To Whom Does Money Matter? Distributional Effects of Private Tutoring on Student Academic Performance

Changhui Kang
Professor of Economics
Chung-Ang University
ckang@cau.ac.kr
http://prof.cau.ac.kr/~ckang

Yoon Soo Park
Research Fellow
Korea Development Institute
yoonpark@kdi.re.kr

June, 2017

Abstract

In order to understand the effectiveness of private tutoring and its implications for educational inequality, this paper examines mean and distributional effects of private tutoring on test scores of students in South Korea. Using a nationally representative sample of middle school students over years 2006-2007, we find that students with higher predetermined test scores tend to benefit more from private tutoring. Given that students with better academic perfomance are more likely to have better family backgrounds, the results suggest that private tutoring is likely to aggravate educational inequality by disproportionately improving test scores of students with better family background and academic quality.

JEL classification: I20, C21

Keywords: private tutoring; test scores; quantile treatment effects; South Korea

I. Introduction

Education is often considered to be one of the most important policy tools to enhance economic growth and social mobility. For this reason, economists have long been interested in measuring the effectiveness of monetary educational investments on student outcomes. Previous studies in this literature have mostly focused on investments made in formal educational sectors such as public schools (Card and Krueger 1996; Krueger 2003; Hanushek 1986, 1997, 2003) and private schools (Evans and Schwab 1995; Neal 1997; Altonji, Elder, and Taber 2005). Recently, a small but growing number of studies are turning attention to measuring the effectiveness of non-formal educational investments, private tutoring. The results are fairly mixed. Some studies (Stevenson and Baker 1992; Tansel and Bircan 2005; Ha and Harphan 2005; Dang 2007; Ono 2007) report strong positive impacts, while others (Briggs 2001; Kang 2007; Gurun and Millimet 2008; Ryu and Kang 2013) present that the effects are close to zero or even negative (Lee, Kim, and Yoon 2004; Cheo and Quah 2005).

Given the lack of previous research and the conflicting views among the existing studies, we attempt to contribute to the literature by estimating the causal effect of private tutoring on achievement test scores of the student. In particular, we are interested in estimating not only the mean effect of private tutoring but also its distributional effects at different locations of a test score distribution. Understanding the distributional effects of private tutoring can offer important social implications for educational equity and intergenerational mobility.[1] It has been well known that family backgrounds are closely related with students' academic performance. For example, low-performing students are more likely to have disadvantageous family backgrounds

[1] As a matter of fact, there exist only a few studies that have examined distributional effects of formal school inputs (Eide et al., 2002; Bedard, 2003; Maasoumi et al., 2005; Lamarche, 2008; Corak and Lauzon, 2009).

than high-performing students. Thus, if private tutoring more strongly affects the lower portion of a test score distribution, it is likely to function as a remedial device for offsetting disadvantages of low-performing students and ameliorating current educational inequalities in the next generation. If private tutoring more strongly affects the upper portion of the distribution, however, current inequality is likely to worsen over time. Understanding distributional effects of private tutoring is likely to help conceive policy responses to fight against long-term inequalities and enhance intergenerational mobility.

As in causal estimation of other educational inputs, the primary difficulty in estimating the causal effect of private tutoring is endogeneity of private tutoring expenditures. In order to deal with such endogeneity, we rely on an empirical model developed by Bonhomme and Sauder (2011) that controls for potential differences in observable (e.g. student, family, and school backgrounds) and time-invariant unobservable (e.g. cognitive ability) educational inputs of students that may affect both test score outcomes and private tutoring decisions. For the empirical analysis we employ longitudinal survey data on nationally representative middle schoolers in South Korea for years 2006 to 2007, focusing on three major academic subjects – Korean, English, and math.

Applying the empirical methods for estimating mean and distributional impacts of private tutoring, we find that private tutoring has no effect on Korean test scores, while it has some effects on English and math test scores. For the Korean subject, both mean and distributional effects of private tutoring are statistically insignificant, which implies the effect of private tutoring in Korean seems to be homogenously close to zero across students at different points of the test score distribution. On the other hand, we find that the mean effects of private tutoring are about 8 and 18 percent of a standard deviation of test scores in English and math, respectively.

Looking at the distributional effects, we find that the effects are positive in the upper half of the test score distribution, while they are statistically insignificant in the lower half of the distribution. At their peaks around the 70th to 80th percentiles, the effects of private tutoring amount to roughly 36 and 39 percent of a standard deviation of English and math test scores, respectively. In sum, these results suggest that private tutoring has some effects on improving test scores on average but the extent of its effectiveness varies substantially across students with different levels of predetermined academic quality. Specifically, it seems that students with higher predetermined test scores tend to benefit more from private tutoring than their peers at the bottom of the test score distribution. This suggests that while private tutoring is not an effective remedial educational measure for students left behind, it reinforces existing inequality of student outcomes by facilitating learning processes of students with better predetermined academic quality.

II. Previous Literature

There are a few studies on the effects of private tutoring on students' academic performance. Stevenson and Baker (1992), Tansel and Bircan (2005), Ha and Harphan (2005), and Ono (2007) examine data from Japan, Turkey, Vietnam, and Japan, respectively, reporting strong positive effects of private tutoring. On the other hand, Briggs (2001) finds a negligible effect in the U.S., while Lee et al. (2004) and Choe and Quah (2005) reports even negative effects of private tutoring in Korea and Singapore, respectively. However, these studies do not explicitly deal with potential endogeneity of private tutoring and hence their results should be interpreted with caution.

More recent studies attempt to address the potential endogeneity of private tutoring in various

ways. Dang (2007) examines the effect of private tutoring expenditures on self-reported academic performance of students using nationally representative household survey data in Vietnam during 1997-1998. In order to address the endogeneity of private tutoring expenditures, Dang estimates a simultaneous equation system consisting of a Tobit model for private tutoring expenditures and an ordered probit model for the self-reported academic performance. He finds that private tutoring has a significant impact on students' academic performance. However, the measure for academic performance used by Dang (2007) is a self-reported variable, which has four ordered responses of " excellent" , " good" , " average" , and " poor." This raises a concern about whether the self-reported measure can reflect a student's true academic performance. Also, it should be noted that the joint Tobit-ordered probit model used by Dang (2007) relies on strong parametric assumptions.

While Dang (2007) deals with the endogeneity of private tutoring by explicitly modeling the process generating private tutoring expenditures and academic performance of the student using a simultaneous framework, Kang (2007) tries to find an exogenous variation that affects private tutoring decisions of parents but does not affect academic outcomes of the student. In particular, Kang uses a student's birth order as an instrumental variable (IV) for private tutoring expenditures. The logic is that parents tend to have more concerns about a first-born child's education and invest more for the student's private tutoring than for the later-borns, while a student's birth order is determined exogenously by nature and is unlikely to directly affect academic performance of the student. Indeed, in the data used by Kang (2007), Korean parents are found to spend on average about 30 percent more tutoring expenditures for their first-borns relative to the later-borns. Using the first-born indicator as an IV for tutoring expenditures, Kang finds evidence on a modest effect of private tutoring on the national college-entrance exam

scores for high school graduates in Korea: a 10 percent increase in tutoring expenditures improves only about 0.56 percentile point in the test score.

Although the IV used by Kang (2007) is fairly strong and novel in the literature, reasonable doubts arise about the validity of the exclusion restriction. If parents indeed are more concerned about their first-born child's education and invest more for the first-born's private tutoring, other parental inputs (e.g., helping children with their homework) would be also greater for the first-born than for the later-borns. To the extent that these parental inputs are related with academic outcomes of the student, the validity of the IV strategy would be questionable. Gurun and Millimet (2008) point out this issue and take a different approach. Given that a valid exclusion restriction is unavailable in their data, Gurun and Millimet tries to assess the importance of the potential endogeneity of private tutoring by using the bivariate probit framework suggested by Altonji et al. (2005, 2008). In particular, they consider

$$Y_i = 1[\beta_1 D_i + X_i\beta_2 + \varepsilon_i > 0], \quad (1)$$

$$D_i = 1[X_i\beta_3 + \lambda_i > 0], \quad (2)$$

where Y_i denotes a dummy variable that takes 1 if student i enters a university and 0 otherwise; D_i is a dummy variable that takes 1 if student i receives private tutoring and 0 otherwise; X_i is a vector of observable characteristics of student i; and ε_i and λ_i are error terms that follow a bivariate normal distribution of zero means, unit variances, and correlation coefficient ρ. The source of the potential endogeneity of D_i in equation (1) is a potential correlation between ε_i and λ_i, which is represented by $\rho \neq 0$. Following Altonji et al. (2005, 2008), Gurun and Millimet (2008) perform the following sensitivity analysis. First, they constrain ρ to be 0 and find a large positive estimate for β_1.[2] Then, they constrain ρ to take

[2] This is basically what Tansel and Bircan (2005) found. Tansel and Bircan used the same data as Gurun and Millimet (2008) did

different values and look at how the corresponding estimates for β_1 change over different values of ρ. They discover that the estimated effect of private tutoring (represented by estimated β_1) becomes statistically insignificant and even falls below zero when only a moderate level of endogeneity (represented by ρ) is allowed. Given these results, they conclude that the strong positive effects of private tutoring often reported in the previous studies may have been driven by the potential endogeneity problem.

Recently, Ryu and Kang (2013) extend Kang (2007)'s study by employing alternative empirical strategies. Specifically, they employ the monotone instrumental variable (MIV) strategy by Manski and Pepper (2000) and try to estimate the *bounds* of the causal effect of private tutoring on test scores. In the standard IV strategy, an IV is not allowed to affect an outcome variable directly (i.e., exclusion restriction). By contrast, in the MIV strategy by Manski and Pepper (2000), an MIV may affect an outcome variable directly but only *monotonically* (i.e., either positively or negatively). A child's first-born status, which was used as an IV in Kang (2007), may not be a valid IV for private tutoring expenditures because the first-born may receive a larger parental support than the later-borns in many other forms than private tutoring. However, the first-born indicator can still be used as an MIV of Manski and Pepper (2000) as long as a child's birth order is *monotonically* related with the student's academic performance. Using a child's first-born status as an MIV for private tutoring expenditures, Ryu and Kang (2013) find that the estimated upper bound for the causal effect of private tutoring expenditures is small, which implies the causal effect is likely to be close to zero.

The empirical strategy of Ryu and Kang (2013) is more advanced than that of Kang (2007) in that it draws the same conclusion while relaxing a restrictive IV assumption to a less restrictive

---

but they assumed that private tutoring was exogenously determined.

MIV assumption. However, in order to derive the sharp bounds, Ryu and Kang (2013) had to employ two additional monotonicity assumptions: (1) the effect of private tutoring expenditures on test scores is non-negative for all students (monotone treatment response: MTR) and (2) the non-random selection into a larger amount of private tutoring expenditures is positive on average (monotone treatment selection: MTS). Given that there are studies reporting evidence of negative effects of private tutoring (Lee et al. 2004; Choe and Quah 2005) and negative selection into a larger amount of private tutoring expenditures (Gurun and Millimet 2008), the MTS and MTR assumptions seem to be restrictive and the validity of the causal inference based on those assumptions will be limited.

The conflicting results from the previous studies on private tutoring may be partly because they use data from different periods and countries. However, a perhaps more important reason may be because it is daunting to deal with the potential endogeneity of private tutoring. Unlike the studies on the effectiveness of educational investments in formal schooling sectors where experimental (e.g., the STAR experiment in Krueger and Whitmore (2001)) and/or quasi-experimental (e.g., private school voucher programs in McEwan (2004)) data are often available, the previous studies on private tutoring have had to rely on observational data for a causal inference. To overcome the challenge given constraints, they try to either model the data generating process econometrically (Dang 2007; Gurun and Millimet 2008) or find an exclusion restriction that affects private tutoring behaviors but not (or in a limited way) students' outcomes (Kang 2007; Ryu and Kang, 2013).

