Department of Economics

Working Paper No. 0704

http://nt2.fas.nus.edu.sg/ecs/pub/wp/wp0704.pdf

Does Money Matter?
The Effect of Private Educational Expenditures on Academic Performance

by

Changhui Kang

Does Money Matter? The Effect of Private Educational Expenditures on Academic Performance

Changhui Kang
Department of Economics
National University of Singapore
1 Arts Link, Singapore 117570
Singapore
E-mail: ecskch@nus.edu.sg
Phone: +65-6516-6830
Fax: +65-6775-2646

Abstract

The causal relationship between educational investments and student outcomes continues to attract attention. The majority of studies have examined the effectiveness of public school expenditures on student outcomes. This paper attempts to shed light on the impacts of educational inputs by examining a private educational investment—private tutoring that is widely employed by South Korean parents as a supplement to public school education. To deal with the endogeneity of private tutoring expenditures, the paper relies on instrumental variables (IV) methods, exploiting a student's birth order as a source of identification. Based on the IV methods, the paper shows that a 10 percent increase in expenditure leads to a 0.56 percentile point improvement in test score. Such an estimated effect is modest and comparable to the effect of public school expenditures on earnings estimated by previous studies.

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

Does Money Matter? The Effect of Private Educational Expenditures on Academic Performance

Abstract

The causal relationship between educational investments and student outcomes continues to attract attention. The majority of studies have examined the effectiveness of public school expenditures on student outcomes. This paper attempts to shed light on the impacts of educational inputs by examining a private educational investment—private tutoring that is widely employed by South Korean parents as a supplement to public school education. To deal with the endogeneity of private tutoring expenditures, the paper relies on instrumental variables (IV) methods, exploiting a student's birth order as a source of identification. Based on the IV methods, the paper shows that a 10 percent increase in expenditure leads to a 0.56 percentile point improvement in test score. Such an estimated effect is modest and comparable to the effect of public school expenditures on earnings estimated by previous studies.

1 Introduction

The causal relationship between educational investments and student outcomes continues to attract the attention of many. Despite decades of intensive study, there is no general consensus regarding the effectiveness of monetary educational inputs for student outcomes.

In particular, papers that summarize the debate on the effects of public school expenditures often advocate conflicting views. For example, Card and Krueger (1996), Greenwald et al. (1996) and Krueger (2003) are in favor of the effectiveness of public school expenditures; Betts (1996) and Hanushek (1986, 1997, 2003) cast doubt on the conclusion of these researchers and suggest several factors that can explain discrepancies in conclusion (e.g., Betts, 1996).

Studies focusing on private schools (e.g., Catholic schools) do not shed much light on the impacts of educational expenditures as well. Evans and Schwab (1995) and Neal (1997) show educational benefits of attending Catholic high school. In contrast, Altonji et al. (2005a), Figlio and Stone (1999) and Goldhaber (1996) find no significant gaps in test scores between public and private schools. Importantly, many studies on private and Catholic schooling seem to suffer from lack of reliable exogenous variation for identifying the causal effect. Altonji et al. (2005b) 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.

In this paper we attempt to shed light on the effectiveness of educational investments by examining private educational expenditures. Specifically, we look into academic effects 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 have little freedom in the choice of their middle and high schools in a school district. Since 1969 student allocation to public and private schools has largely been under the strict control of the government, especially in urban regions. Under this system (labeled 'Leveling Policy'), students are basically assigned—not admitted upon application—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 Ministry of Education office. Moreover, within school ability grouping is rarely implemented due to the government's egalitarian policy on secondary education and parents' objections. Curricula are also controlled by the Ministry of Education nationwide for the most part.

In response to such a rigid public education system, parents in South Korea spend a great deal of money on private tutoring for their children. According to a statistic, South Korean parents spend on private tutoring 85 percent as much as they spend on public schooling (Korean Educational Development Institute, *Survey on Educational Expenditures*, 1998). Given such large expenditures on private tutoring, many including parents as well as educational policy makers are concerned about the effectiveness of private tutoring on student academic performance. From a broader perspective, an examination into the effect of private tutoring serves to illuminate the debates on the impacts of educational inputs on student outcomes.

It is well known that educational expenditures on a student are not exogenously and randomly determined; there is little doubt that private tutoring expenditures are endogenous and correlated with a student's personal, family and academic characteristics. In the absence of a randomized experiment on private tutoring, a causal estimation calls for an extra variable for a student that strongly affects the parents' decision to invest in a child's education, but is independent of educational outcomes of the student (academic performance 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. As long as a student's birth order is exogenously determined by how many older siblings were born before him or her, however, it is unlikely to affect the academic performance of the student.

Using a student's birth order as an IV, this paper shows that a 10 percent increase in private educational expenditure leads to a 0.56 percentile point improvement in test score. Evaluated at the mean value, this amount of the effect is equivalent to a 1.1 percent increase in test score due to a 10 percent increase in expenditure. Our estimated effect of private educational expenditures is modest and fairly comparable to the effect of public school expenditures on earnings estimated by previous studies (e.g., Card and Krueger, 1996).

Our analysis of private educational expenditures reveals that holding other factors constant, birth order significantly influences the amount of educational investment for a student: first-born students receive significantly greater expenditures on private tutoring than do later-born students. Birth order, on the other hand, does not seem to be significantly associated with a student's academic performance according to the over-identification test. Nonetheless, a correlation may exist between a student's birth order and performance along the dimensions that are not revealed by statistical tests. We, however, expect such a correlation to be more likely to overstate the true effect than understate it; thus our IV estimates can be interpreted as the upper limit of the effect of private tutoring.

The rest of the paper is organized as follows. Section 2 presents various theories and empirical evidence on the relationship between birth order, educational investments and outcomes. Section 3 gives a description on private tutoring in South Korea. Data are discussed in section 4; the empirical strategy in section 5. Empirical results are shown in section 6. Potential explanations for the small effect of private tutoring are reviewed in section 7. Section 8 concludes.

2 Birth Order, Educational Investments and Outcomes

2.1 Various Theories

The phenomenon and reasons for parental favor for a certain-parity child in educational investment have been reported in a number of studies. First, studies suggest priorities in education for first-born and last-born children. The resource dilution model in sociology and the quantity/quality tradeoff model in economics argue that parents are faced with time and financial constraints over the life cycle. This makes it impossible to equalize resources over children (Behrman, 1988; Birdsall, 1991; Horton, 1988). First-born and last-born children will benefit because they enjoy more quality time and monetary resources with less competition than do middle-born children.

Second, other scenarios are possible when the availability of a family's financial and time resources is important to a child. Parents may invest more in educating later-borns than earlier-borns, if parents' earnings increase over their life cycle (Parish and Willis, 1993; Powell and Steelman, 1995). And later-born children may benefit from more parental attention, if earliest-born children have moved out of the family's home.

Third, parental preferences may vary by birth order (Behrman and Taubman, 1986). If parents seek their security in old age, they may favor earlier-born children (especially, sons) and devote greater resources to them, as they become economically independent first. If parents are focused on their career, however, they may be willing to spend less time and more money on their children when parents are younger than when they are older and more established professionally.

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

A natal factor may play a formative role in a child's academic 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 will give the oldest children an advantage (Behrman, 1988): the intellectual capability declines with birth order. 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 are known to affect children's educational outcomes such as IQ (Behrman, 1988; Horton, 1988).

A theory in psychology predicts that younger children have worse intelligence than older siblings. The 'confluence theory' (Zajonc, 1976) argues that there exist birth order effects on intelligence that are favorable for older and unfavorable for younger siblings, because the average intellectual environment of the home deteriorates when a higher proportion of household members are young children.

Independent of the natal and early resource allocation factors, the optimal fertility-stopping rule can give rise to birth order effects. When parents employ the optimal childbearing rule where 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 (higher) than expected genetic endowments, then the last-born child has the lowest (highest) quality.

2.2 Empirical Evidence

Although there are many empirical studies on birth order effects, they often confound educational investment by parents with the own quality of a child.[1] 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 pure investment on a child by parents from the own capability of the child. Therefore, empirical evidence on the independent effects of birth order on parents' educational investment and on a child's own quality is somewhat limited. Nevertheless, a trend (albeit arguable) reported by studies is that birth order affects parental allocation of resources that presumably influence educational attainment, but it has no significant (or consistent) effects on a child's capability such as intellectual development and the performance in schools and the labor market (Steelman et al., 2002).[2]

Behrman and Taubman (1986) find empirical evidence that lower birth order or first-borns are favored in schooling, while there exist no statistically significant birth order effects on earnings for young U.S. adults. Powell and Steelman (1995) find that later-born children are more likely to be the beneficiaries of educational resources and receive financial assistance in college. Hauser and Sewell (1985), however, show no significant or systematic effects of birth order on schooling outcomes (high school graduation, postsecondary educational attainments

[1] Empirical analyses of birth order effects are also frequently complicated by the presence of independent sibship size effects. While birth order effects are more or less neglected in theoretical discussions, economic theories on sibship size effects have a long tradition and extensive empirical support in the name of a quantity/quality tradeoff (Becker and Lewis, 1973). Recent evidence, however, suggests that the effects of sibship size on a child's educational attainment may proxy for those of birth order. Black et al. (2005) show that a negative correlation between sibship size and children's education becomes negligible when birth order is added as an explanatory variable along with sibling size. See also Angrist et al. (2005) and Conley and Glauber (2005) for a causal analysis of the effects of sibship size in consideration of birth order.

[2] In psychology, although Sulloway (1996) argues for the presence of birth order effects, recent views seem to be that birth order has no significant effects on personality and psychological characteristics (Harris, 1998; Pinker, 2002).

of those graduates, or educational attainments). Recently, Black et al. (2005) document that for every sibship size up to 10 first-born children have significantly higher years of completed education than later-born children do. 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 effect on earnings is likely to work through education (p.695).

