Dynamic Competition in Parental Investment and Child's Efforts *

Hyunjae Kang

Kyoto University, Japan

First Version: November 2022
Current Version: September 2024

Abstract

Competition for a limited number of seats in prestigious colleges generates a "rat-race" equilibrium effect, leading to increased household investment. I construct and estimate a dynamic tournament model in which each household chooses the quality and hours of private tutoring as well as hours of student self-study. The model rationalizes the high amount of parental investment in secondary school despite its low effects on academic achievement. I use the estimated model to quantify the importance of household choices on the intergenerational persistence of earnings. Removing child efforts amplifies this persistence by 36%, emphasizing the role of self-effort in moderating the intergenerational link. By leveraging the number of seats in the tournament, I find that increasing the number of seats in elite colleges by 50% leads to a decrease in the average private tutoring expenditure by 14%. Declining cohort size leads to a modest decrease in private tutoring expenditure unless there is a significant decrease in the elite college premium.

JEL Classification Codes: D15, D64, I21,I22, I24, I26, J62
Keywords: Parental Investment, Child's Efforts, Intergenerational Mobility, Student Competition

*Institute of Economic Research, Kyoto University, Japan. I am immensely grateful to Steven Stern and Juan Pantano for their invaluable guidance and support. I would also like to thank Meta Brown, David Wiczer, Gabriel Mihalache, Peter Arcidiacono, Michael Keane, Lance Lochner, Olivier De Groote, Nirav Mehta, John Rust, Todd Stinebrickner, Michela Tincani, Michèle Tertilt, Minchul Yum, Sunham Kim, Changhui Kang, Alexis Anagnostopoulos, Arda Aktas, Minji Bang, Hugo Benitez-Silva, Juan Carlos Conesa, Samuele Centorrino, Miguel Poblete-Cazenave, Sunghun Cho, Jihwan Do, Marcos Fernandes, Jeremy Fox, Mateo Velásquez-Giraldo, Yujung Hwang, Fedor Iskhakov, Chinhui Juhn, Shinhyuk Kang, Laura Karpuska, Jiyeon Kim, Alfonso Flores Lagunes, Marc Claveria, Bob Millard, Mark Montgomery, José Alfonso Muñoz, Victor Ronda, Warren Sanderson, Eran Shmaya, Takashi Unayama, Greg Veramendi, Jiayi Wen, Sangha Yoon, Basit Zafar, Weilong Zhang, and Yiyi Zhou for helpful comments. All errors are my own.

1. Introduction

Competition motive is a pivotal driver of parental investments. Graduating from an elite university has a sizeable impact on labor market outcomes (Hoekstra 2009; MacLeod *et al.* 2017; Zimmerman 2019; Anelli 2020; Sekhri 2020; Guo and Leung 2021; Jia and Li 2021; Lee and Koh 2023), but seats for such prestigious colleges are limited. The scarcity of seats in prestigious colleges leads to competition, potentially creating a "rat-race" scenario that drives an increasing trend in parental investment (Ramey and Ramey 2010). In the United States, parents frequently invest considerable time in supporting their children's extracurricular activities. In East-Asian countries, parents often allocate a significant portion of their income to private tutoring, primarily aimed at securing admission to prestigious universities, (Bray 1999, 2022), despite evidence suggesting minimal impact on test score improvement (Ryu and Kang 2013; Kang and Park 2021). Most previous work on parental investment does not include this competition aspect into their framework.[1,2]

On the other hand, parental investment significantly contributes to the intergenerational transmission of earnings (Caucutt and Lochner 2020; Bolt *et al.* 2021a,b; Gayle, Golan and Soytas 2022; Yum 2022). It is natural to conjecture that children who have received more investment from parents are likely to achieve better future outcomes such as better performance in the labor market (Becker and Tomes 1979; Guryan, Hurst and Kearney 2008). Thus, parental investment potentially has important consequences on social mobility. Meanwhile, previous studies report the significant impact of children's own efforts on their educational outcomes (Stinebrickner and Stinebrickner 2004; Del Boca, Monfardini and Nicoletti 2017; Fu and Mehta 2018; Todd and Wolpin 2018; De Groote 2023; Del Boca, Flinn, Verriest and Wiswall 2023). The self-effort of the child is not responsive to parental background as much as parental investment is affected by parental background.[3] Despite the potential relevance, few studies have modeled the interdependence of parental investment and children's self-effort in shaping intergenerational mobility.

