The Determinants of Demand for Private Tutoring in South Korea*

Ji-Ha Kim†
Department of Education, Yonsei University, South Korea

July, 2007

ABSTRACT. Private tutoring (PT) has been a growing economic phenomenon in South Korea for many years. This study investigated the determinants of the demand for PT in South Korea. Data were collected from 45 proportionally stratified South Korean high schools, and 3,605 questionnaires were analyzed using the Heckman two-stage sample selection correction method. Additionally this study exploited the local government regulation of PT to identify participation in PT, and this serves as an identifier of the selection correction term in the second stage outcome equation (expenditure on PT and hours of PT).

Results of the regression analysis showed that among the students and family background characteristics, students' achievement level, household income and parents' education level were positively associated with a higher probability to participate in PT and higher expenditure and spending hours of PT. At the school level, students in schools with higher student-teacher ratio were expected to spend more time on PT. The contextual effect measured by the proportion of classmates receiving PT services were significantly and positively related to expenditure on PT. Residence in urban areas had greater expenditure and hours spent on PT. The implications of these findings are discussed here.

Keywords: Private Tutoring, Demand for Education, Expenditure

* The author would like to thank Professors Henry Levin and Mun Tsang for their helpful comments and suggestions during the preparation of this manuscript.

† Email: jk2791@columbia.edu, porommy@hanmail.net

Ⅰ. Introduction

While PT[1] has had a long history in East Asian regions, it has recently become a thriving industry paralleling the burgeoning demand for PT in many other parts of the world. Bray's investigations (1999, 2003, 2005) on PT in various countries highlights the growing demand for PT. This is noted not just in the East Asian countries of Japan, Hong Kong, Singapore and China, but in large developing countries such as Brazil and Zimbabwe as well. In Eastern Europe, PT has emerged as a major enterprise with the collapse of socialism and the advent of the market economy.[2] The causes which induce the demand for PT have been reported by several researchers (i.e., Dore, 1997; Kwok, 2003; Bray, 2005; Tansel and Fatma, 2006). For instance, in East Asia and some European countries, competitive entrance examinations to prestigious universities, the credential society, and credential inflation have been proposed as primary causes of the increase in the demand for PT. In developing countries, scarce resources within the public educational systems have been suggested to explain the demand for PT.

Despite the widespread phenomena of PT in South Korea, limited quantitative research has been completed investigating the variables related to demand for PT. Even less research has been published illuminating PT and the relationship of PT to formal education.

In South Korean society, PT has been perceived as an economic hardship for the last four decades. According to the Korean Education Development Institute (KEDI) (2003), total expense for PT has increased from 0.34% of the GDP in 1977 to 2.3% of GDP in 2003. This amount is half of the public expenditure on educational institutions, which makes up 4.6% of

[1] PT is classified in two ways according to its purpose. For long term return (effect), PT is used to improve skills in areas such as art, music, English, and gym for primary school students. For short-term return (effect), PT is used to prepare secondary school students for the entrance exam. In this study, the latter purpose of PT is discussed. Thus type of PT is described by its three characteristics: 1) it is separate from formal education and is an extracurricular activity; 2) private tutors are motivated by profit; and 3) students' expectations of the tutor are higher than that of a normal school teacher (Tansel and Fatma, 2006).

There are four major forms of PT that cater to the diverse market demand. First, individual tutoring (one-on-one lessons) is typically provided at the students' home. Second, group tutoring is usually offered at the students' or tutors' home. Third, instruction is provided by private for-profit learning institutions, called 'Hakwon,' where instructors teach in classroom-like settings. Fourth, Internet-based lectures are available through private learning companies. Individual tutoring is the most expensive form of PT, while internet-based lectures are the least expensive.

[2] A recent study of 3,000 primary and secondary school students in England found that 27% had a private tutor (Ireson and Rushforth, 2004). In a survey of nearly 90,000 university students in Eastern Europe and Mongolia, the majority of students (69%) reported having received some types of supplementary PT during their last year in secondary school (Silova and Bray, 2005).

GDP. Although some emphasize the merits of PT for its gains in human capital, others are concerned about the limited access to PT due to household budgetary constraints and the distortion of equity in society. Furthermore, if it consists mainly of cramming and memorization, it may have little long-term value.

This drain can be seen on three different levels, the individual, the household and the nation. On the individual level PT is considered an investment in education with high cost and low benefit. Additionally, it offers low contribution to productivity, and is harmful to creativity, imagination and the self-learning process (Kim and Kim, 2002). On the household level, spending on PT alters the structure of consumption and exerts a harmful influence on the household's economic welfare. A survey conducted by the Ministry of Education (2001) and Korean Education Development Institute (KEDI) (2003) reported that 84% of parents consider PT expenses an economic burden. On the national level, in a country with restrictions on usable resources, PT induces enormous opportunity costs that fail to improve productivity (Paik, 2000). In South Korea, the cost of education has increased at a greater rate than that of manufacturing, and South Korea's educational efficiency is among the lowest ranked of Organization for Economic Co-operation and Development (OECD) **c**ountries (Grundlach & WŐβmann, 2001; Kim, 2002). Consequently, expenditure on PT entails a great number of selection and unnecessary transaction costs which result from the sorting process in selecting the appropriate students for higher education (Park, 1996).

From an economic standpoint, PT is an economic decision made by parents[3] who consider benefits of higher education and the cost of PT, where higher education services are regarded as normal and necessary. Parents purchase PT for their child when the marginal utility of obtaining the PT service is greater than the marginal cost of PT services. However, parents do not have much information on the true utility and effectiveness of PT. Despite this, demand for PT is continually rising in South Korea (Kwak, 2004).

In general, parents choose PT to increase the probability of their children receiving a high score on the university entrance examination thereby, successfully gaining entrance into a

[3] In this study, the demanders of PT are considered to be parents who have purchasing power.

prestigious university which guarantees a greater return in the future. In this respect, the demand for PT may be explained by Demand Theory which considers the individual's demand for goods as a decision made under a restricted budgetary condition and considering the price of PT in order to maximize his or her utility. However, the demand for PT has characteristics that cannot be explained clearly by a traditional economic decision making theory. Therefore a different perspective is warranted. An individual's demand for PT contains characteristics of inter-dependent demand influenced not only by his own utility obtained by consuming PT services, but also determined by other demanders' purchases of PT.[4] In this respect, this study attempts to estimate the effects of parents' demand for PT on other parents' demand for PT, which may be expressed as a contextual effect and has been overlooked in prior studies.

In terms of empirical strategy, most prior studies have utilized the Tobit regression model for analysis. However, Tobit regressions provide inconsistent estimates when the error term is heteroscedastic and not normally distributed. Moreover, the Tobit regression model does not address the possibility that decision making on PT can be divided into two stages, that is parents first decide whether or not to participate in PT, and then they decide how much will be spent on PT. This study attempts to examine the determinants of expenditure on PT after adjusting selection effects on participation in PT. To do this, a two-stage selection correction model (Heckman, 1979) was utilized. The purpose of this study is two-fold. First, it investigates the determinants of participation in PT, the expenditure on PT, and hours of PT. Second, Heckman's two-stage sample selection is employed where the South Korean governmental policy to decrease the demand for PT is used as an identifier in the selection equation. Based on the results of the study, implications of PT expenditure and suggestions for South Korean governmental policies are discussed.

[4] Survey results reveal the main reasons for demand on PT with secondary school students' parents are 1) many other students participate in PT (average rating of 4.12 out of a 5-point scale) and 2) they have to win the competition against others (average rating of 4.25 out of a 5-point scale) (Lee et al., 2003). Over half (52.1%) of secondary school students indicate that the main reason for PT is that they feel uneasy because they cannot catch up to their friends (Kim, 2000). In addition, demands for PT are influenced by the percentage of classmates who are expected to enroll in PT (Paik and Kim, 2003). These results illustrate that demands on PT are determined not only by an individual's budgetary restrictions and preference for PT services, but also by others' demands on PT services.