Empirical evidence on birth order effects on children's endowed qualities and academic capability is more or less mixed. On the one hand, Zajonc (1976) and Zajonc and Mullally (1997) report intellectual test performance declines with birth order. Hanushek (1992) finds that while there is no birth order effect among small families with less than 5 siblings, in large families test scores are higher among last-born and first-born children than among middle-borns. On the other hand, Retherford and Sewell (1991) and Rodgers et al. (2000) find no evidence of birth order effects on intelligence. Olneck and Bills (1979) and Blake (1981) report that there exists a negligible influence of birth order on individuals' educational attainment and test scores. Kessler (1991) presents that birth order fails to significantly influence the level or growth rate of wages.

To summarize, the existing evidence about the effects of birth order on educational investment and outcome measures is mixed. Yet, research shows a trend that birth order matters to educational investment, while having negligible (or inconsistent) impacts on a child's academic performance. More importantly, existing empirical evidence is inherently limited because studies often confound the pure educational investment with the own quality of a child. Given the limitations of existing evidence, the validity of a child's birth order as an IV for private educational expenditures seems to be a matter to be resolved empirically in the context of our main analysis.

3 Private Tutoring in South Korea

There are potentially many forms of private tutoring that a student may receive for various reasons. It varies from a swimming lesson for an exercise to math tutoring for a slow-learning child. Here we focus our attention on a private supplementary instruction of academic subjects that involves financial transactions outside the formal school system. Such private tutoring is generally observed in many countries where public education system is poorly equipped or the existing system fails to satisfy highly motivated parents. While it is apparently most prominent in East Asian societies such as Japan, Hong Kong, Taiwan and South Korea (Time Asia, 2006), studies report private tutoring in a wide range of countries from Egypt, Kenya to India, Romania, Canada and UK (Baker et al., 2001; Bray, 1999).

In South Korea there exist widespread and large-scale markets for private tutoring outside the public education system. For example, Ministry of Education (1999, 2000) reports that about 58.2 percent of all pre-college students experience various kinds of private tutoring in 2000. It also documents that such a proportion is highest (70.7) among elementary school students, medium (59.5) among middle school students, and relatively low (46.8) among college-bound students of general high school. As a consequence of widespread popularity, private tutoring expenditures by parents are also quite large in Korea. Ministry of Education (1999, 2000) shows that private tutoring expenses are about 9 percent of incomes of the households that have school-aged children for all income groups. At the national level, total household expenditures on private tutoring of year 2003 amount to 2.3 percent of the national GDP and 55 percent of the national annual budget for public education (Korean Educational Development Institute, *Media Briefing*, November 19, 2003). A major reason for such widespread private tutoring in the country is that there are virtually no private secondary schools that are independent of the government's control: in Korea private middle and high schools are little different from public schools with respect to school administration, curriculum and student placement, because they are heavily subsidized and controlled by the government.[3]

Among potentially many channels of private instruction, only two broad types of private tutoring are permitted by the government and practiced in the market in Korea. One is a relatively formal instruction offered by *hakwon*s—private for-profit school-like learning institutions. The other is an informal private instruction by individual university students. All other forms of private tutoring including private instruction by full-time school teachers outside the school, that by *hakwon* instructors outside the *hakwon*, and private tutoring through mails, phones and TVs are prohibited by the government. Of these two legal forms of private tutoring, the government maintains a strong control over *hakwon*s, while it has little control over individual tutors:

[3]For an overview of secondary education and private tutoring in South Korea, see Kim and Lee (2001) and OECD (1998).

the government imposes some requirements for establishing a *hakwon*, and exerts administrative controls with respect to pricing, academic qualification of tutors, physical facilities, etc. (Kim and Lee, 2001)

According to our data from the Korean Education and Employment Panel (discussed shortly), the proportion of grade 11 students who receive private tutoring for any subject is 77.8 percent. And their overall average monthly spending on private tutoring is about KRW 285,400—approximately USD$239, which amounts to 9 percent of monthly family income (see Table 1). Among academic subjects, private tutoring is most frequently practiced for mathematics (51.8 percent). Of those students who receive private tutoring for math, 45.4 percent uses *hakwon*s and 47.4 percent employs one-to-one or one-to-many tutoring offered by individual tutors.

4 Data

4.1 Description of the Main Sample

For empirical analysis this study employs the Korean Education and Employment Panel (KEEP). KEEP is a longitudinal study that is conducted from year 2004 by the Korea Research Institute for Vocational Education and Training (KRIVET)—a government-funded research institute. The basic structure of KEEP follows the National Educational Longitudinal Studies (NELS) of the U.S. The beginning cohorts of KEEP consist of 6,000 students from three different populations: 2,000 students each from middle school (grade 9), general high school and vocational high school (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 variety of personal, family and school-related questionnaires. In addition, students' homeroom teachers, school principals and parents are separately surveyed to collect a range of background information on the sampled student.

An important feature of the KEEP data is that the survey collects detailed information on a student's private tutoring experience and expenditures, and the sibling composition from the parent questionnaire. It enables us to construct main explanatory variables and instrumental variables of this study. Also unique in the KEEP data is the availability of the College Scholastic Ability Test (CSAT) scores for high school graduates. CSAT is the national college-entrance examination of Korea that is annually administered under the supervision of the Ministry of Education and whose scores are used as an important factor by colleges and universities to determine the admission of the applicants. Using the resident registration number of the student, the KEEP data are linked to 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). Given the percentile score of each subject, the average of the three individual percentiles is calculated and employed for our main analysis.

Although vocational high school graduates are eligible for CSAT, the majority of the CAST writers are general high school graduates; they are also the majority of students taking private tutoring. Therefore, we restrict our analysis to the general high-school sample of 2,000 students.[5]

For main analysis, the original general high-school sample has been further narrowed down. First, those students whose guardian is not one of the parents are excluded from our main sample. Patterns of private educational investment and academic performance among these students may be far from typical due to the absence of both parents. 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 such a general subject test as CSAT.[6] The number of such students are 168 in

[4]Guidelines of CSAT stipulate that students are free to choose individual subjects for their examination. Nevertheless, the majority of students choose either Korean, 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 select Korean, 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, 98.2 percent of a total of 1,733 CSAT writers in 2004 choose Korean. The corresponding figures are 89.6 and 98.0 percent for math and English, respectively.

[5]For 2004 general high school accounts for 70.5 percent of 1.75 million high school students in the nation; vocational high school accounts for the remainder (*Yearbook of Educational Statistics 2004*, National Statistical Office).

[6]According to the KEEP data, students who attend special high schools for music, fine arts and athletics

total.

Since more than one restriction may be applied to a single student, the preceding two restrictions leave a total of 1,752 students for further analysis. Descriptive statistics of the main sample and their differences between first-born and later-born students are documented in Table 1.

INSERT TABLE 1 HERE.

4.2 Descriptive Statistics

Among CSAT test-writers, the mean percentile score of the Korean language is 49.8; the mean scores of math and English are 49.0 and 49.4, respectively.[7] While mean math score is close between the two groups, mean scores of Korean and English among first-borns significantly exceed those of later-born students. And the mean of the percentile scores averaged over the three subjects is also significantly greater for first-born 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 contemporary variations in educational investment.

As for the amount of spending on private tutoring, first-born students receive larger educational investments from their parents than later-born counterparts do.[8] While the overall average monthly spending on private tutoring is about W285,400—approximately $239, the average spending for first-born students (W323,800) is 35 percent greater than that for later-borns (W240,500). This amount of gap is significantly different from zero. The proportion of

or take private tutoring of these fields expend a monthly average of W513,000—approximately $430. Other students, on the other hand, spend nearly half as much on private tutoring—a monthly average of W285,000. In addition, the average CSAT scores are substantially different between these two groups of students. The average percentile score of the arts and atheletic students is 31.3; that of the non-specialist group is 48.3. The difference is significantly different from zero.

[7]The following procedure is employed to construct the CSAT scores of analysis from the raw CSAT percentile scores: In the KEEP data, there exist students who do not apply for the 2004 CSAT. In addition, some students are absent for a few or all subjects of the examination upon application. We consider a student to have no intention of writing the 2004 CSAT, and treat the CSAT scores as unknown and missing, if the student either does not apply for the 2004 CSAT, or is absent (upon application) for the tests of all three main subjects (Korean, math, and English). If a student misses the test of only one or two subjects and is present for other subjects, on the other hand, the student's CSAT score for the missed subject is set to zero, not missing. (There are 6 students with a zero score for Korean, 4 for math, and 7 for English.) Those who, according to the preceding criteria, are considered as non-writers of the 2004 CSAT account for a total of 235 students—13.4 percent of students in the main sample. In section 6.4, we attempt to take those non-writers of CSAT into consideration. Up to such a point, our main analysis does not include those who miss all the three tests and have their test scores unknown.

[8]The KEEP survey asks the monthly average amount of overall expenditures on private tutoring during the last six months before grade 12—roughly nine to fourteen months prior to the CSAT test.

those who have ever received private tutoring—those with positive monthly spending—is also far higher among first-born students (83.0 percent) than among later-borns (71.6 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 Korean (30.1 percent) in frequency. Weekly tutoring hours are also longest for math (2.5 hours), followed by English (1.7 hours) and Korean (1.3 hours) in terms of duration. Whether in terms of frequency or duration, the preceding statistics suggest that first-born students receive significantly greater educational investments than later-borns do.

Using the questionnaire 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.[9] 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 (11.0 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-born 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. The average age of parents of first-borns is lower than that of later-borns, since a couple gives birth to the first child earlier in life. The parents' 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).

