Private Tutoring and Educational Inequality:*

Evidence from a Dynamic Model of Academic Achievement in Korea

Jaesung Choi †

University of Pennsylvania

December 25, 2012

Abstract

This paper studies the effect of private tutoring on academic achievement and educational inequality in Korea. Korea has the largest system of private tutoring in the world, and maintains an outstanding performance on international academic tests, such as the PISA. Korea's school system is characterized by limited school choice and low variation in quality and curricula across schools, which provides incentives for households to employ private tutoring as an additional educational investment. Prompted by concerns about unequal access to private tutoring and resulting educational inequality, the government has enacted various forms of regulation in the tutoring market. This paper seeks to quantify the effect of private tutoring on academic outcomes and to evaluate the impacts of a range of government policies. It develops and estimates a dynamic discrete choice model of private tutoring and self-study decisions using panel data from the 2005-2011 waves of the Korean Education Longitudinal Study. The data follow 7th graders annually until one year after their high school graduation and contain detailed information on private tutoring use and test scores. Simulations based on the estimated model show that prohibiting private tutoring reduces the achievement gap between higher and lower income households by 57 percent, but at the cost of decreasing average test scores by 0.47 standard deviations. Providing a 50 percent price subsidy for private tutoring to low income households increases average test scores for all students by 0.18 standard deviations and narrows the income achievement gap by 47 percent at the cost of increased government spending. A voucher system funded by tax on private tutoring narrows the income achievement gap by 31 percent, but at the cost of decreasing average test scores by 0.07 standard deviations.

*I am greatly indebted to my advisor Petra Todd for her guidance and encouragement. I am grateful to Jere Behrman, Hyunjoon Park, and Kenneth Wolpin for their advice and support. I also thank Naoki Aizawa, Ilwoo Hwang, Karam Kang, You-Suk Kim, Holger Sieg, and Chamna Yoon for insightful comments and discussions. All errors are mine.

†160 McNeil, 3718 Locust Walk, Philadelphia, PA 19104, USA (Email:jaesungc.econ@gmail.com)

1 Introduction

South Korea has the largest system of private tutoring in the world (Baker and LeTendre 2005; Bray 2009). In 2011, approximately 72 percent of Korean students received tutoring, such as private lessons or for-profit "cram" schools, for an average of 7 hours per week. In 2011, total expenditure on tutoring amounted to US$18 billion, more than 1.7 percent of GDP and equivalent to 80 percent of the government's total expenditure on primary and secondary education (Bray 2009; KNSO 2011). Korean students consistently perform at the top of the rankings on various international tests, such as the Program for International Student Assessment (PISA) test. For instance, in 2009, Korean students ranked 1st in reading and mathematics and 3rd in science among the 34 OECD countries (OECD 2010). This paper studies whether and to what extent the huge extra parental investment in education through private tutoring contributes to students' academic achievement.[1]It also evaluates the effect of a variety of policies that have been introduced or considered for regulating the tutoring market with regard to their effects on tutoring and self-study choices, on academic achievement, and on inequality in achievement.

Why is private tutoring so prevalent in Korea? The country's education system features some unique attributes that contribute to a high demand for tutoring. First, under secondary school equalization policies (introduced in 1968 for middle schools and 1973 for high schools), students entering middle and high school are randomly assigned to schools within their school district, including private schools.[2] Subject to strict government regulations, public and private schools must follow uniform curricula and charge the similar tuition. Hence, in Korea, the scope for improving school quality by choosing to live in a better neighborhood is more limited than in other countries, such as the United States, where there is a wide geographic variation in school quality. Second, high-stakes exams play a vital role in gaining access to a few selective colleges. Under these circumstances, and coupled with a strong societal desire for prestigious university credentials, private tutoring functions as an essential means for families to invest in children in a so-called "hyper-competitive" society.

[1]In fact, each week Korean students spent 3.6 more hours for private tutoring and 2.4 fewer hours for self-study than students in other OECD countries (OECD 2005).

[2]Although some school districts in rural areas and small cities do not implement the random assignment policy, currently the policy is applied to the 70% of academic high school entrants including Seoul (the capital) and the six major metropolitan areas.

The growing importance of the private tutoring industry is being noticed in other developed countries (Bray 2009). Students in East Asia, especially Japan, Taiwan, and South Korea, are heavily involved in various kinds of private tutoring. However, the phenomenon is not limited to Asia. For example, the private tutoring sector in France is estimated to be growing annually at about 10 percent (Bray 2011). In the U.S., spending on tutoring is growing at more than 5 percent, and tutoring rates have surged among the parents who want their children to qualify for gifted programs or be admitted to top public schools (Sullivan 2010; Phillips 2012a; 2012b).[3]

An important policy concern is that because wealthier parents are able to choose more private tutoring services, tutoring activities may widen educational inequality between children of different socioeconomic backgrounds (Buchmann et al. 2010). In 2011 in Korea, families in the highest income quintile spent six times more on private tutoring than those in the lowest (Bank of Korea 2012). Because of concerns about unequal access to private tutoring and resulting educational inequality, during the last four decade the government has enacted policies designed to limit tutoring and to provide more accessible alternatives.[4] Despite such policies, however, the share of expenditure on private tutoring has been rising, and the gap between the household expenditure on private tutoring of high- and low-income families has widened over time (Kim 2011; Nam 2007).

Despite its increasing prevalence and its potential effect on inequality, the role of private tutoring as a determinant of educational achievement has received relatively little attention in academic research. In this paper, I examine participation in private tutoring and quantify its impact on academic outcomes and educational inequality. To this end, I develop a dynamic discrete choice model of private tutoring and self-study decisions by Korean students in middle school and high school (from 7th to 12th grade). The model, which is described in detail below, is estimated by simulated maximum likelihood using data from the Korean Education Longitudinal Study (KELS), a nationally representative panel that follows 7th graders annually until one year after high school graduation. I use the estimated behavioral model to quantitatively evaluate the effects of a range of actual

[3]Along with the rapid growth of the demand for private tutoring demand, the franchise industry for tutoring centers such as Kumon, Huntington, and Sylvan Learning is flourishing in North America (Aurini and Davies 2004).

[4]For example, 7.30 Educational Reform Measures, implemented in 1980, prohibited private tutoring. Recent interventions include restrictions on the hours of operation of private tutoring institutions. The government also established the Educational Broadcasting System in the 1990s to supplement school education.

and potential government policies on individual decisions, academic outcomes, and educational inequality. The policies evaluated include regulating operating hours of private tutoring institutions, providing price subsidies to low income households, and imposing taxes on private tutoring and using resulting tax revenues to provide vouchers for private tutoring to children from low income households.

In my model, each household makes sequential decisions on the hours spent in private tutoring and hours spent in self-study from 7th to 12th grade annually. In each period, a test score is determined by a value-added education production function that includes hours for private tutoring , hours for self-study, and a lagged test score as inputs (e.g., Hanushek 1986; Todd and Wolpin 2003). At the end of the senior year of high school, each student's score on the national college entrance exam is realized, which determines the model's terminal value–the discounted life-cycle expected payoff upon high school graduation. A household's utility depends on consumption, academic performance of a child, private tutoring hours, and self-study hours. The model also incorporates unobserved heterogeneity by allowing for permanent differences in preferences, in production of educational outcome, and in parental income using unobservable types (Heckman and Singer 1984; Keane and Wolpin 1997). There are time-varying stochastic shocks, and future shocks are unknown at the time decisions are made. Initial conditions include parental education and age, location, number of children, and a performance at 6th grade.

As noted, the model is estimated using KELS data, which follow a sample of 6,908 7th graders annually since 2005. The KELS contains scores for mathematics, English, and Korean tests administered in school as part of data collection at the end of each academic year during middle school, and national college entrance exam scores taken at the end of 12th grade. The KELS provides key demographic and family characteristics and rich longitudinal information about students' participation in private tutoring and self-study hours.

I use the estimated model to simulate the effects of a range of government policies. The model incorporates decisions about both private tutoring hours and self-study hours, so policies that regulate tutoring can also influence self-study decisions, with the effect depending on whether these inputs are substitutes or complements. When there is a ban on private tutoring, as actually occurred in the 1980s, average test scores decrease by 0.47 standard deviations, but inequality in achievement also decreases: The score gap between children from the top household income quartile and those from the bottom quartile decreases by 57 percent. Providing a 50 percent price subsidy for private tutoring to households with below-median income increases average test scores by 0.18 standard deviations and decreases the income achievement gap by 47 percent at the cost of increased government spending. Finally, a voucher system funded by a tax on private tutoring also narrows the income achievement gap by 31 percent at the cost of decreasing average test scores by 0.07 standard deviations.

Related Literature There is a growing literature on private tutoring. Bray's extensive research (2009; 2011; 2012) provides a comprehensive description of private tutoring from a cross-national perspective. Dang and Rogers (2008) survey the literature to study the effects of private tutoring and conclude that evidence regarding the effects of private tutoring is limited and inconsistent.[5] However, most of the previous investigations have not controlled for endogeneity bias, partially due to data limitations. There are few recent studies which address the potential endogeneity of self-selection into private tutoring and account for unobservables (Choi 2012; Kim 2010; Ryu and Kang 2012).[6] Furthermore, despite of a strong demand for effective policies to better respond to growing concerns arising from private tutoring, there has not been a study to evaluate different policies implemented or to design a policy based on quantitative analysis. This paper is the first attempt to conduct policy experiments to derive meaningful policy implications based on a behavioral model which incorporates individual unobserved heterogeneity.

There is an interesting research to investigate a policy relevance of private tutoring in developing country setting where for-profit tutoring is provided to their own students in schools. Jayachandran (2012) models teacher incentives and student achievement in the presence of school provided tutoring. After deriving testable implication from the

[5] Results depend on the country studied, the definitions of private tutoring outcomes used, and the statistical methods applied. Buchmann (2002), Tansel and Bircan (2005), Dang (2007), and Ono (2007) find strong positive effects. In contrast, Briggs (2001), Domingue and Briggs (2009), and Kuan (2011) find positive but modest effects at best. Baker et al (2001) and Elbadawy (2009) report that the effects of private tutoring are insignificant. Although there have been a growing number of studies using Korean datasets to study the effects of private tutoring, the results of these studies are also inconsistent (e.g., Park et al. (2011); Byun (2011); Kang and Ryu (2011); Kim (2010);Park et al. (2008); Shin et al. (2008); Choi (2007)).

[6] Ryu and Kang (2012) and Kim (2010) analyze first three waves of KELS and find positive but modest effects on test scores. Choi (2012) analyzes the first two waves of the Seoul Education Longitudinal Study and finds that private tutoring had a positive impact on English and math test scores and that the effectiveness of private tutoring as compared with that of self-study decrease as the level of schooling increases.

model, she empirically tests whether there is an adverse effect on teacher's effort level and students' academic performance when schools offer for-profit tutoring to their own students in developing countries. My study can complement her study in the sense that Korean case is more relevant to the developed country setting where for-profit tutoring in schools is prohibited and well-structured private tutoring market exists.

Finally, my paper contributes to the literature that uses discrete choice dynamic programming (Keane, Todd and Wolpin 2011) to study the policy impacts in education research. Todd and Wolpin (2006) investigates the effects of a conditional cash transfer program in Mexico using data from the PROGRESA program. They use the estimated model to suggest an alternative subsidy schedule that would induce a greater impact on average school attainment at similar cost to the existing program. Bravo et al. (2010) study the effects of national wide school voucher program in Chile and investigate how the school voucher reform influenced sorting among different types of schools, educational attainment, earnings, and labor market participation. More recently, Duflo et al. (2012) study whether monitoring and financial incentives can reduce teacher absence and increase learning using data from a randomized experiment in India. They formulate and estimate a dynamic teacher labor supply model and use the estimated behavioral model to propose cost-minimizing compensation policies.

The rest of the paper is organized as follows. Section 2 explains Korean educational background. Section 3 lays out the model. The solution method and the estimation method are provided in Section 4. Section 5 describes the data. Section 6 presents empirical results, including parameter estimates and model fit, and discusses results of policy simulations. The last section concludes the paper.

