
Family income	0.004 (0.003)	0.002 (0.003)	0.002 (0.003)	0.021 (0.002)**	0.021 (0.002)**	-0.012 (0.010)	-0.010 (0.009)
Parents' avg edu	1.410 (0.263)**	1.134 (0.239)**	1.111 (0.239)**	0.900 (0.168)**	0.890 (0.167)**	0.724 (0.498)	0.604 (0.451)
Parents' avg age	0.059 (0.174)	0.033 (0.154)	0.141 (0.169)	0.189 (0.126)	0.152 (0.125)	0.067 (0.189)	0.054 (0.167)
Age	-2.797 (1.037)**	-2.778 (0.922)**	-2.830 (0.925)**	0.198 (0.687)	0.409 (0.681)	-3.011 (1.132)**	-3.063 (1.003)**
Male	-4.182 (1.654)*	-0.463 (1.522)	-0.356 (1.521)	-0.236 (1.080)	-0.724 (1.081)	-3.863 (1.843)*	0.057 (1.697)
Only child	-2.576 (2.371)	-2.540 (2.198)	-3.295 (2.237)	-1.606 (1.633)	-1.520 (1.617)	-2.418 (2.514)	-2.429 (2.277)
Number of siblings	0.004 (1.029)	0.250 (0.931)	0.392 (0.940)	-1.846 (0.679)**	-1.672 (0.673)*	1.540 (1.434)	1.345 (1.243)
Intercept	60.04 (19.80)**	44.78 (17.69)*	40.13 (17.85)*	-6.566 (13.26)	-8.729 (13.14)	58.38 (22.03)**	45.11 (19.17)*
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-square	0.209	0.375	0.375	0.229	0.247	0.032	0.267
Number of sample	1,494	1,494	1,494	1,494	1,494	1,494	1,494

Note: Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

Table 3: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Performance: Birth Order as an IV

	Reduced form models (OLS)		Structural models (2SLS)	
Dependent variable:	Spending on tutoring (log)		Average test score	
	(1)	(2)	(3)	(4)
Tutoring spending			0.774 (0.406)	0.644 (0.378)
Middle child	-1.788 (1.696)	-1.736 (1.680)		
Last-born child	-3.514 (0.855)**	-3.322 (0.850)**		
Hours of self-study	0.032 (0.035)	0.019 (0.035)	0.522 (0.060)**	0.352 (0.053)**
Prior quality (Q2)		2.954 (1.301)*		11.548 (2.299)**
Prior quality (Q3)		1.504 (1.280)		20.373 (2.019)**
Prior quality (Q4)		1.244 (1.318)		33.137 (2.024)**
Prior quality missing		-2.831 (1.251)*		19.713 (2.091)**
Single father	-1.534 (2.216)	-1.937 (2.200)	-5.146 (3.708)	-5.690 (3.292)
Single mother	-2.816 (1.731)	-2.711 (1.714)	4.835 (3.160)	3.695 (2.783)
Books at home	0.001 (0.002)	0.001 (0.002)	0.003 (0.003)	0.003 (0.003)
Family income	0.021 (0.002)**	0.021 (0.002)**	-0.010 (0.009)	-0.009 (0.008)
Parents' avg edu	0.903 (0.168)**	0.893 (0.167)**	0.777 (0.478)	0.624 (0.433)
Parents' avg age	0.210 (0.127)	0.171 (0.126)	0.067 (0.192)	0.054 (0.169)
Age	0.200 (0.687)	0.410 (0.681)	-2.994 (1.152)**	-3.053 (1.025)**
Male	-0.118 (1.086)	-0.612 (1.088)	-3.887 (1.807)*	0.037 (1.636)
Only child	-2.101 (1.705)	-1.977 (1.690)	-2.430 (2.661)	-2.433 (2.338)
Number of siblings	-2.203 (0.766)**	-2.000 (0.759)**	1.423 (1.421)	1.303 (1.232)
Intercept	-3.529 (13.15)	-5.840 (13.04)	58.50 (21.82)**	45.10 (19.19)*
School characteristics	Yes	Yes	Yes	Yes
F (IVs excluded from the 2nd stage)	8.45	7.65		
P-value for overid test			0.591	0.814
R-square	0.229	0.247	0.058	0.275
Number of sample	1,494	1,494	1,494	1,494

Note: Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

Table 4: OLS and 2SLS Estimates of the Effect of Weekly Tutoring Hours on Subject Test Scores