[9]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 support objectively this conversion in the KEEP data set.

As the KEEP survey collects information on private tutoring during the second semester of grade 11 (as mentioned in footnote 8), an assumption required in our analysis is that it takes some time for private tutoring to take effect in a student's performance. Otherwise, our measure of a student's pre-treatment quality is error-ridden, because it contains part of short-term effects of private tutoring.

In our main sample, 53.9 percent is first-born students that include only children. 31.1 percent is first-born boys; 22.8 percent is first-born girls. In the sample, only children account for 7.5 percent; the mean number of siblings in a family is 2.2 and the maximum is 6.[10]

5 Empirical Framework

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

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

where Y_i is the (average) percentile test score of student i; T_i is the monthly spending on tutoring for i (in natural log)[11]; $\tilde{Y}_i$ is i's pre-tutoring performance in grade 11; 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 yield a causal and consistent estimate for β_1, because of the endogeneity problem (i.e., $E(T_iu_i) \neq 0$). For example, 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 upward. And if parents determine the expenditure on private tutoring according to the pre-tutoring performance of the child, then the estimate can also be biased: if parents tend to spend more on lower performers and less on higher performers, the estimate will be biased downward; if parents spend less on lower performers and more on higher performers, the estimate will be biased upward.

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

$$T_i = \gamma_0 + First_i\gamma_1 + \tilde{Y}_i\gamma_2 + X_i\gamma_3 + \epsilon_i \quad (2)$$

[10]In the empirical analysis we include only children and control for their status with a dummy variable. Excluding them from the sample does not yield qualitative differences in results.

[11]In order to minimize outlier problems and make the interpretation convenient, the raw values of tutoring spending are converted into the natural log metric. And the log-transformed values are multiplied by 10 in order to examine the changes in Y_i associated with 10 percent changes in tutoring spending. 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 positive unit of money reported in the survey (₩10,000) and it is about 3.5 percent of the mean expenditure on private tutoring. Whether a smaller value (e.g., 1) is added to every expenditure or the level of raw values are employed rather than the log, the results are qualitatively similar. Appendix Table 1 reports estimates under the two different measures of the expenditure.

We use two sets of variables for the first-born indicator $First_i$. One set is a dummy variable that takes 1 if i is a first-born child in the family and 0 otherwise regardless of the sex. The other set consists of two dummy variables: first-born boy and first-born girl indicators. The reference group is the group of later-born boys and girls. Given that Korean parents show son preference (Park and Cho, 1995), educational investments may vary between first-born son and first-born daughter. In addition, the second set of $First_i$ enables us to test more formally the exogeneity of birth order by means of the over-identification test.[12]

Although the over-identification test sheds light on potential exogeneity of birth order to u_i, failure to reject the null hypothesis does not necessarily imply the IV is exogenous. As discussed in section 2.2, some studies report birth order effects on academic capability such as IQ, though their presence remains controversial. Furthermore, given that parents favor a certain-parity child (e.g., first-born) over others with respect to observable educational investments, it is possible that they may favor the same child along unobservable dimensions, too: for example, while spending more money on tutoring for the first-born, parents may provide better emotional and non-financial supports for the first-born than for other siblings. If this is the case, our IV estimates will be biased. Nevertheless, we expect such a bias to be more likely to be upward than downward.

There are three reasons for this expectation. First, if parents favor, say, the first-born with respect to monetary educational investments, they will tend to support the same child more over unobservable dimensions as well; thus a correlation between the first-born indicator and u_i is more likely to be positive than negative.[13] Second, studies reporting significant birth order

[12] This method of creating an overidentified model is similar to that of Angrist and Evans (1998). In an estimation of causal effects of sibship size on women's labor supply, they introduce a mixed sibling-sex composition as an IV for sibship size. To produce an overidentified model they decompose the *same sex* instrument into *two boys* and *two girls* indicators.

[13] In a simple model expressed by $y = \beta_0 + \beta_1 x + u$, if x is endogenous and z is a potential IV, the IV estimator for β_1 is given by $\widehat{\beta}_{IV} = \frac{Cov(y,z)}{Cov(x,z)}$, whose probability limit is equal to $\beta_1 + \frac{Cov(u,z)}{Cov(x,z)}$. Thus, given that $Cov(x, z)$ is positive, the direction of bias in $\widehat{\beta}_{IV}$ is determined by the sign of $Cov(u, z)$.

In the KEEP data there are not many variables available by which we can examine such unobservable supports of parents for a certain-parity child. Nevertheless, we attempt an indirect method to address how differently parents treat a first-born child relative to other siblings by focusing on early childhood. Using the information on how often parents read books to a child in pre-school period (1 if very often and 0 otherwise) and how often they help the child's homework and examination in elementary school period (1 if very often or often and 0 otherwise), we run linear probability models in which the student's personal and family backgrounds are controlled for together with the first-born indicator. The results show that parents more often read to a first-born child in pre-school period (coefficient 0.125; SE 0.027), and help the first-born child's homework and examination in elementary school period (coefficient 0.073; SE 0.029). In addition, parents have greater educational aspiration for a first-born child (a post-graduate degree as opposed to a university degree or below) than for other siblings (coefficient 0.052; SE 0.026).

effects on intelligence usually show negative rather than positive effects of birth order: older siblings have higher intelligence than younger siblings. Thus there will be a positive if any correlation between the first-born indicator and u_i. Third, according to Table 1, the average education and family income of parents are higher among first-borns than among later-born students. Although our regressions control for parents' education and income, they may not fully capture a potential association between birth order and u_i. To the extent that parents of higher education and income tend to affect positively a child's performance, we also expect the correlation between the first-born indicator and u_i to be positive rather than negative. Therefore, given that first-borns are favored in private tutoring expenditures, as empirically found later, using the first-born indicator as an IV is more likely to overstate the effect of private tutoring expenditures than understate it.[14] Our IV estimates can be viewed as the upper limit of the effect of private tutoring.

In addition to the potential exogeneity of an IV, weak IVs have been another major concern in IV estimations. It is now well-known that if instruments are weak, then the sampling distribution of a conventional 2SLS estimator is nonnormal; standard 2SLS point estimates, hypothesis tests, and confidence intervals are unreliable. According to a criterion proposed by Stock et al. (2002, Table I) for a single endogenous variable, a single IV is weak if the first-stage F statistic is less than 8.96; two IVs are jointly weak if the F statistic is less than 11.59. In the face of potential weakness of our IV, we employ two standard sets of methods for point estimation and testing that are robust to the presence of weak IVs.[15]

The first set is three testing methods that are fully robust to weak IVs: the Anderson-Rubin (AR) test (Anderson and Rubin, 1949), the Lagrangian Multiplier (LM) test (Kleibergen, 2002), and the Conditional Likelihood Ratio (CLR) test (Moreira, 2003). Based on the test statistic of each test, we construct the 95 percent confidence interval for testing. The second set is two point estimation methods that are partially robust to weak IVs: Limited-Information Maximum Likelihood (LIML) and Fuller-k estimator with $\alpha = 1$ (Fuller, 1977). Unlike 2SLS estimators,

[14] An optimal fertility-stopping rule (Ejrnaes and Portner, 2004) may imply either downward or upward bias in our IV estimates of β_1. For example, if parents stop fertility when an above-average quality child is born and they invest more on the less endowed first-born, then we observe higher tutoring expenditures on the first-born and downward bias in the IV estimates of β_1; on the other hand, if parents stop fertility when a below-average quality child is born and they invest more on the more endowed first-born, then we observe higher tutoring expenditures on the first-born and upward bias in the IV estimates of β_1. Empirical evidence on fertility-stopping rule and its connection with educational investment is, however, scarce.

[15] There are several survey papers on the IV estimation under weak IVs: Andrews and Stock (2005), Dufour (2003), Hahn and Hausman (2003) and Stock et al. (2002).

conventional normal asymptotic approximations can be applied to these two estimators for inference and testing.

6 Estimation Results

6.1 OLS and IV Results

6.1.1 Reduced-form Results by OLS

Basic OLS results of the effect of private tutoring expenditures on student performance are presented in columns (1) and (2) of Table 2. The associations between tutoring expenditures and percentile test scores (averaged over the three main subjects) are positive but quite small, although significantly different from zero. A 10 percent greater monthly expenditure on private tutoring is related to no more than a 0.1 percentile point higher test score. As explained previously, such an association that is estimated by OLS may not be consistent and causal. Depending on the correlation between T_i and u_i, the estimate may be biased upward or downward.

INSERT TABLE 2 HERE.

In column (2) 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 private tutoring expenditures, this finding is suggestive of the possibility that birth order can be exogenous to u_i in (1).

As for other explanatory variables, higher pre-tutoring performance leads to higher post-tutoring test scores. Single-fatherhood (relative to the presence of both parents) and parents' average education level are significantly related to the test score of a student. In contrast, single motherhood, the number of books at home, family income and parents' average age are not associated with test scores. While a student's age have a strong connection with test scores, sex, being an only child and the number of siblings fail to have a significant bearing with test scores.

6.1.2 First-stage Results

The results of the first-stage regression of tutoring expenditures are presented in columns (3) and (4) of Table 2. In column (3), being first-born significantly and positively affects private tutoring expenditures for a student. First-born students receive about 30 percent greater expenditures on tutoring than later-born students. (Recall that a log of a monthly tutoring expenditure is multiplied by 10. See footnote 11.) The F statistic for the first-born indicator is 12.9. According to a criterion of Stock et al. (2002), whether a student is first-born is a strong—or not-weak—IV for tutoring expenditures for the student. Provided that being first-born has no direct association with test scores, this variable can serve as a legitimate instrument for spending on tutoring.