2 Korean Education

2.1 The Basic Structure of the School System

Education in Korea is divided into elementary school (grades 1 to 6), middle school (grades 7 to 9) and high school (grades 10 to 12). Almost all children complete middle school, after which they can continue their education at either academic high schools (76%) or vocational high schools (24%). More than 75 percent of high school graduates enter either two- or three-year junior colleges or four-year universities right after their high school graduation.

Before 1973, Korean high schools could select their students based on students' performance on entrance examinations administered by the individual high schools. It caused a sorting by family background and substantial gaps in academic performance across schools along with a severe competition to get into prestigious high schools. Out of concerns about between-school inequality and academic pressure on students to do well on high school entrance examinations, the Korean government introduced a national educational reform known as the High School Equalization Policy (Park, Behrman and Choi 2012). Under this policy middle school graduates are randomly assigned to academic high schools within their school districts in most urban areas. Along with the earlier implementation of the No Middle School Entrance Examination Policy in 1968, these two major government policies have limited school choice in secondary education.[7] This equalization policy has has been maintained in most cities with only small modifications.[8]

Because random assignment is applied regardless of whether schools are public or private, students cannot choose to attend private secondary schools.[9] The Korean government also imposes uniform curricula and tuition requirements on public and private schools. It has been documented that students attending private and public schools do not differ significantly in terms of socioeconomic background (Park 2010)

In Korea, educational credentials play a critical role not only in labor market but also in marriage market. Thus, students face intense competition in the national college entrance exam used for admission to top tier colleges. Under the limited school choice and low variation in quality and curricula across schools, the majority of students and parents rely on private tutoring as a means to improve their chance of obtaining educational credentials from prestigious institutions (Kim 2012).

[7]Elementary school students are assigned to schools closest to their homes. 98 percent of elementary schools are public, and there are no substantial differences in school resources, teacher quality, and curriculum across schools.

[8]Some school districts have relaxed the equalization policy to respond to growing concerns about limited school choice. They allow students to list a couple of schools that they prefer to enter. Then, school districts randomly assign 30 to 40 percent of their enrollment from among those students who reveal their preferences for the school; the remaining students are allocated entirely by lottery without considering students' preferences. (Park et al. 2012)

[9]A possible way to avoid random assignment is to apply for specialized high schools, including foreign language and science schools, which are not subject to the high school equalization policy and thus can select students on the basis of their own criteria. However, the number of students attending such specialized high schools is fairly small, about 2 percent of the total student enrollment in academic high schools in 2009. Hence, it is not a feasible option for the majority of high school students.

2.2 Private Tutoring

In 2011, the private tutoring participation rate was highest among elementary school students, 84.6 percent. For middle school students, it was 71 percent, and for high school students, 58.7 percent. Private tutoring takes various forms - e.g., one-on-one or few-on-one tutoring, commercial cram schools, worksheet programs sent to home, and on-line tutoring. It is most common in math and English given their importance in school grades and on college entrance exams. The participation rate in 2011 was 50.2 percent for math and 49.2 percent for English. 77.6 percent of students participated in private tutoring in academic subjects for the purpose of 'make-up for previous materials' and 52.1 percent for 'study in advance' (KNSO 2012). In terms of supply side of private tutoring, the number of employees in private tutoring industry has increased by 7% annually between 2001 and 2006, and this industry was the largest employer of college graduates in 2009 (Kim and Park 2012).

Table 1, based on the data from the Private Education Survey of 2010, provides a general overview of the prevalence of private tutoring in Korea. It focuses on academic areas, especially English and mathematics. The amount of money spent on private tutoring generally increases as students advance from elementary school to high school. Hours for private tutoring peaks in middle school. On average, each household spent about $100 for English and another $100 for math private tutoring per student in 2010. A middle school student participated in private tutoring about 7 hours a week in academic subjects. The hourly price of private tutoring rises as the student's grade level increases.

Choi (2012) confirms earlier findings that parental education and household income significantly affect the participation decision and the amount spent on private tutoring (Byun 2012, Kuan 2011). Also, the better a student is doing in school, the more likely he/she is to use private tutoring. Tutoring is more common in large cities. In terms of total spending on private tutoring, for instance, it is reported that the highest income quintile spent six times more than the lowest income quintile (Bank of Korea 2012). This finding suggests that private tutoring may play an important role in the intergenerational mobility in Korea by linking family background to children's academic achievement. (Byun 2012)

2.3 Government Reaction to Private Tutoring

Lee et al (2010) classify patterns of policy responses to private tutoring in Korea into four types: equalization of schools, prohibition of private tutoring, enhancement of public education, and provision of supplementary tutoring programs to increase access. As mentioned earlier, the No Middle School Entrance Exam Policy of 1968 and the High School Equalization Policy of 1973 represent the main legislative attempts to equalize schools.

The 7.30 Educational Reform Measures implemented in 1980 prohibited students from participating in private tutoring. However, in 2000, the Constitutional Court ruled that government's comprehensive prohibition of private tutoring was unconstitutional, and, after that, the private tutoring market grew dramatically as shown in Figure 1. Recently, the government has intervened in the private tutoring market by forcing private institutions to close by 10 pm in metro areas, by requiring them to post on the Internet the tuition for each program, and by imposing strict requirements on the establishment and operation of private institutions.

As for public education, the government has consistently attempted to improve its quality by, for example, reducing student-teacher ratios and assessing teacher performance. Recently, it tried to broaden households' school choice options and to give limited autonomy to schools in matters such as designing curricula and implementing within-school tracking. The government also has tried to provide alternative free or lower-priced supplementary education by offering after-school programs and educational broadcasting services on TV or the Internet. Since 2006 vouchers have been provided to children of low-income families to use for after-school programs held in the schools after regular school hours.

During the 1980s and 1990s the government reformed the college entrance exam system to reduce expenditure on private tutoring, basing this move on the assumption that the difficulty of the exam and the burden of materials to be covered in the exam were the chief reasons why students used private tutoring (Park 2011).

As mentioned above, over the last few decades there have been consistent policy responses to the problems created by private tutoring. However, the effectiveness of such policies and their long-term outcomes have not yet been carefully investigated.

2.4 Student Time Allocation

On weekdays most middle school students spend 10 hours at school, from 8 am to 4 pm. High school students spend an additional one and a half hours a day at school. When they finish their regular school day, students use their remaining time either for leisure, self-study, or private tutoring. Although some students attended private institutions until as late as 1 am before the recent government regulation; most students usually finish their study by 11 pm. Hence, students typically have 6-8 hours after school to allocate during weekdays.[10]

Using 2004 Korean Time Use Survey (KTUS) data, Oh (2010) provides detailed summaries of study time allocation for middle and high school students.[11] Related to learning activities, middle school students spent about 370 minutes on weekdays and 230 minutes on Saturday in school. The corresponding figure for high school students was 520 minutes on weekdays and 280 minutes on Saturday. In school, middle school students used 39 minutes for self-study on weekdays and 26 minutes on Saturday. High school students used 111 minutes for self-study on weekdays and 58 minutes on Saturday.

Study time outside school can be decomposed into time spent in private tutoring and in self-study. For private tutoring, middle school students spent 88 minutes on weekdays and 27 minutes during the weekend. For self-study, they spent 57 minutes on weekdays and 72 minutes during the weekend. For high school students, they participated in private tutoring for 30 minutes on weekdays and 31 minutes during the weekend. For self-study, they spent 60 minutes on weekdays and 128 minutes during the weekend.

3 Model

My model represents a dynamic decision problem of a household with a child in the secondary school. The model begins at the beginning of the middle school, 7th grade. There are six periods from 7th to 12th grade. At the end of 12th grade, each student's outcome on college entrance exam is realized and this test score determines the terminal value of the model, which is the discounted life-cycle expected payoff.

[10] Students go to school every other Saturday and some students also participate in private tutoring during the weekend.

[11] Unfortunately, KTUS does not distinguish academic high school and vocational high school student. Hence, the provided study hours may underestimate study hours of academic high school students who are the main focus of this study.

At each time period t, each household determines hours spent in private tutoring, x_t, and hours spent in self-study, s_t. The choice set, D_t is $\{x_t, s_t\}$. I allow 6 discretized choices for private tutoring and 9 discretized choices for self-study, and each choice corresponds to 4 hours interval. That is, I allow 54(=6*9) exclusive alternatives in each period. Although I observe study hours, grouping every certain range of hours to one choice is necessary to make computation manageable.

In the model, preference shocks for private tutoring and self-study and an income shock realize at the beginning of each period. Given these shocks, each household makes private tutoring and self-study hours decisions. Afterward, an achievement shock in education production function realizes, then it determines the period test score given the study hours decision made by household at the beginning of each period.

Preference

A household's utility depends on household consumption (c), on child's test score performance (q), on hours spent in private tutoring (x) and hours spent in self-study (s), and on preference shocks (ϵ_x, ϵ_s). The utility function also incorporates permanent household heterogeneity in the form of discrete permanent types. The precise functional form of the utility function is

$$
\begin{aligned}
U_t \quad = \quad & c_t * (1 + \alpha_1(n-1)) \\
& + [\alpha_2 + \sum_{k=2}^{K} \alpha_{3k} I(type = k) + \alpha_5 I(E_p \geq 16)] * (q_t + \alpha_6 * q_t^2) \\
& + [\alpha_7 + \sum_{k=2}^{K} \alpha_{8k} I(type = k) + \alpha_{10} c_t + \epsilon_{xt}] * x_t + \alpha_{11} x_t^2 \\
& + [\alpha_{12} + \sum_{k=2}^{K} \alpha_{13k} I(type = k) + \epsilon_{st}] * s_t + \alpha_{15} s_t^2 \\
& + \alpha_{16} I\left(x_t = 0\right) I\left(x_{t-1} \neq 0\right)
\end{aligned}
$$

The utility function is linear in consumption. Because I only observe in the data a focal child's test score and hours of private tutoring, the linear utility functional form restriction is necessary to circumvent complication arising from parental decision of optimal resource allocation over multiple children with heterogeneous ability. However, to consider difference in financial burden depending on the number of children (n), I allow the marginal utility of consumption to change depending on the number of children in the household.

I allow marginal utility from a child's test score performance (q_t) to vary nonlinearly and to depend on unobserved type of the household and on parental education level. The utility from private tutoring and self-study hours is also assumed to be quadratic and to depend on the unobserved type. Utility from private tutoring incorporates a consumption value of private tutoring which comes in addition to the direct utility from test scores. Engaging in private tutoring and spending time together with friends at private tutoring institutions can have psychological benefits beyond test score increases. Increases in self-study and private tutoring hours represent a loss of utility from decreases in leisure. The last term in utility function captures the transition cost of changing study styles, which is incurred anytime a student who used private tutoring last period stops using it. The exact specification of the utility function was determined in part using model fit criteria.

Budget and Time Constraint

The annual household income (y) depends on parental characteristics, residence location, number of children (n) and household unobserved type. There is also a stochastic shock ϵ_y to capture additional source of heterogeneity and uncertainty in income process.

$$\begin{aligned} ln(y_t) &= \gamma_1 + \sum_{k=2}^{K} \gamma_{2k} I(type = k) + \gamma_4 E_f + \gamma_5 E_m \\ &\quad + \gamma_6 a_t + \gamma_7 a_t^2 + \gamma_8 * I(metro) + \gamma_9 * n + \epsilon_{yt} \end{aligned}$$

Household consumption (c) equals income net of private tutoring expenditure.

$$c_t + p_t^x x_t \leq y_t$$

where p_t^x is the price for tutoring at period t.

In addition to the budget constraint, each student faces time constraints. There is a maximum time allocation available for private tutoring, $\bar{t}_{pt}$, and a maximum time allocation for self-study, $\bar{t}_{ss}$.

$$x_t \leq \bar{t}_{pt}$$
$$s_t \leq \bar{t}_{ss}$$

Education Production Function

An important component of the model is the education production function which determines a test score in each period. I introduce a value added specification so that current period test score (q_t) depends on the lagged test score (q_{t-1}), on hours of private tutoring (x_t), and on hours of self study (s_t), and on a test score shock (ϵ_t^q).