Ⅱ. Prior Studies on PT and South Korean Governmental Policy for PT

1. Literature Review

Existing empirical studies on the determinants for PT usage differ in their conclusions because of various factors, including the type of PT usage, the choice of dependent variables, the operational definition of PT, and the statistical techniques applied. This section classifies and reviews prior studies by their dependent variables: 1) determinants of PT participation and 2) determinants of PT expenditure.

Participation in PT is affected by numerous factors. Stevenson and Baker (1992) found that parents' education, income, and students' academic standings were positively related to participation in all types of PT in Japan. Aside from those taking correspondence courses, students living in urban areas showed high probability of PT participation. Additionally, the higher the reputation of the students' high school, the more likely students participated in correspondence courses. In addition, boys were more likely to participate in PT than girls.

Moreover, Assaad and El-Badawy (2004) found that at the individual level, significant determinants in participating in private and group tutoring in Egypt were students' age, whether the student is the eldest child, and whether the student is in a diploma year. At the household level, parents' educational level, parental absence, total household expenditure, and residential location were found to be significant in determining if one participates in PT. In contrast, at the community level, pupil-teacher ratio was negatively associated with likelihood to participate in PT. That is, the higher the student-teacher ratio, the more likely a student would be involved in PT.

In a survey of parents of elementary and secondary school-aged children in Canada, Davies (2004) investigated current and future PT participation of parents. Results showed that older and more educated parents are more likely to participate in PT. Thus, older children are more likely to receive PT. Parents who wanted their children to attend a private school were almost four times more likely to hire private tutors compared to parents who did not.

In recent decades, empirical studies on PT in South Korea have focused on PT and PT expenditure, answering such questions as why do parents choose PT or what factors influence

parents' decision making process. It has been found that significant student level variables in PT participation include student gender, parents' educational level, educational capital of the family, whether the household is a single parent family, student's attitude toward learning, academic pressure from the mother, and student's academic performance (Yang, 2003).

Determinants of PT expenditure have been analyzed by several researchers. For instance, Tansel and Fatma (2006) analyzed the determinants of PT expenditure in Turkey. PT had unitary income elasticity. Age of the head of the household and mother's education were positively associated with expenditure on PT. In addition, single mothers spent significantly more on PT and urban households showed more expenditure on PT than rural households.

A study by Kanellopoulos and Psacharopoulos (1997) found that PT is a luxury item in Greece. Factors that negatively affected the probability of private spending on education included household size and number of children under six years of age. Positively significant factors included head of household's years of education and income. However, Psacharopoulos and Papakonstantinou (2005) reported that PT is a necessity for the household with 12^{th} grade students in Greece.

Yang (2004) found at the household level, income, number of children, social status of the head of the household, and amount of social capital (such as having a relative in a high class) positively influenced PT expenditure. In terms of regional variables, the areas of Seoul, Kyunggi, and other metropolitan cities had greater expenditure on PT than other areas. Moreover, Lee and Hong (2001) and Lee and Kim (2002) analyzed similar data using the South Korean Household Expenditure Survey, focusing on the effect of the High School Equalization Policy[5] on expenditure on PT while controlling demographic and regional variables. However, neither of them showed a significant influence of the Equalization policy on PT expenditure.

In terms of yearly average expenditure on PT per child, as yearly income and mother's years of education increased, the yearly average expenditure on PT per child also increased (Park, 1998).

[5] In 1969, elementary education was compulsory in Korea, which resulted in increased demand for secondary education. Furthermore, economic development in the 1960's and 1970's increased household incomes, making secondary education affordable. As a result, entrance into highly ranked secondary schools became severely competitive resulting in 'entrance examination purgatory,' Consequently, the High School Equalization Policy was enacted in 1974 to replace individually administered entrance exams targeted to metropolitan cities to, instead, randomly allocate students within their school district.

Notably, spending on public education per child had a positive influence on yearly tutoring expense. This indirectly undermined the policy to raise government expenditures on public education to improve the quality of public education and thus, to reduce the demand for PT.

2. South Korean Governmental Policy to Decrease PT

Education has been suggested as the fastest impartial conduit to upward social mobility as individuals with higher credentials often receive higher return in the labor market and occupy more prominent societal positions**.** Parents invest in their children's education with the intention of enhancing their future. This may be the only strategy for parents to gain upward mobility for their children in the absence of good access to the capital market (Tsang & Levin, 1985). Typically in many societies, graduation from a prestigious university ensures future attainment of a prestigious position in the labor market. This suggests the main motivation for expenditure on PT. The competition for acceptance into prestigious universities is strong due to admission quotas. Students' scores on the College Scholastic Ability Test have been a predominant factor for admission to prestigious universities. The belief that utilizing PT services will foster greater success on entrance examinations has had a great impact on the demand for PT.

It has been pointed out that PT is a social ill (KEDI, 2003). Thus, the South Korean government has attempted to intervene through legal and policy regulations. The South Korean government has implemented various policies in the past four decades. Related government policies fall into three categories. First, the government defined specific types of PT as illegal based on related laws and punished illegal PT in order to add economic entities' financial burden to the payoff. Second, the government developed alternatives to PT in order to decrease expected cost by absorbing the demand for PT into the public sector. An example of this was the broadcast of lectures for the national university exam by the Education Broadcasting Station (EBS), which began in April, 2004 and separated course levels based on students' achievement. Third, the government attempted to decrease expected profit from PT by weakening the impact of PT on the university entrance exams. One example was to diversify

the standards of selecting proper students with the intention to nullify the importance of cramming through PT. The most direct attempt to hinder PT demand was to mandate restrictions on PT as illegal. These regulations impacted office hours of institution for PT, fees for PT, and the number of students allowed per PT class. Since 2001, the policy regulating institution for PT after 10:00 P.M. has been a source of debate. Typically, South Korean high school students participate in complementary classes after the regular school day. Considering regular school hours and complementary classes, the initial intention of this regulation was to reduce students' usable time allocated for PT services. However, the effects of this regulation have not been evaluated. The regulation of PT by the South Korean government has been perceived by the public and media as a failure. This study attempts to evaluate the influence of the local governmental regulation of PT by employing it as an identifier to distinguish participation in PT from expenditure and hours of PT.

Ⅲ. Data and Variables

1. Data

Questionnaires were distributed to parents of third-year high school students in the spring of 2005. A proportional stratified sampling within five metropolitan areas (Seoul, Busan, Incheon, Taeku, Taejeon) and five provinces (Kyunggi[6], Kyungbuk, Chungbuk, Kangwon, Jeonnam) of South Korea was selected. Within each geographic region, a convenience sample of schools was identified. Within each of the 40 general high schools and five Special Purpose High Schools, three classes in each school were randomly selected. Within each of these classes, 30 students were randomly selected, for a total sample of 90 participants per school. A total of 4,045 questionnaires were distributed to parents with a return rate of 90.2%. Questionnaires with incomplete or missing data were not included in the analysis, resulting in a total sample of 3,605 questionnaires.

[6] In Kyunggi province, among the 25 district offices of education, seven district offices are under the Equalization Policy (EP) and 18 are under the non-Equalization Policy (non-EP). In this study, five districts under EP and five districts under non-EP were selected.

2. Measures

A validated questionnaire used in past research (Paik, 1999) was modified and utilized in this study. The questionnaire is composed of four parts: 1) demographic characteristics (student gender, parents' years of schooling, income, number of children), 2) student achievement (high school record), 3) parents' perceptions on PT, and 4) general questions about PT usage within the last year. In addition, school information gathered through communications with administrative staff and websites at each school.

3. Variables

3.1. Dependent Variables

Using the definition of demand as "the maximum quantity of goods that one wants to buy with one's purchasing power" (Mankiw, 2000), the current study uses two variables as demand for PT. First, 'expenditure on PT' is the dependent variable. Most prior empirical studies (i.e., Lee and Kim, 2002; Tansel and Fatma, 2006) use this variable as a proxy for demand for PT. Policymakers and those who consume PT are most concerned with issues related to PT expenditure**.** Thus, this was used as a dependent variable for practical applications and implications**.**

Second, monthly hours of PT are selected as demand for PT to examine the need for PT from the aspect of quantity of PT actually received. The concept of monthly expenditure on PT includes both price and quantity and is expressed as the following:

'Monthly expenditure on PT = Hourly price of PT × Quantity of PT within a month'

In these terms, monthly hours of PT is used as an appropriate proxy variable for quantity of PT. Monthly expenditure on PT and monthly hours of PT are transformed to the logarithmic scale to reduce heteroscedasticity from the large variation in expenditure and hours of PT.