When the first-born boy and first-born girl indicators are employed as IVs, the amount of spending for first-born boys is about 26 percent greater, and that for first-born girls is about 34 percent greater than that of later-borns. Although parents appear to spend more for first-born girls than for first-born boys relative to later-borns, difference in size between them is not significantly larger than zero (p-value>0.1). While the individual significance differs between the two IVs, the first-born boy and first-born girl indicators are jointly significant predictors of educational expenditures. In column (4) the F statistic for the joint significance of these two variables is 6.59; the p-value is less than 0.001. According to Stock et al. (2002), however, the two IVs fail to be jointly sufficiently strong so that conventional 2SLS inference and testing methods can be applied. Thus, the over-identified case calls for estimation and testing methods that are robust to weak IVs.

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 to 50th percentiles) than on those with the lowest prior performance (below the 25th percentile), the expenditure starts 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 negative, though not significant—negative 0.011 (SE 0.016). This inverse relationship between pre-tutoring performance and tutoring expenditures suggests that the OLS estimates of β_1 in (1) are more likely to be biased downward than upward, because parents seem to spend educational funds for siblings in a compensating way.

Other variables, which significantly affect expenditures on 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 sibship size and tutoring expenditures is consistent with a quality/quantity tradeoff in fertility.[16] 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.

6.1.3 Second-stage Results

The IV estimates of the effect of tutoring expenditures are shown in columns (5) and (6) of Table 2. The estimates in column (5) are based on the first-born indicator as an IV, while those in column (6) on the first-born boy and first-born girl indicators as IVs. Estimates from weak-IV-robust estimation and testing methods are also reported at the bottom.

According to 2SLS estimates, a 10 percent increase in expenditure enhances a student's performance by 0.54-0.56 percentile points. Evaluated at the mean percentile score (49), they imply a 1.1 percent increase in test score due to a 10 percent increase in expenditure on private tutoring. The LIML and Fuller-k estimates imply that a 10 percent increase in expenditure leads to a 0.51-0.56 percentile point improvement in test score. Tests based on weak-IV-robust methods, however, suggest that the estimated effects are statistically indistinguishable from zero. All the 95% confidence bands of three weak-IV-robust testing methods contain zero.

Although they are statistically indistinguishable from zero, our IV estimates are over 5 times greater than the OLS estimates. This implies that the OLS estimates are severely biased downward. Such a bias arises probably because parents tend to spend more for low-performing siblings than for high-performing siblings within a family.

The over-identification test suggests that the first-born boy and first-born girl indicators are not necessarily invalid instruments. The p-value of the test exceeds 0.1 at the bottom of column (6): the two variables of birth order seem to be uncorrelated with the error term (u_i)

[16] In a quality/quantity model of fertility, the number of siblings is determined simultaneously with the human-capital investment in the children (Becker and Lewis, 1973). Consequently, 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). Excluding the number of siblings from X_i, however, yields little qualitative difference in results.

and correctly excluded from the main equation (1). However, failure to reject the hypothesis of a zero correlation between birth order variables and u_i does not necessarily imply that such a hypothesis is true. Nonetheless, as discussed earlier, a potential correlation between birth order and u_i is more likely to overstate the true effect of private tutoring than understate it.

Given such a modest effect of private tutoring expenditures, it would be instructive to compare our IV estimates with corresponding estimates of previous studies, precision of the estimates set apart, in order to gain some perspective of our results. Unfortunately, however, 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 for comparison.

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 3 leads to a 0.67-0.88 percentile point 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 accompanying 10 percent increase in per-pupil expenditure.[17] Compared with Krueger's (1999) estimates, our estimated effect of private tutoring expenditures is no more than a half.

In terms of earnings in the labor market, Card and Krueger (1996, p.37) summarize that a 10 percent increase in public 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.[18] 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 that a 10 percent increase in mean spending per student leads to a 0.7 percent increase in wages. Our estimated effect of private tutoring expenditures is fairly comparable to the estimated effects of public school expenditures on earnings, although ours are on a slightly higher side.

[17]Krueger (2003, F55-F56) infers that a one percent decrease in class size will be approximately converted into a one percent increase in annual per pupil cost.

[18]Our estimated effect of tutoring expenditures is dwarfed by a meta-analysis of Hedges et al. (1994) that 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 15.9 percentile point improvement in test score in our metric. The summary of Hedges et al. (1994) are, however, criticized by Hanushek (1997) for being biased in favor of large positive effects of school expenditures.

6.2 Effects of Tutoring Hours on Subject Test Scores

In the preceding section we examine the effect of tutoring expenditures on a student's test score that is averaged over three subjects—math, Korean and English. An implicit assumption was that these three subjects account for the majority of the educational expenditures. In fact, however, the tutoring expenditures that are reported in KEEP and used by previous regressions are the monthly average of total expenditures that are paid to all tutoring-related activities for *any* academic subject; total tutoring expenditures may include money spent on other subjects such as social studies and science subjects. As a result, the expenditures may be overstated and contaminated by measurement errors in an examination that focuses only on math, Korean and English. To the extent that OLS estimates are biased toward zero under measurement errors, using total expenditures on tutoring may be a reason for our OLS estimates of the effect of tutoring expenditures being close to zero.

In this section, instead of using undifferentiated overall expenditures on private tutoring, we employ information specific to each subject: weekly average tutoring hours and percentile test scores for each subject. The OLS and IV estimation results for each subject are documented in Table 3.[19] While all the regressions are based on the value-added specification, odd columns use only the first-born indicator as a single IV; even columns use the first-born boy and first-born girl indicators as two IVs. The upper panel shows OLS estimates of the first-stage regressions, and the F statistics for examining weak IVs; the lower-panel reports the point estimates of the effect of expenditures that are produced by OLS, 2SLS, LIML and Fuller-k methods, the 95 percent CIs from Wald, AR, LM and CLR tests, and the p-values of the over-identification tests.

INSERT TABLE 3 HERE.

For all three subjects, a student's birth order is related to weekly tutoring hours. Degree of the association, however, varies according to the subject. Whichever variable is used in the first stage, birth order is a strong predictor of tutoring hours only for English, and a weak predictor for math and Korean. First-born students receive about 20 percent greater investments in English tutoring than later-borns do; the former receive about 14 and 10 percent more tutoring for math and Korean, respectively, than the latter do. Although the estimates of the first-born

[19] As in footnote 11, we convert the raw values of weekly tutoring hours into the natural log metric and multiply them by 10. To deal with zero hours in the log transformation, a value of 1 is added to every student's raw value of tutoring hours. When a smaller value (e.g., 0.1) is added instead of 1, the results are qualitatively similar.

indicators are significantly different from zero, F-statistics for them reveal different degrees of strength as an IV.

In addition to its potential relevance as an IV, the exogeneity of birth order in the main equation is supported by the over-identification test, when two birth order variables are employed. For all subjects, the test fails to reject the hypothesis of IV exogeneity. Thus, for each subject birth order seems to serve as a legitimate IV for educational expenditures for a student.

Similar to previous results based on overall expenditures, for all three subjects, the OLS estimates for the effect of tutoring hours show quite a small association between tutoring hours and test scores. A 10 percent longer tutoring is related to only a 0.03-0.3 percentile point higher subject test score. When we examine the causal effect of tutoring hours, however, different trends emerge for different subjects.

First, tutoring for math seems to be ineffective in enhancing the test score. The IV estimates show a negative impact of math tutoring on the test score, although they are statistically indistinguishable from a zero effect according to robust methods.

Second, tutoring for English seems relatively effective in improving the performance. The IV estimates imply that a 10 percent increase in weekly tutoring hours leads to a 1.3-1.4 percentile point improvement in the English test score. However, while the LM and CLR tests reject the null hypothesis of a zero effect, the estimated effect fails to be significantly different from zero according to other four robust methods.

Third, it is difficult to decide whether tutoring for Korean is effective in improving the test score. Although the IV point estimates suggest causal impacts that are larger than for math and English (i.e., 1.9 to 2.3 points), the tests fail to reject the null hypothesis of a zero effect; all the 95 percent confidence bands include zero. Thus a conservative view will be that tutoring for Korean does not improve the test score as well.

6.3 Heterogeneity in the Effect of Private Tutoring

When there is heterogeneity in effects, the treatment effects that are identified in an IV analysis is nothing but local (Local Average Treatment Effects, LATE). Angrist et al. (1996) show that the treatment effects estimated by 2SLS are applicable only to the group or groups whose behavior is influenced by the instruments (i.e., *compliers*). Thus the effect of private tutoring estimated in this study does not necessarily represent the average treatment effect among students randomly chosen in the population, but is valid only for those students whose private tutoring expenditures are likely to be affected by birth order.

Although it is difficult to distinguish a priori between those students whose tutoring expenditures are likely to be influenced by birth order and those that are not, it would be illuminating to examine whether there exists heterogeneity in effects at least along observable dimensions of student and family characteristics. In addition, previous research suggests that the effects of educational resources may vary according to sex, race, family income and the ability level of the student (Bedard, 2003; Krueger, 1999; Rivkin et al., 2005). In order to address such heterogeneity, we disaggregate the main sample by the following three measured dimensions of students: the level of pre-tutoring performance, sex, and family income. In subsequent empirical analysis, because of concerns for sample size, we return to the variables that are undifferentiated for individual subjects: the monthly average of total tutoring expenditures and the percentile test scores averaged over math, Korean and English.