I allow parameters of education production function differ for 7th grade, for 8th and 9th grades, and for high school, 10th-12th grades. I allow 7th grade has a separate education production function from 8th and 9th grades because I use self-reported 6th grade academic performance as a lagged test score for the estimation of 7th grade production function.[12]

I allow the effect of private tutoring and self study to vary by grade level and effect of self-study depends on parental education level to incorporate possible difference in parental academic guidance at home. To allow the marginal effect of private tutoring and self-study to vary nonlinearly, the functional form includes square terms of time spent in private tutoring and time spent in self-study. Finally, to consider the possible relationship between private tutoring and self-study hours as substitutes or complements, I also include an interaction term of private tutoring and self-study hours. The precise functional form of education production function is

- For Middle School (t=1,2,3)

$$\begin{aligned} q_t &= \beta_1^{M_i} + \sum_{k=2}^{K} \beta_{2k}^{M_i} I(type = k) \\ &+ \beta_4^{M_i} q_{t-1} + \beta_{5t}^{M_i} * x_t + \beta_6^{M_i} x_t^2 \\ &+ [\beta_{7t}^{M_i} + \beta_8^{M_i} I(E_p = 14) + \beta_9^{M_i} I(E_p \geq 16)] * s_t + \beta_{10}^{M_i} s_t^2 + \beta_{11}^{M_i} x_t * s_t + \epsilon_t^q \end{aligned}$$

- For High School (t=4,5,6)

$$\begin{aligned} q_t &= \beta_1^{H} + \sum_{k=2}^{K} \beta_{2k}^{H} I(type = k) \\ &+ \beta_4^{H} q_{t-1} + \beta_{5t}^{H} * x_t + \beta_6^{H} x_t^2 \\ &+ [\beta_{7t}^{H} + \beta_8^{H} I(E_p = 14) + \beta_9^{H} I(E_p \geq 16)] * s_t + \beta_{10}^{H} s_t^2 + \beta_{11}^{H} x_t * s_t + \epsilon_t^q \end{aligned}$$

Unobserved Heterogeneity

The unobserved heterogeneity takes the form of discrete unobserved types, as in Heckman and Singer (1984). The probability of being a particular type k, μ_k, depends on demographic and family background variables that include parental education, number of

[12] I find that correlation between lagged test score and current test score and productivity of private tutoring and self-study are distinguishable between 7th grade and the other middle school grades.

children in the 7th grade and metro or non-metro area. These variables constitute the model's initial conditions, z_i

The type probability is specified as a multinomial logit:

$$\mu_k = Pr(type = k|z_i) = \begin{cases} \frac{exp(z_i\omega_k)}{1+\sum_{j=2}^{K} exp(z_i\omega_k)} & \text{for k=2,3,..,K,} \\ 1 - \sum_{k=2}^{K} Pr(type = k) & \text{for k=1.} \end{cases}$$

$$\text{where } z_i\omega_k = \omega_{k1} + \omega_{k2}E_f + \omega_{k3}E_m + \omega_{k4}n + \omega_{k5}Metro$$

4 Model Solution and Estimation

In each period, households maximize the expected discounted present value of their lifetime utility subject to budget and time constraints, the production function and other laws of motion. Conditional on the state space Ω_t, consisting of all the relevant factors affecting current or future utility, a household chooses the optimal sequence of choices, d_t^j, where d_t^j=1 if jth alternative is chosen at t, and 0 otherwise. Hence, the solution to the optimization problem is a set of grade-level-specific decision rules. The maximized present discounted value of lifetime utility at time t, the value function, is given by

$$V_t(\Omega_t) = maxE\left[\sum_{\tau=t}^{T}\sum_{j=1}^{J} \delta^{\tau-t}u_\tau^j d_\tau^j|\Omega_t\right]$$

$$\text{where } \Omega_t=\{E_f, E_m, age_p, q_{t-1}, metro, x_{t-1}, \epsilon_x, \epsilon_s\}$$

Here, J denotes the set of alternatives available to the household at period t. x_{t-1} is hours of private tutoring in the last period. The expectation is taken over the distribution of preference, test score, and income shocks. I assume that ϵ's are normal and are jointly serially independent, conditioning on type. In the above equation δ is the discount factor, which is set to 0.95.

Recasting the problem in a dynamic programming framework, the value function can be written as the maximum over alternative-specific value functions, $V_t^j(\Omega_t)$, i.e., the expected discounted value of alternative $j \in J$ that satisfies the Bellman equation,

$$V_t(\Omega_t) = \max_{j\in J}[V_t^j(\Omega_t)]$$

$$V_t^j(\Omega_t) = \begin{cases} U_t^j + \delta E(V_{t+1}(\Omega_{t+1})|j \in J, \Omega_t) & \text{for } t < 6, \\ U_t^j + \delta E(R|j \in J, \Omega_6) & \text{for t=6.} \end{cases}$$

Terminal value, R, depends on the college entrance exam score, q_6

$$R(q_6) = \begin{cases} 0 & \text{if } q_6 < q_l, \\ \zeta * (q_6 - q_l)^2 & \text{if } q_6 \in [q_l, q_u], \\ \zeta * (q_u - q_l)^2 & \text{if } q_6 > q_u. \end{cases}$$

where $Pr(q_6 < q_L) = 0.01$ and $Pr(q_6 > q_U) = 0.99$

The model does not have an analytical solution, but the finite horizon dynamic programming model can be numerically solved by backward recursion. The solution consists of values of $E(V_{t+1}(\Omega_{t+1})|j \in J, \Omega_t)$ for all j and elements of Ω_t. I refer to this function as Emax. The solution method proceeds by backward recursion beginning with the last period, 12th grade. The multivariate integrations are necessary to calculate the expected value of the maximum of the alternative-specific value functions at each state point and are performed by Monte Carlo integration over the shocks. The state space includes a continuous variable, q_{t-1}, so I use the Emax approximation method developed in Keane and Wolpin (1994) in which the Emax functions are evaluated at a random subset of the state points, and the values are used for interpolation at non-evaluated points. The backward recursion continues until the period 1, 7th grade. The set of Emax functions fully describes the solution to the optimization problem.

I estimate the parameters of the model by simulated maximum likelihood, using the model solution as an input for the likelihood function. The Emax function is needed to calculate the alternative-specific value functions, $V_t^j(\Omega_t)$ for j=1,...,J. The likelihood function specifies the joint probability of observing a sequence of choices and outcomes in the data. The observed choices of a household are time spent in private tutoring (x) and time spent in self-study (s). The observed outcomes are test scores (q) and annual household income (y).

At any period t, I denote the vector of observed outcomes as $O_t = \{x_t, s_t, q_t, y_t\}$. Suppose we observe these outcomes for a sample of N households from t=1 to 6. Then, the likelihood for these households is

$$\prod_{n=1}^{N} Pr(O_1, O_2, O_3, O_4, O_5, O_6|\overline{\Omega}_1)$$

$$\text{where } \overline{\Omega}_1 = \{E_f, E_m, age_p, n, metro, q_0\}$$

Let $\overline{\Omega}_1$ be the observable component of household characteristics at the first wave in the data, that is, the initial state space net of the household's type and stochastic shocks. The observable part of the state space at the first wave consists of parent education and age, number of children, location and 6th grade self-reported test score.

Given the assumption of jointly serial independence of the vector of shocks, the likelihood for each individual can be written as the product of within-period outcome probabilities, integrating over the unobserved type conditional on the corresponding state space. The initial conditions are assumed to be exogenous conditional on type.

$$L(\Theta) = \prod_{n=1}^{N}\sum_{k=1}^{K} Pr(O_{1n}, O_{2n}, O_{3n}, O_{4n}, O_{5n}, O_{6n}|\overline{\Omega}_1, type = k) \cdot Pr(type = k|\overline{\Omega}_1)$$

The outcome probability at a point in time can be calculated from the choice probability conditional on the state space and the probability of observing a test score and household income. I compute the conditional choice probability using a kernel smoothed frequency simulator (McFadden 1989; Eckstein and Wolpin 1999). For each observed household n, I draw error vectors to replicate 1,000 hypothetical households with the same state points as the household n. For each simulated household, I compute the alternative(j)-specific value functions in each period, using the solution of the model – the Emax function – for all possible alternatives j. Then, the kernel of the integral for each of the draws is:

$$\frac{exp\left[\frac{V(d_t^j(\Omega_t)) - max_{d_t \in D_t} V(d_t(\Omega_t))}{\tau}\right]}{\sum_{k=1}^{J} exp\left[\frac{V(d_t^k(\Omega_t)) - max_{d_t \in D_t} V(d_t(\Omega_t))}{\tau}\right]}$$

The conditional choice probability at the given period is obtained by the integration of the kernel. The expectation is obtained by Monte Carlo integration by averaging the above object over 1,000 draws.

If state variables are observed at all interview waves and respondents, the above likelihood can be obtained using the kernel smoothed frequency estimator and laws of motion.

However, the state variables, such as test scores at period 4 (10th grade) and period 5 (11th grade) are not observed for all households in the data. Hence, I also integrate over the distribution of unobserved test scores at missing periods to construct the outcome probabilities by Monte Carlo integration. I simulate 3,000 draws for each individual's missing scores. (See Appendix for more information)

As noted, the model is estimated by simulated maximum likelihood. In implementation, there are assumed to be three types. The model parameters enter the likelihood through the choice probabilities that are computed from the solution of the dynamic programming problem. Subsets of parameters also enter through the education production function and income process. The maximization of the likelihood function iterates between solving the dynamic program and calculating the likelihood. Standard errors of the parameter estimates are obtained by the inverse of the average of the product of the score matrices, where the derivatives of the log likelihood are evaluated numerically. This is known as the BHHH estimator (Berndt et al., 1974).[13]

5 Data

The estimation is based on a nationally representative sample of 7th graders in middle schools, the Korean Education Longitudinal Study (KELS) that are followed until one year after their high school graduation. This survey is conducted annually by the Korean Educational Development Institute (KEDI), a government-funded educational research agency. KELS is the longest and richest longitudinal education survey to provide student achievements, educational investments and school characteristics from questionnaires that are answered by students, parents, teachers, and school principals. KELS implemented a two-stage stratified sampling design. After randomly sampling middle schools within each of four types of regions proportionate to the population, approximately 50 students in the 7th grade were randomly sampled within selected schools (Kim et al. 2006).

This survey started with a total of 6,908 7th graders in 150 schools across the nation. In the sixth wave (2010), 5,346 (77.4%) of the original 6,908 respondents were resurveyed.

[13] To obtain numerical derivatives, I use a step size equal to 1% of the parameter estimate. For a description of methods used to estimate discrete choice dynamic programming models, see Keane and Wolpin (1994). Solving the model and optimizing over the 81 parameters is computationally fairly intensive. For this reason, computation was done on a parallel linux cluster with using the asynchronous parallel pattern search algorithm to optimize the function (HOPSPACK; see Plantenga, 2009).

In addition to the students, separate surveys were administrated each year to students' families, teachers, and school principals to collect a wide range of family, class, and school information. KELS also administered achievement tests in Korean, English, and math in each stage of data collection for the first three years before students spread out into many different high schools. Of particular relevance to our analysis are the test scores and the annually reported data on private tutoring use and hours spent in self-study as well as information about parental education and age, household income, number of children, and location.

Samples

For the analysis, I restrict the sample to students who consistently answered these annual surveys. Since the students who chose vocational schools could have different characteristics, such as being more inclined to seek a career after high school graduation and being exposed to different academic environments, I focus on students who attended academic high schools. Because my model assumes both parents are in the household, imposing this condition further restricts the sample used for estimation. I also omit observations that have missing information on major variables such as academic achievement and private tutoring use. My final analysis sample consists of 2,192 students who were 7th graders in 2005 and who remained in the survey over 6 years. I have a total of 13,152 person-year observations on these students.