3.2. Independent Variables

The variables in this study are classified into seven groups: 1) student characteristics, 2) family background, 3) school characteristics, 4) governmental policy, 5) contextual effect, 6)

price of PT, and 7) residential area. Student characteristics include 1) gender, which is a binary variable and 2) academic achievement level, which is dummy coded into four variables that represent the upper two quintiles of student rank in a class and the lower two quintiles of student rank in a class. The reference group is the middle quintile of student rank (40-60%).

Family background variables include 1) household income (transformed by natural logarithm to reduce heteroscedasticity and prevent violating the normality assumption of independent variables) 2) parents' education level (calculated by the average years of the father's and mother's schooling), and 3) number of children.

High school characteristics include 1) type of school (private or public); 2) size, with 3 year high school enrollment of less than 300 (the reference group), between 300-500, and more than 500; 3) location (whether the school is under the High School Equalization Policy (EP)); 4) school characteristics, that is, whether the school is a Special Purpose High School (SPHS)[7]; and 5) student-teacher ratio (school mean).

The policy variable includes local government policies regarding the regulation of institutions for PT. Some local governments have prohibited PT after 10:00 PM but others have no policies regulating hours for PT. This policy was included as a dummy variable, indicating whether the local governments have the regulation or not.

This variable is defined as a contextual effect[8] selected to show evidence that parents' demand for PT relies on the distribution of other parents' consumptions**.** In addition, price of PT was included to demonstrate the price elasticity of PT. However, this variable was included only when the dependent variable is monthly hours of PT because price of PT is derived from monthly expenditure on PT. The variable for location of residence includes two dummies indicating whether the residential area is in Seoul or other metropolitan areas. The reference

[7] Recognizing differences in students' cognitive abilities, largely ignored in the High School Equalization Policy, the Ministry of Education established Special Purpose High Schools for the gifted. In 1987, each province had these high schools, such as science schools and foreign language high schools, with superior facilities and faculty. These schools have a higher level of autonomy, set their own tuition levels, design their instruction formats, and select students from across the country using independently designed tests.

[8] Manski (2000) describes the framework for the systemic analysis of social interaction (1) endogenous effect (2) contextual effect (3) correlation effect. Contextual effect is the propensity of an economic agent to behave in some way varied with exogenous characteristics of the group members.

group is suburban areas.

The sample used in this study consists of 3,605 observations with non-missing values on all variables. Tables 1 presents descriptive statistics of the explanatory and outcome variables by students' participation in PT service. It is clear that there are systematic differences in student characteristics by whether they are participating in PT or not. Those receiving PT are likely to be male, high achievers, and individuals from high socio-economic brackets. There are also differences in school characteristics of size and location. For instance, students receiving PT tend to attend Special Purpose High Schools or schools under the EP policy. In addition, they are likely to attend schools with a higher proportion of students receiving PT.

Ⅳ. Empirical Strategy

Few empirical studies exist that examine the determinants of PT participation and expenditure. The empirical strategy of previous studies uses the Tobit regression model to correct the censored distribution of expenditure on PT. However, the maximum likelihood estimator of the Tobit model is susceptible to misspecification errors (Deaton, 1997). Therefore, researchers have developed several misspecification tests for omitted variables, heteroscedasticity, and non-normality. In this study, conditional moments tests proposed by Pagan and Vella (1989) are used to demonstrate the misspecification of the Tobit model.

Additionally, in the Tobit model, the same set of variables and coefficients determines both the probability that an observation will be censored and the value of the dependent variable. However, Heckman's sample selection models allow for greater theoretical development because the observations are censored by some other variables. Thus, this study makes use of Heckman's two-stage sample selection model.

1. Tobit Regression Model

The standard Tobit model is generally written as:

$$y_i^* = x_i'\beta + u_{i,} \qquad i=1,2,\ldots,n$$

$$y_i = 0 \qquad \text{if } y_i^* \le 0 \qquad (1)$$
$$y_i = y_i^* \qquad \text{if } y_i^* > 0$$

where y_i^* is a latent variable; y_i; x_i is a k-dimensional vector of known regressors; β is a k-dimensional vector of unknown parameters; and the disturbances, u_i, are assumed to be $N(0,\sigma^2)$.

The likelihood function of this model is:

$$L(\theta) = \Pi_o \left[1 - \Phi(\frac{x_i'\beta}{\sigma}) \right] \Pi_i \sigma^{-1} \phi \left[\frac{(y_i - x_i'\beta)}{\sigma} \right] \qquad (2)$$

where Π_0 denotes the product over values of i such that $y_i^* \le 0$, Π_i denotes the product over values of i such that $y_i^* > 0$; $\theta = (\beta', \sigma^2)$; and $\Phi(\bullet)$ and $\phi(\bullet)$ are, respectively, the cumulative distribution and probability density function of the standard normal variate.

2. Specification Test for the Tobit Model

This study takes advantage of conditional moments tests developed by Pagan and Vella (1989). Conditional moments tests include three steps. First, they identify a set of population conditional moment restrictions, $m(w_i, \theta)$, where w_i includes endogenous as well as exogenous variables. These restrictions should hold if the model is properly specified. Second, they calculate the sample analogue of the restrictions, $m(w_i, \hat{\theta})$, and the scores of the log-likelihood function. Third, they verify a model inadequacy by regressing each $m(w_i, \hat{\theta})$ against unity and the scores of the log-likelihood function.

These tests consider three misspecifications. The first misspecification is when the variable z has not been incorrectly omitted from the model. The second is when the disturbances in the model are homoscedastic and the third misspecification is when the underlying disturbances in the model are normally distributed. By analyzing the structure of the model, inferred potential errors in the moments of y_i^* were deduced and the following conditional moment restrictions for the observed dependent variables are suggested:

$$n^{-1}\Sigma E[z_i(E(y_i^* \mid y_i) - x'\beta)] = 0 \quad (3)$$

$$n^{-1}\Sigma E[z_i(E(u_i^2 \mid y_i) - \sigma^2)] = 0 \quad (4)$$

$$n^{-1}\Sigma E[z_i(E(u_i^3 \mid y_i)] = 0, \quad n^{-1}\Sigma E[z_i(E(u_i^4 \mid y_i) - 3\sigma^4] = 0 \quad (5)$$

where the summations range from one to n and z_i is a q-dimensional vector that has not been erroneously omitted from the model. In equations (3) and (4), the first expectation operator refers to an expectation calculated using the joint density of z_i and y_i whereas the second expectation operator relates to the density of y_i^*. In equation (5), the double expectation operator refers to the unconditional third and fourth moment of u_i. For observations where $y_i = 0$, the expected value of y_i^* is $x_i'\beta - \sigma\lambda_i$, where the inverse-Mills ratio is $\lambda_i = \phi(x_i'\beta/\sigma)/\left[1 - \Phi(x_i'\beta/\sigma)\right]$. For positive observations, $y^* = y$. Equation (3) is the diagnostic test for omitted variables and equation (4) is applied as a test for heteroscedasticity. Finally, equation (5) is employed to test for normality of the error terms in the censored model (Pagan and Vella, 1989; Green, 2003).