Dependent variable:	Weekly hours of tutoring		Subject test scores		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
		A. Mathematics ($N = 1,407$)			
Tutoring hours			-0.231 (0.452)	-0.220 (0.444)	0.176 (0.035)**
First-born child	3.401 (1.128)**				
Middle child		-3.938 (2.330)			
Last-born child		-3.316 (1.174)**			
Hours of self-study	0.076 (0.049)	0.076 (0.049)	0.387 (0.077)**	0.386 (0.074)**	0.355 (0.064)**
F (Excluded IVs)		4.58			
P-value of overid test				0.780	
		B. English ($N = 1,493$)			
Tutoring hours			0.616 (0.303)*	0.591 (0.295)*	0.017 (0.033)
First-born child	4.935 (1.067)**				
Middle child		-6.408 (2.193)**			
Last-born child		-4.701 (1.110)**			
Hours of self-study	-0.015 (0.046)	-0.014 (0.046)	0.433 (0.065)**	0.433 (0.064)**	0.427 (0.059)**
F (Excluded IVs)		10.99			
P-value of overid test				0.601	
		C. Korean language ($N = 1,481$)			
Tutoring hours			1.238 (0.724)	1.116 (0.679)	0.068 (0.036)
First-born child	2.489 (0.991)*				
Middle child		-3.380 (2.030)			
Last-born child		-2.346 (1.031)*			
Hours of self-study	0.033 (0.043)	0.033 (0.043)	0.233 (0.083)**	0.237 (0.078)**	0.274 (0.059)**
F (Excluded IVs)		3.28			
P-value of overid test				0.368	

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator,number of siblings, prior performances, family background (single father and mother indicators, family income, parents' average education and age), and school characteristics. Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

Table 5: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Performance by Ability Level

Dependent variable:	Spending on tutoring (log)		Average test score		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
		A. Low-level prior performance ($N = 415$)			
Tutoring spending			0.690 (0.420)	0.623 (0.411)	0.172 (0.079)*
First-born child	5.064 (1.358)**				
Middle child		-6.764 (3.237)*			
Last-born child		-4.883 (1.395)**			
Hours of self-study	0.056 (0.061)	0.060 (0.062)	0.218 (0.100)*	0.222 (0.097)*	0.244 (0.094)**
F (Excluded IVs)		7.11			
P-value of overid test				0.288	
		B. Medium-level prior performance ($N = 442$)			
Tutoring spending			1.298 (1.394)	0.537 (0.734)	0.190 (0.071)**
First-born child	1.613 (1.341)				
Middle child		2.107 (2.945)			
Last-born child		-2.125 (1.387)			
Hours of self-study	0.074 (0.063)	0.079 (0.063)	0.276 (0.149)	0.334 (0.103)**	0.360 (0.084)**
F (Excluded IVs)		1.73			
P-value of overid test				0.373	
		C. High-level prior performance ($N = 249$)			
Tutoring spending			0.537 (0.660)	0.478 (0.617)	0.164 (0.073)*
First-born child	3.828 (2.012)				
Middle child		-2.406 (4.225)			
Last-born child		-4.006 (2.069)			
Hours of self-study	0.065 (0.108)	0.064 (0.108)	0.505 (0.122)**	0.508 (0.133)**	0.525 (0.106)**
F (Excluded IVs)		2.87			
P-value of overid test				0.831	

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performances, family income and school characteristics. Other family background variables (single father and mother indicators, and parents' average education and age) are not controlled in order to gain sufficient degrees of freedom. Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

Table 6: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Performance by Sex

Dependent variable:	Spending on tutoring (log)		Average test score		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
		A. Male students ($N = 852$)			
Tutoring spending			0.911 (0.571)	0.692 (0.482)	0.092 (0.049)
First-born child	3.164 (1.136)**				
Middle child		1.806 (3.920)			
Last-born child		-3.359 (1.145)**			
Hours of self-study	0.056 (0.048)	0.056 (0.048)	0.295 (0.086)**	0.309 (0.078)**	0.349 (0.066)**
F (Excluded IVs)		4.76			
P-value of overid test				0.333	
		B. Female students ($N = 642$)			
Tutoring spending			0.253 (0.592)	0.282 (0.591)	0.121 (0.056)*
First-born child	2.782 (1.170)*				
Middle child		-2.525 (1.936)			
Last-born child		-2.866 (1.275)*			
Hours of self-study	-0.020 (0.053)	-0.020 (0.053)	0.366 (0.075)**	0.367 (0.077)**	0.362 (0.073)**
F (Excluded IVs)		2.84			
P-value of overid test				0.471	

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performances, family background (single father and mother indicators, family income, parents' average education and age), and school characteristics. Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

Table 7: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Performance by Family Income