Test Scores

The test score variables used in the analysis are constructed in the following way. For the measure of cognitive achievement, I use test scores administered by KEDI and college entrance exam scores (CSAT) extracted from administrative data collected by KEDI in collaboration with the Korean Ministry of Education.[14] In each of the first three waves of KELS, student academic performance was measured by scores on achievement tests in three subjects: Korean, English and mathematics. The raw test score of each subject is scaled from 0 (lowest) to 100 (highest). I calculate an average of the mathematics and English scores and normalize the score to have a mean of 5 and a variance of 1.5 for all test

[14] At the end of grade 12, students take the College Scholastic Aptitude Test (CSAT) exam administered by the government in November. The CSAT score is presumably the most reliable measure of academic achievement in Korea. Given the importance of the CSAT score for college admission, high school curricula and learning are heavily oriented toward preparing students for the exam.

takers in the dataset. For the 12th grade CSAT score, I construct normalized scores on the same scale, using a population sample of CSAT takers, and match the corresponding numbers to each sample's score. Initial cognitive ability at the beginning of the model is a 6th grade self-reported measure on a scale of 1 to 9. I normalized it to have the same scales as other scores.

Private Tutoring and Self-Study Hours

A unique feature of the KELS data is the availability of detailed information on students' private tutoring experience and the number of self-study hours by subject. Students can use different types of private tutoring such as participating in private tutoring institutions (Hakwon, also called "cram" schools), individual or group tutoring provided by individual tutors, paper correspondences, and tutoring via the Internet and broadcasting media, etc. In this study, I focus on the use of Hakwon and private lessons because these two types are the dominant way of private tutoring in secondary education and the question in the survey explicitly asked hours for these two types. In each survey students were asked to answer the following questions separately for both math and English: "How many hours per week do you participate in Hakwon and private lessons?", "How many hours per week do you study by yourself excluding school and tutoring hours?" Using the answers from these two questions I aggregate hours for mathematics and English for private tutoring and self-study.

To construct a choice set, I discretize the hours in the following way: If the aggregated hours are zero then the choice is zero. I then group next every 4 hours to be in the same category and assign a middle point as a representative choice. For private tutoring, students can choose among the options 0, 2.5, 6.5, 10.5, 14.5 and 18.5 hours, and for self-study they can choose among 0, 2.5, 6.5, 10.5, 14.5, 18.5, 22.5, 26.5 and 30.5 hours. Considering the distribution of study hours observed in the data, this choice set is comprehensive enough to include all possible alternatives given the physical time available in a week. [15]

Other Characteristics

In the baseline survey, parents reported their highest levels of educational attainment. Both parents' educational attainments are classified as one of four levels: high school or less,

[15] The difference between a middle point in each category and an average of students' observed hours who belong to each category is small.

two-year junior college, four-year university and graduate degree or above. Parents also reported their monthly household income and total number of children in the household. I distinguish location between metro and non-metro areas. Metro area includes Seoul and six other large cities and medium-sized cities in the Seoul metropolitan area.

Summary Statistics

Table 2 presents summary statistics from the KELS for the main variables regarding students' demographic characteristics and family backgrounds. It also contains summary statistics about hours of private tutoring and hours of self study at each grade level. Average hours of participation in private tutoring is around six hours at 8th grade. With higher grades, students use less private tutoring, the average is about three hours at 12th grade. In the sample, 51% of students are boys, and 61% of students live in metro areas. Average number of children in a household is slightly above two, and average annual household income is about 48 million won (KRW), about US$43,600.

Table 3 describes the pattern of private tutoring and self-study hours by income and by parental education when students are in the 8th grade. About 67% of children from the household in the highest income quartile participated in private tutoring more than 6.5 hours. However, the figure is 50% among the children from the household in the lowest income quartile. A similar pattern is observed by parental education level, however, the difference is less striking than the pattern by income level. As for self-study hours, we can see that there is a positive association between parental education and self-study hours. Figure 2 provides the distribution of hours for private tutoring and self-study at 8th and 11th grade. One noticeable change between the 8th and 11th grades in the figure is that at 11th grade, the use of private tutoring decreases dramatically and self-study hours increase substantially.

6 Empirical Results

The estimated model has 81 parameters. The full set of parameters of the model with their asymptotic standard errors are provided in Tables 4 and 5. Table 4 contains estimation results for utility function, income process, type probability, coefficient for terminal value function, and standard errors of stochastic shocks. Table 5 contains estimation results of education production functions.

In Table 5, we can see that private tutoring has a positive impact on academic achievement across all school levels and that the effects are stronger in the early stages of academic life. We can also see that the productivity of one additional hour of self-study increases slightly as schooling stages advance. Hence, the relative effectiveness of private tutoring compared to that of self-study decreases as schooling stages progress.[16] The coefficients in front of the interaction terms between private tutoring and self-study are all positive, which implies that in all grades they are complements in achievement production. This complementary relationship is important because it functions as a channel through which the government's intervention on private tutoring also impacts on self-study decision.

6.1 Model Fit

Table 6 compares the actual and predicted values of key summary measures of final test scores and hours for private tutoring and self-study. Overall, the model fits these unconditional means quite well. However, the model is less successful in fitting the variation observed in the data.

Table 7 provides the mean of test scores, private tutoring hours, and self-study hours by income group. The model slightly understates the gap between different level income groups in private tutoring hours in high school and overstates the gap in middle school. For self-study hours, the model overstates the gap by income group. However, the differences are less than 10% in magnitude. Similar patterns appear in Table 8, where the conditional means for the same variables are provided by parental education. Figure 3 and Figure 4 provide additional evidence of model fit by comparing the choice distribution by income level and by parental education.

6.2 Policy Simulations

I use the estimated behavioral model to conduct policy simulations, involving policies recently enacted by the Korean government. Each experiment investigates how a suggested government policy affects the overall test score distribution and whether the policy can improve test scores or affect test score inequality. To evaluate the impact of the each

[16] Although we normalize test scores to compare in a standard unit, results should be interpreted carefully because a test in one period is not directly comparable to that of another period.

policy, I simulate private tutoring and self-study choices and final test score distributions under the new policy environment. Each household's behavior is simulated 1,000 times (i.e., for 1,000 sets of draws of the model shocks); the reported results are averages from those simulations. Specifically, I perform experiments by 1) restricting the choice set of private tutoring; 2) providing a subsidy to a selected group of households by income level; and 3) levying a tax on private tutoring and using those tax revenues to provide vouchers to low income households.

6.2.1 Direct Regulation of Private Tutoring

The government has some powers to directly regulate private tutoring markets. In fact, the Korean government prohibited private tutoring in 1980 and only gradually lifted restrictions until 2000, when the Constitutional Court ruled that strict government regulation of private tutoring is unconstitutional. Recently, since 2010, cram schools have been required to close by 10 p.m. in Seoul and other metro areas.

In the model I allow each student to use up to 18.5 hours of private tutoring ($x \in \{0, 1, 2, 3, 4, 5\}$) in each period. For this policy simulation, I restrict the maximum hours of private tutoring use. The first scenario (CF1) allows each student to use up to 6.5 hours of private tutoring at most ($x \in \{0,1,2\}$). This experiment mimics the policy of regulating business hours as implemented in 2010 by the Korean government. The second scenario (CF2) is more restrictive, allowing up to 2.5 hours each week ($x \in \{0,1\}$). The third scenario (CF3) is the most restrictive, prohibiting use of private tutoring at all ($x=0$).

In Table 9, columns 2 to 4 provide the mean of time spent in private tutoring and in self-study and test scores under the imposed policy environments. For all three scenarios, the mean number of hours of using private tutoring substantially decreases due to the imposed time regulations. Self-study hours also decrease, because private tutoring and self-study hours are complements in the education production function. We can see that test scores in each period and final college entrance exam scores also drop, which implies a decrease in average academic performance. For example, in the third scenario (CF3), a decrease of 0.71 points from 5.28 to 4.57 amounts to about 0.47 standard deviation decrease.

Despite the loss in overall achievement, simulations based on the estimated model show that regulation of private tutoring hours narrows the achievement gap by income groups. Table 10 presents the mean college entrance exam score by income groups under each scenario. The gap between the top and bottom 25% income groups decreases by 57 percent, from 0.79 to 0.34 standard deviations, in CF3. If we look in Table 11 at the backgrounds of students who are in the upper 25% of the test score distribution, we can see an increase in the representation of students from the bottom income group from 16.3% in baseline to 21.2% in CF3; correspondingly, we note a decreasing share of students from the top income group.

6.2.2 Subsidies for Private Tutoring

One of the most serious concerns about private tutoring is the possibility that it could undermine equal opportunities in learning. If private tutoring is an effective tool for improving cognitive development of children, subsidies or vouchers to use for private tutoring can be provided to children in low income households.

I consider policies to provide private tutoring resources to students who are from the less affluent families. Specifically, I implement a policy in which the government provides a 50 percent price subsidy to low income households. If the annual household income is lower than the median income, the household is eligible for the subsidy (CF4). In another scenario (CF5), I implement a policy that provides free private tutoring to children of households whose annual income is lower than the 25th percentile of income distribution.

When the actual price that eligible households pay for private tutoring decreases, private tutoring becomes more accessible to children in low income households. They participate in more private tutoring and their overall performance improves. In Table 9, average test scores increase by about 0.18 standard deviations in CF4 and 0.13 standard deviations in CF5, and these average score improvements mostly come from test score gains among the children in low income households (Table 10). Cost of these policies is increases in the government spending on education. In case of CF4, it requires 15% increases in the government spending on secondary education given that the Korean government spent about US$10,000 per student on secondary education in 2009.

Hence, if either subsidy scenario occurs, the income achievement gap narrows by 47 percent between the top and bottom 25% income groups. We can also see that the composition of student family backgrounds becomes more diverse in the upper tails test score distribution as share of students from the low income group increases (Table 11). The proportion of students who belong to the top 25% score distribution increases from 16.7%

to 19.4% for the bottom income group, and, in CF4, it also increases from 24.2% to 26.1% for children in the next lowest level income group.

6.2.3 The Budget Balanced Policy: Vouchers Funded by Taxes on Private Tutoring

In this policy simulation, I impose taxes on private tutoring and use resulting tax revenues to provide vouchers for private tutoring to children from low income households. In Korea, vouchers have in fact been provided to children of low income families since 2006. The vouchers can be used for after-school programs that are held in schools after the official school day is over.

To balance the government budget, I set a voucher eligibility requirement at the 25th percentile of the annual household income distribution, and I allow eligible households to obtain up to 2.5 hours of private tutoring using a provided voucher (CF6). Simulation results suggest that to keep the government's budget balanced, a tax rate of approximately 20% for private tutoring is required.[17] Because the first 2.5 hours of private tutoring can be covered by vouchers, private tutoring is more accessible to eligible households for the first 2.5 hours. However, the increased tutoring price caused by the tax on private tutoring works as a burden for additional use of private tutoring over the initial 2.5 hours.

The simulation predicts that about 92% of eligible households will actually use their vouchers in each period given the stochastic shocks. However, due to the increased price of private tutoring due to the tax on it, average hours of private tutoring decrease, and average scores for all students drop by 0.07 standard deviations (Table 9). As a result of the government intervention in favor of low income households, the income achievement gap narrows by 31 percent, from 0.79 to 0.55, between the top and bottom 25% income groups (Table 10). We can also see that the share of students from the low income group in the top 25% score distribution increases from 16.7 percent to 19.2 percent.

[17] In my behavior model, households optimally decide their actual usage of vouchers and hours for private tutoring given stochastic shocks in each period. Thus, I cannot perfectly predict their usage of vouchers and their tutoring hours before conducting the simulation. By repeating simulations under different tax rates, I find that an approximately 20% tax rate is required to cover the government's expenditure for the suggested voucher scheme.