3. Heckman Two-Stage Sample Selection Model

The underlying theory of this model is that the correction terms derived from the PT choice model can control unobservable characteristics of the selected sample in the outcome equation (Heckman, 1979). In this study, they are expenditure on PT and hours of PT. As shown in equation (6), a dichotomous variable z_i, denoting student i's probability of participating in PT determines whether or not y_i is observed. y_i, monthly expenditure on PT and monthly hours of PT are observed only if z=1, and we estimated a model with a vector of independent variables w_i that affect the level of parents' subjective probability and obtain coefficients γ, with an error term, u_i which is assumed to have stochastic disturbance normally distributed with a mean 0 and a constant variance σ_u^2.

x_i is a vector of independent variables that affect the expenditure on PT or hours of PT. ε_i

is a stochastic error term.

$$\text{Selection mechanism:} \quad z_i^* = w_i'\gamma + u_i, \ \ z_i = 1 \text{ if } z^* > 0 \text{ and 0 otherwise} \qquad (6)$$

$$\text{Regression model:} \quad y_i = x_i'\beta + \varepsilon_i \text{ observed only if } z_i = 1, \qquad (7)$$

$$E(y_i \mid z_i = 1, x_i) = x_i'\beta + \rho\sigma_u\sigma_\varepsilon\lambda(w_i'\gamma) \qquad (8)$$

The first step of Heckman's procedure entails the estimation of the participation equation by probit, which gives us an estimate of $\lambda = \frac{\phi(w_i'\gamma)}{\Phi(w_i'\gamma)}$. The second step consists of a least squares regression (for participants in PT) of y_i on x' and $\lambda(w_i'\gamma)$.

It is important that a number of variables are common to the vector w in equation (6) and x in equation (7). If the vector w is exactly the same or a subset of the vector x, then the combination of the variables in the vector w may lead to non-identification of the regression in equation (8) (Green, 2003). A solution is to add at least one variable in the vector z, but not in the vector x. This study takes advantage of a variation in the local government regulation on PT. They are expected to decrease the probability of choosing PT but they are not expected to decrease the expenditure on PT. Moreover, no prior empirical studies exist that show local government regulations on PT affect the parents' expenditure on PT.

Ⅴ. Result

1. Participation Equation Estimation

Following existing literature, this study models parents' decision on PT participation as a function of student characteristics (gender and academic achievement level), family background (income, parents' education, number of children), school characteristics (private, size, student-teacher ratio, schools under EP, SPHS), residential area (Seoul, metropolitan cities, suburban) and local government policy regarding regulation of PT (See Table 2).

Table 2 highlights gender differences in PT participation where male students are less likely to participate in PT by 8%. This result contradicts findings of Stevenson and Baker (1992) and Assaad and El-Badawy (2004). This may reflect cultural differences and changes in time because their data were collected in 1980 and 1998, respectively.

Students' achievement is a significant determinant of participation in PT. When compared to the reference group, students in the 2nd quintile from the top have a higher probability of participating in PT whereas those in the 5th quintile have a lower probability of participating in PT. Parents with higher education and higher income show higher probability to participate in PT. For example, an additional year in parents' education increases the probability to participate in PT by about 3%. These findings are similar to most prior studies that show positive association between participation in PT and parents' SES. The number of children in the family is a significant determinant of participation in PT, where an additional child decreases the probability to participate in PT by 3%. Table 2 suggests that school-level variables also play an important role in students' participation in PT. In particular, student-teacher ratio and schools under the EP and SPHS are shown to be the most significant predictors of PT participation.

Students in schools with higher student-teacher ratios are more likely to participate in PT. This implies a trade-off between teacher-student quality and quantity, where the fewer students the teacher has to allocate his/her labor, efforts, and motivation, the more resources the student can obtain from the teacher. This may result in less need to seek additional resources for learning.

Students in areas under the EP are 20% more likely to participate in PT than those under the non-EP. Those who attend Special Purpose High Schools (SPHS) are 10% more likely to participate in PT than those attending general high schools. Considering the distinct characteristics of these schools and students, students in schools under non-EP are selected by entrance examinations and there are formal high school rankings in areas under the non-EP. Therefore, students who attend schools with good reputations are generally highly motivated to achieve. Teachers in schools under non-EP are more likely to have higher standards and internal incentives themselves, thus meeting parents' and students' expectations. This may lead to higher

school satisfaction and lower motivation to participate in PT. In contrast, unobserved characteristics of SPHS increase the propensity of students' participation in PT. Special Purpose High Sschools, like schools under non-EP, select students by entrance examination. However, tuition, curriculum, teachers, and geographic region are different from schools under the non-EP. The higher probability to participate in PT in SPHS may be attributed to the geographical location of SPHS, that is, metropolitan cities which allow for convenient access to the PT market and enhances internet based PT.

Less likely to participate in PT are students in large-sized schools and private schools, although under the standardized regimen of formal schooling in South Korea, private and public schools are indistinguishable in terms of student fees, overall school finance, and curriculum. In areas under the EP, students are even randomly assigned between public and private schools. A possible reason as to why private school students participate less in PT may be explained by the key difference between private and public schools in South Korea. That is, public school teachers and principals serve a limited term in one school before rotating to another. This rotation system hinders a sense of belonging and makes it difficult for principals to establish authority. Results reveal that students at large-sized schools are less likely to participate in PT than those at small-sized schools. On the one hand, this suggests the possibility that students in large schools are more polarized, with a group of active participants at one end of the continuum and a large group of students who did not participate in PT at the other. On the other hand, it is possible that parents with children at small-sized schools are able to interact and network more with other parents and share their experiences about PT, generating more interest, thereby promoting the probability to participate in PT. Households in Seoul and other metropolitan areas are more likely to purchase PT than those in suburban areas. Parents residing in Seoul are 22% more likely to participate in PT and those in metropolitan areas are 9% more likely to choose PT than those living in suburban areas.

It appears that local government regulations of PT, as designed, have a strong impact on parents' choice of PT. Local government regulations decrease the probability of students' participation in PT by 19%. This significant relationship between PT regulation and parents'

consumption of PT serves as an identifier of the selection correction term in the second stage outcome equation.

2. Demand Equation Estimation

Several studies have examined the determinants or predictors of household expenditure on PT (i.e., Lee and Kim, 2002; Davis, 2004; Tansel and Fatma, 2006). Based on the literature, this study posits that parents' expenditure on PT is dependent upon student, family, school, and community characteristics, as well as contextual effects. There has been no prior studies that employed the hours of PT as a dependent variable. Thus, as explanatory variables for hours of PT, this study utilizes the same covariate with those used for expenditure on PT by adding hourly price of PT.

Table 3 shows results from the OLS, Tobit, and Heckman two-step sample selection methods. The dependent variable is the log monthly expenditure on PT. Results of these different specifications show similar patterns in significance, although there is quite a large difference in the marginal effects. While the coefficients from the Heckman model are somewhat larger than those produced by OLS, those of Tobit are much larger than the other two estimates.

To check the efficiency and consistency of the estimates from the Tobit model, specification tests were performed and results highlight problems with the omitted variables, heteroscedasticity, and non-normality of the disturbances. Details will be discussed in the next section.

Based on the results of the specification tests, this study considers the estimates from Heckman's sample selection model are more robust and consistent than those from the Tobit model. Results of the sample selection model are discussed below.

Significant λ shows that unobserved influences that contribute to participation in PT cause an increase in expenditure. As shown in Table 3, it is clear that students' academic achievement level is associated with parents' expenditure on PT. Compared to the reference group (those achieving in the 3rd quintile), students in the first and second quintiles spend more on PT while students in the 5th quintile spend less on PT. This result is consistent with results of Stevenson and Baker (1992) and Lee and Kim (2002). There is a significant relationship

between family background and parents' expenditure on PT. The coefficient on household income can be interpreted as the income elasticity of PT. Elasticity is less than 1, indicating that PT is a necessary good for 3rd year high school students. As household income increases 10%, PT expenditure increases 6.3%. Lee and Kim (2002) and Psacahropoulos and Papakonstantinou (2005) had a similar finding, however Kenellopoulos and Psacharopoulos (1997) and Tansel and Fatma (2006) reported different results. The difference of income elasticity may be impacted by students' age and the urgency of PT. The studies which included students from elementary school to high school showed higher income elasticity than those with samples of only high school students. This implies that parents view PT with more discretion at lower education levels.

The coefficient on average years of parents' schooling suggests that one-year increase in parents' education increases PT expenditure by 8%. Noticeably, all school characteristic variables are not significant, which is quite different from results of the Probit estimates.