This paper investigates these two interrelated aspects of parental investment. First, the intensity of household competition is affected by the limited number of college

[1] On the other hand, the competition between private and public schools is extensively studied in the literature. See Epple and Romano (1998); Epple, Figlio and Romano (2004); Epple and Romano (2008); Epple, Romano and Urquiola (2017, 2021).

[2] In Section 3.3, I state the reasons why I do not model competition between secondary schools.

[3] I show related empirical evidence in Section 4.

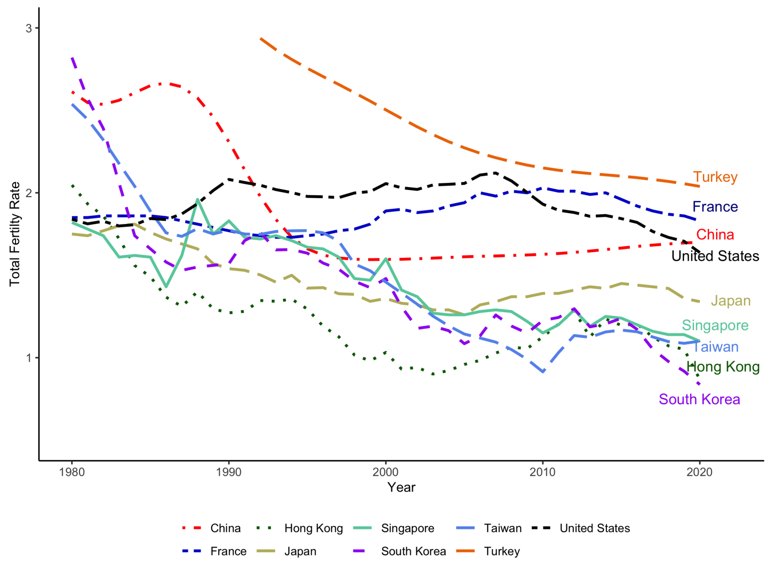
seats with different level of quality and the number of competitors. If there are significant changes in the number of competitors, the decision of parental investment and child effort is likely to be affected. Relatedly, many developed countries face a drastic shift in demographic structure caused by a declining fertility rate, as shown in Figure 1. At the same time, empirical evidence suggests that colleges tend to not adjust the seats to accommodate for increasing cohort size (Bound and Turner 2007). Little is known about the impact of the number of seats in elite colleges and the shift in demographic structure on parental investment. Second, using the college admission competition set-up, this paper seeks to shed light on the role of parental investment in intergenerational persistence or earnings. The inclusion of self-effort of the child, which is often overlooked in the literature, might amplify or offset the link between the two generations.

To answer these questions, this paper builds and estimates a dynamic tournament model using a unique longitudinal dataset that contains information on parental investment, the child's time allocation, and administrative test scores. I first document the descriptive evidence regarding parental investments and the self-efforts of the child. Second, motivated by empirical evidence, I build a dynamic model of a tournament which approximates the college admission competition among households. I estimate the tournament model using Maximum Simulated Likelihood. In a series of graphs, I show that the model fits reasonably well with the data. Finally, I perform counterfactual exercises using the estimated structural model. I quantify the impact of parental investment and the child's self-efforts on intergenerational mobility. Then I simulate the model to investigate the impacts of relaxing elite college constraints on parental investments.

This study uses Korean datasets and is based on institutional features of the country.[4] Students are assigned to the middle schools within the residential education district by lottery. As the distribution of school quality of secondary school is relatively homogeneous, the private tutoring expenditure of parents stands out as a primary contribution to the child's future outcomes. The importance of the final test score in college admissions helps to link the test score measure to the child's labor market outcomes. Such institutional characteristics offer a transparent environment in which household income is translated into the educational outcome of the child.

[4]A number of countries share the institutional features, which I explain in Section 3.

Figure 1: Shrinking Cohort



Source: World Bank for data for China, Hongkong, Japan, Singapore and South Korea. Data for Taiwan is drawn from United Nations World Population Prospects.