6.3.1 Level of Pre-tutoring Performance

At first, based on the pre-tutoring performance of a student, we split the main sample into three sub-samples: bottom-third, middle-third and top-third samples. The OLS and IV results for each sub-sample are documented in Table 4. In this table and subsequent tables, odd columns are based on the first-born indicator as a single IV; even columns on the first-born boy and first-born girl indicators as two IVs. The upper panel shows OLS estimates of the first-stage regressions, and the F statistics for examining weak IVs; the lower-panel reports the point estimates of the effect of expenditures that are produced by OLS, 2SLS, LIML and Fuller-k methods, the 95 percent CIs from Wald, AR, LM and CLR tests, and the p-values of the over-identification tests.

INSERT TABLE 4 HERE.

For each third of pre-tutoring performance, the OLS estimates imply a small association between tutoring expenditures and test scores. A 10 percent larger expenditure on tutoring is related to only a 0.2 percentile point higher test score. The size of the association varies little by the level of pre-tutoring performance.

For students whose pre-tutoring performance is in the bottom third—columns (1) and (2), birth order has a fairly strong association with tutoring expenditures. Although each of the dummy variables is individually significant at the 5 percent level, the first-born indicator is a strong IV for tutoring expenditures, but the first-born boy and first-born girl indicators are jointly weak IVs. The IV estimates for the effects of private tutoring show modest improvement in test score due to the increase in tutoring expenditures. A 10 percent increase in spending raises a low-ability student's test score by a 0.69-0.74 percentile points. Nevertheless, robust testing methods suggest that private tutoring fails to have significant effects on the performance of low ability students.

For other students whose pre-tutoring performance is in the middle—columns (3) and (4)—and top third—columns (5) and (6), birth order fails to be significantly associated with tutoring expenditures. Although testing results are not informative because of the weakness of IVs, the size of IV estimates implies modest improvement in test score due to the increase in tutoring expenditures. For middle-third students, a 10 percent increase in spending on tutoring raises the test score by 0.42-0.81 percentile points. For top-third students, a 10 percent increase in spending enhances the test score by 0.51-0.69 percentile points.

In sum, there seems to exist no substantial heterogeneity in the effect of private tutoring across students of different pre-tutoring abilities; for all levels, private tutoring fails to have significant causal effects on the test outcome.

6.3.2 Sex

The estimation and testing results for each sex are documented in Table 5. Note that the same overidentified model as earlier is not defined in this case. For both male and female students, birth order significantly affects tutoring expenditures. In both cases, however, birth order fails to serve as a strong IV.

INSERT TABLE 5 HERE.

Although point estimates are imprecisely estimated, the IV results show a stronger positive effect of private tutoring for boys than for girls. While a 10 percent larger spending on tutoring enhances a girl's test score only by 0.12 percentile points, the same amount of spending increases a boy's score by 0.81-0.94 percentile points. Again, since all of the 95 percent CIs from robust methods contain zero for each sex, we fail to find evidence that tutoring expenditures have significant causal effects on test outcomes for boys and girls alike.

6.3.3 Family Income

Using 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.[20] The estimation results for each group of family income are presented in Table 6.

INSERT TABLE 6 HERE.

Birth order has a strong association with tutoring expenditures only among mid-income families. In contrast, among low-income and high-income families, first-born children do not receive significantly greater investments in tutoring than later-born siblings; the estimates are insignificant.

The IV estimates for mid-income students imply that a 10 percent increase in tutoring expenditures leads to 0.37-0.42 percentile point improvement in test score. The size of improvement seems modest. Again, since all of the 95 percent CIs from robust methods contain zero, we fail to find evidence that tutoring expenditures have significant causal effects on the test outcome for mid-income students.

Although point estimates are imprecisely estimated, the effect of tutoring expenditures is slightly weaker among low-income students than among mid- and high-income students. A 10 percent increase in spending on tutoring either lowers a low-income student's test score or enhances it by at most 0.06 percentile points; the same increase in spending on tutoring raises a high-income student's score by 0.63-1.22 points. No such estimates, however, are significantly different from zero. Overall, the causal effect of private tutoring expenditures does not seem to vary substantially according to the family income of a student. In fact, it is close to zero for all groups of family income.

To summarize, the examination of heterogeneity in effects along measurable dimensions of characteristics reveals no considerable differences in the effect of tutoring expenditures. Al-

[20]Total family income includes average monthly labor and financial incomes of all family members residing in the household. The low-income families are those whose total monthly income is less than 2 million Won (approximately $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 and 4 million Won ($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.

though we find some differences in effects between boys and girls, the size of the difference in average effects across groups is within the range of statistical errors along all observable dimensions examined; the IV estimates reported in Table 2 do not seem to represent a very narrow group of students. For all the groups of students identified by observable characteristics, the causal effect of tutoring expenditures are modest and insignificant.

6.4 Robustness to Missing Test Scores

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 observable and/or unobservable characteristics, then our IV 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 for test-writing are shown in Table 7.

INSERT TABLE 7 HERE.

The OLS estimate in column (5) suggests that private tutoring expenditures are positively associated with test-writing decisions. The size of the association, though, is negligible and indistinguishable from zero. The 2SLS estimates in columns (3) and (4) also suggest no strong relationship between private tutoring expenditures and test-writing behavior. The estimation results suggest that missing test scores of some students, which may take place in association with private tutoring expenditures, are unlikely to cause a serious bias in our IV estimates for the effect of educational expenditures.

In contrast to tutoring expenditures, 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 IV estimates when they are uncorrelated with private tutoring expenditures. However, it is hard to believe this is so. As long as students who have lower pre-tutoring performance and shorter hours of self-study, hence requiring greater private tutoring expenditures 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 yield a bias in our IV estimates.

Due to the lack of an appropriate variable that explains the selection process, while not affecting one's test score, the Heckman's method for correcting a sample-selection problem is not well applied in our context. Instead, we attempt to address a potential sample-selection bias by replacing missing test scores with those that are randomly generated in reference to a student's pre-tutoring performance.[21] When missing test scores are replaced, the causal effect of private tutoring expenditures on performance is more precisely estimated, but remains at the similar level. For the whole sample, a 10 percent increase in expenditure enhances a student's performance by 0.8-0.9 percentile points. If the whole sample is divided by the pre-tutoring performance of a student, the results of the low-ability group are generally similar to those in Table 4; the estimates of the mid-ability group, however, slightly fall to a 0.05 to 0.1 point level, but those of the high-ability group rise to a 1.3 to 1.9 point level. Again, these amounts of the effect are relatively modest and statistically insignificant.

7 Potential Explanations for the Modest Effect

Our main conclusion of the preceding analysis is that private educational expenditures have modest, if any, causal effects on student test performance. Such a finding may be contrary to expectations of many. Here we review a few potential explanations for the finding.

First, monetary educational investments may not always matter in children's educational outcomes. Although controversial, some researchers summarize that public school resources are

[21] The simulation procedures are as follows: 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) that is randomly generated by the formula given by:

$$Y_i = \tilde{Y}_i + 5 \times \upsilon_i \tag{3}$$

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 test score and pre-tutoring performance of a student are not available in the data, we employ the following assignment process:

$$Y_i = \tau_i \tag{4}$$

where τ_i is a random number from a uniform distribution between 0 and 25. The reason for using the range between 0 and 25 is because Table 7 suggests that those not taking the test are likely to be low-performers than medium- or high-performers. (Here, we use values of 5 in (3) and 25 in (4) for generating a hypothetical test score for a student whose test score is unknown. When we employ other values than 5 and 25, however, there are no substantial qualitative differences in the results. The results under different values are available upon request.) According to the preceding procedures, a total of 235 missing test scores are substituted for. Their average value and standard deviation are 41.1 and 31.7, respectively.

not an important determinant of student outcomes. Our main results lend further support to the claim in light of private educational expenditures.

As a potential explanation about small effects of public educational investments, researchers suggest poor management of educational resources due to large bureaucracies in the public sector (Anderson et al., 1991) and unionization of public school teachers (Hoxby, 1996). Our study offers an answer to such explanations: private use of educational funds also has no more significantly positive impacts on student outcomes than public management does. In Korea most of private educational institutions such as *hakwon*s are small-sized and non-unionized. Thus public mis-management due to bureaucracies and unionization is unlikely to be responsible for small effects of public educational expenditures.

Second, in response to the lack of significant effects of public school resources in the U.S., researchers pose a possibility that if added resources have diminishing effects on student achievement, current level of public school expenditures may be on the flat portion of the production function. The expenditures may be more effective at lower levels than at higher levels (Betts, 1996; Hanushek, 1997). International educational statistics show that South Korea is placed below median among OECD countries with respect to educational expenditures: OECD (2004) reports that the educational expenditures per student (private and public in total) of Korea is 20th largest among 26 OECD countries that have valid statistics. In 2001 Korea spent $5,159 per student for secondary education, and $8,873 for both primary and secondary education; in contrast, the U.S. spent $8,779 for secondary, and $16,339 for both primary and secondary education. Educational resources do not seem to be more effective at lower levels than at higher levels.

Third, as shown in section 6.3, birth order is a significant determinant of private tutoring expenditure only for the groups of low-performing students and those from mid-income families. Thus, our IV estimates are more likely to show the average effect of private tutoring for these groups of students than for other groups. Although we fail to find strong evidence of heterogenous effects, it is possible that the effect of educational investment is weaker among low-performers than among other higher-performers (e.g., Bedard, 2003). In order to address heterogeneous effects of private tutoring over student quality, we estimate quantile treatment effects using an IV quantile regression (IVQR) method proposed by Chernozhukov and Hansen

(2005).[22] (Here we exclude a student's pre-tutoring performance from X_i.) The IVQR results show stronger effects in lower quantiles—1.92 (SE 2.13) for the 0.1 quantile; 0.78 (SE 0.47) for the 0.25 quantile; 0.93 (SE 0.43) for the 0.5 quantile, while weaker effects in higher quantiles—0.47 (SE 0.53) for the 0.75 quantile; 0.45 (SE 0.75) for the 0.9 quantile. Although the interpretation is limited by the lack of precision of the estimates, it seems unlikely that the effect of private tutoring is larger for the group of higher-performing students.