7 Conclusion

In this paper, I develop and structurally estimate a dynamic discrete choice model of private tutoring and self-study decisions using panel data from the 2005-2011 waves of the Korean Education Longitudinal Study. I use the model to investigate the effect of private tutoring on academic achievement and educational inequality, and to evaluate the impacts of a range of government policies on educational outcomes. This study is the first attempt in the research on private tutoring to conduct policy simulations with the goal of deriving policy implications based on a behavioral model.

My empirical analyses suggest that private tutoring has a positive impact on academic achievement, and that the effects are stronger in the early stages of academic life. It also shows that the productivity of self-study increases with grade level. The relative importance of self-study to academic achievement grows over time. I also find that private tutoring and self-study are complements in achievement production.

Simulations based on the estimated model show that prohibiting private tutoring reduces the achievement gap between higher and lower income households, but at the cost of decreasing average test scores by 0.47 standard deviations. Thus, private tutoring is an important determinant of Korean students' test score performance. Providing a 50 percent price subsidy for private tutoring to households under the median household income increases average test scores by 0.18 standard deviations and narrows the income achievement gap by 47 percent, at the cost of increased government spending. A voucher system funded by a tax on private tutoring also narrows the income achievement gap by 31 percent, but at the cost of decreasing average test scores by 0.07 standard deviations.

In future work, I will include the labor force participation decision of mothers in the model. Given the importance of parental involvement on child development, a mother's decision to work can be greatly affected by the presence of children and by the children's ages. I am choosing to focus on the mother's workforce participation decision only, as paternal workforce participation is generally a given at this point in time in Korean society. When a mother works, she increases household income, which contributes to increased utility through consumption and the greater availability of resources for private tutoring. However, when a mother does not work, she can provide more support to her children both academically and emotionally.

Another extension to the model that I plan to implement is to incorporate school quality as an additional input in the education production function. This will allow us to investigate important policy questions such as the effect of improved school quality on academic achievement. We can also examine substitutability between formal schooling and private tutoring by studying the extent to which increasing school quality changes demand for private tutoring. (Das et. al 2011; Pop-Eleches and Urquiola 2011)[18]

[18]I examined whether observed school characteristics–teacher's schooling attained, experience, and student-teacher ratio–have explanatory power in the education production function estimation. However, these variables are statistically insignificant. To incorporate school quality in the model, I need to acquire more detailed information about the school name, location, and other characteristics from KEDI.

References

[1] Baker, David P. and Gerald K. LeTendre. 2005. *National Differences, Global Similarities: World Culture and the Future of Schooling.* Stanford, CA: Stanford University Press.

[2] Bravo, David, Sankar Mukhopadhyay, and Petra E Todd. 2010. "Effects of school reform on education and labor market performance: Evidence from Chile's universal voucher system." *Quantitative economics*, 1(1):47-95, July

[3] Bray, Mark. 2007. *The shadow education system: private tutoring and its implications for planners.* 20(6):608-609

[4] Bray, Mark. 2011. *The Challenge of Shadow Education: Private tutoring and its implications for policy makers in the European Union.* Education. European Commission.

[5] Buchmann, Claudia, Dennis J Condron, and Vincent J Roscigno. 2010. "Shadow Education , American Style : Test Preparation , the SAT and College Enrollment." *Social Forces*, 89: 435-462

[6] Byun, Soo-yong. "Shadow Education and Academic Success in South Korea." 2012. *Working paper*

[7] Choi, Jaesung. "Unequal Access to Private Tutoring and Its Impacts on Academic Outcomes: Evidence from Korea." 2012. *Working paper*

[8] Dang, H.-a. and F. H. Rogers. 2008. "The Growing Phenomenon of Private Tutoring: Does It Deepen Human Capital, Widen Inequalities, or Waste Resources?" *The World Bank Research Observer*, 23(2):161-200

[9] Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2011. School Inputs, Household Substitution, and Test Scores. NBER Working Paper.

[10] Duflo, Esther, Rema Hanna, and Stephen P Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review*, 102(4): 1241-1278.

[11] Eckstein, Zvi and Kenneth I. Wolpin. 1999. "Why youths drop out of high school: The impact of preferences,opportunities, and abilities." *Econometrica*, 67(6):1295-1339

[12] Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature*, 24 (3):1141-1177

[13] Heckman, James and Burton Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, 52:271-320.

[14] Jayachandran, Seema. 2012. "Incentives to Teach Badly : After-School Tutoring in Developing Countries." *Working paper*

[15] Jung, Jin Hwa and Kyung Hee Lee. 2010. "The determinants of private tutoring participation and attendant expenditures in Korea." *Asia Pacific Education Review*, 11(2):159-168.

[16] Keane, Michael P., Petra E Todd, and Kenneth I Wolpin. 2011. "The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications." volume 4. Elsevier Inc.

[17] Keane, Michael P. and Kenneth I Wolpin. 1994. "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *Review of Economics and Statistics*, 76(4):648-672

[18] Keane, Michael P. and Kenneth I Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy*, 105(3):473-522.

[19] Kim, Hisam. 2010. "Analysis of the Effects of Private Education on Academic Achievement and Labor Market Performance." *KDI Report 2010-05.*

[20] Kim, Kyung-Min and Daekwon Park. 2012. "Impacts of Urban Economic Factors on Private Tutoring Industry." *Asia Pacific Education Review*, 13:273-280.

[21] Kim, Misook. 2012. "South Koreas Educational Sedative: Private Institutes." In *No Alternative?: Experiments in South Korean Education*, edited by Nancy Abelmann, Jung-ah Choi, and So Jin Park, 128-154. Berkeley: Univ. of California Press.

[22] Kim, Sunwoong and JuHo Lee. 2010. "Private Tutoring and Demand for Education in South Korea." *Economic Development and Cultural Change*, 58(2):259-296.

[23] Korea National Statistical Office. 2012. "The survey of private education expenditures in 2011." Seoul: Statistics Korea.

[24] Kim, Y., Kim, S., Kang, S., Kim, H., Shin, J., & Park, S. 2006. "Korean educational longitudinal study 2005 (II)." Seoul, Korea: Korean Educational Development Institute. (in Korean)

[25] Kim, Yangboon and Sookyung Yang. 2011. "Trend and Prediction of Private Tutoring Expenditure." KEDI Pisition Paper. (in Korean)

[26] Kuan, P. Y. 2011. "Effects of Cram Schooling on Mathematics Performance: Evidence from Junior High Students in Taiwan." *Comparative Education Review*, 55(3):342-368.

[27] Lee, Chong Jae, Heesook Lee, and Hyo-Min Jang. "The history of policy responses to shadow education in South Korea: implications for the next cycle of policy responses." *Asia Pacific Education Review*, 11(1):97-108, February 2010.

[28] McFadden, Daniel. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica*, 57: 995-1026.

[29] Oh, Mahn Seug. 2010. "A Comparative Study on the Time USE of Korean Students. *Korean Journal of Comparative Education*, 20: 299-330. (in Korean)

[30] Park, Cheol-sung. 2011. "A Study on the Determinants of Parental Satisfaction with Public Education and Demand for Private Tutoring for the Child." *Analysis of Korean Economy*, 17(1): 73-119.

[31] Park, Hyunjoon, Soo-yong Byun, and Kyung-keun Kim. 2011. "Parental Involvement and Students' Cognitive Outcomes in Korea: Focusing on Private Tutoring." Sociology of Education, 84(1):3-22.

[32] Park, Hyunjoon, Jere R Behrman, and Jaesung Choi. Forthcoming. "Causal Effects of Single-Sex Schools on College Entrance Exams and College Attendance : Random Assignment in Seoul High Schools." *Demography*

[33] Park, Hyunjoon. 2010. Japanese and Korean high schools and students in comparative perspective. In J. Dronkers (Ed.), Quality and inequality of education: A cross-national attempt to unravel it (pp. 255-273)Dordrecht, The Netherlands: Springer.

[34] Park, So Jin. 2012. "The Private After-School Market: South Korean Mothers" Anxious Education Management." In *No Alternative?: Experiments in South Korean Education*, edited by Nancy Abelmann, Jung-ah Choi, and So Jin Park, 155-183. Berkeley: Univ. of California Press.

[35] Phillips, Anna M. 2012. After Number of Gifted Soars, a Fight for Kindergarten Slots. *New York Times*. 14 April.

[36] Phillips, Anna M. 2012. Tutoring Surges With Fight for Middle School Spots. *New York Times*. 16 April.

[37] Pop-eleches, Cristian, and Miguel Urquiola. 2011. "Going to a Better School: Effects and Behavioral Responses." NBER Working Paper No. 16886.

[38] Plantenga, T. D. 2009. "HOPSPACK 2.0 User Manual." Sandia National Laboratories, Albuquerque, NM and Livermore, CA, SAND2009-626.

[39] Ryu, Deockhyun and Changhui Kang. Forthcoming. "Do Private Tutoring Expenditures Raise Academic Performance? Evidence from Middle School Students in South Korea." *Asian Economic Journal.*

[40] Safarzynska, K. 2011. "Socio-economic Determinants of Demand for Private Tutoring." *European Sociological Review.*

[41] Tansel, A and F Bircan. 2006. "Demand for education in Turkey: A tobit analysis of private tutoring expenditures." *Economics of Education Review* , 25(3):303-313.

[42] Todd, Petra E. and Kenneth I Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *The Economic Journal*, 113(485):3-33.

[43] Todd, Petra E. and Kenneth I.Wolpin. 2010. "Structural Estimation and Policy Evaluation in Developing Countries. *Annual Review of Economics*, 2(1):21-50

A1. Appendix

Likelihood Details

As noted in the data section, test scores in 10th and 11th grades are not available in the data. Hence, likelihood components for period 5 and 6 include integration over possible paths of test scores in those periods.

$$Pr(O_t|\overline{\Omega}_t) = Pr(q_t, x_t, s_t, y_t|\overline{\Omega}_t) = Pr(q_t|x_t, s_t, y_t, \overline{\Omega}_t) \cdot Pr(x_t, s_t|y_t, \overline{\Omega}_t) \cdot f(y_t|\overline{\Omega}_t)$$

$$\text{where } \overline{\Omega}_t = \{E_f, E_m, age_p, n, metro, q_{t-1}\}$$

- Period 5: q_4 is not observed

$$\begin{aligned} Pr(O_5|\overline{\Omega}_5) =& Pr(x_5, s_5, y_5|\overline{\Omega}_5) \\ =& \sum_{\tilde{q}_4 \in Q_4} f(\tilde{q}_4|q_3, x_4, s_4) \cdot Pr(x_5, s_5, y_5|\tilde{\Omega}_5) \end{aligned}$$

$$\text{where } \tilde{\Omega}_t = \{E_f, E_m, age_p, n, metro, \tilde{q}_{t-1}\}$$

- Period 6: q_5 is not observed

$$\begin{aligned} Pr(O_6|\overline{\Omega}_6) =& Pr(q_6, x_6, s_6, y_6|\overline{\Omega}_6) \\ =& \sum_{\tilde{q}_5 \in Q_5} f(\tilde{q}_5|\tilde{q}_4, x_5, s_5) \cdot Pr(q_6, x_6, s_6, y_6|\tilde{\Omega}_6) \\ =& \sum_{\tilde{q}_5 \in Q_5} \sum_{\tilde{q}_4 \in Q_4} \underbrace{f(\tilde{q}_5|\tilde{q}_4, x_5, s_5) \cdot f(\tilde{q}_4|q_3, x_4, s_4)}_{\equiv \Psi} \cdot Pr(q_6, x_6, s_6, y_6|\tilde{\Omega}_6) \\ =& \sum_{\tilde{q}_5 \in Q_5} \sum_{\tilde{q}_4 \in Q_4} \Psi \cdot Pr(q_6|x_6, s_6, y_6, \tilde{\Omega}_6) \cdot Pr(x_6, s_6, y_6|\tilde{\Omega}_6) \end{aligned}$$

$$\text{where } \tilde{\Omega}_t = \{E_f, E_m, age_p, n, metro, \tilde{q}_{t-1}\}$$