Dependent variable:	Spending on tutoring (log)		Average test score		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
		A. Low-income Family ($N = 315$)			
Tutoring spending			0.561 (0.958)	0.556 (0.842)	0.159 (0.085)
First-born child	2.767 (1.847)				
Middle child		-0.350 (3.297)			
Last-born child		-3.311 (1.948)			
Hours of self-study	0.069 (0.080)	0.065 (0.080)	0.297 (0.129)*	0.297 (0.134)*	0.328 (0.107)**
F (Excluded IVs)		1.51			
P-value of overid test				0.992	
		B. Mid-income Family ($N = 773$)			
Tutoring spending			0.439 (0.345)	0.428 (0.337)	0.058 (0.051)
First-born child	4.771 (1.149)**				
Middle child		-3.373 (2.366)			
Last-born child		-4.996 (1.197)**			
Hours of self-study	0.008 (0.051)	0.009 (0.051)	0.360 (0.076)**	0.360 (0.073)**	0.365 (0.073)**
F (Excluded IVs)		8.84			
P-value of overid test				0.845	
		C. High-income Family ($N = 406$)			
Tutoring spending			0.780 (1.177)	0.624 (0.796)	0.141 (0.084)
First-born child	2.044 (1.472)				
Middle child		2.472 (3.599)			
Last-born child		-2.586 (1.522)			
Hours of self-study	0.012 (0.060)	0.010 (0.060)	0.339 (0.097)**	0.340 (0.093)**	0.345 (0.089)**
F (Excluded IVs)		1.91			
P-value of overid test				0.839	

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performances, family background (single father and mother indicators, family income, parents' average education and age), and school characteristics. Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

Table 8: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Performance by Rating on Tutoring

Dependent variable:	Spending on tutoring (log)		Average test score		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
		A. Students Rating Tutoring as Useful ($N = 948$)			
Tutoring spending			0.327 (0.544)	0.328 (0.549)	0.060 (0.046)
First-born child	2.566 (0.959)**				
Middle child		-2.637 (2.013)			
Last-born child		-2.555 (0.997)*			
Hours of self-study	0.028 (0.040)	0.028 (0.040)	0.270 (0.062)**	0.270 (0.061)**	0.279 (0.057)**
F (Excluded IVs)		3.58			
P-value of overid test				0.872	
		B. Students Rating Tutoring as Not Useful ($N = 546$)			
Tutoring spending			1.292 (0.869)	1.050 (0.623)	0.106 (0.058)
First-born child	3.039 (1.478)*				
Middle child		0.897 (2.918)			
Last-born child		-3.669 (1.530)*			
Hours of self-study	-0.075 (0.070)	-0.070 (0.070)	0.620 (0.152)**	0.600 (0.124)**	0.522 (0.093)**
F (Excluded IVs)		3.34			
P-value of overid test				0.602	

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performances, family background (single father and mother indicators, family income, parents' average education and age), and school characteristics. Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

Table 9: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Test-writing

Dependent variable:	Spending on tutoring (log)		Take Test (No=1)		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
Tutoring spending			0.001 (0.005)	0.002 (0.005)	-0.0011 (0.0006)
First-born child	3.509 (0.768)**				
Middle child		-2.865 (1.546)			
Last-born child		-3.614 (0.799)**			
Hours of self-study	0.044 (0.034)	0.044 (0.034)	-0.003 (0.001)**	-0.003 (0.001)**	-0.003 (0.001)**
Prior quality (Q2)	2.891 (1.186)*	2.886 (1.186)*	-0.123 (0.037)**	-0.124 (0.033)**	-0.115 (0.032)**
Prior quality (Q3)	1.988 (1.169)	1.987 (1.169)	-0.131 (0.034)**	-0.132 (0.031)**	-0.126 (0.031)**
Prior quality (Q4)	2.192 (1.205)	2.206 (1.205)	-0.087 (0.036)*	-0.087 (0.032)**	-0.081 (0.033)*
Prior quality missing	-2.079 (1.127)	-2.071 (1.127)	-0.072 (0.032)*	-0.072 (0.029)*	-0.077 (0.032)*
Single father	-3.586 (2.049)	-3.627 (2.051)	-0.006 (0.052)	-0.005 (0.052)	-0.013 (0.048)
Single mother	-2.968 (1.527)	-2.923 (1.530)	0.079 (0.047)	0.080 (0.041)	0.071 (0.043)
Books at home	0.000 (0.002)	0.000 (0.002)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Family income	0.019 (0.002)**	0.019 (0.002)**	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Parents' avg edu	0.907 (0.155)**	0.905 (0.155)**	-0.005 (0.007)	-0.006 (0.006)	-0.003 (0.004)
Parents' avg age	0.209 (0.118)	0.218 (0.120)	-0.008 (0.002)**	-0.008 (0.003)**	-0.008 (0.002)**
Age	0.129 (0.644)	0.133 (0.645)	-0.006 (0.016)	-0.006 (0.016)	-0.005 (0.016)
Male	-0.252 (0.999)	-0.201 (1.005)	0.002 (0.026)	0.002 (0.024)	0.001 (0.026)
Only child	-1.886 (1.494)	-2.091 (1.555)	0.044 (0.037)	0.045 (0.036)	0.043 (0.037)
Number of siblings	-1.113 (0.620)	-1.261 (0.693)	0.028 (0.017)	0.028 (0.017)	0.024 (0.016)
Intercept	-9.653 (12.41)	-6.311 (12.32)	0.845 (0.302)**	0.845 (0.297)**	0.843 (0.301)**
School characteristics	Yes	Yes	Yes	Yes	Yes
F (Excluded IVs)		10.55			
P-value for overid test				0.527	
R-square	0.255	0.255	0.091	0.088	0.100
Number of sample	1,727	1,727	1,727	1,727	1,727