I start by documenting the descriptive evidence that provides the empirical basis of the dynamic tournament model. Two empirical facts show that competition with respect to getting into a more prestigious college is the primary motivation for parental investment. First, college ranking positively affects the growth of alumni's income. Using the Korean Labor Income and Panel Study, I estimate the effects of college-tier, a categorization of colleges in Korea based on their quality measured by alumni's income growth. Pooled OLS results suggest that there is a significant variation in lifetime income based on tier of the college from which workers graduate. The effects are economically and statistically significant, controlling for CSAT score. This evidence is consistent with the empirical studies on the effects of elite colleges on labor market outcomes (Zimmerman 2019; Sekhri 2020; Jia and Li 2021; Lee and Koh 2023).[5] Second, the amount of parental investment drops substantially as students finish the college admission process. This suggests that the purpose of parental investment is for their child to do well in the college admission competition rather than enhancing the child's human capital.

Also, another empirical fact suggests that parental investment and the child's self-efforts potentially have different implications for intergenerational mobility. As

[5]Using data on students majoring in science at University of California campuses, Arcidiacono, Aucejo and Hotz (2016) show the mismatch between minority students' preparedness and higher ranked campuses decreases the likelihood of graduation.

expected, data show that parental background, especially household income, generates a significant variation in parental investment. On the other hand, self-efforts of the child, measured by hours of self-study, do not vary as much as parental investment with different levels of parental income. At the same time, both parental investment and the child's self-efforts are expected to affect the child's outcome. If parental investment and self-efforts are technological substitutes, an income-constrained household can compensate for the lack of parental investment by increasing hours of self-study. Thus, omitting self-efforts of the child might result in an exaggeration of intergenerational persistence of earnings. This suggests the importance of modeling both parental investment and the child's self-efforts in understanding the contribution of educational investments to the intergenerational persistence of earnings.

Motivated by the empirical evidence, I develop and estimate an equilibrium dynamic tournament model of college admission competition. The model builds upon the rank-order tournament model introduced by Lazear and Rosen (1981). The tournament structure is embedded into the model of altruistic households. The household cares about the future outcome of the child, which is the result of the college admission tournament. In every period, based on its state variable, each household makes decisions regarding parental investment and the level of the child's self-efforts, which produces a subsequent test score. The model structure repeats until the final test score is produced, and students are assigned to the college tiers based on the ranking of the final test score and the number of seats for each college tier. The college tier is the sole determinant of the child's lifetime income.[6]

Two main features of the dynamic tournament model help answer the research question of this paper. First, the rich heterogeneity of state variables and the choice set, along with the specification of the test score production function, help disentangle the source of intergenerational persistence of earnings. Second, the rank-order feature of the tournament model enables me to study the effects of the changes in the number of seats in colleges and the disparity in college quality on parental investment.

I estimate the model using Maximum Simulated Likelihood. The estimation results suggest that hours of self-study have stronger average marginal effects on the subsequent test score than hours of tutoring, while the marginal effects of both investments decline over time. Additionally, the estimate of the substitution parameter of the pro-

[6]This is an arguably reasonable assumption. In Section 5.2, using confidential job offers data provided by a conglomerate in South Korea, I show that the effects of college-tier on earnings are economically and statistically significant controlling for effort during the college period.

duction function suggests that parental investments and hours of self-study are close to perfect substitutes. Unlike disutility from hours of tutoring, parental education has significant effects in reducing disutility from hours of self-study. Using local linear regression, I show that the estimated model fits the sample reasonably well.

Using the estimated structural model, I quantify the impact of household heterogeneity and choices by alternatively shutting down each channel one at a time. Removing heterogeneity in the parental income during the adolescent period decreases the rank-rank slope by 47.2%. When the channel of self-study is shut downThe rank-rank slope increases by 30.2%. The result of the quantification suggests that parental investment reinforces the intergenerational persistence of earnings and the self-study of the child mitigates it.

Next, I use the model to investigate the consequences of relaxing the seat constraint in elite colleges on the choices of households. First, I simulate the model under the alternative scenario where the number of seats in elite colleges expands.[7] A 50% increase in the number of seats of elite colleges leads to a 14% decrease in the average private tutoring expenditure. Second, I simulate the model under the scenario where the size of the cohort is reduced by half, which is based on the projected number of 12th-grade students in Korea in 2033. Shrinking cohort size does not lead to a decrease in private tutoring expenditure unless there is a significant decrease in the elite college premium. As the cohort-to-seat ratio decreases, students have an increased chance of going to a better college tier, which increases their incentives for spending on private tutoring.

The rest of the paper is organized as follows. I discuss the related literature and contributions of this paper in Section 2. I describe the institutional features in Section 3. In Section 4, I document empirical facts that motivate the dynamic tournament model. Section 5 introduces the tournament model. Section 6 explains the estimation procedure, source of identification, and results. I present the counterfactual exercises in Section 7 and conclude in Section 8.