As in the previous studies, we do not find an alternative plausible source of exogenous variation in students' (or their parents') private tutoring decisions from our data. Given this limitation, we attempt to estimate the mean and distributional effects of private tutoring, relying

on a longitudinal nature of the data. Specifically, we employ the econometric model developed by Bonhomme and Sauder (2011) that controls for students' observable and time-invariant unobservable educational inputs to deal with endogeneity of private turoing. This empirical model will be described in section IV.

III. Data

The data for this study are from the Korea Education Longitudinal Study (KELS). The KELS is an annual longitudinal survey whose basic structure is similar to that of the National Educational Longitudinal Studies of the U.S. (Ryu and Kang 2013). A nationally representative sample of 6,908 seventh graders (age 13) was first surveyed in 2005 and followed every year since then.[3] The KELS also surveyed parents, teachers, and school principals of each of the 6,908 students in order to collect information on family and school characteristics of the student.

We measure academic performance of the student by test scores of three academic subjects – Korean, English, and math. These test scores, which take a value between 0 and 100, are available in the first three waves (years 2005, 2006, and 2007) of the KELS data.[4] Among the three waves, information on private tutoring experience of the student is available in the second and the third waves (years 2006 and 2007).[5] We mainly use these two waves for this paper. In

[3] In the Korean educational system, seventh grade is the first year of middle school.

[4] Concerned about a situation where every student's test score may simultaneously rise as all of them take private tutoring but an empirical analysis based on scores normalized by the mean and standard deviation fails to capture a positive effect of private tutoring, we rely on raw test scores that take a value between 0 and 100. If normalized test scores are used instead of raw scores, however, the primary results of this paper are not affected. They are available upon request.

[5] In fact, the first wave (year 2005) of the KELS data also includes information on private tutoring. However, a careful reading of the parental questionnaires of the KELS shows that the reference period of the questions on private tutoring has changed

each year of 2006 and 2007, the KELS asked parents of the 6,908 students whether or not the student had received private tutoring during the survey year for each subject of Korean, English, and math, separately. Based on the parents' responses to this question, we divide students into two groups – treatment and control groups – based on whether they have received private tutoring in 2007 (treatment group) or not (control group). In order to compare test scores of students in the treatment and control groups at a common baseline, we control for students' predetermined academic quality as well as other predetermined characteristics. Specifically, we control for test scores and private tutoring status of the student as well as their individual, parental, and school characteristics that are observed in 2006.

[INSERT TABLE 1 HERE]

[INSERT TABLE 2 HERE]

[INSERT TABLE 3 HERE]

Tables 1, 2, and 3 show summary statistics of our Korean, English, and math samples, respectively. After removing observations with missing information on the variables we use in this study, we have 4,073, 4,464, and 4,574 valid observations for the Korean, English, and math samples, respectively.[6] In each of the three tables, columns (1) to (6) show summary statistics of

---

between the first wave (year 2005) and the other two waves (years 2006 and 2007). In 2005, the KELS asks a student's private tutoring experience during the survey month (Oct. 2005). On the other hand, in 2006 and 2007, it asks a student's private tutoring experience during the entire survey year (2006 or 2007). Since pupils may receive different amounts of tutoring in different seasons (Bray and Kwok 2003), we choose to focus on the second and third waves of the KELS during which the survey collected the information on private tutoring in a consistent way.

[6] In addition to the observations that have missing information on key variables, we remove students whose subject scores between 2006 and 2007 tests are more than 50 points apart, since they seem to have neglected one of the tests. The number of these students is 53, 153 and 123 for the Korean, English and math sample, respectively. While empirical results of this paper are drawn without them, the primary findings of this paper are not altered qualitatively if we include them in the analysis samples.

the entire sample, the treatment group, and the control group, respectively. In all of the three subjects, students in the treatment group tend to report better academic performance than those in the control group. In the Korean achievement test, students who receive private tutoring score slightly higher on average than those who do not receive private tutoring by about 1.68 points. This is about 8 percent of a standard deviation of Korean test scores of the treatment group. By contrast, students in the treatment group achieve on average 18.0 and 18.4 points higher in the English and math tests than those not receiving private tutoring, which amount to about 69 and 72 percent of a standard deviation of the treatment group, respectively. The tables also show that the treatment and control groups are also very different from each other in terms of other dimensions as well. For example, even before the treatment is realized, students in the treatment group tend to score higher in the achievement tests and are more likely to receive private tutoring than those in the control group. The two groups also show a large difference in their student, parental, and school characteristics. In the following section, we discuss how to control for these differences in observable characteristics and potential differences in *time-invariant* unobservable characteristics between the two groups.

IV. Empirical Analysis

A. Mean Effects of Private Tutoring

For each of the Korean, English, and math samples, we have data on

$$(Y_{i1}, Y_{i2}, D_i, X_i), \tag{3}$$

where Y_{i1} and Y_{i2} denote test scores of student i in 2006 (denoted by period 1 hereafter)

The results are available upon request.

and 2007 (denoted by period 2 hereafter) periods, respectively; D_i is our treatment indicator that takes 1 if student i receives private tutoring in period 2 and 0 otherwise; and X_i is a vector of individual, parental, and school characteristics of student i.

In order to examine whether the improvement in test scores is greater for students in the treatment group than those in the control group, we begin by estimating the following equation by the ordinary least square (OLS) method:

$$Y_{i2} - Y_{i1} = \beta_0 + \beta_1 D_i + X_i \beta_2 + \varepsilon_i, \quad (4)$$

where the dependent variable is a change in test score of student i which is calculated by subtracting a pre-treatment test score from a post-treatment test score of student i, $(Y_{i2} - Y_{i1})$. The error term, ε_i, denotes unobserved and unmeasured characteristics of student i that might affect academic achievement of the student. If private tutoring (D_i) is as good as randomly assigned among students who share the same observable baseline characteristics (X_i) and hence D_i is uncorrelated with ε_i conditional on X_i, the OLS estimation of equation (4) will yield a consistent estimate for β_1.

[INSERT TABLE 4 HERE]

Table 4 shows the OLS estimation results of equation (4) for the Korean, English, and math sample, respectively. In column (1), as a benchmark, we do not control for any observable characteristics, and report the simple mean difference in changes in test scores between the treatment and control groups. In columns (2), (3) and (4), we progressively add student, parental and school characteristics to the list of control variables. Regardless of the choice of the covariate specifications, we find evidence that students benefit from receiving private tutoring in English and math but not in Korean. When the individual, parental, and school characteristics of the student are controlled for, test scores of students who receive private tutoring tend to improve

more than those of students who do not by about 2.8 and 4.1 points in English and math, respectively. However, we find no statistically significant difference in the changes in test scores between the two groups for Korean.

The OLS estimator in Table 4 has some weaknesses that can be improved on. First, the estimator requires a parametric functional form of equation (4). More importantly, in order for the OLS estimate for β_1 to have a causal interpretation, we need the so-called selection-on-observables assumption which presumes that students who receive private tutoring are comparable to those who do not as long as differences in observable characteristics between the two groups are controlled for. Although this assumption has been used extensively in the previous studies (Stevenson and Baker 1992; Briggs 2001; Lee et al. 2004; Tansel and Bircan 2005; Ha and Harphan 2005; Choe and Quah 2005; Ono 2007), it is easy to think of why this assumption may not hold. Students receiving private tutoring are likely to be different from their peers in many unobservable, but perhaps very important, ways. For example, it is possible that students with higher motivation are also more likely to use private tutoring to enhance their academic achievement further. On the other hand, it is also possible that students with a lower level of cognitive ability are more likely to seek private tutoring lest they should be left behind in their school classes. To the extent that there are unobservable characteristics that are correlated with academic performance of the student but unbalanced between the treatment and control groups, the OLS estimates in Table 4 would not reflect a causal relationship.

To address, at least partly, these problems, we employ a semiparametric estimation technique developed by Bonhomme and Sauder (2011).[7] We begin by explicitly specifying the average causal effect of interest using the potential outcomes framework by Rubin (1974). Note that our

[7] Our illustration of the method is heavily drawn from sections II and III of Bonhomme and Sauder (2011).

treatment (D_i) is realized in period 2. Hence, the observed test score of student i in period 2 (Y_{i2}) would be either of two potential test scores depending on whether the student receives the treatment or not. Specifically, Y_{i2} can be written as

$$Y_{i2} = D_i Y_{i2}^1 + (1 - D_i) Y_{i2}^0 \quad (5)$$

where Y_{i2}^1 (or Y_{i2}^0) denote the potential test score of student i in period 2 had they received (or not received) private tutoring during that period.

The causal effect of receiving private tutoring on the achievement test score of student i can be measured by $Y_{i2}^1 - Y_{i2}^0$, the difference between the two potential test score outcomes of the student. Since we can only observe either Y_{i2}^1 or Y_{i2}^0 for a given student, however, the causal effect is fundamentally unidentifiable at the individual level. Instead, we try to identify the average value of the causal effects for students who receive private tutoring in 2007, which is often referred to as the average treatment effect on the treated or ATT:

$$ATT = E[Y_{i2}^1 - Y_{i2}^0 \mid D_i = 1] \quad (6)$$

$$= E[Y_{i2} \mid D_i = 1] - E[Y_{i2}^0 \mid D_i = 1] \quad (7)$$

In equation (7), $E[Y_{i2} \mid D_i = 1]$ is empirically observable, while $E[Y_{i2}^0 \mid D_i = 1]$ is counterfactual and hence unobservable. To identify the counterfactual $E[Y_{i2}^0 \mid D_i = 1]$, we model the potential post-treatment test score of student i in the state of world where she does not receive private tutoring (Y_{i2}^0) and the baseline test score of student i (Y_{i1}) as the sum of the following three components (Bonhomme and Sauder 2011):

$$Y_{i1} = f_1(X_i) + \eta_i + v_{i1} \quad (8)$$

$$Y_{i2}^0 = f_2^0(X_i) + \eta_i + v_{i2}^0, \quad (9)$$

where X_i denotes student, parental, and school characteristics of student i whose effects on test scores are flexibly modeled as arbitrary functions of $f_1(\cdot)$ and $f_2^0(\cdot)$; η_i represents

unobservable characteristics of student i which are fixed between the two periods (e.g. cognitive ability); v_{i1} and v_{i2}^0 represent time-varying unobservable shocks to test scores (e.g. physical conditions on the exam day) that are allowed to be correlated with each other in an arbitrary way. These equations may be viewed as an educational production function (Hanushek 1986) that relates observable (X_i) and unobservable ($\eta_i,\ v_{i1},\ v_{i2}^0$) educational inputs to test score outputs ($Y_{i1},\ Y_{i2}^0$) in each period. Except for its additive structure, the educational production function is fairly flexible in that it does not impose any distributional or functional-form restrictions on its three components.

Given that the educational production function is modeled by equations (8) and (9), we impose the following assumption in order to identify the ATT:

Assumption 1: (v_{i1}, v_{i2}^0) are independent of D_i conditional on X_i

Assumption 1 requires that there is no systematic difference in time-varying unobservables (v_{i1}, v_{i2}^0) of test scores between those in the treatment and those in the control group as long as they have the same observable characteristics (X_i). Note that the educational production functions in equations (8) and (9) divide unobservable educational inputs into two parts: time-invariant unobservables (η_i) and time-varying ones ($v_{i1},\ v_{i2}^0$). In this respect, assumption 1 may be referred to as *selection-on-observable-and-time-invariant-unobservable* assumption in that it restricts the determinants of potential non-random selection into private tutoring to be observables or time-invariant unobservables or both.