For the past 40 years, it has been suggested that the greatest factor contributing to the increased demand for PT is the low quality of public schools. Thus, South Korean governmental policies have focused on investing schools with the intention of diminishing spending on PT. However, according to this study's results, covariates related to schools fail to explain the variance of parents' expenditure on PT.

Rather, the proportion of classmates receiving PT within a school is significantly associated with the demand for PT. This variable can be interpreted as a contextual effect, meaning that parents' consumption of PT depends on the distribution of other parents' participation in PT. In terms of marginal effect, as the proportion of classmates receiving PT increases 10%, the expenditure on PT increases 8%.

Geographic location is a significant predictor of PT expenditure and this association continued after controlling for all other characteristics. Households in Seoul are more likely to spend on PT than those in the suburbs by 66%. In addition, parents in metropolitan areas spend 19% more on PT compared to those in suburban areas. Lee and Kim (2002), Yang (2004) and Tansel and Fatma (2006) found similar results.

Table 4 presents results of the OLS, Tobit, and Heckman two-step models that estimate the determinants which affect the log monthly hours of PT. Different patterns in significance and marginal effects are revealed in each of the regression models. Results from the Heckman sample selection model are discussed below due to results of the Tobit specification test, which show serious heteroscedasticity and non-normality of disturbance (see Table 6).

In the Heckman model, $\hat{\rho}$ and $\hat{\lambda}$ coefficients are significantly different from zero indicating that there is a sample selection. If non-participants of PT are removed from the sample, the resulting estimates would be biased. In particular, since the correlation is positive, the estimates would be biased upwards.

Of the student characteristic variables, students in the 1st and 2nd quintiles of school performances show significantly more likelihood in spending more time on PT than the reference group (those in the 3rd quintile range). In terms of family characteristics, while the number of children is negatively associated with the time spent on PT, household income and the parents' average years of schooling are positively associated with spending more time on PT.

Among school characteristics, teacher-student ratio is positively associated with students' hours spent on PT. The coefficient denotes that when the teacher-student ratio increases by one, students spend 3% more hours per month on PT. Teacher-student ratio is one of indicator of quality of schooling. Therefore, if quality of schooling declines, students' learning from their teacher becomes relatively smaller because the teacher has to divide his or her time and attention towards more students. Students attending SPHS are expected to spend less time on PT by 13% than students attending general high schools. From Tables 2 and 4, it is inferred that students in SPHS have a higher probability of participation PT but spend less time on PT. In contrast to its effect on the expenditure on PT, the proportion of classmates receiving PT is not significant.

Log hourly price of PT measures the degree of relationship between changes in quantity of goods and changes in its price by deriving the price elasticity of demand. The marginal effect of price of PT indicates that there is negative own-price elasticity, denoting that a 10% increase in

the average price of PT reduces hours spent on PT about 5%. Thus, PT demand is not very price responsive. In addition, students in Seoul and those in metropolitan cities are likely to spend 54% and 16% more time, respectively, on PT than those from suburban areas.

3. Results of the Conditional Moments Tests

Results of the conditional moments tests are shown in Table 5 and Table 6. Three misspecification tests were conducted. First, for the conditional moment test for omitted variables, the dependent variables are the product of squared, cubed values and the generalized residuals. The independent variables consist of an intercept term and transformed score vectors. The test of significance on the intercept term indicates values of 44.37 and 45.46, respectively. Hence, the model has problems associated with the omitted variables. The second and third types of misspecification tests focus on potential problems intrinsic to the disturbance of the model. The test for heteroscedasticity involves a separate least squares regression on each of the independent variables in the model. As presented in Table 5, the intercept terms in all of the regression models are significant at a significance level of 0.01, which implies that the variables for the Tobit model have considerable heteroscedasticity.

To test for normality in the distribution of disturbance, two independent regressions were conducted using the moment restriction functions. Results of the significance test on the intercept term show values of 7.39 and 5.38. This demonstrates that the error distribution assumed for the Tobit model does not conform to the assumption of normality.

Table 6 presents results of the conditional moments tests for monthly hours of PT. Aside from the test result for omitted variables, the significances of the intercept terms exhibit serious problems of heteroscedasticity and non-normality of estimates of the Tobit model.

Ⅵ. Conclusion

This paper examines determinants of demand for PT by using the Heckman two-stage sample selection method. The analysis compares estimates from OLS and Tobit estimators and the Heckman two-stage sample selection correction method. The Tobit estimate shows a misspecification problem. Therefore, estimates from the Heckman method are considered more efficient values. In the outcome equation, λ was very significant, meaning that if non-participants of PT are removed from the sample, the resulting estimates would be biased. The analysis also makes use of local government regulations of PT to identify participation in PT. Many of these findings replicate results from prior studies with respect to student and family characteristics.

In terms of students' characteristics the results from the first and second stages show similar patterns in the significance of the regression coefficient. Based on these results, students who are ranked in the upper-middle part of class in terms of academic achievement should be the target group of government policies which attempt to decrease the demand for PT. To decrease the demand for PT of this group, supplementary classes after the regular school day, which are provided by the schools, should be organized according to students' achievement level. Moreover, the academic need for students in this upper-middle academically achieving group should be analyzed more closely by experienced teachers within the school so that the school can absorb these students' demand for PT more effectively. In addition, support for students with low academic achievement is required so that they do not feel disadvantaged and isolated.

Family background has the most significant influence on both the selection and the outcome equations. Higher income households show more demand for PT and parents' average years of schooling are positively related to PT demands. Such evidence suggests that demand for PT is not a remedial strategy used predominantly by students who have difficulty meeting the academic standard, but rather, is used by students who have already accrued advantages in the formal school system (Stevenson & Baker, 1992). Based on this result, it should be noted that PT utilization may be inefficient as well as inequitable. For poor parents who have high latent

demand for PT because of the high ability of their children, their demand for PT is not met because it is not affordable. This partial distribution of PT may lead to inefficient allocation of their household budget.

According to the results based on school characteristics, the determinants of the selection mechanism and those of the outcomes (expenditure and hours of PT) are different. In addition, the quality of schools and unobserved characteristics of schools under non-EP decreased participation in PT. This highlights the possibility that parents or students in schools under non-EP are satisfied with the formal school education more than those in schools under EP. However, significantly lower demand for PT in schools under non-EP cannot be completely attributed to parents' greater satisfaction with formal school education. It is possible that a less active PT market, due to economic disadvantages typical of non-EP cities, impacts participation in PT. Additionally, to clarify the association between the SPHS and demand for PT, more targeted sampling and further study to estimate the effect of SPHS should be conducted.

Student-teacher ratio was an important factor to determining the participation in PT and monthly hours of PT. This result not only justifies the governmental investment in public schools to diminish the class size but also implies that a decrease of class size may contribute to the higher quality of teaching-learning condition, contributing to decreasing the demand for PT.

One important factor that determines parents' participation in PT is the local government policies on the regulation of PT. Parents in areas under the local governmental regulation of PT show less likelihood to participate in PT than those in the areas where the government has no input in their operating hours. This suggests that even though the mass media in South Korea often criticizes the government policy to decrease the demand for PT, these policies of the local government have a significant influence on parents' decision making in the consumption of PT.

One powerful predictor in determining expenditure is the contextual effect, which was evaluated using the proportion of classmates receiving PT. This signifies that parents' demand for PT is consumptive and is a defensive behavior influenced by other parents' purchase of PT, which confirms the public belief in South Koran society that a strong sense of rivalry increases demand for PT service (Yang, 2004).

Those living in Seoul or other metropolitan areas of high density residential developments are found to have more demand for PT after controlling for other factors. This may be interpreted in two ways. First, environmental factors such as parents' relationship with neighbors and cultural aspects of big cities may be associated with parents' greater anxiety for the comparative ranking of their child in the competition to enter prestigious universities. Second, large-scale PT markets create greater exposure to PT, accelerating parents' purchase of PT services. These environmental factors may increase the demand for PT in larger cities.