A2. Tables

Table 1: Summary Statistics from Survey of Private Education 2010

	Variable	Obs	Mean	Std.	Min	Max
Elementary School	Spending in Academics	22068	194.6	193.0	0	1453.3
	(if spending>0)	15974	268.8	177.6	7.9	1453.3
	Spending in English	22068	93.5	113.8	0	941.8
	(if spending>0)	12591	163.9	105.6	2.3	941.8
	Spending in Mathematics	22068	49.6	69.6	0	800.1
	(if spending>0)	12350	88.7	72.1	2.3	800.1
	Hours of PT in Academics	22068	4.9	5.0	0	30
	(if spending>0)	15974	6.8	4.7	1	30
Middle School	Spending in Academics	23033	281.4	273.9	0	2622.8
	(if spending>0)	15793	410.5	237.5	7.9	2622.8
	Spending in English	23033	107.6	123.9	0	1159.5
	(if spending>0)	14297	173.3	115.5	2	1159.5
	Spending in Mathematics	23033	106.1	123.1	0	1192.2
	(if spending>0)	14548	168.0	116.5	2	1192.2
	Hours of PT in Academics	23033	7.1	6.6	0	42
	(if spending>0)	15793	10.3	5.6	1	42
High School	Spending in Academics	36083	247.0	338.9	0	3988.8
	(if spending>0)	18758	475.2	335.5	3.2	3988.8
	Spending in English	36083	83.6	145.8	0	2087.2
	(if spending>0)	12177	247.8	149.5	0.8	2087.2
	Spending in Mathematics	36083	118.2	175.0	0	2319.1
	(if spending>0)	15184	280.8	164.8	0.8	2319.1
	Hours of PT in Academics	36083	3.5	4.6	0	40
	(if spending>0)	18758	6.8	4.3	1	40

Note: Weekly hours for private tutoring are reported in the table. Unit of spending is 1,000 Korean Won (KRW), which is equivalent to US$0.9.

Table 2: KELS Summary Statistics

	Mean	Std	Min	Max
Education Level of Father	13.88	2.07	12.0	18.0
Education Level of Mother	13.03	1.70	12.0	18.0
Age of Father	43.42	3.61	32.0	68.0
Age of Mother	40.77	3.26	30.0	54.0
Gender of Student	0.51	0.50	0.0	1.0
Living in Metro	0.61	0.49	0.0	1.0
Num. of Children	2.15	0.53	1.0	3.0
Household Income at 8th	4627	2123	996	14400
Household Income at 10th	5064	2359	1080	14400
Household Income at 12th	4795	2285	1080	14400
Hours for Private Tutoring at 7th	5.37	4.51	0.0	14.5
Hours for Private Tutoring at 8th	6.14	4.59	0.0	14.5
Hours for Private Tutoring at 9th	5.81	4.66	0.0	14.5
Hours for Private Tutoring at 10th	3.85	3.99	0.0	14.5
Hours for Private Tutoring at 11th	3.33	3.78	0.0	14.5
Hours for Private Tutoring at 12th	2.73	3.77	0.0	14.5
Hours for Self-Study at 7th	2.17	3.18	0.0	14.5
Hours for Self-Study at 8th	4.42	3.76	0.0	14.5
Hours for Self-Study at 9th	4.87	4.03	0.0	14.5
Hours for Self-Study at 10th	5.78	4.23	0.0	14.5
Hours for Self-Study at 11th	6.44	4.51	0.0	14.5
Hours for Self-Study at 12th	9.97	6.89	0.0	22.5
Standardized Test Score at 6th	5.49	1.38	1.8	7.5
Standardized Test Score at 7th	5.64	1.32	1.8	8.0
Standardized Test Score at 8th	5.60	1.40	2.2	8.1
Standardized Test Score at 9th	5.65	1.38	2.5	7.9
Standardized Test Score at 12th	5.29	1.41	2.1	8.9
Num. of Sample	2192			

Note: Weekly hours are reported. Unit of income is 10,000 Korean Won (KRW), which is equivalent to US$9.

Table 3: Hours of Private Tutoring and Self Study by Income and Parental Education

Income 8th Grade	**Private Tutoring**					**Self-Study**				
	0	2.5	6.5	10.5	14.5	0	2.5	6.5	10.5	14.5
Bottom	29.01	20.62	25.36	19.53	5.47	14.96	57.85	15.33	8.39	3.47
2nd	21.24	20.46	27.99	22.01	8.3	13.9	54.44	18.92	8.88	3.86
3rd	15.25	18.79	30.14	23.4	12.41	10.99	52.13	22.52	8.69	5.67
Top	12.1	18.86	32.56	23.13	13.35	9.07	45.37	25.44	12.1	8.01

Income 10th Grade	**Private Tutoring**					**Self-Study**				
	0	2.5	6.5	10.5	14.5	0	2.5	6.5	10.5	14.5
Bottom	59.27	18.15	16.41	4.83	1.35	8.11	38.8	27.8	15.44	9.85
2nd	47.2	23.6	22.03	5.59	1.57	8.92	34.27	26.57	16.61	13.64
3rd	39.22	22.84	26.94	9.48	1.51	7.54	35.99	22.84	19.83	13.79
Top	26.18	26.8	31.35	11.13	4.55	5.33	30.88	27.27	19.75	16.77

Parental Edu 8th Grade	**Private Tutoring**					**Self-Study**				
	0	2.5	6.5	10.5	14.5	0	2.5	6.5	10.5	14.5
12	23.33	20.04	29.14	19.65	7.84	15.2	56.63	17.72	7.94	2.52
14	15.36	22.29	26.81	26.81	8.73	11.45	54.82	20.48	7.83	5.42
16+	15.84	18.14	29.87	23.1	13.06	8.71	46.07	24.3	12.21	8.71

Parental Edu 10th Grade	**Private Tutoring**					**Self-Study**				
	0	2.5	6.5	10.5	14.5	0	2.5	6.5	10.5	14.5
12	50.05	21.39	20.43	6.68	1.45	9.49	38.04	26.23	16.65	9.58
14	43.67	22.89	23.19	7.83	2.41	7.83	39.16	20.48	16.57	15.96
16+	31.92	25.27	29.99	9.31	3.51	4.59	28.78	28.66	20.07	17.9

Note: Numbers in each cell presents the share of students who spent a certain number of hours in each subgroup.

Table 4: Model Estimates

Function		Param	Est.	Std
Utility Function	Num. Children	α_1	0.073173	0.00110157
	Constant	α_2	104.278032	1.72064435
	Type=2	α_3	26.98908	0.55808865
	Type=3	α_4	28.034345	0.31468845
	$E_p \geq 16$	α_5	37.873841	0.19789393
	q^2	α_6	0.225522	0.00094942
	PT	α_7	-16.442136	0.33355433
	PT-Type=2	α_8	12.238903	0.22370775
	PT-Type=3	α_9	26.808385	0.41785522
	$PT * C$	α_{10}	0.675897	0.00001004
	x^2	α_{11}	-1.544009	0.01090736
	SS	α_{12}	-65.151022	0.17526502
	SS-Type=2	α_{13}	-9.337376	0.06847887
	SS-Type=3	α_{14}	-7.796231	0.00263951
	s^2	α_{15}	-2.093324	0.01097525
	Habit-Persistence	α_{15}	-8.466454	0.14656128
Income Process	Constant	γ_1	6.575002	0.00305775
	Type=2	γ_2	0.455494	0.0037551
	Type=3	γ_3	0.901637	0.00515619
	Father's Edu	γ_4	0.028991	0.00035739
	Mother's Edu	γ_5	0.039886	0.00051408
	Father's Age	γ_6	0.01002	0.00014291
	Age^2	γ_7	-0.000055	0.00000146
	Metro	γ_8	-0.000001	0.00000076
	Num. Children	γ_9	0.039138	0.0005972
Type Probability	Constant	ω_{21}	-3.74539	0.07268062
Type=2	Father's Edu	ω_{22}	0.163442	0.00368641
	Mother's Edu	ω_{23}	0.209143	0.00331028
	Num. Children	ω_{24}	-0.163398	0.00001263
	Metro	ω_{25}	0.187984	0.00476695
Type=3	Constant	ω_{31}	-6.344735	0.0878548
	Father's Edu	ω_{32}	0.194093	0.00371274
	Mother's Edu	ω_{33}	0.308658	0.00440124
	Num. Children	ω_{34}	-0.253905	0.00086692
	Metro	ω_{35}	0.459565	0.00841465
Termina Value		ζ	744.821524	0.48833968
Shocks	Private Tutoring	ϵ_x	48.87358	0.42383424
	Self-Study	ϵ_s	15.581895	0.10039517
	EPF	ϵ_q	0.936903	0.00631652
	Income	ϵ_y	0.287154	0.0012836

Table 5: Model Estimates (Continued)

Function		Param	Est.	Std
EPF-7th	Constant	β_1	2.475331	0.02807592
	Type=2	β_2	0.409273	0.00636533
	Type=3	β_3	0.663742	0.01123612
	Lagged Score	β_4	0.455717	0.00483908
	x^2	β_6	-0.001692	0.0000572
	$E_p = 14$	β_8	0.001118	0.00002941
	$E_p \geq 16$	β_9	0.001374	0.00009376
	s^2	β_{10}	-0.000001	0.00000006
	$x * s$	β_{11}	0.00007	0.00000457
	PT 7th	β_{51}	0.052001	0.00086333
	SS 7th	β_{71}	0.036819	0.00025578
EPF-8th & 9th	Constant	β_1	1.041562	0.01937125
	Type=2	β_2	0.000004	0.00114618
	Type=3	β_3	0.019416	0.00221089
	Lagged Score	β_4	0.742022	0.00415971
	x^2	β_6	-0.001637	0.00005061
	$E_p = 14$	β_8	0.000203	0.00001322
	$E_p \geq 16$	β_9	0.000023	0.00004698
	s^2	β_{10}	-0.000116	0.00000037
	$x * s$	β_{11}	0.000311	0.0000094
	PT 8th	β_{52}	0.052369	0.00095721
	PT 9th	β_{53}	0.048697	0.00095251
	SS 8th	β_{72}	0.042172	0.00034528
	SS 9th	β_{73}	0.040639	0.00036289
EPF-High	Constant	β_1	0.581666	0.01456256
	Type=2	β_2	0.065169	0.0003368
	Type=3	β_3	0.006379	0.00022997
	Lagged Score	β_4	0.784386	0.00324778
	x^2	β_6	-0.001309	0.00005922
	$E_p = 14$	β_8	0.00111	0.00002834
	$E_p \geq 16$	β_9	0.000004	0.0000306
	s^2	β_{10}	-0.000002	0.00000003
	$x * s$	β_{11}	0.0001	0.00000596
	PT 10th	β_{54}	0.051819	0.00129527
	PT 11th	β_{55}	0.046681	0.00130501
	PT 12th	β_{56}	0.039048	0.00126071
	SS 10th	β_{74}	0.041308	0.00042094
	SS 11th	β_{75}	0.040848	0.00044577
	SS 12th	β_{76}	0.042983	0.0006474

Table 6: Model Fit - Mean and Std. of PT & SS Hours and Test Scores

	Data		**Model Fit**	
	Mean	**Std**	**Mean**	**Std**
PT Hours (7th)	5.37	4.51	5.32	4.52
PT Hours (8th)	6.14	4.59	6.04	4.66
PT Hours (9th)	5.81	4.66	5.76	4.45
PT Hours (10th)	3.85	3.99	3.95	4.22
PT Hours (11th)	3.33	3.78	3.34	3.88
PT Hours (12th)	2.73	3.77	2.65	3.47
SS Hours (7th)	2.17	3.18	2.07	2.80
SS Hours (8th)	4.42	3.76	4.63	3.56
SS Hours (9th)	4.87	4.03	5.08	3.60
SS Hours (10th)	5.78	4.23	6.03	3.95
SS Hours (11th)	6.44	4.51	6.66	4.01
SS Hours (12th)	9.97	6.89	9.98	6.54
Score (7th)	5.64	1.32	5.62	1.19
Score (8th)	5.60	1.40	5.64	1.34
Score (9th)	5.65	1.38	5.64	1.42
Score (12th)	5.29	1.41	5.28	1.75