Under assumption 1, the ATT in equation (7) can be identified as (Abadie 2005; Bonhomme and Sauder 2011)

$$ATT = \frac{1}{\Pr[D_i=1]} E\left[\left\{\frac{D_i-\Pr[D_i=1 \mid X_i]}{1-\Pr[D_i=1 \mid X_i]}\right\}(Y_{i2}-Y_{i1})\right], \quad (10)$$

for which we need a usual common support assumption:

$$\textbf{Assumption 2}: \Pr[D_i = 1] > 0 \; and \; Pr[D_i = 1 \mid X_i] < 1 \; with \; probability \; 1$$

Details on how to derive equation (10) is provided in Appendix A1. Following Bonhomme and Sauder (2011), we estimate the ATT in equation (10) by

$$\widehat{ATT} = \frac{1}{\frac{1}{N}\sum_{i=1}^{N} D_i} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{D_i - \widehat{\Pr}[D_i = 1 \mid X_i]}{1 - \widehat{\Pr}[D_i = 1 \mid X_i]} (Y_{i2} - Y_{i1}) \right] \quad (11)$$

Since X_i consists of many covariates including continuous variables, we estimate $\widehat{\Pr}[D_i = 1 \mid X_i]$ in equation (11) by a logit regression of D_i on X_i in order to avoid the curse of dimensionality problem. When computing the $\widehat{ATT}$ in equation (11), we restrict our estimation sample to observations with $.05 < \widehat{\Pr}[D_i = 1 \mid X_i] < .95$ in order to make sure that the common support assumption (assumption 2) would hold. We compute standard errors of the $\widehat{ATT}$ by bootstrapping with 2000 iterations.

[INSERT TABLE 5 HERE]

Table 5 reports estimation results for $\widehat{ATT}$ in equation (11). For Korean, the estimated mean effects are statistically insignificant regardless of the choice of covariate specifications. This suggests that students on average do not benefit from receiving private tutoring for the subject. For English and math, when all of the student, parental, and school characteristics are controlled for, we find mean effects of 2.00 and 4.64 points, which are roughly 8 and 18 percent of a standard deviation of the test score, respectively. This suggests that private tutoring has a positive effect on academic performance of students in those subjects.[8]

[8] Evaluated at the mean value of the test score, such estimates imply that a 10-percent increase in expenditures raises the average test score of English and math by 0.36 and 0.88 percent, respectively. These amounts of the effect generally agree with the findings of Kang (2007) and Ryu and Kang (2013). Such magnitudes are, however, much smaller than the amount of

B. Distributional Effects of Private Tutoring

The empirical model in section IV.A evaluates the *average* effects of receiving private tutoring on test scores. We find that private tutoring improves test scores of students on average by about 2.00 and 4.64 points in English and math, respectively, while it exerts no effect in Korean. Although estimating the average effect of private tutoring is crucial for understanding how private tutoring affects academic performance of the student, the estimates may miss some important features of the effects of private tutoring. For example, private tutoring may be more helpful to students who are left behind in their school classes than those who are already in good standing for themselves. It is also possible that more advanced students can make better use of private tutoring to enhance their academic performance further and, hence, private tutoring may be most effective for students at the top of a test score distribution. In order to account for this potential heterogeneity of how private tutoring affects academic achievement of the student, we employ the model by Bonhomme and Sauder (2011) and estimate the *distributional* effect of receiving private tutoring on test scores at each percentile point of a test score distribution. In particular, our object of interest in this section is the quantile treatment effect on the treated (QTT) which is defined as:

$$QTT(\tau) = F^{-1}_{Y_{i2}|D_i=1}(\tau) - F^{-1}_{Y^0_{i2}|D_i=1}(\tau),\ \tau \in (0,1), \tag{12}$$

where $\tau \in (0,1)$ represents a percentile point of a test score distribution and $F^{-1}_W(\cdot)$ denotes the inverse of the cumulative distribution function (CDF) of a random variable W. Note that the distribution of $Y_{i2}|D_i = 1$ – post-treatment test scores of students who receive private tutoring –

improvement in the test score (2.8 to 3.6 percent) due to a 10-percent increase in public school expenditures in the U.S. summarized by Krueger (2003). Our estimates are more analogous to the effect sizes suggested by Guryan (2001) in terms of test scores (0.77 to 1.15 percent), and by Card and Krueger (1996) in terms of labor market earnings (0.7 to 1.1 percent).

is empirically observable and hence it is straightforward to estimate $F^{-1}_{Y_{i2}|D_i=1}(\cdot)$ nonparametrically. The key issue is how to estimate $F^{-1}_{Y^0_{i2}|D_i=1}(\cdot)$ because $Y^0_{i2}|D_i = 1$ – *potential* post-treatment test scores of students who receive private tutoring had they not received private tutoring – is counterfactual and hence unobservable. Bonhomme and Sauder (2011) provide conditions under which the counterfactual distribution of $Y^0_{i2}|D_i = 1$ can be identified and estimated. Following their approach, we maintain all the assumptions that we made in section IV.A (the educational production functions of equations (8) and (9) and assumptions 1 and 2). On top of those assumptions, we further assume that:

Assumption 3: (v_{i1}, v^0_{i2}) are independent of η_i conditional on X_i and D_i.

This condition presumes that the temporal shocks to test scores (v_{i1} and v^0_{i2}) are independent of time-invariant unobservables (η_i) among students who share the same observable characteristics (X_i and D_i). For example, this assumption excludes the possibility that students with high levels of motivation and cognitive ability (represented by η_i) face systematically higher (or lower) patterns of physical conditions on their exam days (represented by v_{i1} and v^0_{i2}) conditional on their observable characteristics (X_i and D_i). Given that the educational production function takes an additive structure of equations (8) and (9) and that assumptions 1, 2, and 3 hold, the probability density function (PDF) of the counterfactual $Y^0_{i2}|D_i = 1$ is identified as (Bonhomme and Sauder 2011)

$$f_{Y^0_{i2}|D_i=1}(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-jty)\frac{1}{\Pr[D_i=1]}E[\omega(t \mid X_i)(1-D_i)\exp(jtY_{i2})]dt, \quad (13)$$

where $j = \sqrt{-1}$, $t \in \mathbf{R}$, and

$$\omega(t \mid X_i) \equiv \frac{E[D_i exp(jtY_{i1}) \mid X_i]}{E[(1-D_i) exp(jtY_{i1}) \mid X_i]} \quad (14)$$

The details on the identification procedure are presented in Appendix A2. Following Bonhomme and Sauder (2011), we estimate the counterfactual density in equation (13) with

$$\hat{f}_{Y_{i2}^0|D_i=1}(y) = \frac{1}{2\pi}\int_{-T_N}^{T_N} \exp(-jty)\frac{1}{\frac{1}{N}\sum_{i=1}^{N} D_i}\left(\frac{1}{N_i}^{N} \widehat{\omega}(t \mid X_i)(1-D_i)\exp(jtY_{i2})\right)dt \qquad (15)$$

where $j = \sqrt{-1}$, $t \in \mathbf{R}$, and

$$\widehat{\omega}(t \mid X_i) = \frac{\hat{E}[D_i exp(jtY_{i1}) \mid X_i]}{\hat{E}[(1-D_i) exp(jtY_{i1}) \mid X_i]} \qquad (16)$$

Since X_i consists of many covariates including continuous variables, we approximate the conditional expectations in the numerator and denominator of equation (16) with linear projections of $D_i exp(jtY_{i1})$ and $(1-D_i)exp(jtY_{i1})$ onto X_i, respectively, in order to avoid the curse of dimensionality. We choose the trimming parameter T_N in equation (15), which is analogous to choosing a bandwidth in nonparametric density estimation, by the rule of thumb method suggested by Diggle and Hall (1993).[9] We compute the integration using the trapezoid rule with 200 equidistant nods.

The estimation results of equation (15) for each of the three subjects are presented in the left columns of Figures 1, 2, and 3. In each of the three figures, we report estimation results for three different covariate specifications. In plot A, we include only student characteristics in X_i of equation (15), while, in plots B and C, we augment parental and school characteristics to the list of covariates. In all of the plots, solid lines represent the kernel density estimates for the realized test score distribution of students who receive private tutoring ($Y_{i2}|D_i = 1$).[10] Dashed lines represent the counterfactual test score distribution of students who receive private tutoring had

[9] Details on how we determine the values of T_N are given in Appendix A3.

[10] When estimating the density of $Y_{i2}|D_i = 1$, we use the Gaussian kernel with the rule of sum bandwidth suggested by Silverman (1986).

they not received it ($Y_{i2}^0|D_i = 1$). Since we compare a realized test score distribution of students receiving private tutoring with a counterfactual test score distribution of the same group of students, any difference between the two distributions can be attributable to the causal effect of private tutoring.

[INSERT FIGURE 1 HERE]

[INSERT FIGURE 2 HERE]

[INSERT FIGURE 3 HERE]

By integrating the estimated densities in Figures 1, 2, and 3 over the range of test scores, we calculate the estimated CDFs of $Y_{i2}|D_i = 1$ and $Y_{i2}^0|D_i = 1$ for each of the three subjects. Given the results, we estimate the quantile treatment effects on the treated (QTT) in equation (12) by

$$\widehat{QTT}(\tau) = \hat{F}^{-1}_{Y_{i2}|D_i=1}(\tau) - \hat{F}^{-1}_{Y_{i2}^0|D_i=1}(\tau), \tau \in (0,1), \quad (17)$$

where $\hat{F}_W^{-1}(\cdot)$ denotes the inverse of the estimated CDF of a random variable W. We compute standard errors of the $\widehat{QTT}(\tau)$ by bootstrapping with 2000 iterations.[11]

The estimation results for equation (17) are in the right columns of Figures 1, 2, and 3. As in the mean effects, the patterns of the distributional effects for Korean are different from those for English and math. Recall that the mean effects of receiving private tutoring on Korean test scores are statistically indistinguishable from zero. Similarly, we find no statistically significant effect throughout the Korean test score distribution, either. This suggests that private tutoring has negligible effects on Korean test scores homogenously across students with different levels of academic quality.

[11] Following Hall (1992) and Horowitz (2001), when estimating the bootstrap standard errors, we use a four-time larger trimming parameter (i.e. undersmoothing) than the one chosen to compute the point estimates in equation (15).

As opposed to the results for Korean, the estimation results for English and math reveal some important heterogeneity in the effects of private tutoring which is not captured by simply looking at its mean effects. Regardless of the choice of the covariate specifications, the quantile treatment effects are at most modest or statistically insignificant at lower tails of the test score distribution. However, the effects rise to be positive and statistically significant in the middle of the distribution and become largest around the 70th to 80th percentiles where they amount to about 10 points. After reaching their peaks, the effects revert to modest or statistically insignificant levels as moving up to upper tails of the test score distribution. This may be because students at the top of the distribution have already scored almost 100 points, the maximum possible points of the achievement tests, before they receive the treatment, and hence any potential positive treatment effects for these top students cannot be captured by their achievement test scores. For example, the 90th percentiles of the baseline test scores of the students in the treatment group are 94 and 90 points in English and math, respectively. In sum, the QTT estimation results imply that private tutoring is helpful for students at the upper half of the test score distribution but not for those at the lower half of the distribution. This suggests that private tutoring mainly facilitates learning processes of students in good standing rather than serving as a remedial educational measure for students who are left behind.

C. Falsification Test

In order to confirm that our main results in Table 5 and Figures 1, 2, and 3 are not mistakenly drawn by a model misspecification, we perform the following falsification test. We estimate the effect of receiving private tutoring in 2007 on test scores in 2005 which is determined *before* the

treatment is realized and hence should not be affected by the treatment. In particular, we compute the $\widehat{ATT}$ in equation (11) and the $\widehat{QTT}(\tau)$ in eqution (17) using the pre-determined test scores in 2005 as a new outcome variable instead of test scores in 2007.

Table 6 and Figures 4, 5, and 6 show the results of the falsification test. For all the subjects and specifications, we do not find any statistically significant effects. These results suggest that the estimated ATT and QTT results in Table 5 and Figures 1, 2, and 3 do not seem to be driven by model misspecifications, but reflect the causal effects of private tutoring that we intend to measure.

[INSERT TABLE 6 HERE]

[INSERT FIGURE 4 HERE]

[INSERT FIGURE 5 HERE]

[INSERT FIGURE 6 HERE]

D. Extensions

The QTT estimation results in Figures 1, 2, and 3 indicate that distributional effects of private tutoring, if any, are positive at upper part of a test score distribution but statistically insignificant from zero at lower part of the distribution. We interpret these results as evidence suggesting that the effect of private tutoring varies substantially across students with different levels of pre-determined academic quality. However, such observed patterns of distributional effects could also emerge simply because students at upper part of test score distribution tend to receive a larger amount of private tutoring while the effect of private tutoring is indeed homogenous across students with varying levels of academic quality.