Finally, the fundamental reason of high demand for PT is attributed to social structure in which the benefits of graduation from a prestigious university assure a preferred position in the job market. Therefore, the government should not encourage employers to look solely at whether an applicant graduated from a prestigious university or not in evaluating her ability and productivity. One approach is to persuade private firms and public institutions to recruit those from less prestigious universities and give employers incentives such as tax reductions.

Another possible method is to create more alternatives in college choice which would give the consumer the same utility. Furthermore, the Korean higher education system should be more diversified and learn from the academic models of the U.S. For example, American Ivy League universities are not the only prestigious universities. Besides them, there are small-scale liberal art colleges which provide a similar highly qualified education. Through increasing the number of choices for consumers, the competition for top-class universities might be appeased.

References

Assaad, Ragui, and Asmaa El-Badawy. 2004. Private and Group Tutoring in Egypt: Where Is the Gender Inequality? Paper presented at the Workshop on Gender, Work, and Family in the Middle East and North Africa, University of Minnesota.

Bray, Mark. 1999. The Shadow Education System: Private Tutoring and its Implication for Planners. Fundamentals of Educational Planning No 61. Paris: UNESCO International Institute for Educational Planning (IIEP).

Bray, Mark. 2003. Adverse Effects of Private Supplementary Tutoring: Dimensions, Implications, and Government Responses. Series: Ethics and Corruption in Education. Paris: UNESCO International Institute for Educational Planning (IIEP).

Bray, Mark. 2005. Private Supplementary Tutoring; Comparative Perspectives on Patterns and Implication. Oxford International Conference on Education and Development. 'Learning and Livelihood', 13-15th September, 2005.

Bray, Mark, and Percy Kwok. 2003. "Demand for supplementary tutoring: Conceptual consideration, and socio- economic patterns in Hong Kong." *Economics of Education Review* 22 (6): 611-620.

Davies, Scott. 2004. "School Choice by Default? Understanding the Demand for Private Tutoring in Canada." *American Journal of Education* 110 (3): 233-255.

Deaton, S. Angus. 1997. *The Analysis of Household Survey: A microeconometric Approach to Development Policy*. Baltimore, MD: The Johns Hopkins University Press.

Dore, Ronald. 1997. "Reflections on the diploma disease twenty years later." *Assessment in Education* 4 (1):189-206.

Greene, H. William. 2003. *Econometrics Analysis*, ed. NY: Prentice-Hall.

Grundlach, Erich, and Ludger WÖβmann. 2001. "Fading productivity of schooling in East Asia." *Journal of Asian Economics* 12 (3): 401-417.

Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1): 153-161.

Ireson, Judith, and Rushforth, Katie. 2004. Mapping the nature and extent of private tutoring at transition points in education. Paper presented at the British Educational Research Association conference, UMIST, Manchester, 16-18th September, 2004.

Kanellopoulos, Costas, and George Psacharopoulos. 1997. "Private Education Expenditure in a Free Education Country: The Case of Greece." *International Journal of Educational Development* 17 (1): 73-81.

Kim, Kee Seung. 2002. "Educational investment and efficiency in South Korea from the standpoint of international comparison." *LG economics*: 22-25.

Kim, Yang Boon. 2000. *Research on the actual conditions of private education in Korea.* CR2000-3. Seoul: Korean Educational Development Institute.

Kim, Yang Boon, and Mee Sook Kim. 2002. *Research on actual conditions of cramming institutions for entrance exams.* Seoul: Korean Educational Development Institute.

Korean Education Development Institute. 2001. *Actual study for household' expenditure on private tutoring and the implication for policy*. Seoul: Author.

Korean Education Development Institute. 2003. *Monitoring the private tutoring and analyzing the cost of private tutoring*. CR 2003-19. Seoul: Author.

Kwak, Byung Sun. 2004. Struggle against private tutoring in Korean Education Context. Paper presented at the 28th annual conference of the Pacific Circle Consortium. Hong Kong Institute of Education, 21-23th April, 2004.

Kwok, Percy. 2003. "Examination-Oriented Knowledge and Value Transformation in East Asian Cram Schools." *Asian Pacific Education Review* 5 (1): 64-75.

Lee, Chong Jae. 2005. "Korean Education Fever and PT." *KEDI Journal of Educational Policy* 2 (1): 99-107.

Lee, Joo Ho, and Sung Chang Hong. 2001. "Schools vs. Private Tutoring, Choice and Equity in Korean Education." *Research on Economics* 49 (1): 37-56.

Lee, Joo Ho, and Sun Woong Kim. 2002. "Economic Analysis of the School Policy and private tutoring in South Korea." *Journal of Analysis of Korean Economy* 8 (2): 1-51.

Lee, Jong Jae, Sang Geun Choi, & Hyun Jin Kim. 2003. *The Counterplan for the Problem of Private Tutoring: Centoring on Enhancing the Competitiveness of Public Education*. RM 2003-32-5. Seoul: Korea Education Development Institute.

Lee, Jong Tae, Yang Boon Kim, and Cho-Hee Yoon. 2004. "The Effects of Pre-Class Tutoring in Student Achievement: Challenges and Implication for Public Education in Korea." *KEDI journal of Educational Policy* 1 (1): 25-42.

Mankiw, Greg. 2000. *Principles of Economics 2nd ed*. OH: Thompson.

Manski, Charles F. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives* 64 (3): 115-136.

Ministry of Education. 2001. *Survey on Private Tutoring*. Seoul: Author.

Pagan, Adrian, and Francis Vella. 1989. "Diagnostic Tests for Models Based on Individual Data: A Survey." *Journal of Applied Econometrics* 4 (4): 29-59.

Paik, Ill-Woo. 1999. "Understanding the Behavior of PT with the Educational Economics Approaches." *South Korean Journal of Educational Research* 37 (4): 291-306.

Paik, Ill-Woo. 2000. *Economics of Education.* Seoul: Hakjisa.

Paik, Ill-Woo, and Ji-Ha Kim. 2003. "Analysis of Regression Models for Estimating Demand for PT." *The Journal of Economics and Finance of Education* 12 (1): 81-100.

Park, Jeong Soo. 1996. "The expansion of private tutoring and the equity of educational opportunity in South Korea." *The Journal of Economics and Finance of Education* 5 (2): 515-538.

Park, Gi Baik. 1998. "Tax, Public and Private Cost of Education." *Public Finance* 5 (1): 1-21.

Psacharopolous, George, and George Papakonstantinou. 2005. "The real university cost in a "free" higher education country." *Economics of Education Review* 24 (1): 103-108.

Silova, Iveta, and Mark Bray. 2005. *Shadow Education: Private tutoring and its implication in the nine post-socialist countries*. Summary Report. The Education Support Program of the Open Society Institute Implemented by the Network of Education Policy Centers.

Stevenson, David Lee, and David P. Baker. 1992. “Shadow education and allocation of formal schooling transition to university in Japan.” *American Journal of Sociology 97 (6)*: 1639-1657.

Tansel, Aysit, and Bircan Fatma. 2006. “Demand for Education in Turkey: A Tobit analysis of private tutoring expenditure.” *Economics of Education Review* 25 (3): 303-313.

Tsang, Mun C., and Henry M. Levin. 1985. “The Economics of Overeducation.” *Economics of Education Revie*w 4 (2): 93-104.

Yang, Jeong Ho. 2003. “Middle School Students’ Determinants of Participation in PT.” *Journal of Korean Education* 30 (2): 261-282.

Yang, Jeong Ho. 2004. “High School Students’ Determinants of Participation in PT.” [Online] http://youthpanel.work.go.kr/docs/f_8/3/2-2.pd.