Note: Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

Table 10: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Performance: Missing Test Scores Replaced

Dependent variable:	Spending on tutoring (log)		Average test score		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
		Total Sample ($N = 1,727$)			
Tutoring spending			0.948 (0.428)*	0.978 (0.424)*	0.026 (0.043)
First-born child	3.509 (0.768)**				
Middle child		-2.865 (1.546)			
Last-born child		-3.614 (0.799)**			
Hours of self-study	0.044 (0.034)	0.044 (0.034)	0.245 (0.067)**	0.244 (0.069)**	0.288 (0.056)**
F (Excluded IVs)		10.55			
P-value of overid test				0.493	
		A. Low-level prior performance ($N = 463$)			
Tutoring spending			0.641 (0.472)	0.614 (0.465)	0.140 (0.084)
First-born child	5.116 (1.286)**				
Middle child		-5.986 (3.083)			
Last-born child		-5.019 (1.325)**			
Hours of self-study	0.094 (0.058)	0.096 (0.059)	0.257 (0.118)*	0.260 (0.116)*	0.302 (0.109)**
F (Excluded IVs)		7.95			
P-value of overid test				0.455	

(Continued at the next page.)

Table 10: OLS and 2SLS Estimates of the Effect of Tutoring Spending on Performance: Missing Test Scores Replaced (Continued)

Dependent variable:	Spending on tutoring (log)		Average test score		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
		(Continued) B. Medium-level prior performance ($N = 492$)			
Tutoring spending			0.508 (0.714)	0.455 (0.558)	0.126 (0.064)*
First-born child	2.411 (1.306)				
Middle child		1.229 (2.781)			
Last-born child		-2.915 (1.348)*			
Hours of self-study	0.105 (0.061)	0.113 (0.061)	0.245 (0.107)*	0.251 (0.101)*	0.287 (0.077)**
F (Excluded IVs)		2.81			
P-value of overid test				0.905	
		C. High-level prior performance ($N = 312$)			
Tutoring spending			1.636 (1.328)	1.628 (1.365)	-0.049 (0.108)
First-born child	3.437 (1.883)				
Middle child		-3.469 (3.717)			
Last-born child		-3.433 (1.937)			
Hours of self-study	0.108 (0.106)	0.108 (0.106)	-0.230 (0.266)	-0.229 (0.288)	-0.080 (0.145)
F (Excluded IVs)		1.66			
P-value of overid test				0.317	

Note: Each model of Panel A, B and C includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performances, family income and school characteristics. Other family background variables (single father and mother indicators, and parents' average education and age) are not controlled in order to gain sufficient degrees of freedom. Standard errors are reported in parentheses. * and ** indicate that the estimate is significant at the 0.05 and 0.01 levels, respectively.

References

Altonji, J. G., T. E. Elder, and C. R. Taber (2002). An evaluation of instrumental variable strategies for estimating the effects of Catholic schools. *NBER Working Papers 9358.*

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy 113*(1), 151–184.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Becker, G. S. and N. Tomes (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy 84*(4), S143–S162.

Behrman, J. R. (1988). Nutrition, health, birth order and seasonality: Intrahousehold allocation among children in rural India. *Journal of Development Economics 28*(1), 43–62.

Behrman, J. R. and P. Taubman (1986). Birth order, schooling, and earnings. *Journal of Labor Economics 4*(3), S121–S145.

Betts, J. R. (1995). Does school quality matter? evidence from the national longitudinal survey of youth. *Review of Economics and Statistics 77*(2), 231–250.