To check whether the observed patterns of distributional effects reflect the *heterogeneity of*

treatment effect or are simply driven by the *heterogeneity of treatment intensity*, we divide our treatment group into halves by relative treatment intensity and re-estimate the QTT in equation (17) by using each of the high-intensity and low-intensity groups as a new treatment group. In particualr, we divide students who receive priate tutoring into those whose private tutoring expenditures are greater than the median level (high-expenditure group) and those whose expenditures are smaller than or equal to the median (low-expenditure group).[12] We then compare each of the high-expenditure and low-expenditure groups with the no-expenditure group. Figure 7 summarizes the estimation results. Regradless of the choice of treatment intensity, we find a similar pattern of distributional effects to those in Figures 1, 2, and 3. For Korean, we do not find statistically significant effects. For English and math, we find that the effect of private tutoring tend to be larger at upper percentiles of the test score distribution, although error bounds become larger probably due to smaller sample size.[13] These results suggest that the observed patterns of distributional effects in Figures 1, 2, and 3 largely demonstrate heterogeneity of treatment effects across students with different levels of academic quality rather than being simply driven by heterogeneity of treatment intensity.

[INSERT FIGURE 7 HERE]

[12] The median values of private tutoring expenditures among those who receive private tutoring are 127, 205, and 197 (in 1,000 Korean Won) for Korean, English, and math samples, respectively.

[13] This is mainly because we use only half of those who receive private tutoring (i.e. students with above-median tutoring expenditures and those with below-median expenditures) as a treatment group in this section. Another reason for the reduction of the sample size is that many parents did not report the detail amount of expenditures on private tutoring for their children. Students with missing information on private tutoring expenditures are about 18, 15, and 15 percent in our Korean, English, and math samples, respectively.

V. Summary and Conclusion

We estimate the causal effects of receiving private tutoring on academic performance of nationally representative middle school students in Korea during 2006-2007. We use a binary indicator that takes 1 if a student receives private tutoring and 0 otherwise as our treatment variable. In this respect, our estimation results may be understood as measuring the overall effect of receiving private tutoring. Our measures for the academic performance are test scores for three major academic subjects: Korean, English, and math.

Using the data, we first examine the average effect of private tutoring. The raw mean test score gaps between students who receive private tutoring and those who do not are 1.68, 18.0, and 18.4 points for Korean, English, and math, respectively, when the maximum possible point of each test is 100. When the differences in observables (student, parental, and school characteristics) and time-invariant unobservables (e.g., cognitive ability) between the two groups are controlled for, the estimated mean effects reduce to 0.579 (statistically insignificant), 2.00, and 4.64 points for Korean, English, and math, respectively. These results suggest that the observed test score gaps between students who receive private tutoring and those who do not are largely driven by a positive selection into private tutoring in terms of both observable and time-invariant unobservable characteristics.

We also attempt to contribute to the literature by estimating distributional effects of private tutoring, which has rarely been discussed in the previous studies. For Korean, we cannot reject no effect throughout the entire test score distribution, which implies that the effect of private tutoring is homogenously negligible across students at different points of the test score distribution. For English and math, however, we find positive effects in the upper half of the test score distribution but no effect in the lower half of the distribution. The effects reach their peaks

around the 70th to 80th percentiles of the distribution where they amount to roughly 10 points for both subjects. We find statistically insignificant effects around the top of the distributions. However, this seems to be simply because the top students have already scored close to 100, the maximum possible point of the achievement test, at their baselines and hence their test scores have a narrow margin to be improved.

These results contribute to the literature by finding that the extent of the effectiveness of private tutoring varies substantially across students with their predetermined academic quality. In particular, we find that students with better academic quality tend to benefit more from private tutoring. Given that students with better academic perfomance are more likely to have better family backgrounds, the results suggest that private tutoring is likely to aggravate educational inequality by disproportionately improving test scores of students with better family background and academic quality.

Appendix

A1. Details on deriving equation (10)

This section illustrates how we apply the model by Bonhomme and Sauder (2011) to our study and derive equation (10). We draw on section II.B of Bonhomme and Sauder (2011) for the following illustration of their method.

The ATT can be written as

$$ATT = E[Y_{i2}^1 - Y_{i2}^0 \mid D_i = 1] = \int E[Y_{i2}^1 - Y_{i2}^0 \mid X_i, D_i = 1] dP(X_i \mid D_i = 1)$$

$$= \int \{E[Y_{i2} \mid X_i, D_i = 1] - E[Y_{i2}^0 \mid X_i, D_i = 1]\} \, dP(X_i \mid D_i = 1), \quad \text{(A1)}$$

where $E[Y_{i2}^0 \mid X_i, D_i = 1]$ in equation (A1) needs to be identified. By the additive structure of the educational production function in equations (8) and (9) and the selection on observables and time-invariant unobservables assumption (assumption 1), $E[Y_{i2}^0 \mid X_i, D_i = 1]$ is identified as

$$E[Y_{i2}^0 \mid X_i, D_i = 1] = E[Y_{i2} | X_i, D_i = 0] + E[Y_{i1} | X_i, D_i = 1] - E[Y_{i1} | X_i, D_i = 0] \quad \text{(A2)}$$

Substituting (A2) into (A1), the ATT is identified as the following difference-in-differences estimand:

$$ATT = \int \{E[Y_{i2} - Y_{i1} \mid X_i, D_i = 1] - E[Y_{i2} - Y_{i1} \mid X_i, D_i = 0]\} dP(X_i \mid D_i = 1) \quad \text{(A3)}$$

Estimating equation (A3) nonparametrically is infeasible due to the curse of dimensionality problem. To proceed, we use the Lemma 3.1 in Abadie (2005) following Bonhomme and Sauder (2011). Under the selection on observables and time-invariant unobservables assumption (assumption 1) and the common support assumption (assumption 2), Abadie (2005) has shown that

$$E[Y_{i2}^1 - Y_{i2}^0 \mid X_i, D_i = 1] = E[\omega_i (Y_{i2} - Y_{i1}) \mid X_i], \quad \text{(A3)}$$

where

$$\omega_i = \frac{D_i - P[D_i=1 \mid X_i]}{P[D_i=1 \mid X_i]P[D_i=0 \mid X_i]} \tag{A4}$$

Summing (A3) over the conditional distribution of $X_i \mid D_i = 1$, the ATT can be written as (Abadie 2005)

$$\begin{aligned} ATT &= \int E[\omega_i(Y_{i2} - Y_{i1}) \mid X_i] dP(X_i \mid D_i = 1) \\ &= \int E[\omega_i(Y_{i2} - Y_{i1}) \mid X_i] \frac{\Pr[D_i=1 \mid X_i]}{\Pr[D_i=1]} dP(X_i) \\ &= \frac{1}{\Pr[D_i=1]} \int E[\omega_i \Pr[D_i = 1 \mid X_i](Y_{i2} - Y_{i1}) \mid X_i] dP(X_i) \end{aligned} \tag{A5}$$

Substituting (A4) into (A5) yields equation (10).

A2. Details on deriving equation (13)

This section illustrates how we apply the model of Bonhomme and Sauder (2011) to our study and derive equation (13). We draw on section II.B of Bonhomme and Sauder (2011) for the following illustration of their method.

For any real-valued random variable W, it is known that its probability density function can be obtained by the inverse Fourier transformation of its characteristic function, $\Psi_W(t) \equiv E[\exp(jtW)]$:

$$f_W(w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-jtw) \Psi_W(t) dt, \tag{A6}$$

where $j = \sqrt{-1}$ and $t \in R$. Hence, identifying the characteristic function of $Y_{i2}^0 | D_i = 1$ suffices the identification of its density.

Let $\Psi_{Y_{i2}^0 | D_i=1}(t)$ denote the characteristic function of $Y_{i2}^0 | D_i = 1$. By definition,

$$\Psi_{Y_{i2}^0 | D_i=1}(t) = E[\exp(jtY_{i2}^0) | D_i = 1] = \int \Psi_{Y_{i2}^0 | D_i=1, X_i}(t|X_i) dP(X_i|D_i = 1) \tag{A7}$$

Bonhomme and Sauder (2011; Theorem 2) has shown that, when educational production

functions take the form of equations (8) and (9), and assumptions 1, 2, and 3 hold, the conditional characteristic function of the counterfactual $Y_{i2}^0|D_i = 1$ is identified as a function of three conditional characteristic functions of the realized $Y_{i2}|D_i = 0$, $Y_{i1}|D_i = 1$, and $Y_{i1}|D_i = 0$:

$$\Psi_{Y_{i2}^0|D_i=1,X_i}(t|x) = \frac{\Psi_{Y_{i1}|D_i=1,X_i}(t|x)}{\Psi_{Y_{i1}|D_i=0,X_i}(t|x)}\Psi_{Y_{i2}|D_i=0,X_i}(t) \tag{A8}$$

Substituting (A8) into (A7) yields:

$$\Psi_{Y_{i2}^0|D_i=1}(t) = \int \frac{\Psi_{Y_{i1}|D_i=1,X_i}(t|x)}{\Psi_{Y_{i1}|D_i=0,X_i}(t|x)}\Psi_{Y_{i2}|D_i=0,X_i}(t|x)dP(X_i|D_i = 1)$$

$$= \int \frac{\Psi_{Y_{i1}|D_i=1,X_i}(t|x)}{\Psi_{Y_{i1}|D_i=0,X_i}(t|x)}\Psi_{Y_{i2}|D_i=0,X_i}(t|x)\frac{P(D_i=1|X_i)}{P(D_i=1)}dP(X_i)$$

$$= \frac{1}{P(D_i=1)}\int \frac{P(D_i=1|X_i)}{P(D_i=0|X_i)}\frac{\Psi_{Y_{i1}|D_i=1,X_i}(t|x)}{\Psi_{Y_{i1}|D_i=0,X_i}(t|x)}\Psi_{Y_{i2}|D_i=0,X_i}(t|x)P(D_i = 0|X_i)dP(X_i) \tag{A9}$$

Note that it holds that

$$\Psi_{Y_{i2}|D_i=0,X_i}(t|x)P(D_i = 0|X_i) = E[\exp(jtY_{i2})|D_i = 0, X_i]P(D_i = 0|X_i)$$

$$= E[(1 - D_i)\exp(jtY_{i2})|X_i] \tag{A10}$$

Similarly,

$$\Psi_{Y_{i1}|D_i=0,X_i}(t|x)P(D_i = 0|X_i) = E[(1 - D_i)\exp(jtY_{i1})|X_i] \tag{A11}$$

$$\Psi_{Y_{i1}|D_i=1,X_i}(t|x)P(D_i = 1|X_i) = E[D_i\exp(jtY_{i1})|X_i] \tag{A12}$$

Substituting (A10), (A11), and (A12) into (A9) yields

$$\Psi_{Y_{i2}^0|D_i=1}(t) = \frac{1}{p_D}E[\omega(t|X_i)(1 - D_i)\exp(jtY_{i2})], \tag{A13}$$

where

$$\omega(t|X_i) \equiv \frac{E[D_i exp(jtY_{i1})|X_i]}{E[(1-D_i)exp(jtY_{i1})|X_i]}$$

Substituting equation (A13) into equation (A6) yields equation (13).

A3. Details on choosing the truncation parameter in equation (15)

According to Diggle and Hall (1993), an optimal T_N must satisfy (Bonhomme and Sauder 2011)

$$\log\left|\widehat{\Psi}_{Y_{i2}^0|D_i=1}(T_N)\right| = -\frac{1}{2}\log N$$

Figure A1 plots the estimated $\log\left|\widehat{\Psi}_{Y_{i2}^0|D_i=1}(t)\right|$ against t^2 for the math sample when student, parental, and school characteristics are used as covariates. Following Bonhomme and Sauder (2009, 2011), we extrapolate the (almost) linear part of $\log\left|\widehat{\Psi}_{Y_{i2}^0|D_i=1}(t)\right|$ and find the value of t where the extrapolated line crosses $-\frac{1}{2}\log N$. This yields $t \approx \sqrt{.010} = .100$, which we use as the T_N. For other subject samples and covariate specifications, we determine T_N in a similar way.