Table 7: Model Fit - Mean of PT & SS Hours and Test Scores by Income

Income	**Data**				**Model Fit**			
	Bottom	**2nd**	**3rd**	**Top**	**Bottom**	**2nd**	**3rd**	**Top**
PT Hours (7th)	4.24	5.09	5.61	6.48	3.98	5.05	5.66	6.60
PT Hours (8th)	5.01	5.85	6.69	6.95	4.61	5.70	6.47	7.40
PT Hours (9th)	4.62	5.56	6.04	6.81	4.44	5.47	6.09	7.02
PT Hours (10th)	2.65	3.14	4.11	5.09	2.78	3.65	4.25	5.12
PT Hours (11th)	2.22	2.84	3.54	4.54	2.37	3.07	3.55	4.35
PT Hours (12th)	1.53	2.44	3.13	3.69	1.78	2.40	2.82	3.58
SS Hours (7th)	1.70	2.08	2.26	2.63	1.48	1.88	2.23	2.69
SS Hours (8th)	3.83	4.08	4.50	5.22	3.71	4.39	4.88	5.54
SS Hours (9th)	4.14	4.25	5.13	5.83	4.26	4.88	5.28	5.89
SS Hours (10th)	5.07	5.27	6.23	6.41	5.40	5.82	6.19	6.71
SS Hours (11th)	5.83	6.31	6.47	7.05	6.32	6.45	6.70	7.19
SS Hours (12th)	8.64	9.84	10.21	11.02	8.94	9.64	10.17	11.15
Score (7th)	5.23	5.48	5.79	6.01	5.27	5.56	5.73	5.94
Score (8th)	5.21	5.42	5.66	6.10	5.28	5.57	5.75	5.97
Score (9th)	5.27	5.53	5.69	6.03	5.29	5.57	5.73	5.96
Score (12th)	4.84	5.16	5.36	5.72	4.93	5.24	5.38	5.57

Table 8: Model Fit - Mean of PT & SS Hours and Test Scores by Parental Education

Education	**Data**			**Model Fit**		
	12	**14**	**16+**	**12**	**14**	**16+**
PT Hours (7th)	4.71	5.44	6.16	4.88	5.16	5.93
PT Hours (8th)	5.60	6.38	6.71	5.54	5.74	6.79
PT Hours (9th)	5.14	5.77	6.66	5.31	5.59	6.38
PT Hours (10th)	3.20	3.87	4.64	3.52	3.73	4.58
PT Hours (11th)	2.77	3.25	4.07	2.97	3.18	3.86
PT Hours (12th)	2.28	2.76	3.27	2.24	2.51	3.21
SS Hours (7th)	1.67	1.96	2.89	1.45	1.79	2.95
SS Hours (8th)	3.77	4.31	5.28	3.96	4.16	5.65
SS Hours (9th)	4.21	4.54	5.83	4.43	4.65	6.06
SS Hours (10th)	5.07	6.09	6.54	5.31	6.00	6.94
SS Hours (11th)	5.79	6.36	7.28	6.00	6.66	7.50
SS Hours (12th)	9.03	9.83	11.20	8.70	10.01	11.56
Score (7th)	5.26	5.62	6.12	5.42	5.63	5.88
Score (8th)	5.24	5.56	6.07	5.44	5.60	5.91
Score (9th)	5.30	5.63	6.09	5.44	5.57	5.91
Score (12th)	4.92	5.19	5.78	5.05	5.27	5.58

Table 9: Simulation Results - Mean of PT & SS Hours and Test Scores

	Baseline	**CF1**	**CF2**	**CF3**	**CF4**	**CF5**	**CF6**
PT Hours (7th)	5.3	4.0	1.9	0.0	6.7	6.4	4.7
PT Hours (8th)	6.0	4.4	2.0	0.0	7.5	7.1	5.3
PT Hours (9th)	5.8	4.3	2.0	0.0	7.1	6.8	5.1
PT Hours (10th)	4.0	3.1	1.6	0.0	6.3	5.7	3.0
PT Hours (11th)	3.3	2.8	1.4	0.0	5.5	5.0	2.6
PT Hours (12th)	2.7	2.2	1.2	0.0	4.6	4.2	2.1
SS Hours (7th)	2.1	2.0	1.8	1.7	2.2	2.1	2.0
SS Hours (8th)	4.6	4.3	3.8	3.4	4.9	4.8	4.5
SS Hours (9th)	5.1	4.8	4.3	3.8	5.4	5.3	4.9
SS Hours (10th)	6.0	5.9	5.6	5.2	6.4	6.2	5.9
SS Hours (11th)	6.7	6.6	6.3	5.9	7.0	6.9	6.5
SS Hours (12th)	10.0	9.7	8.7	7.5	10.9	10.7	9.6
Score (7th)	5.62	5.60	5.51	5.42	5.67	5.66	5.60
Score (8th)	5.64	5.57	5.39	5.20	5.73	5.70	5.60
Score (9th)	5.64	5.55	5.31	5.06	5.75	5.72	5.58
Score (12th)	5.28	5.18	4.91	4.57	5.55	5.47	5.17

Table 10: Simulation Results - Mean Test Scores by Income Group

	Baseline	**CF1**	**CF2**	**CF3**	**CF4**	**CF5**	**CF6**
		$x \leq 6.5$	$x \leq 2.5$	$x = 0$	50%	100%	Voucher
Bottom	4.82	4.77	4.60	4.35	5.26	5.40	4.87
2nd	5.27	5.18	4.91	4.58	5.62	5.40	5.12
3rd	5.42	5.32	5.02	4.67	5.65	5.46	5.27
Top	5.62	5.47	5.10	4.69	5.66	5.62	5.42
Top-Bottom	0.79	0.70	0.50	0.34	0.40	0.22	0.55
Change (%)		-13%	-37%	-57%	-50%	-72%	-31%

Table 11: Percentage of Each Income Group in the Top 25% of College Entrance Exam

	Baseline	**CF1**	**CF2**	**CF3**	**CF4**	**CF5**	**CF6**
		$x \leq 6.5$	$x \leq 2.5$	$x = 0$	50%	100%	Voucher
Bottom	16.7	17.5	18.9	20.7	19.4	23.4	19.2
2nd	24.2	24.4	24.9	24.8	26.1	23.6	23.9
3rd	27.7	27.6	27.4	27.0	27.0	24.9	26.9
Top	31.4	30.5	28.9	27.5	27.5	28.1	30.0

Note

CF1 - Allow maximum two units for private tutoring: x∈{0,1,2}

CF2 - Allow maximum one unit for private tutoring: x∈{0,1}

CF3 - Ban private tutoring: x=0

CF4 - 50% price subsidy to household with less than median income

CF5 - 100% price subsidy to household under the 25th percentile household income distribution

CF6 - Voucher for 2.5 hours of PT to the 25th percentile household income distribution (Government spending is supported by 20% of tax on private tutoring)

Table 12: Simulation Results: Private Tutoring and Self-Study Hours by Income Group

	Baseline	**CF1**	**CF2**	**CF3**	**CF4**	**CF5**	**CF6**
By Income Level: Private Tutoring							
PT Hours (8th)							
Bottom	4.6	3.6	1.7	0.0	7.0	8.9	4.5
2nd	5.7	4.3	2.0	0.0	8.1	5.7	4.8
3rd	6.5	4.7	2.1	0.0	7.4	6.5	5.5
Top	7.4	5.1	2.2	0.0	7.4	7.4	6.4
PT Hours (11th)							
Bottom	2.4	2.0	1.1	0.0	5.9	9.0	2.8
2nd	3.1	2.6	1.4	0.0	6.9	3.1	2.0
3rd	3.6	2.9	1.5	0.0	4.8	3.6	2.4
Top	4.4	3.5	1.7	0.0	4.4	4.4	3.0
By Income Level: Self-Study							
SS Hours (8th)							
Bottom	3.7	3.5	3.2	2.9	4.2	4.4	3.7
2nd	4.4	4.1	3.7	3.2	4.9	4.4	4.2
3rd	4.9	4.5	4.0	3.5	5.1	4.9	4.7
Top	5.5	5.1	4.5	3.9	5.6	5.5	5.3
SS Hours (11th)							
Bottom	6.3	6.3	6.1	5.8	6.9	7.0	6.3
2nd	6.4	6.4	6.1	5.7	6.9	6.5	6.3
3rd	6.7	6.6	6.3	5.9	7.0	6.7	6.5
Top	7.2	7.0	6.6	6.1	7.3	7.2	6.9

Note

CF1 - Allow maximum two units for private tutoring: x∈{0,1,2}

CF2 - Allow maximum one unit for private tutoring: x∈{0,1}

CF3 - Ban private tutoring: x=0

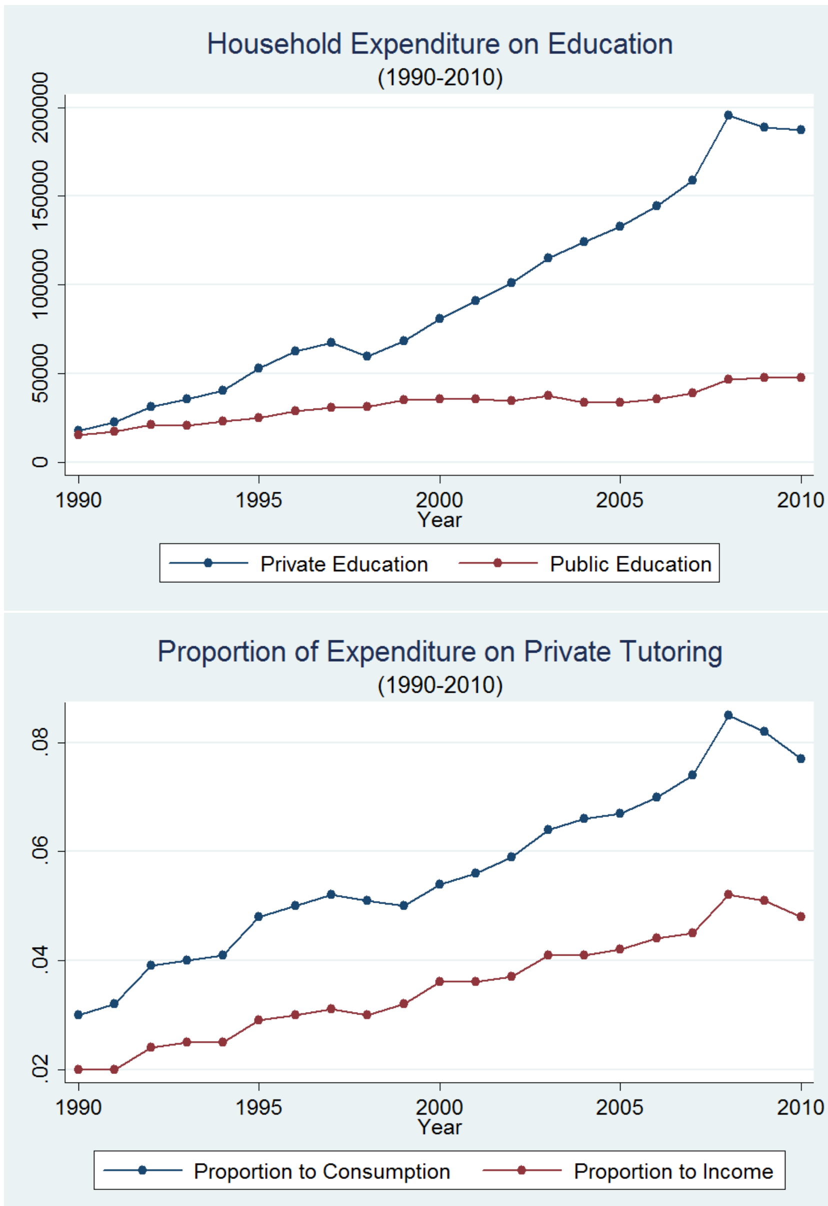
CF4 - 50% price subsidy to household with less than median income