Fourth, as Hanushek (2003) implies, overall quality of teachers in the private sector may be responsible for small effects of private expenditures. In Korea, full-time public school teachers are tenured up to 62 years of age and enjoy the same employment benefits as government officials. In addition, school teachers enjoy public respect. In contrast, contracts of *hakwon* instructors are usually short-term in nature and fairly unstable as in other private small firms. This will cause teachers' quality in the private sector to be worse than that in the public sector. To the extent that teachers' quality can make differences in student outcomes, private investment of educational funds may have small effects.

Fifth, peer pressure among parents may explain the lack of the effect. According to the recent theory of group socialization, there are cultural factors in parenting practices, and their attitudes and beliefs about children; they are shared by parents' peer groups (Harris, 1995; 1998). When private tutoring is a norm in parents' peer groups, the decision to invest in children's tutoring may not be based on an objective cost-benefit analysis; they may be based on a subjective/cultural belief about the effectiveness of private tutoring, or the concern about their being viewed by the peers as neglectful of children's education. If the decision about tutoring is based on peer pressure, small effects of private tutoring will not be a big surprise.

8 Concluding Remarks

In order to shed light on the effectiveness of educational investments on student outcomes, this paper examines the effect of private educational expenditures (private tutoring expenditures 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 educational investments, while the child's academic

[22] Computer codes for performing IVQR are obtained from the web site "http://www.gsb.uchicago.edu/fac/christian.hansen /research/." We thank Christian Hansen for sharing the computer codes.

capability can be little affected by birth order in the family.

The analysis of private tutoring expenditures reveals that holding other factors constant, birth order significantly influences the amount of educational investment received by a student. The causal estimates based on IV methods imply that a 10 percent increase in expenditure on private tutoring leads to a 0.56 percentile point improvement in test score. Evaluated at the mean value, this amount of effect is equivalent to a 1.1 percent increase in test score. Given that a correlation may exist between a student's birth order and performance along the dimensions that are not revealed by statistical tests, the estimated effect is more likely to be the upper limit of the true effect than the lower limit. Nevertheless, our estimated effect is fairly comparable to the effects of public school expenditures on earnings estimated by previous studies.

Disaggregating the entire sample by the level of pre-tutoring performance, sex and family income of students reveals no considerable differences in the effect of tutoring expenditures. While there exist differences in estimated effect between boy and girl students, the size of the differences is within the range of statistical error. As potential explanations for modest effects of private tutoring, we propose lower teacher quality in the private education sector of Korea and peer pressure of parents. Yet, whether monetary educational investments raise student educational outcomes in different contexts remains to be further examined.

References

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

Altonji, J.G., Elder, T.E., Taber, C.R., 2005b. An evaluation of instrumental variable strategies for estimating the effects of Catholic schooling, Journal of Human Resources 40 (4), 791-821.

Anderson, G.M., Shughart, W.F., Tollison, R., 1991. Educational achievement and the cost of bureaucracy, Journal of Economic Behavior & Organization 15 (1), 29-45.

Anderson, T.W., Rubin, H., 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations, Annals of Mathematical Statistics 20 (1), 46-63.

Andrews, D.W.K., Stock, J.H., 2005. Inference with weak instruments, NBER Techinical Working Paper Series T0313.

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

Angrist, J.D., W.N. Evans, Children and their parents' labor supply: Evidence from exogenous variation in family size, American Economic Review 88 (3), 450-477.

Angrist, J.D., Lavy, V., Schlosser, A., 2005. New evidence on the causal link between the quantity and quality of children, NBER Working Paper 11835.

Argys, L.M, Rees, D.I., Averett, S.L., Witoonchart, B., 2006. Birth Order and Risky Adolescent Behavior, Economic Inquiry 44 (2), 215-233.

Baker, D.P., Akiba, M., LeTendre, G.K., Wiseman, A.W., 2001. Worldwide shadow education: Outside-school learning, institutional quality of schooling, and cross-national mathematics achievement, Educational Evaluation and Policy Analysis 23 (1), 1-17.

Becker, G.S., Lewis, H.G., 1973. On the interaction between the quantity and quality of children, Journal of Political Economy 81 (2, Pt 2), S279-S288.

Bedard, K., 2003. School quality and the distribution of male earnings in Canada, Economics of Education Review 22, 395-407.

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., Taubman, P., 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: Burtless, G. (Eds.), Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, Brookings, Washington, DC, pp. 141-191.

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

Black, S.E., Devereux, P.J., Salvanes, K.G., 2005. The more the merrier? The effect of family size and birth order on children's education, Quarterly Journal of Economics 120 (2), 669-700.

Blake, J., 1981. Family size and the quality of children, Demography 4, 421-442.

Bray, M., 1999. The shadow education system: Private tutoring and its implications for planners, UNESCO International Institute for Educational Planning, Paris.

Card, D., Krueger, A.B., 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., Krueger, A. B., 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.

Chernozhukov, V., Hansen, C., An IV model of quantile treatment effects, Econometrica 73 (1), 245-262.

Conley, D., Glauber, R., 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.

Dufour, J.-M., 2003. Identification, weak instruments, and statistical inference in econometrics, Canadian Journal of Economics 36 (4), 767-808.

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

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

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

Fuller, W.A., 1977. Some properties of a modification of the limited information estimator, Econometrica 45 (4), 939-954.

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., Hedges, L.V., Laine, R.D., 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.

Hahn, J., Hausman, J., 2003. Weak instruments: Diagnosis and cures in empirical econometrics, American Economic Review 93 (2), 118-125.

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., 2003. The failure of input-based schooling policies, Economic Journal 113 (485), F64-F98.

Harris, J.R., 1995. Where is the child's environment? A group socialization theory of development, Psychological Review 102 (3), 458-489.

Harris, J.R., 1998. The Nurture assumption: Why children turn out the way they do, New York, NY:Touchstone.

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

Hedges, L.V., Laine, R.D., Greenwald, R., 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.

Hoxby, C.M., 1996. How teachers' unions affect education production, Quarterly Journal of Economics 111 (3), 671-718.

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

Kim, S., Lee, J.-H., 2001. Demand for education and developmental state: Private tutoring in South Korea, Social Science Research Network Electronic Paper Collection: http://ssrn.com/abstract=268284.

Kleibergen, F., 2002. Pivotal statistics for testing structural parameters in instrumental variables regression, Econometrica 70 (5), 1781-1803.

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.

Ministry of Education, 1999, 2000. Sakyoyuk Siltae Josa [Survey on Private Tutoring in 1999 (2000)], Seoul, South Korea (in Korean).

Moreira, M.J., 2003. A conditional likelihood ratio test for structural models, Econometrica 71 (4), 1027-1048.

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

OECD (Organization for Economic Co-Operation and Development), 1998. Reviews of National Policies for Education: Korea, OECD Publications, Paris.

OECD (Organization for Economic Co-Operation and Development), 2004. Education at a glance 2004, OECD Publications, Paris.

Olneck, M.R., Bills, D.B., 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., Willis, R.J., 1993. Daughters, education, and family budgets Taiwan experiences, Journal of Human Resources 28 (4), 863-898.

Park, C.B., N.H. Cho, Consequences of son preference in a low-fertility society: Imbalance of the sex ratio at birth in Korea, Population and Development Review 21 (1), 59-84.

Pinker, S., 2002. The blank slate: The denial of human nature and modern intellectual life, Penguin Books, New York.

Powell, B., Steelman, L.C., 1995. Feeling the pinch: Child spacing and constraints on parental economic investments in children, Social Forces 73 (4), 1465-1486.

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

Rivkin, S.G., Hanushek, E.A., Kain, J.F., 2005. Teachers, schools, and academic achievement, Econometrica 73 (2), 417-458.

Rodgers, J.L., Cleveland, H.H., Oord, E. V.D., Rowe, D.C., 2000. Resolving the debate over birth order, family size, and intelligence, American Psychologist 55 (6), 599-612.

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

Steelman, L.C., Powell, B., Werum, R., Carter, S., 2002. Reconsidering the effects of sibling configuration: Recent advances and challenges, Annual Review of Sociology 28,243-269.

Stock, J.H., Wright, J.H., Yogo, M., 2002. A survey of weak instruments and weak identification in Generalized Method of Moments, Journal of Business and Economic Statistics 20 (4), 518-529.

Sulloway, F.J., 1996. Born to rebel: Birth order, family dynamics, and creative lives, Pantheon, New York.

Time Asia, 2006. Asia's overscheduled kids, 167 (12), 48-55.

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

Zajonc, R.B., Mullally, P.R., 1997. Birth order reconciling conflicting effects, American Psychologist 52 (7), 685-699.