[INSERT FIGURE A1 HERE]

References

Abadie, Alberto. 2005. Semiparametric difference-in-differences estimators. The Review of Economic Studies 72, no. 1: 1-19.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools. Journal of Political Economy 113, no. 1: 151-84.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2008. Using selection on observed variables to assess bias from unobservables when evaluating Swan-Ganz catheterization. The American Economic Review 98, no. 2: 345-50.

Bedard, Kelly. 2003. School quality and the distribution of male earnings in Canada. Economics of Education Review 22, no. 4: 395-407.

Bonhomme, Stéphane and Ulrich Sauder. 2009. Accounting for unobservables in comparing selective and comprehensive schooling. Working paper no. 0906. CEMFI. Madrid, Spain.

Bonhomme, Stéphane and Ulrich Sauder. 2011. Recovering distributions in difference-in-differences models: a comparison of selective and comprehensive schooling. The Review of Economics and Statistics 93, no. 2: 479–94.

Bray, Mark and Percy Kwok. 2003. Demand for private supplementary tutoring: conceptual considerations, and socio-economic patterns in Hong Kong. Economics of Education Review 22, no. 6: 611-20.

Briggs, Derek. C. 2001. The effect of admissions test preparation: evidence from NELS: 88. Chance 14, no. 1: 10-8.

Card, David and Alan B. Krueger. 1996. School resources and student outcomes: an overview of the literature and new evidence from North and South Carolina. The Journal of Economic Perspectives 10, no. 4: 31-50.

Cheo, Roland and Euston Quah. 2005. Mothers, maids and tutors: an empirical evaluation of their effect on children's academic grades in Singapore. Education Economics 13, no. 3: 269–85.

Corak, Miles and Darren Lauzon. 2009. Differences in the distribution of high school achievement: The role of class-size and time-in-term. Economics of Education Review 28, no. 2: 189-198.

Dang, Hai-Anh. 2007. The determinants and impact of private tutoring classes in Vietnam. Economics of Education Review 26, no. 6: 683–98.

Dang, Hai-Anh and F. Halsey Rogers. 2008. The growing phenomenon of private tutoring: does it deepen human capital, widen inequalities, or waste resources? The World Bank Research Observer 23, no. 2: 161–200.

Diggle, Peter J. and Peter Hall. 1993. A Fourier approach to nonparametric deconvolution of a density estimate. Journal of the Royal Statistical Society. Series B (Methodological) 55, no. 2: 523-31.

Eide, E.R., M.H. Showalter and D.P. Sims. 2002. The effects of secondary school quality on the distribution of earnings. Contemporary Economic Policy 20, no. 2: 160-170.

Evans, William N. and Robert M. Schwab. 1995. Finishing high school and starting college: do Catholic schools make a difference? The Quarterly Journal of Economics 110, no. 4: 941–74.

Guryan, Jonathan. 2001. Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts. National Bureau of Economic Research Working Paper No. 8269. National Bureau of Economic Research, Cambridge: MA.

Gurun, Ayfer and Daniel L. Millimet. 2008. Does private tutoring payoff? IZA discussion paper no. 3637. The Institute for the Study of Labor (IZA). Bonn. Germany.

Ha, Tran Thu and Trudy Harpham. 2005. Primary education in Vietnam: extra classes and outcomes. International Education Journal 6, no. 5: 626–34.

Hall, Peter. 1992. Effect of bias estimation on coverage accuracy of bootstrap confidence intervals for a probability density. Annals of Statistics 20, no. 2: 675-94.

Hanushek, Eric. A. 1986. The economics of schooling: production and efficiency in public schools. Journal of Economic Literature 24, no. 3: 1141–77.

Hanushek, Eric. A. 1997. Assessing the effects of school resources on student performance: an update. Educational Evaluation and Policy Analysis 19, no. 2: 141–64.

Hanushek, Eric. A. 2003. The failure of input-based schooling policies. The Economic Journal 113: F64–F98.

Horowitz, Joel L. 2001. The bootstrap. In Handbook of econometrics vol. 5, ed. James J. Heckman and Edward E. Leamer.: Elsevier.

Kang, Changhui. 2007. Does money matter? The effect of private educational expenditures on academic performance. Working paper no. 0704, Department of Economics, National University of Singapore, Singapore.

Krueger, Alan B. 2003. Economic considerations and class size. The Economic Journal 113: F34–F63.

Krueger, Alan B. and Diane M. Whitmore. 2001. The effect of attending a small class in the early grades on college-test taking and middle school test results: evidence from project STAR. The Economic Journal 111: 1–28.

Lamarche, Carlos. 2008. Private school vouchers and student achievement: A fixed effects quantile regression evaluation. Labour Economics 15, no. 4: 575-590.

Lee, Jong-Tae, Yang-Boon Kim, and Cho-Hee Yoon. 2004. The effects of pre-class tutoring on student achievement: challenges and implications for public education in Korea. KEDI Journal of Educational Policy 1, no. 1: 25–42.

Manski, Charles F. and John V. Pepper. 2000. Monotone instrumental variables: with an application to the returns to schooling. Econometrica 68, no. 4: 997–1010.

Maasoumi, E., D.L. Millimet and V. Rangaprasad. 2005. Class size and educational policy: Who benefits from smaller classes?. Econometric Reviews 24, no. 4: 333-368.

McEwan, Patrick J.. 2004. The potential impact of vouchers. Peabody Journal of Education 79, no. 3: 57–80.

Neal, Derek A. 1997. The effects of Catholic secondary schooling on educational achievement. Journal of Labor Economics 15, no. 1: 98–123.

Ono, Hiroshi. 2007. Does “examination hell” pay off? A cost–benefit analysis of “ronin” and college education in Japan. Economics of Education Review 26, no. 3: 271–84.

Park, Yoon Soo. 2013. Essays in Empirical Microeconomics. PhD diss., Department of Economics, Princeton University.

Rubin, Donald B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology 66, no. 5: 688-701.

Ryu, Deockhyun and Changhui Kang. 2013. Do private tutoring expenditures raise academic performance? Evidence from middle school students in South Korea. Asian Economic Journal 27, no. 1: 59–83.

Silverman, Bernard W. 1986. Density estimation for statistics and data analysis. London, U.K.: Chapman & Hall.

Stevenson, David Lee. and David .P. Baker. 1992. Shadow education and allocation in formal schooling: transition to university in Japan. American Journal of Sociology 97, no. 6: 1639–57.

Tansel, Aysıt and Fatma Bircan. 2005. Effect of private tutoring on university entrance examination performance in Turkey. IZA discussion paper no. 1609, The Institute for the Study of Labor (IZA), Bonn, Germany.

Tansel, Aysıt and Fatma Bircan. 2006. Demand for education in Turkey: a Tobit analysis of private tutoring expenditures. Economics of Education Review 25, no. 3: 303–13.

Table 1. Summary Statistics (Korean)

	Total		Treatment		Control	
	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)	Mean (5)	S.D. (6)
Treatment (in 2007)						
Private tutoring (yes=1)	.570	.495	1.00	0.00	0.00	0.00
Outcome (in 2007)						
Test scores	57.9	20.6	58.7	20.2	57.0	21.0
Baseline characteristics (in 2006)						
Student characteristics						
Test scores	60.1	18.7	60.8	18.6	59.3	18.8
Private tutoring (yes=1)	.542	.498	.727	.445	.296	.456
Female (yes=1)	.482	.500	.430	.495	.550	.498
First born (yes=1)	.499	.500	.528	.499	.461	.499
Number of siblings	1.21	.724	1.18	.679	1.26	.777
Disabled (yes=1)	.019	.138	.019	.136	.020	.140
Parental characteristics						
Average age	42.3	4.07	42.1	3.66	42.6	4.55
Average years of education	12.8	2.22	13.0	2.10	12.6	2.35
Married (yes=1)	.896	.305	.931	.253	.850	.357
Monthly income (1000 KRW)	3378	2394	3642	2461	3027	2256
Having a religion (yes=1)	.685	.464	.698	.459	.670	.471
School characteristics						
Large city (yes=1)	.465	.499	.474	.499	.453	.498
Medium city (yes=1)	.447	.497	.458	.498	.433	.496
Rural area (yes=1)	.088	.284	.068	.252	.115	.319
Private school (yes=1)	.205	.404	.199	.399	.213	.409
Coed school (yes=1)	.640	.480	.643	.479	.637	.481
Boy-only school (yes=1)	.188	.390	.204	.403	.166	.372
Girl-only school (yes=1)	.172	.378	.153	.360	.197	.398
Grade size (# of students)	299	149	313	142	280	155
Class size (# of students)	35.4	5.47	35.8	5.01	35.0	5.99
Number of observations	4073		2321		1752	

Table 2. Summary Statistics (English)

	Total		Treatment		Control	
	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)	Mean (5)	S.D. (6)
Treatment (in 2007)						
Private tutoring (yes=1)	.756	.429	1.00	0.00	0.00	0.00
Outcome (in 2007)						
Test scores	56.0	26.5	60.4	26.1	42.4	23.0
Baseline characteristics (in 2006)						
Student characteristics						
Test scores	57.1	24.1	60.9	23.8	45.2	20.9
Private tutoring (yes=1)	.746	.435	.869	.338	.365	.482
Female (yes=1)	.489	.500	.478	.500	.523	.500
First born (yes=1)	.504	.500	.532	.499	.417	.493
Number of siblings	1.20	.706	1.17	.652	1.31	.843
Disabled (yes=1)	.019	.135	.018	.132	.021	.144
Parental characteristics						
Average age	42.3	3.99	42.1	3.65	42.7	4.88
Average years of education	12.9	2.21	13.2	2.13	12.0	2.21
Married (yes=1)	.903	.297	.937	.243	.796	.403
Monthly income (1000 KRW)	3427	2320	3727	2363	2494	1900
Having a religion (yes=1)	.689	.463	.699	.459	.657	.475
School characteristics						
Large city (yes=1)	.469	.499	.491	.500	.401	.490
Medium city (yes=1)	.447	.497	.440	.496	.468	.499
Rural area (yes=1)	.084	.278	.069	.254	.131	.337
Private school (yes=1)	.200	.400	.200	.400	.201	.401
Coed school (yes=1)	.641	.480	.648	.478	.620	.486
Boy-only school (yes=1)	.186	.389	.186	.389	.188	.391
Girl-only school (yes=1)	.173	.378	.167	.373	.192	.394
Grade size (# of students)	304	149	318	144	261	158
Class size (# of students)	35.6	5.42	36.0	4.98	34.2	6.42
Number of observations	4464		3377		1087	

Table 3. Summary Statistics (Math)

	Total		Treatment		Control	
	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)	Mean (5)	S.D. (6)
Treatment (in 2007)						
Private tutoring (yes=1)	.761	.426	1.00	0.00	0.00	0.00
Outcome (in 2007)						
Test scores	52.9	26.0	57.3	25.7	38.9	22.0
Baseline characteristics (in 2006)						
Student characteristics						
Test scores	52.6	24.3	56.2	24.2	41.1	20.9
Private tutoring (yes=1)	.743	.437	.861	.346	.368	.483
Female (yes=1)	.493	.500	.484	.500	.521	.500
First born (yes=1)	.507	.500	.532	.499	.427	.495
Number of siblings	1.20	.699	1.16	.643	1.31	.844
Disabled (yes=1)	.020	.139	.019	.136	.022	.147
Parental characteristics						
Average age	42.3	3.98	42.1	3.60	42.7	4.97
Average years of education	12.9	2.22	13.2	2.13	12.0	2.26
Married (yes=1)	.903	.295	.936	.245	.799	.401
Monthly income (1000 KRW)	3458	2314	3767	2359	2473	1851
Having a religion (yes=1)	.694	.461	.705	.456	.657	.475
School characteristics						
Large city (yes=1)	.468	.499	.492	.500	.390	.488
Medium city (yes=1)	.446	.497	.442	.497	.460	.499
Rural area (yes=1)	.086	.280	.065	.247	.150	.357
Private school (yes=1)	.201	.401	.200	.400	.202	.402
Coed school (yes=1)	.639	.480	.646	.478	.614	.487
Boy-only school (yes=1)	.186	.389	.186	.389	.187	.390
Girl-only school (yes=1)	.175	.380	.167	.373	.200	.400
Grade size (# of students)	305	149	320	143	255	158
Class size (# of students)	35.6	5.43	36.1	4.97	34.0	6.45
Number of observations	4574		3482		1092	