TABLE 1.
Descriptive Statistics of the Sample

Variables	All			Non PT student			PT student		
	Mean	Std.Dev	Freq	Mean	Std.Dev	Freq	Mean	Std.Dev	Freq
Student Characteristics									
Male(=1)	0.5345	0.4989	0=1,678 1=1,927	0.5157	0.4999	0=526 1=560	0.5427	0.4983	0=1,152 1=1,367
School Ranking 1st quintile(=1)	0.2649	0.4413	955	0.1823	0.3863	0=888 1=198	0.3005	0.4586	0=1,762 1=757
School Ranking 2nd quintile(=1)	0.1989	0.3992	717	0.1427	.34995	0=931 1=155	0.2231	0.4164	0=1,957 1=562
School Ranking 4rd quintile(=1)	0.0935	0.2911	337	0.0958	.29440	0=982 1=104	0.0925	0.2898	0=2,286 1=233
School ranking 5th quintile(=1)	0.1226	0.3280	442	0.1842	0.3878	0=886 1=200	0.0961	0.2947	0=2,277 1=242
Family Background									
Natural logarithm of monthly income	15.0349	0.5295	3,605	14.8157	0.4935	1,086	15.1294	0.5164	2,519
Average years of parents' schooling	13.6781	2.3026	3,605	12.7762	2.2136	1,086	14.0669	2.2305	2,519
Number of children in a family	2.1059	0.5701	3,605	2.1473	0.6443	1,086	2.0881	0.5342	2,519
School Characteristics									
Private school(=1)	0.4832	0.4998	0=1,863 1=1,742	0.5147	0.5	0=527 1=559	0.4696	0.4992	0=1,336 1=1,183
Student-teacher ratio	16.9475	2.4355	3605	16.9685	2.4988	3,605	16.9384	2.4081	3,605
School size: 300-500 in 3rd year (=1)	0.5864	0.4925	0=1,491 1=2,114	0.5994	0.4902	0= 435 1= 651	0.5808	0.4935	0=1,056 1=1,463
School size: 500 or more in 3rd year (=1)	0.3467	0.4759	0=2,355 1=1,250	0.3582	0.4797	0= 697 1= 389	0.3418	0.4744	0=1,658 1= 861
High Schools under Equalization Policy(=1)	0.8513	0.3558	0=536 1=3,069	0.7799	0.4144	0=239 1= 847	0.8821	0.3226	0= 297 1=2,222
Special Purpose High School(=1)	0.0965	0.2953	0=3,257 1=348	0.0506	0.2194	0=1,031 1= 55	0.1163	0.3207	0=2,226 1=293
Governmental Policy									
Regulation (=1)	0.5975	0.4905	0=1,451 1=2,154	0.5608	0.4965	0=477 1=609	0.6133	0.4871	0=974 1=1,545
Contextual effect									
Proportion of Classmates receiving PT service	69.8765	19.5273	3,605	57.2235	17.2093	1,086	75.3316	17.8702	2,519
Price of PT									
Log price of PT	9.4759	0.4141	3,605	-	-	-	9.4759	0.4141	2,519
Region									
Seoul	0.4172	0.4932	0=2,101 1=1,504	0.2781	0.4483	0=784 1=302	0.4772	0.4996	0=1,317 1=1,202
Metropolitan	0.3908	0.4880	0=2,196 1=1,409	0.3987	0.4899	0=653 1=433	0.3875	0.4873	0=1,543 1=976
Natural logarithm of monthly expenditure on PT	8.6385	5.7480	3605	0	0	1,086	12.3628	1.1087	2,519
Natural logarithm of monthly hours of PT	2.4052	1.6722	3605	0	0	1,086	3.4422	0.6569	2,519
Observations(N)		3,605			1,086			2,519	

TABLE 2.

Participation Equation: Probit Estimation

Variables	Coefficient	Marginal effect
Student Characteristics		
Gender	-0.2505***	-0.083***
	(0.0563)	(0.0185)
school ranking 1st quintile	0.07026	0.0232
	(0.0681)	(0.0223)
School ranking 2nd quintile	0.1629**	0.0527**
	(0.0703)	(0.0221)
School ranking 4th quintile	-0.0045	-0.0015
	(0.0858)	(0.0286)
School ranking 5th quintile	-0.2062***	-0.0717***
	(0.0755)	(0.0272)
Family Background		
Log monthly income	0.4241***	0.1414***
	(0.0539)	(0.0179)
Average years of parents' schooling	0.0817***	0.0272***
	(0.0124)	(0.004)
Number of children	-0.0871**	-0.0291**
	(0.0405)	(0.0135)
School characteristics		
Private(=1)	-0.11*	-0.0367*
	(0.0604)	(0.0202)
Student-teacher ratio	0.0595***	0.0198***
	(0.0149)	(0.005)
School size:300-500 3rd grade(=1)	-0.1854	-0.0612
	(0.157)	(0.0512)
School size:500 or more 3rd grade(=1)	-0.3197*	-0.1093*
	(0.1837)	(0.0641)
High school under Equalization Policy (=1)	0.6045***	0.221***
	(0.0843)	(0.0324)
Special Purpose High School(=1)	0.3236**	0.0987***
	(0.1287)	(0.0353)
Local government regulation of PT		
Regulation of PT(=1)	-0.5992***	-0.191***
	(0.0969)	(0.0291)
Region		
Seoul	0.6921***	0.2201***
	(0.0788)	(0.0236)
Metropolitan cities	0.2846***	0.0929***
	(0.072)	(0.023)
Constant	-7.9263***	
	(0.8116)	
Observations	3,605	
Pseudo R^2	0.1283	
Wald chi^2	566.22	
Log likelihood	-1922.85	

Note. Standard error in parentheses are adjusted for clustering on schools

***p<.01 , **p<.05, * p<.1

TABLE 3.

Ordinary Least Square, Tobit, and Heckman sample selection model estimation

Dependent Variable: Log monthly expenditure on PT

	Variables	OLS[1]	Tobit Marginal Effect	Heckman Two-stage[2]
Student Characteristics	Gender	0.0408 (0.0483)	0.0481 (0.2524)	0.0096 (0.0451)
	High school record 1st quintile	0.1312** (0.0652)	0.4262 (0.3472)	0.155** (0.0606)
	High school record 2nd quintile	0.1062 (0.0698)	0.91** (0.3566)	0.1591** (0.0643)
	High school record 4th quintile	-0.0201 (0.0884)	-0.2575 (0.4504)	-0.0186 (0.0787)
	High school record 5th quintile	-0.0912 (0.0758)	-1.7552*** (0.4181)	-0.1812** (0.0832)
Family Background	Log monthly income	0.5102*** (0.0582)	1.8593*** (0.2751)	0.6343*** (0.0662)
	Average years of parents' schooling	0.0549*** (0.0123)	0.2519*** (0.0649)	0.0801*** (0.0147)
	Number of children	-0.0407 (0.0438)	-0.3547* (0.2144)	-0.0706* (0.0401)
School characteristics	Private(=1)	0.0742 (0.0678)	-0.0797 (0.2955)	0.0603 (0.0506)
	Student-teacher ratio	-0.0132 (0.0147)	0.0367 (0.0703)	0.0008 (0.013)
	School size:300-500 3rd year(=1)	-0.0004 (0.169)	0.3639 (0.7121)	-0.0263 (0.1173)
	School size:500 or more 3rd year(=1)	-0.0464 (0.1834)	0.2359 (0.7968)	-0.0628 (0.1298)
	High schools under Equalization Policy (=1)	-0.0706 (0.072)	0.3589 (0.4206)	0.0671 (0.0907)
	Special Purpose High School(=1)	-0.1606 (0.119)	-0.4679 (0.5523)	-0.1085 (0.0913)
Contextual effect	Proportion of classmates receiving PT service	0.0077*** (0.002)	0.1502*** (0.0082)	0.0087*** (0.0014)
Region	Seoul	0.4862*** (0.1307)	0.7407* (0.4325)	0.6644*** (0.1023)
	Metropolitan cities	0.0239 (0.1006)	0.672* (0.3918)	0.1902** (0.0941)
Constant		3.3487*** (0.8166)	-35.5392*** (4.0779)	0.3082*** (1.3275)
Lambda(λ)		-	-	0.7268*** (0.2672)
Rho (ρ)		-	-	0.6758
Observations		2,519	3,605	3,605
R^2 / Pseudo R^2		0.2453	0.0511	-
LR chi^2		-	1011.17***	759.67***

Note. 1. Standard error in parentheses are adjusted for clustering on schools.
2. Standard errors in parentheses are adjusted for clustering on schools.
***p<.01 , **p<.05, * p<.1

TABLE 4.