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	0.135	0.342	0.133	0.339	0.003	0.016	0.17
Average score of three tests	1503	49.00	22.68	50.65	22.36	47.07	22.90	3.571	1.170	3.05
Test score of Korean	1490	49.75	25.82	51.72	25.59	47.42	25.93	4.302	1.338	3.21
Test score of math	1419	48.97	26.33	49.22	26.31	48.67	26.36	0.549	1.404	0.39
Test score of English	1490	49.37	26.21	51.77	25.86	46.57	26.37	5.195	1.356	3.83
Tutoring Expenditure ($W1,000$)	1749	285.4	341.1	323.8	375.6	240.5	289.6	83.31	16.25	5.13
Any tutoring (Yes=1)	1749	0.778	0.416	0.830	0.376	0.716	0.451	0.114	0.020	5.79
Tutoring hours for Korean	1752	1.295	2.504	1.448	2.607	1.116	2.368	0.331	0.120	2.76
Tutoring for Korean (Yes=1)	1752	0.301	0.459	0.338	0.473	0.258	0.438	0.080	0.022	3.64
Tutoring hours for math	1752	2.464	3.136	2.746	3.251	2.134	2.963	0.612	0.150	4.09
Tutoring for math (Yes=1)	1752	0.518	0.500	0.576	0.495	0.451	0.498	0.125	0.024	5.24
Tutoring hours for English	1752	1.732	2.627	1.956	2.520	1.471	2.725	0.485	0.125	3.86
Tutoring for English (Yes=1)	1752	0.410	0.492	0.473	0.500	0.337	0.473	0.136	0.023	5.82
Prior quality	1285	46.12	26.68	45.13	26.18	47.30	27.23	-2.169	1.493	-1.45
Hours of self-study	1752	11.34	10.22	11.62	10.35	11.01	10.06	0.613	0.490	1.25
Age	1752	17.74	0.517	17.72	0.521	17.75	0.512	-0.026	0.025	-1.05
Male (Yes=1)	1752	0.580	0.494	0.577	0.494	0.585	0.493	-0.008	0.024	-0.34
Only child (Yes=1)	1752	0.075	0.263	0.139	0.346	0.000	0.000	0.139	0.012	11.39
Number of Siblings	1752	2.192	0.663	1.997	0.540	2.420	0.719	-0.423	0.030	-14.05
Parents' average age	1751	46.06	3.178	44.91	2.816	47.40	3.050	-2.493	0.140	-17.78
Parents' average education	1748	12.10	2.538	12.46	2.471	11.68	2.553	0.781	0.120	6.48
Both parents present (Yes=1)	1752	0.922	0.269	0.921	0.270	0.923	0.266	-0.003	0.013	-0.20
Single father (Yes=1)	1752	0.027	0.162	0.030	0.170	0.024	0.152	0.006	0.008	0.79
Single mother (Yes=1)	1752	0.051	0.221	0.050	0.218	0.053	0.225	-0.004	0.011	-0.34
Books at home	1752	184.1	211.3	196.0	217.3	170.2	203.2	25.75	10.110	2.55
Family income ($W1,000$)	1730	318.8	196.7	332.4	200.9	302.9	190.6	29.47	9.465	3.11
School characteristics										
Private school (Yes=1)	1752	0.549	0.498	0.553	0.497	0.544	0.498	0.009	0.024	0.40
Coed school (Yes=1)	1752	0.464	0.499	0.463	0.499	0.465	0.499	-0.001	0.024	-0.05
Boys-only school (Yes=1)	1752	0.307	0.461	0.314	0.464	0.299	0.458	0.016	0.022	0.71
Girls-only school (Yes=1)	1752	0.229	0.420	0.222	0.416	0.237	0.425	-0.014	0.020	-0.72
Rural area (Yes=1)	1752	0.193	0.395	0.176	0.381	0.214	0.411	-0.039	0.019	-2.05
Medium city (Yes=1)	1752	0.330	0.470	0.329	0.470	0.331	0.471	-0.002	0.023	-0.08
Metropolitan city (Yes=1)	1752	0.477	0.500	0.495	0.500	0.455	0.498	0.040	0.024	1.69
First-born child (Yes=1)	1752	0.539	0.499							
First-born boy (Yes=1)	1752	0.311	0.463							
First-born girl (Yes=1)	1752	0.228	0.420							

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

	Reduced form models (OLS)				Structural models (2SLS)	
Dependent variable:	Average test score		Tutoring Expenditure (log)		Average test score	
	(1)	(2)	(3)	(4)	(5)	(6)
Tutoring Expenditure	0.097 (0.036)**	0.093 (0.037)*			0.564 (0.393)	0.536 (0.387)
First-born child		1.390 (1.101)	2.952 (0.821)**			
First-born boy				2.611 (1.054)*		
First-born girl				3.351 (1.127)**		
Hours of self-study	0.354 (0.049)**	0.353 (0.049)**	0.018 (0.036)	0.019 (0.036)	0.345 (0.051)**	0.345 (0.051)**
Prior quality (Q2)	13.723 (1.575)**	13.565 (1.583)**	3.042 (1.304)*	3.045 (1.305)*	12.133 (2.275)**	12.227 (2.253)**
Prior quality (Q3)	21.764 (1.582)**	21.617 (1.586)**	1.526 (1.284)	1.544 (1.285)	20.898 (1.953)**	20.949 (1.938)**
Prior quality (Q4)	34.124 (1.673)**	34.036 (1.677)**	1.269 (1.321)	1.269 (1.321)	33.438 (1.953)**	33.479 (1.939)**
Prior quality missing	19.289 (1.605)**	19.169 (1.608)**	-2.717 (1.260)*	-2.721 (1.260)*	20.449 (2.030)**	20.380 (2.012)**
Single father	-5.707 (2.383)*	-5.894 (2.380)*	-1.613 (2.251)	-1.615 (2.252)	-5.135 (3.216)	-5.168 (3.195)
Single mother	2.181 (2.334)	2.278 (2.338)	-3.019 (1.728)	-3.035 (1.728)	3.699 (2.756)	3.610 (2.733)
Books at home	0.004 (0.002)	0.004 (0.002)	0.001 (0.002)	0.001 (0.002)	0.004 (0.003)	0.004 (0.003)
Family income	0.002 (0.003)	0.002 (0.003)	0.021 (0.002)**	0.021 (0.002)**	-0.008 (0.009)	-0.007 (0.009)
Parents' avg edu	1.067 (0.236)**	1.051 (0.237)**	0.871 (0.169)**	0.871 (0.169)**	0.640 (0.431)	0.666 (0.425)
Parents' avg age	0.061 (0.152)	0.146 (0.166)	0.125 (0.126)	0.122 (0.126)	0.087 (0.165)	0.085 (0.164)
Age	-2.601 (0.912)**	-2.647 (0.916)**	0.402 (0.683)	0.407 (0.684)	-2.837 (0.986)**	-2.823 (0.979)**
Male	0.166 (1.503)	0.266 (1.504)	-0.698 (1.089)	-0.307 (1.326)	0.595 (1.579)	0.570 (1.569)
Only child	-2.624 (2.181)	-3.214 (2.221)	-1.200 (1.635)	-1.095 (1.648)	-2.648 (2.258)	-2.647 (2.244)
Number of siblings	-0.068 (0.920)	0.046 (0.929)	-1.647 (0.676)*	-1.644 (0.677)*	0.822 (1.211)	0.769 (1.199)
Intercept	39.69 (17.44)*	36.18 (17.52)*	15.72 (13.19)	15.52 (13.20)	28.78 (20.58)	29.42 (20.41)
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Estimates for Expenditure						
LIML					0.564 (0.393)	0.544 (0.391)
Fuller-k					0.530 (0.376)	0.512 (0.374)
95 % CI						
Wald					[-0.207, 1.335]	[-0.222, 1.294]
AR					[-0.394, 2.332]	[-0.598, 2.940]
LM					[-0.166, 1.698]	[-0.196, 1.693]
CLR					[-0.236, 1.980]	[-0.417, 2.095]
F (IVs excluded from the 2nd stage)			12.92	6.59		
P-value for overid test						0.620
R-square	0.377	0.378	0.244	0.244	0.301	0.310
Number of sample	1,480	1,480	1,480	1,480	1,480	1,480

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 IV Estimates of the Effect of Weekly Tutoring Hours on Subject Test Scores

	A. Mathematics ($N = 1,398$)		B. English ($N = 1,467$)		C. Korean language ($N = 1,467$)	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:						
Weekly hours of tutoring (log)						
First-born child	1.447 (0.526)**		2.039 (0.482)**		1.035 (0.445)*	
First-born boy		0.781 (0.672)		1.883 (0.617)**		1.165 (0.572)*
First-born girl		2.238 (0.724)**		2.221 (0.662)**		0.885 (0.609)
F (Excluded IVs)	7.57	5.05	17.91	9.03	5.42	2.77
Subject test scores						
Estimates for Tutoring hours (log)						
OLS	0.319 (0.076)**	0.319 (0.076)**	0.032 (0.074)	0.032 (0.074)	0.101 (0.082)	0.101 (0.082)
2SLS	-0.420 (1.028)	-0.374 (0.887)	1.376 (0.722)	1.320 (0.713)	2.300 (1.610)	2.224 (1.572)
LIML	-0.420 (1.028)	-0.375 (0.887)	1.376 (0.722)	1.370 (0.731)	2.300 (1.610)	2.261 (1.595)
Fuller-k	-0.333 (0.959)	-0.312 (0.842)	1.304 (0.695)	1.297 (0.704)	2.128 (1.460)	1.925 (1.394)
95 % CI						
Wald	[-2.435, 1.595]	[-2.112, 1.364]	[-0.038, 2.790]	[-0.077, 2.716]	[-0.854, 5.455]	[-0.857, 5.305]
AR	[-8.945, 2.990]	[-8.116, 2.965]	[-0.206, 4.334]	[-0.402, 4.987]	[-1.106, ∞]	[-2.312, ∞]
LM	[-4.045, 1.859]	[-2.990, 1.457]	[0.121, 3.387]	[0.088, 3.485]	[-0.226, 15.226]	[-5.327, 15.578]
CLR	[-19.372, 2.864]	[-5.854, 2.513]	[0.088, 4.073]	[-0.141, 3.942]	[-0.603, 19.246]	[-4.246, ∞]
P-value of overid test		0.928		0.406		0.754

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performance, 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 (in the usual z-test sense) at the 0.05 and 0.01 levels, respectively.