Table 4. OLS Estimation Results: Private Tutoring and Change in Test Scores

Dependent variable:	Specifications			
Change in test scores	(1)	(2)	(3)	(4)
A. Subject: Korean				
Private tutoring	.246	.576	.483	.492
(S.E.)	.509	.564	.569	.568
Covariates:				
Student characteristics	No	Yes	Yes	Yes
Parental characteristics	No	No	Yes	Yes
School characteristics	No	No	No	Yes
R-squared	.000	.004	.006	.014
Number of observations	4073	4073	4073	4073
B. Subject: English				
Private tutoring	2.33	2.87	2.69	2.80
(S.E.)	.560	.676	.692	.694
Covariates:				
Student characteristics	No	Yes	Yes	Yes
Parental characteristics	No	No	Yes	Yes
School characteristics	No	No	No	Yes
R-squared	.004	.007	.009	.013
Number of observations	4464	4464	4464	4464
C. Subject: Math				
Private tutoring	3.35	4.05	4.00	4.14
(S.E.)	.649	.765	.786	.783
Covariates:				
Student characteristics	No	Yes	Yes	Yes
Parental characteristics	No	No	Yes	Yes
School characteristics	No	No	No	Yes
R-squared	.006	.008	.008	.021
Number of observations	4574	4574	4574	4574

Note. The outcome variable is a change in achievement test scores of students between December 2006 and November 2007. The treatment variable (private tutoring) is an indicator that takes 1 if a student has ever received private tutoring in 2007 and 0 otherwise. Covariates include (1) student characteristics: a dummy for having ever received private tutoring in 2006, a dummy for female, a dummy for being handicapped, number of siblings; (2) parental characteristics: parents' average age, parents' average years of education, a dummy for being married, parents' average monthly income, and a dummy for having a religion; and (3) school characteristics: a dummy for being located in a metropolitan area, a dummy for being located in a suburban area, a dummy for private school, a dummy for boy-only school, a dummy for girl-only school, logarithm of grade size, and class size.

Table 5. Mean Effects of Private Tutoring

Dependent variable:	Specifications		
Test scores in 2007	(1)	(2)	(3)
A. Subject: Korean			
Estimated ATT	.511	.616	.579
(S.E.)	.654	.657	.583
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4073	4073	4073
B. Subject: English			
Estimated ATT	2.58	1.96	2.00
(S.E.)	.774	.863	.863
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4464	4464	4464
C. Subject: Math			
Estimated ATT	4.22	4.41	4.64
(S.E.)	.915	1.02	1.00
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4574	4574	4574

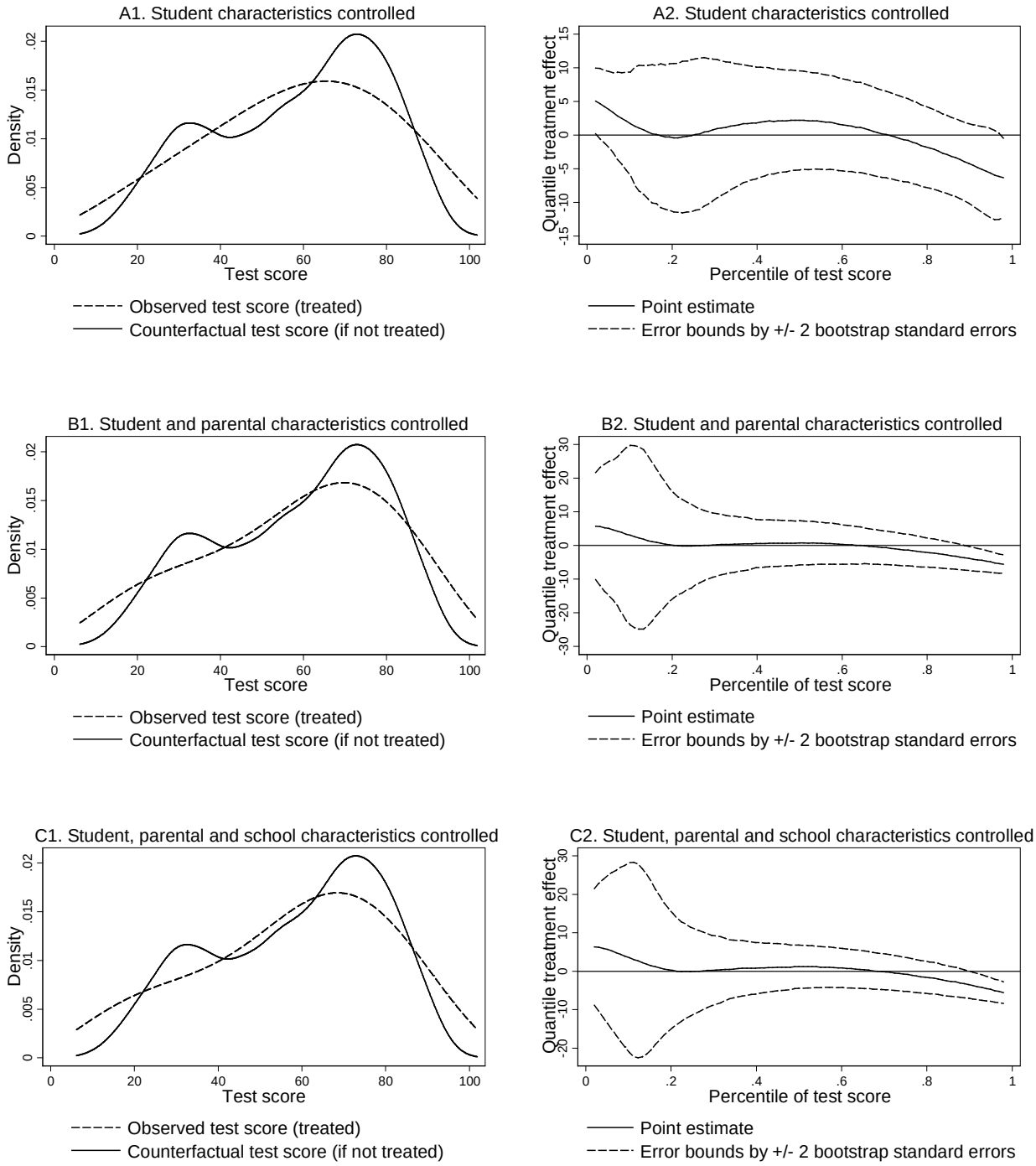
Note. The outcome variable is achievement test scores of students measured in November 2007. The treatment variable (private tutoring) is an indicator that takes 1 if a student has ever received private tutoring in 2007 and 0 otherwise. Covariates include (1) student characteristics: a dummy for having ever received private tutoring in 2006, a dummy for female, a dummy for being handicapped, number of siblings; (2) parental characteristics: parents' average age, parents' average years of education, a dummy for being married, parents' average monthly income, and a dummy for having a religion; and (3) school characteristics: a dummy for being located in a metropolitan area, a dummy for being located in a suburban area, a dummy for private school, a dummy for boy-only school, a dummy for girl-only school, logarithm of grade size, and class size. Standard errors are computed by bootstrap of 2000 iterations.

Table 6. Falsification Test Results: Mean Effects of Private Tutoring

Dependent variable:	Specifications		
Test scores in 2005	(1)	(2)	(3)
A. Subject: Korean			
Estimated ATT	.178	.110	.090
(S.E.)	.626	.654	.640
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4016	4016	4016
B. Subject: English			
Estimated ATT	-.286	-.043	-.308
(S.E.)	.776	.888	.904
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4442	4442	4442
C. Subject: Math			
Estimated ATT	-.899	-1.19	-1.14
(S.E.)	.911	1.01	1.03
Covariates:			
Student characteristics	Yes	Yes	Yes
Parental characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Number of observations	4515	4515	4515

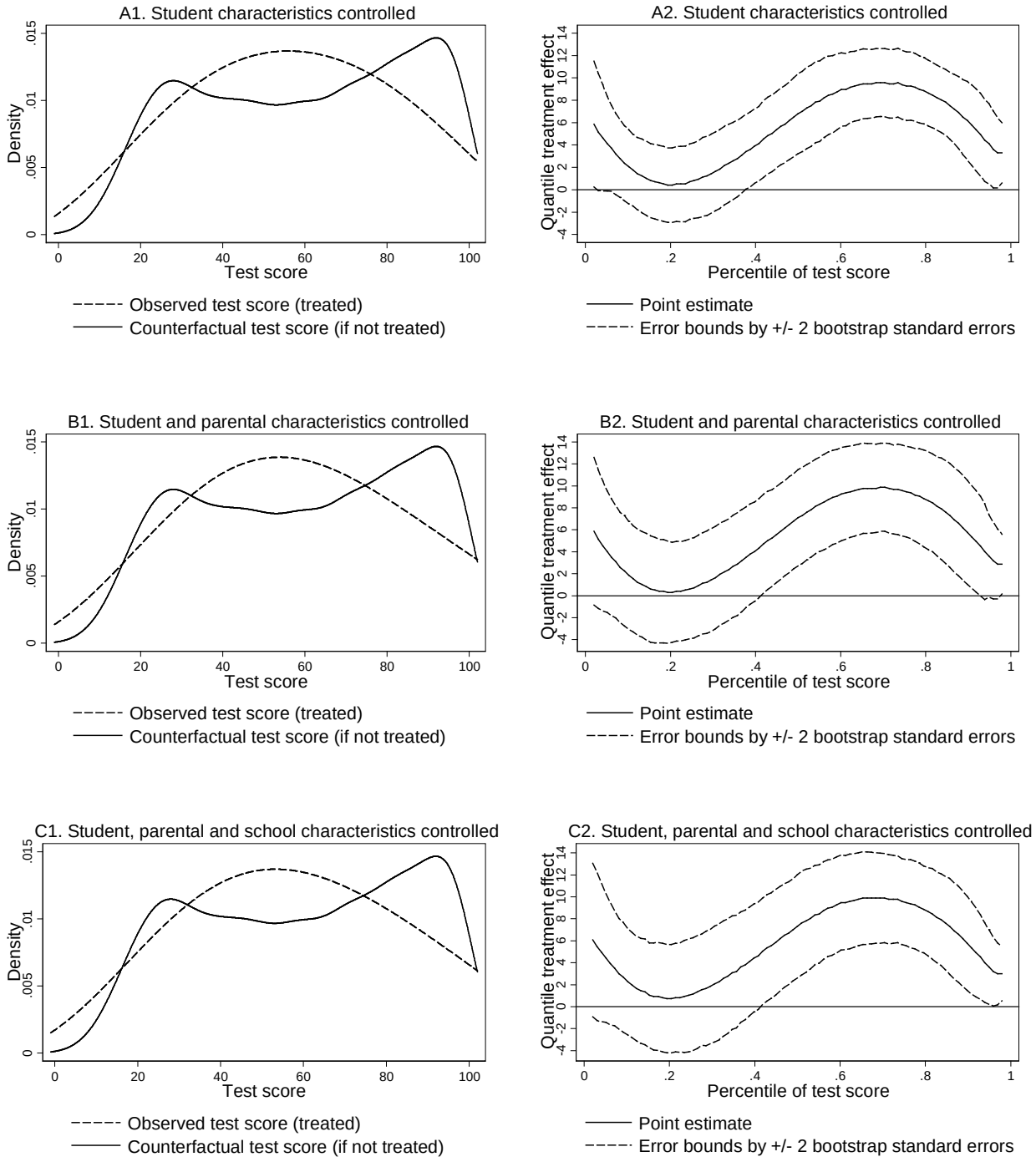
Note. The outcome variable is achievement test scores of students measured in December 2005. The treatment variable (private tutoring) is an indicator that takes 1 if a student has ever received private tutoring in 2007 and 0 otherwise. Covariates include (1) student characteristics: a dummy for having ever received private tutoring in 2006, a dummy for female, a dummy for being handicapped, number of siblings; (2) parental characteristics: parents' average age, parents' average years of education, a dummy for being married, parents' average monthly income, and a dummy for having a religion; and (3) school characteristics: a dummy for being located in a metropolitan area, a dummy for being located in a suburban area, a dummy for private school, a dummy for boy-only school, a dummy for girl-only school, logarithm of grade size, and class size. Standard errors are computed by bootstrap of 2000 iterations.