Betts, J. R. (1996). Is there a link between school inputs and earnings? Fresh scrutiny of an old literature. In G. Burtless (Ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, Washington, DC, pp. 141–191. Brookings.

Birdsall, N. (1991). Birth order effects and time allocation. In T. P. Schultz (Ed.), *Research in Population Economics. A Research Annual*, Greenwich, CT and London, pp. 191–213. JAI Press.

Black, S. E., P. J. Devereux, and K. G. Salvanes (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.

Blake, J. (1981). Family size and the quality of children. *Demography 4*, 421–442.

Blake, J. (1989). *Family Size and Achievement.* Berkeley and Los Angeles, CA: University of California Press.

Bound, J., D. A. Jaeger, and R. M. Baker (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association 90*(430), 443–450.

Card, D. and A. B. Krueger (1992). Does school quality matter? returns to education and the characteristics of public schools in the united states. *Journal of Political Economy 100*(1), 1–40.

Card, D. and A. B. Krueger (1996). School resources and student outcomes: An overview of the literature and new evidence from North and South Carolina. *Journal of Economic Perspectives 10*(4), 31–50.

Conley, D. and R. Glauber (2005). Parental educational investment and children's academic risk: Estimates of the impact of sibship size and birth order from exogenous variations in fertility. *NBER Working Paper 11302*.

Ejrnaes, M. and C. C. Portner (2004). Birth order and the intrahousehold allocation of time and education. *Review of Economics and Statistics 86*(4), 1008–1019.

Evans, W. N. and R. M. Schwab (1995). Finishing high school and starting college: Do Catholic schools make a difference? *Quarterly Journal of Economics 110*(4), 941–974.

Figlio, D. N. and J. A. Stone (1999). Are private schools really better? In S. Polachek (Ed.), *Research in Labor Economics*, Volume 18, pp. 115–140. Stamford, Connecticut: JAI Press.

Goldhaber, D. D. (1996). Public and private high schools: Is school choice an answer to the productivity problem? *Economics of Education Review 15*(2), 93–109.

Greenwald, R., L. V. Hedges, and R. D. Laine (1996). The effect of school resources on student achievement. *Review of Educational Research 66*(3), 361–396.

Grogger, J. (1996). School expenditures and post-schooling earnings: Evidence from High School and Beyond. *Review of Economics and Statistics 78*(4), 628–637.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature 24*(3), 1141–1177.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy 100*(1), 87–117.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis 19*(2), 141–164.

Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis 21*(2), 143–163.

Hanushek, E. A. (2003). The failure of input-based schooling policies. *Economic Journal 113*(485), F64–F98.

Hauser, R. M. and W. H. Sewell (1985). Birth order and educational attainment in full sibships. *American Educational Research Journal 22*(1), 1–23.

Hedges, L. V., R. D. Laine, and R. Greenwald (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher 23*(3), 5–14.

Horton, S. (1988). Birth order and child nutritional status: Evidence from the Philippines. *Economic Development and Cultural Change 36*(2), 341–354.

Kessler, D. (1991). Birth order, family size, and achievement: Family structure and wage determination. *Journal of Labor Economics 9*(4), 413–426.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics 114*(2), 497–532.

Krueger, A. B. (2003). Economic considerations and class size. *Economic Journal 113*(485), F34–F62.

Neal, D. (1997). The effects of Catholic secondary schooling on educational achievement. *Journal of Labor Economics 15*(1), 98–123.

Olneck, M. R. and D. B. Bills (1979). Family configuration and achievement: Effect of birth order and family size in a sample of brothers. *Social Psychology Quarterly 42*, 135–148.

Parish, W. L. and R. J. Willis (1993). Daughters, education, and family budgets Taiwan experiences. *Journal of Human Resources 28*(4), 863–898.

Retherford, R. D. and W. H. Sewell (1991). Birth order and intelligence: Further tests of the confluence model. *American Sociological Review 56*(2), 141–158.

Rosenzweig, M. R. (1986). Birth spacing and sibling inequality: Asymmetric information within the family. *International Economic Review 27*(1), 55–76.

Royer, H. N. (2004). What all women (and some men) want to know: Does maternal age affect infant health? *Center for Labor Economics Working Paper 68*. University of California, Berkeley.

Strobino, D., M. Ensminger, Y. Kim, and J. Nanya (1995). Mechanisms for maternal age differences in birthweight. *American Journal of Epidemiology 142*(5), 504–514.

Zajonc, R. B. (1976). Family configuration and intelligence. *Science 192*(4236), 227–236.