Table 4: OLS and IV Estimates of the Effect of Tutoring Expenditures on Performance by Ability Level

	A. Low-level prior performance ($N = 415$)		B. Medium-level prior performance ($N = 437$)		C. High-level prior performance ($N = 249$)	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:						
Tutoring Expenditure						
First-born child	4.940 (1.358)**		1.623 (1.353)		3.511 (2.015)	
First-born boy		4.010 (1.927)*		1.588 (1.757)		2.706 (2.648)
First-born girl		5.740 (1.843)**		1.554 (2.033)		4.495 (2.910)
F (Excluded IVs)	13.24	6.77	1.44	0.68	3.04	1.62
Average test score						
Estimates for Expenditure						
OLS	0.168 (0.078)*	0.168 (0.078)*	0.184 (0.070)**	0.184 (0.070)**	0.164 (0.072)*	0.164 (0.072)*
2SLS	0.741 (0.429)	0.726 (0.423)	0.578 (1.100)	0.591 (1.137)	0.689 (0.710)	0.585 (0.664)
LIML	0.741 (0.429)	0.734 (0.427)	0.578 (1.100)	0.813 (1.501)	0.689 (0.710)	0.631 (0.708)
Fuller-k	0.699 (0.410)	0.693 (0.408)	0.416 (0.821)	0.477 (0.944)	0.558 (0.592)	0.511 (0.592)
95 % CI						
Wald	[-0.100, 1.582]	[-0.103, 1.555]	[-1.578, 2.733]	[-1.638, 2.820]	[-0.702, 2.081]	[-0.716, 1.886]
AR	[-0.256, 2.698]	[-0.447, 3.397]	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
LM	[-0.035, 2.015]	[-0.045, 2.015]	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
CLR	[-0.095, 2.397]	[-0.146, 2.769]	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
P-value of overid test		0.659		0.473		0.562

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

Table 5: OLS and IV Estimates of the Effect of Tutoring Expenditures on Performance by Sex

	A. Male students ($N = 840$) (1)	B. Female students ($N = 640$) (2)
Dependent variables:		
Tutoring Expenditure		
First-born boy	2.772 (1.147)*	
First-born girl		2.944 (1.167)*
F (Excluded IVs)	5.84	6.37
Average test score		
Estimates for Expenditure		
OLS	0.088 (0.049)	0.123 (0.056)*
2SLS	0.936 (0.641)	0.118 (0.542)
LIML	0.936 (0.641)	0.118 (0.542)
Fuller-k	0.811 (0.568)	0.119 (0.504)
95 % CI		
Wald	[-0.320, 2.193]	[-0.943, 1.180]
AR	[-0.462, ∞]	[-4.322, 4.724]
LM	[-0.121, 5.176]	[-1.508, 1.910]
CLR	[-0.377, 14.49]	[-19.00, 18.60]

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performance, 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 (in the usual z-test sense) at the 0.05 and 0.01 levels, respectively.

Table 6: OLS and IV Estimates of the Effect of Tutoring Expenditures on Performance by Family Income

	A. Low-income Family ($N = 309$)		B. Mid-income Family ($N = 767$)		C. High-income Family ($N = 404$)	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:						
Tutoring Expenditure						
First-born child	2.845 (1.857)		4.526 (1.157)**		1.937 (1.480)	
First-born boy		2.823 (2.469)		3.956 (1.482)**		2.280 (1.889)
First-born girl		2.867 (2.468)		5.196 (1.588)**		1.515 (2.065)
F (Excluded IVs)	3.13	1.17	15.31	7.84	1.71	0.90
Average test score						
Estimates for Expenditure						
OLS	0.185 (0.082)*	0.185 (0.082)*	0.047 (0.051)	0.047 (0.051)	0.141 (0.083)	0.141 (0.083)
2SLS	0.008 (0.858)	0.016 (0.858)	0.415 (0.350)	0.377 (0.343)	0.921 (1.219)	1.067 (1.244)
LIML	0.008 (0.858)	-0.089 (1.103)	0.415 (0.350)	0.388 (0.349)	0.921 (1.219)	1.215 (1.405)
Fuller-k	0.062 (0.716)	0.024 (0.837)	0.392 (0.338)	0.366 (0.337)	0.632 (0.900)	0.792 (0.963)
95 % CI						
Wald	[-1.674, 1.691]	[-1.665, 1.697]	[-0.271, 1.101]	[-0.296, 1.050]	[-1.469, 3.311]	[-1.371, 3.505]
AR	$[-\infty, \infty]$	$[-\infty, \infty]$	[-0.457, 1.774]	[-0.638, 2.035]	$[-\infty, \infty]$	$[-\infty, \infty]$
LM	$[-\infty, \infty]$	$[-\infty, \infty]$	[-0.256, 1.352]	[-0.296, 1.332]	$[-\infty, \infty]$	$[-\infty, \infty]$
CLR	$[-\infty, \infty]$	$[-\infty, \infty]$	[-0.276, 1.553]	[-0.477, 1.633]	$[-\infty, \infty]$	$[-\infty, \infty]$
P-value of overid test		0.326		0.479		0.602

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performance, 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 (in the usual z-test sense) at the 0.05 and 0.01 levels, respectively.

Table 7: OLS and IV Estimates of the Effect of Tutoring Expenditures on Test-writing

Dependent variable:	Tutoring Expenditure (log)		Take Test (No=1)		
Estimation method:	OLS-I (1)	OLS-II (2)	2SLS-I (3)	2SLS-II (4)	OLS (5)
Tutoring Expenditure			0.001 (0.005)	0.002 (0.005)	-0.0011 (0.0006)
First-born child	3.382 (0.766)**				
First-born boy		3.129 (0.973)**			
First-born girl		3.698 (1.072)**			
Hours of self-study	0.043 (0.034)	0.043 (0.034)	-0.003 (0.001)**	-0.003 (0.001)**	-0.003 (0.001)**
Prior quality (Q2)	2.952 (1.186)*	2.950 (1.186)*	-0.123 (0.033)**	-0.125 (0.033)**	-0.116 (0.032)**
Prior quality (Q3)	2.048 (1.169)	2.059 (1.170)	-0.131 (0.031)**	-0.132 (0.031)**	-0.126 (0.031)**
Prior quality (Q4)	2.224 (1.205)	2.216 (1.206)	-0.086 (0.032)**	-0.087 (0.032)**	-0.081 (0.033)*
Prior quality missing	-2.047 (1.127)	-2.056 (1.128)	-0.073 (0.029)*	-0.072 (0.029)*	-0.077 (0.032)*
Single father	-3.559 (2.050)	-3.552 (2.050)	-0.007 (0.052)	-0.005 (0.052)	-0.014 (0.048)
Single mother	-3.002 (1.528)*	-3.021 (1.529)*	0.079 (0.041)	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	0.910 (0.155)**	0.911 (0.155)**	-0.005 (0.006)	-0.005 (0.006)	-0.003 (0.004)
Parents' avg age	0.199 (0.118)	0.197 (0.118)	-0.008 (0.003)**	-0.008 (0.003)**	-0.008 (0.002)**
Age	0.122 (0.645)	0.125 (0.645)	-0.006 (0.016)	-0.006 (0.016)	-0.005 (0.016)
Male	-0.259 (1.000)	0.046 (1.234)	0.003 (0.024)	0.003 (0.024)	0.002 (0.026)
Only child	-1.863 (1.496)	-1.782 (1.509)	0.045 (0.035)	0.045 (0.035)	0.043 (0.037)
Number of siblings	-1.123 (0.620)	-1.117 (0.621)	0.027 (0.017)	0.028 (0.017)	0.024 (0.016)
Intercept	14.04 (12.41)	13.88 (12.42)	0.817 (0.318)*	0.806 (0.319)*	0.867 (0.301)**
School characteristics	Yes	Yes	Yes	Yes	Yes
F (Excluded IVs)	19.50	9.83			
P-value for overid test				0.360	
R-square	0.254	0.254	0.092	0.089	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.

Appendix Table 1: OLS and IV Estimates of the Effect of Tutoring Expenditures on Performance with Different Functional Forms of Expenditure

Functional forms:	A. Log (expenditure+1)		B. Level of expenditure (unit: W10,000)	
	(1)	(2)	(3)	(4)
Dependent variables:				
Tutoring Expenditure				
First-born child	4.830 (1.305)**		4.720 (1.800)**	
First-born boy		4.289 (1.675)*		3.915 (2.311)
First-born girl		5.463 (1.790)**		5.660 (2.470)*
F (Excluded IVs)	13.70	6.98	6.88	3.59
Average test score				
Estimates for Expenditure				
OLS	0.058 (0.023)*	0.058 (0.023)*	0.034 (0.019)	0.034 (0.019)
2SLS	0.345 (0.240)	0.328 (0.236)	0.353 (0.262)	0.323 (0.251)
LIML	0.345 (0.240)	0.333 (0.239)	0.353 (0.262)	0.336 (0.259)
Fuller-k	0.325 (0.230)	0.314 (0.229)	0.312 (0.239)	0.297 (0.236)
95 % CI				
Wald	[-0.125, 0.815]	[-0.134, 0.791]	[-0.161, 0.866]	[-0.170, 0.816]
AR	[-0.236, 1.382]	[-0.359, 1.698]	[-0.261, 4.945]	[-0.377, ∞]
LM	[-0.095, 1.030]	[-0.113, 1.013]	[-0.085, 1.638]	[-2.035, 1.734]
CLR	[-0.131, 1.153]	[-0.166, 1.329]	[-0.191, 4.804]	[-14.698, 14.849]
P-value of overid test		0.624		0.577

Note: Each model includes controls for a constant, a student's age, sex, only-child indicator, number of siblings, prior performance, 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 (in the usual z-test sense) at the 0.05 and 0.01 levels, respectively.