Figure 1. Distributional Effects of Private Tutoring (Korean)



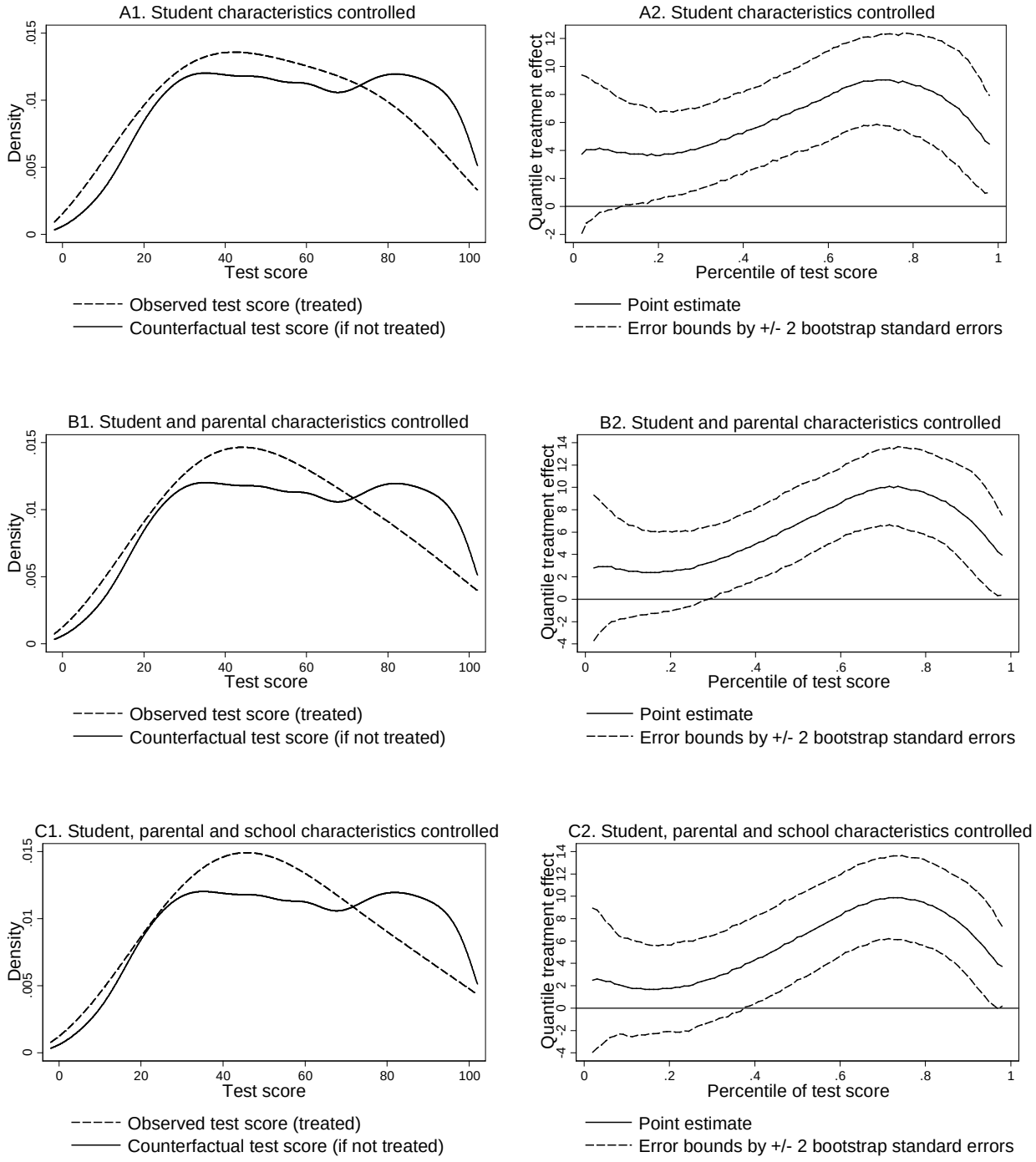
Note. The outcome variable is achievement test scores measured in November 2007. The treatment variable is an indicator for having ever received private tutoring in 2007. The same covariates used in table 5 are controlled for. Plots in the left column compare density estimates for the realized test score distribution of the treatment group (students receiving private tutoring in 2007) with those for the counterfactual test score distribution of the same students had they not received private tutoring. Plots in the right column report estimated quantile treatment effect on the treated (QTT). Standard errors for the QTT estimates are computed by bootstrap of 2000 iterations.

Figure 2. Distributional Effects of Private Tutoring (English)



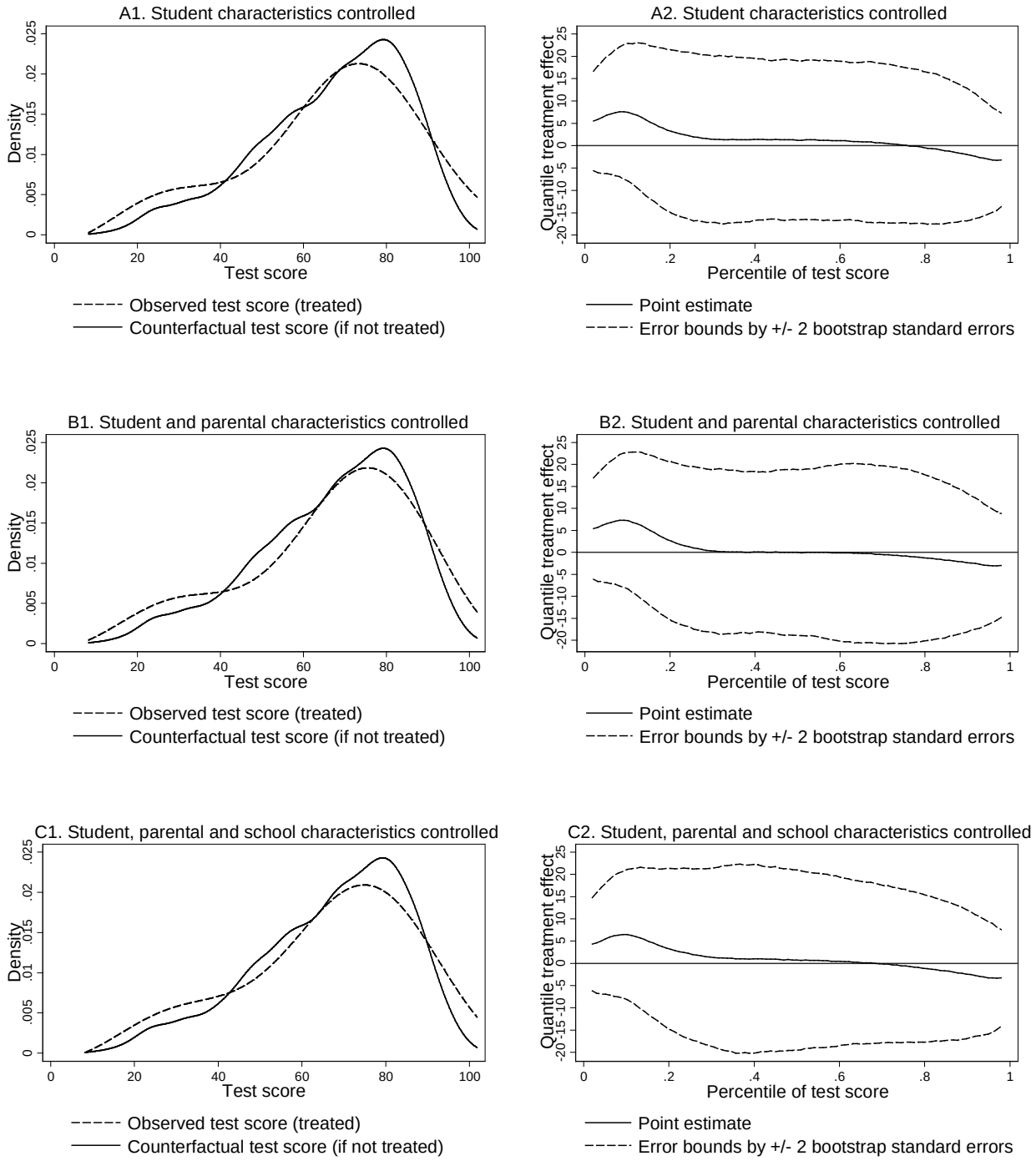
Note. The outcome variable is achievement test scores measured in November 2007. The treatment variable is an indicator for having ever received private tutoring in 2007. The same covariates used in table 5 are controlled for. Plots in the left column compare density estimates for the realized test score distribution of the treatment group (students receiving private tutoring in 2007) with those for the counterfactual test score distribution of the same students had they not received private tutoring. Plots in the right column report estimated quantile treatment effect on the treated (QTT). Standard errors for the QTT estimates are computed by bootstrap of 2000 iterations.

Figure 3. Distributional Effects of Private Tutoring (Math)



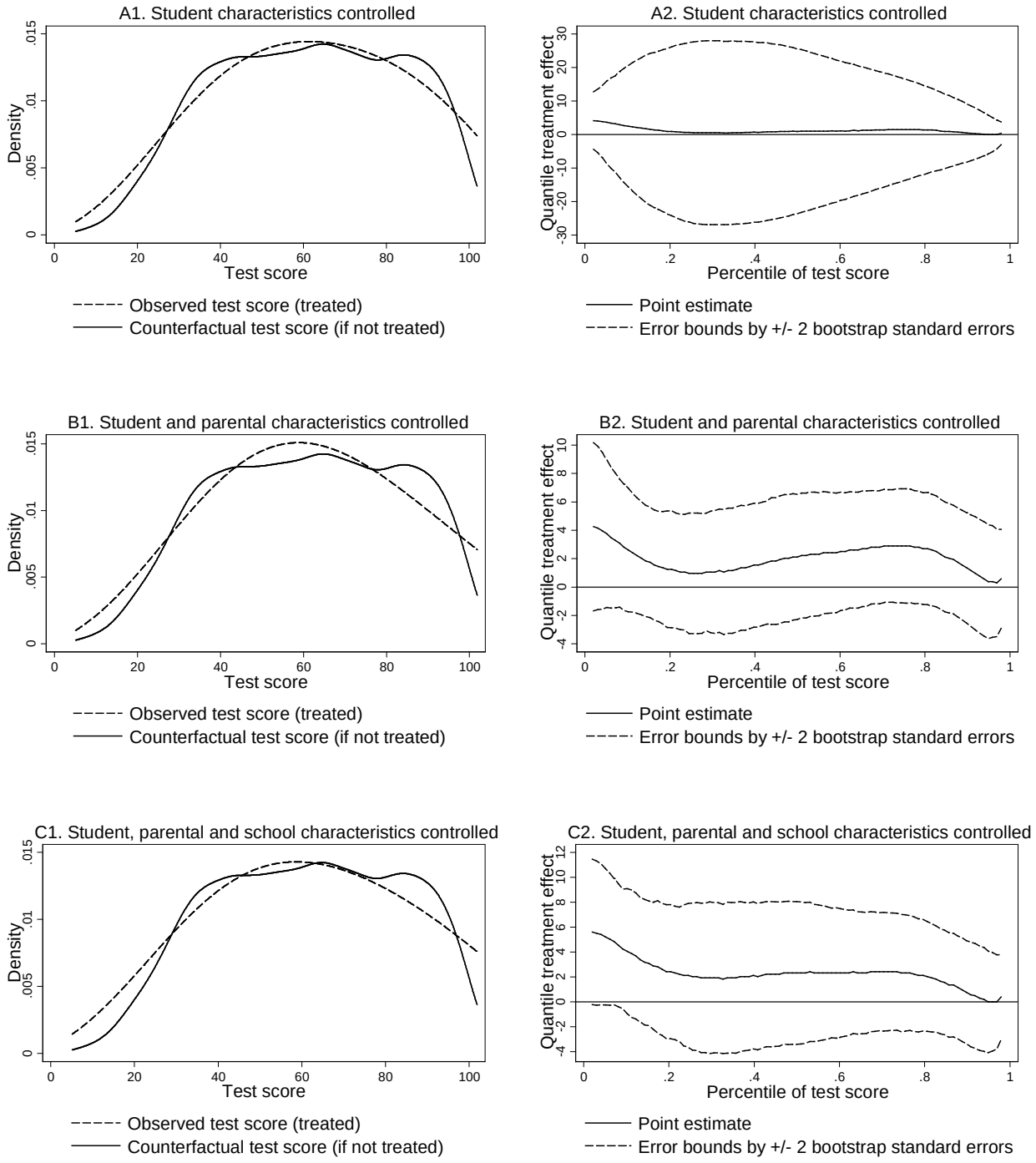
Note. The outcome variable is achievement test scores measured in November 2007. The treatment variable is an indicator for having ever received private tutoring in 2007. The same covariates used in table 5 are controlled for. Plots in the left column compare density estimates for the realized test score distribution of the treatment group (students receiving private tutoring in 2007) with those for the counterfactual test score distribution of the same students had they not received private tutoring. Plots in the right column report estimated quantile treatment effect on the treated (QTT). Standard errors for the QTT estimates are computed by bootstrap of 2000 iterations.