CF5 - 100% price subsidy to household under the 25th percentile household income distribution

CF6 - Voucher for 2.5 hours of PT to the 25th percentile household income distribution (Government spending is supported by 20% of tax on private tutoring)

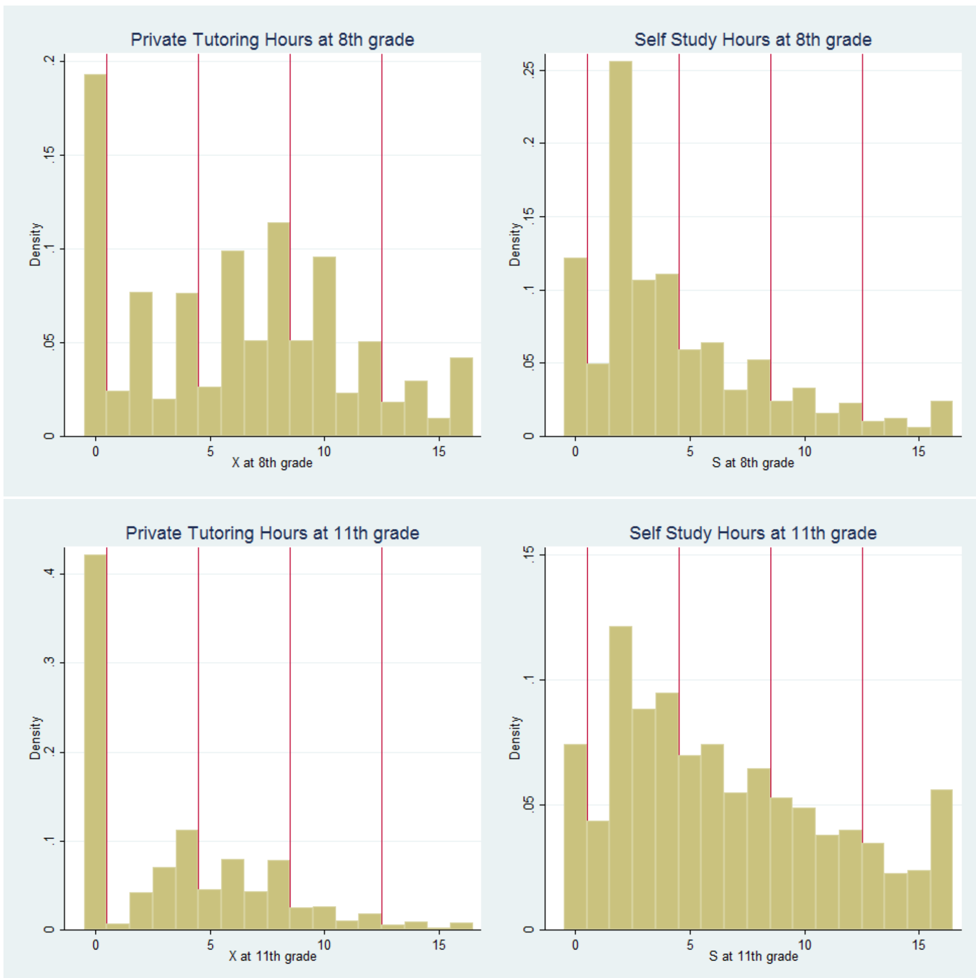
A3. Figures

Figure 1: Private Tutoring Time Trend



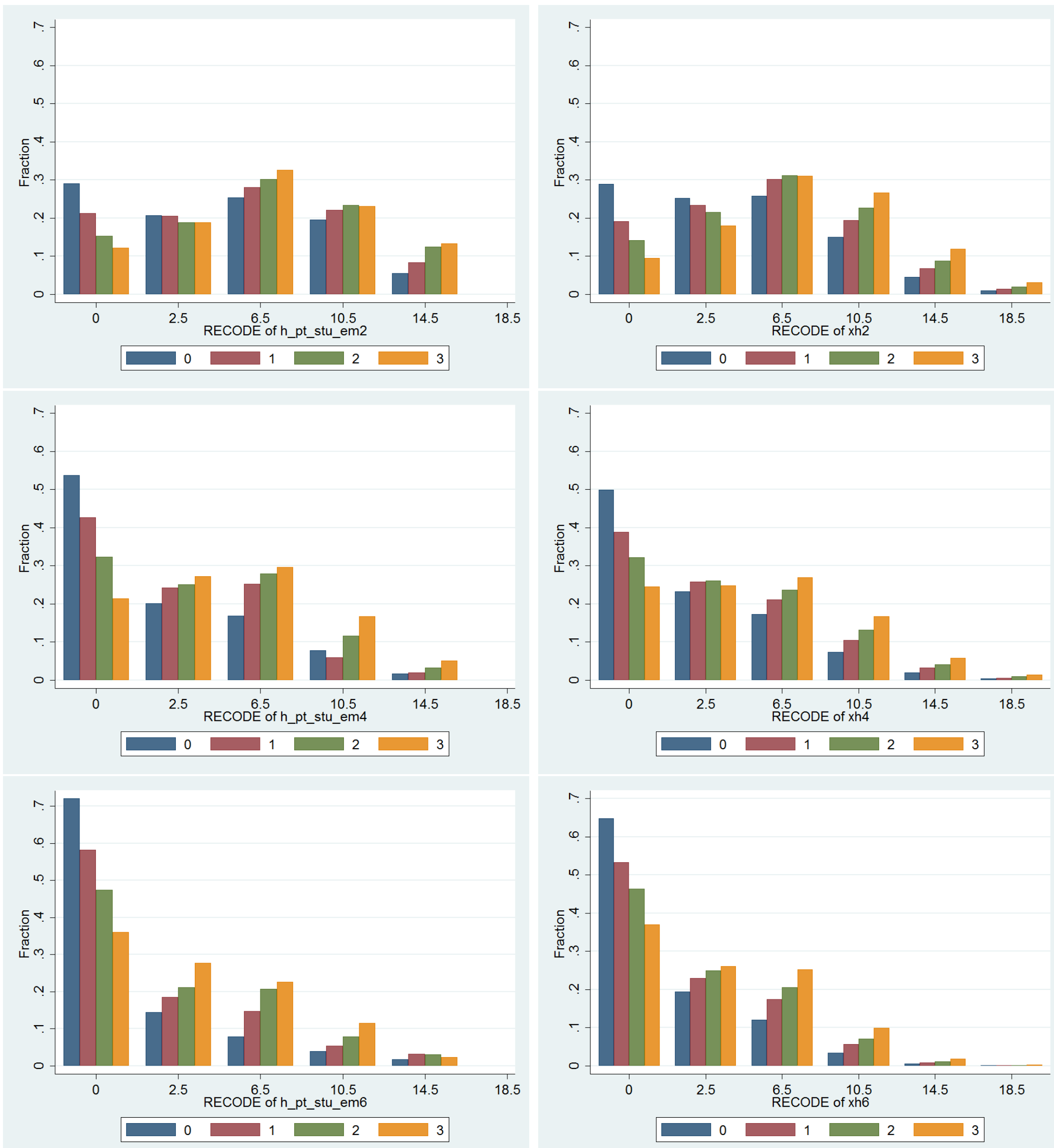
Note: First figure plots the household expenditure on private tutoring and public education over 20 years. Bottom figure presents the proportion of expenditure on private tutoring out of total consumption and out of household income. Source: Kim (2011).

Figure 2: Private Tutoring and Self Study Hours Distribution



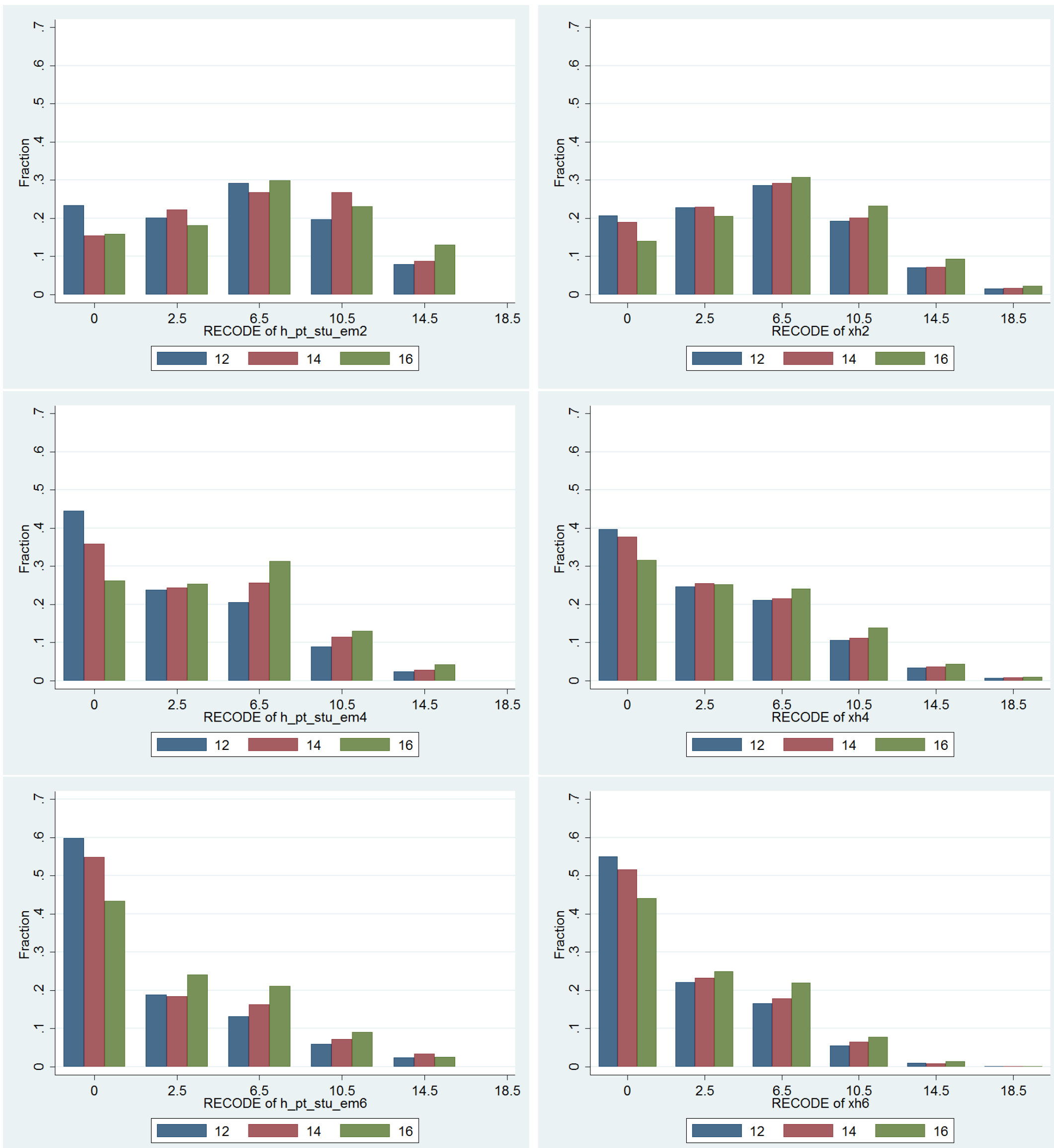
Note: This figure presents the distribution of time spent in private tutoring and in self-study at 8th and 11th grades. Source: KELS.

Figure 3: PT Distribution by Income Group at 8th, 10th, 12th - Data & Model Fit



Note: This figure compares distribution of private tutoring participation by income group. Graphs in the left side are distribution from data, and graphs in the right side are distribution from the simulation at each corresponding grade. Source: KELS.

Figure 4: PT Distribution by Parental Education at 8th, 10th, 12th - Data & Model Fit



Note: This figure compares distribution of private tutoring participation by parental education. Graphs in the left side are distribution from data, and graphs in the right side are distribution from the simulation at each corresponding grade. Source: KELS.