Ordinary Least Square, Tobit, and Heckman sample selection model estimation

Dependent Variable: Log monthly hours of PT

	Variables	OLS[1]	Tobit Marginal Effect	Heckman Two-stage[2]
Student Characteristics	Gender	-0.0252	-0.0406	-0.0453*
		(0.0368)	(0.0749)	(0.0273)
	school ranking 1st quintile	0.0599*	0.1454	0.0746**
		(0.03)	(0.1029)	(0.0367)
	School ranking 2nd quintile	0.0773**	0.2985***	0.1107***
		(0.0349)	(0.1057)	(0.0389)
	School ranking 4th quintile	0.0579	-0.0427	0.0589
		(0.0509)	(0.1331)	(0.0477)
	School ranking 5th quintile	0.0013	-0.4759***	-0.0561
		(0.0389)	(0.1241)	(0.0503)
Family Background	Log monthly income	0.1975***	0.6219***	0.2769***
		(0.0328)	(0.0826)	(0.0403)
	Average years of parents' schooling	0.0136**	0.0775***	0.0298***
		(0.0064)	(0.0193)	(0.0089)
	Number of children	-0.0288	-0.1145*	-0.0479**
		(0.0226)	(0.0637)	(0.0243)
School characteristics	Private (=1)	-0.0243	-0.0611	-0.0325
		(0.043)	(0.0875)	(0.0307)
	Student-teacher ratio	0.0176**	0.0252	0.0267***
		(0.0078)	(0.0208)	(0.0079)
	School size: 300-500 3rd year (=1)	-0.0641	0.0498	-0.0819
		(0.0998)	(0.2104)	(0.0713)
	School size: 500 or more 3rd year (=1)	-0.0248	0.0477	-0.0363
		(0.1111)	(0.2356)	(0.0788)
	High school under Equalization Policy (=1)	-0.0056	0.14301	0.0827
		(0.0601)	(0.1248)	(0.0549)
	Special Purpose High school (=1)	-0.1621**	-0.2419	-0.129**
		(0.0638)	(0.163)	(0.0553)
Contextual effect	Proportion of classmates receiving PT service	0.0006	0.0411***	0.0013
		(0.0013)	(0.0024)	(0.0009)
Price	Log hourly price of PT	-0.5053***	-0.6313***	-0.5075***
		(0.0241)	(0.0885)	(0.0258)
Region	Seoul	0.4268***	0.4559***	0.5419***
		(0.0742)	(0.1287)	(0.0623)
	Metropolitan cities	0.0538	0.2197*	0.1609***
		(0.0648)	(0.1168)	(0.0572)
Constant		4.6137***	-5.8062***	2.6856***
		(0.4235)	(1.33)	(0.8166)
Lambda(λ)		-	-	0.4923***
				(0.1597)
Rho (ρ)		-	-	0.7429
Observations		2,515	3,562	3,601
R^2 / Pseudo R^2		0.2248	0.0721	-
LR chi^2		-	964.86	1077.88

Note. 1. Standard error in parentheses are adjusted for clustering on schools.
2. Standard errors in parentheses are adjusted for clustering on schools.
***p<.01 , **p<.05, * p<.1

TABLE 5.

Conditional Moments Tests Results

Dependent Variable: Log monthly expenditure on PT

Test	Moment Restrictions	t-statistics
Omitted Variables	$E(PRED^2) = 0$	44.37
	$E(PRED^3) = 0$	45.46
Heteroscedasticity	$E(Z_1(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	30.19
	$E(Z_2(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	22.36
	$E(Z_3(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	16.78
	$E(Z_4(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	10.74
	$E(Z_5(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	7.69
	$E(Z_6(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	52.06
	$E(Z_7(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	51.45
	$E(Z_8(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	47.63
	$E(Z_9(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	25.60
	$E(Z_{10}(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	35.00
	$E(Z_{11}(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	23.23
	$E(Z_{12}(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	51.36
	$E(Z_{13}(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	47.47
	$E(Z_{14}(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	10.25
	$E(Z_{15}(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	53.36
	$E(Z_{16}(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	35.79
	$E(Z_{17}(E(u^2 \mid y) - \hat{\sigma}^2)) = 0$	18.07
Non-normality	$E(E(u^3 \mid y)) = 0$	7.39
	$E(E(u^4 \mid y) - 3\sigma^4) = 0$	5.58

Note 1. PRED^2= $(x_i'\hat{\beta})^2 \times \hat{\eta}$), PRED^3= $(x_i'\hat{\beta})^3 \times \hat{\eta}$), η is the generalized residuals.

Note 2. Absolute values

Note 3. Z_1=Gender, Z_2= school ranking 1st quintile, Z_3= School ranking 2nd quintile, Z_4=School ranking 4th quintile, Z_5= School ranking 5th quintile, Z_6=Log monthly income, Z_7=Average years of parents' schooling, Z_8=Number of children, Z_9=Private school, Z_{10}=School size:300-500 3rd grade, Z_{11}= School size:500 or more 3rd grade, Z_{12}= Teacher-student ratio, Z_{13}=High school under Equalization Policy, Z_{14}=Special Purpose High School, Z_{15} =Proportion of classmates receiving PT service, Z_{16}= Seoul, Z_{17}= Metropolitan cities

TABLE 6.

Conditional Moments Tests Results

Dependent Variable: Log monthly hours of PT

Test	Moment Restrictions	t-statistics
Omitted Variables	$E(PRED^2)=0$	0.01
	$E(PRED^3)=0$	5.81
Heteroscedasticity	$E(Z_1(E(u^2 \mid y)-\hat{\sigma}^2))=0$	10.19
	$E(Z_2(E(u^2 \mid y)-\hat{\sigma}^2))=0$	10.48
	$E(Z_3(E(u^2 \mid y)-\hat{\sigma}^2))=0$	9.87
	$E(Z_4(E(u^2 \mid y)-\hat{\sigma}^2))=0$	5.99
	$E(Z_5(E(u^2 \mid y)-\hat{\sigma}^2))=0$	1.35
	$E(Z_6(E(u^2 \mid y)-\hat{\sigma}^2))=0$	19.39
	$E(Z_7(E(u^2 \mid y)-\hat{\sigma}^2))=0$	20.09
	$E(Z_8(E(u^2 \mid y)-\hat{\sigma}^2))=0$	18.11
	$E(Z_9(E(u^2 \mid y)-\hat{\sigma}^2))=0$	10.34
	$E(Z_{10}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	14.40
	$E(Z_{11}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	10.51
	$E(Z_{12}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	19.15
	$E(Z_{13}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	19.31
	$E(Z_{14}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	2.83
	$E(Z_{15}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	22.20
	$E(Z_{16}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	18.63
	$E(Z_{17}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	4.05
	$E(Z_{18}(E(u^2 \mid y)-\hat{\sigma}^2))=0$	18.05
Non-normality	$E(E(u^3 \mid y))=0$	9.15
	$E(E(u^4 \mid y)-3\sigma^4)=0$	13.14

Note 1. $PRED^2=(x_i'\hat{\beta})^2\times\hat{\eta}$), $PRED^3=(x_i'\hat{\beta})^3\times\hat{\eta}$), η is the generalized residuals.

Note 2. Absolute values

Note3. Z_1=Gender, Z_2=school ranking 1st quintile, Z_3= School ranking 2nd quintile, Z_4=School ranking 4th quintile, Z_5=School ranking 5th quintile, Z_6=Log monthly income, Z_7=Average years of parents' schooling, Z_8=Number of children, Z_9=Private school, Z_{10}=School size:300-500 3rd grade, Z_{11}= School size:500 or more 3rd grade, Z_{12}=Student-teacher ratio, Z_{13}=High school under Equalization Policy, Z_{14}=Special Purpose High School, Z_{15}=Proportion of classmates receiving PT service, Z_{16}= Seoul, Z_{17}=Metropolitan cities, Z_{18} =Price of PT