Figure 4. Falsification Test Results: Distributional Effects of Private Tutoring (Korean)



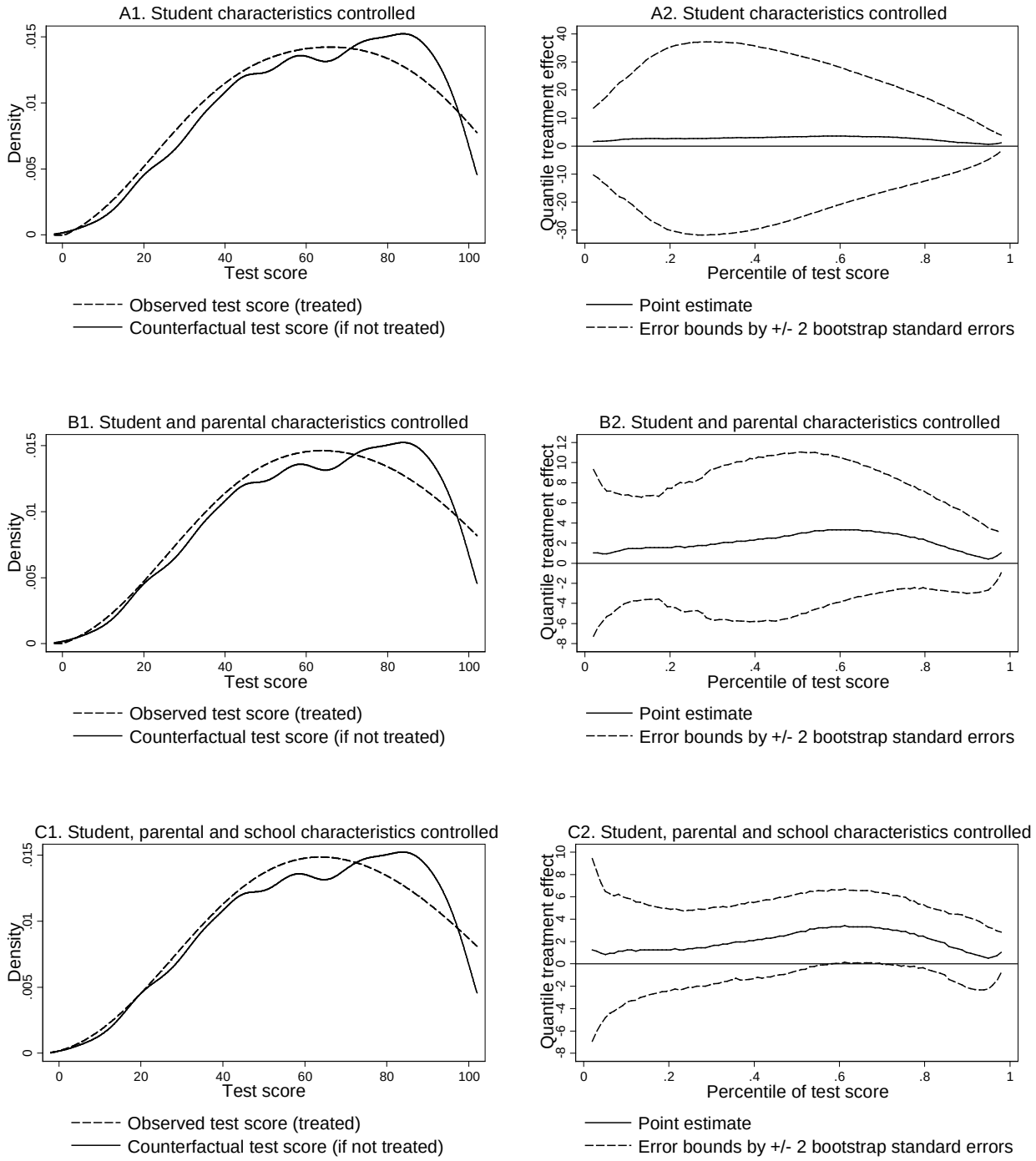
Note. The outcome variable is achievement test scores measured in December 2005. The treatment variable is an indicator for having ever received private tutoring in 2007. The same covariates used in table 5 are controlled for. Plots in the left column compare density estimates for the realized test score distribution of the treatment group (students receiving private tutoring in 2007) with those for the counterfactual test score distribution of the same students had they not received private tutoring. Plots in the right column report estimated quantile treatment effect on the treated (QTT). Standard errors for the QTT estimates are computed by bootstrap of 2000 iterations.

Figure 5. Falsification Test Results: Distributional Effects of Private Tutoring (English)



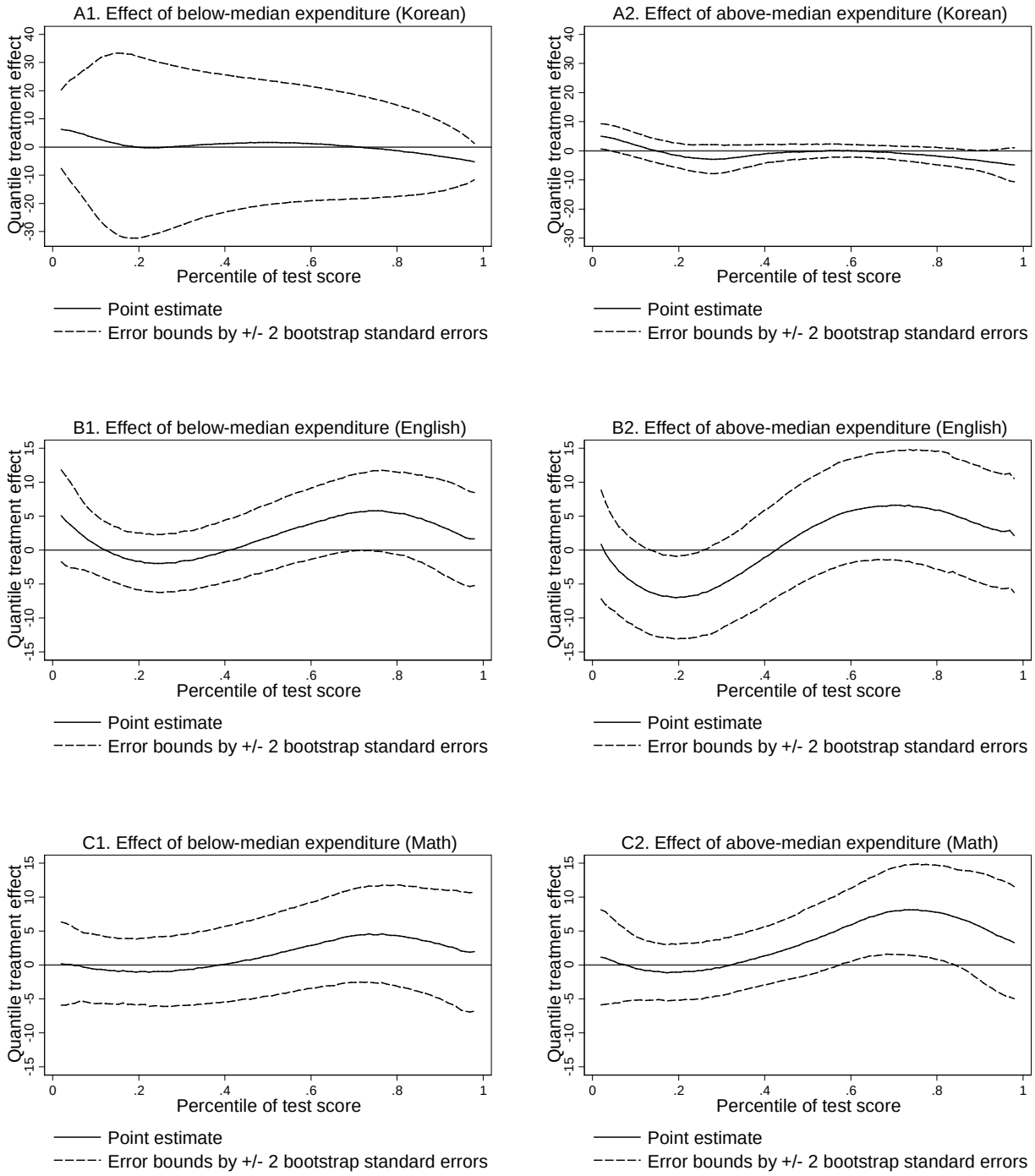
Note. The outcome variable is achievement test scores measured in December 2005. The treatment variable is an indicator for having ever received private tutoring in 2007. The same covariates used in table 5 are controlled for. Plots in the left column compare density estimates for the realized test score distribution of the treatment group (students receiving private tutoring in 2007) with those for the counterfactual test score distribution of the same students had they not received private tutoring. Plots in the right column report estimated quantile treatment effect on the treated (QTT). Standard errors for the QTT estimates are computed by bootstrap of 2000 iterations.

Figure 6. Falsification Test Results: Distributional Effects of Private Tutoring (Math)



Note. The outcome variable is achievement test scores measured in December 2005. The treatment variable is an indicator for having ever received private tutoring in 2007. The same covariates used in table 5 are controlled for. Plots in the left column compare density estimates for the realized test score distribution of the treatment group (students receiving private tutoring in 2007) with those for the counterfactual test score distribution of the same students had they not received private tutoring. Plots in the right column report estimated quantile treatment effect on the treated (QTT). Standard errors for the QTT estimates are computed by bootstrap of 2000 iterations.

Figure 7. Distributional Effects of Private Tutoring by Levels of Tutoring Expenditure



Note. The outcome variable is achievement test scores measured in November 2007. The treatment variable is an indicator for having ever received private tutoring in 2007. Student, parental, and school characteristics as in column 3 of table 5 are controlled for. Plots in the left column report estimated quantile treatment effect on the treated (QTT) when the treatment group is restricted to students whose amount of private tutoring expenditure is below the median level. Plots in the right column report the estimated QTTs when the treatment group is restricted to students whose amount of private tutoring expenditure is greater than or equal to the median level. Standard errors for the QTT estimates are computed by bootstrap of 2000 iterations.

Figure A1. Log of the Absolute Value of the Estimated Characteristic Function