2. Related Literature and Contributions

This paper relates to the large body of literature on post-birth parental choice (Becker and Tomes 1979; Del Boca, Flinn and Wiswall 2014; Doepke and Zilibotti 2017; Bolt,

[7]The elite colleges refer to a group of colleges which I define as Tier 1 in Section 3.

French, Maccuish and O'Dea 2021b). Papers in this literature have recently started incorporating competition into parental choices.[8] Ramey and Ramey (2010) are the first paper that rationalizes the increase of parental time investment in the United States using a theoretical model of competition for elite colleges. Bodoh-Creed and Hickman (2019) build a static structural model of an admission contest to study returns to pre-college human capital investment in the United States and estimate their model.[9,10] A closely related paper on the same context is by Kim, Tertilt and Yum (2024), which studies the cause of the low fertility problem of South Korea. They propose a model of "status externality" based on the assumption that parents care about the relative position of their children's human capital compared to that of other children. The tournament model of my paper complements their study by formally modeling the dynamic competition with respect to getting into prestigious colleges, which rationalizes underlying source of the status externalities in their paper.[11] Also, the number of seats and the equilibrium cutoffs of the tournament model enable me to investigate the impact of the college seats' constraint on the demand for parental investments.

This paper naturally relates to the theoretical and empirical literature on parental investment and its intergenerational implications (Lee and Seshadri 2019; Caucutt and Lochner 2020; Bolt, French, Maccuish and O'Dea 2021b; Daruich 2022; Gayle, Golan and Soytas 2022; Yum 2022). In particular, Del Boca, Flinn and Wiswall (2014) build and estimate a dynamic model of parental investment and cognitive development, which allows them to separately identify the different effects of parental time and monetary investments. Subsequently, emphasizing the role of child's self-investment, Del Boca, Flinn, Verriest and Wiswall (2023) build a Stackelberg model of parent-child interaction and investigate the effects of conditional cash transfers on child outcomes. De Groote (2023) quantifies the role of students' efforts in the academic tracking system using a dynamic model. My paper contributes to the literature by jointly modeling the dynamic decisions of parental investment and the self-efforts

[8]A group of papers associates parental choices with social interactions (Agostinelli 2018; Agostinelli, Doepke, Sorrenti and Zilibotti 2023; Boucher, Bello, Panebianco, Verdier and Zenou 2022)

[9]While abstracting from the notion of parental investment, Grau (2018) builds and estimates a static tournament model to study the college competition in Chile.

[10]Outside the broad literature of economics of education, a handful of papers build and estimate structural tournament models (Vukina and Zheng 2007; Chen and Shum 2010; Vukina and Zheng 2011).

[11]Gu and Zhang (2024) is another recent paper modeling the college admission competition using heterogeneous agents framework. However, they do not model the thresholds for the college admission as equilibrium objects of the students' competition.

of the child. The estimated model quantifies the impact of parental investment and the child's self-efforts on the intergenerational persistence of earnings, providing novel insights about the mechanism that generates intergenerational correlations in earnings.

Finally, this paper contributes to the literature on childhood investments and skill development by estimating the age-specific effects of parental investment and the self-efforts of the child during adolescence.[12] Most previous work focuses on estimating the effects of parental investment on child outcomes alone (for example, Cunha and Heckman 2007; Cunha, Heckman and Schennach 2010; Del Boca, Flinn and Wiswall 2014).[13] These studies find declining effects of parental time investment over age. Several studies estimate the effects of hours of self-study on academic achievements (e.g., Cooper, Robinson and Patall 2006; Stinebrickner and Stinebrickner 2008; Fu and Mehta 2018; Todd and Wolpin 2018). The production function estimates in my paper add to this literature by providing age-specific estimates of the effects of parental investment and self-efforts (Del Boca, Monfardini and Nicoletti 2017; Del Boca, Flinn, Verriest and Wiswall 2023), and offer novel evidence on their substitutability.

3. Key Institutional Features

As this paper utilizes Korean datasets, the theoretical framework and the identification strategy are based on the country's institutional features. In this section, I explain the key institutional features of the country: the high-stakes college entrance exam, hierarchical college structure, homogeneous secondary schools, and an established private tutoring market. While these institutional characteristics offer several advantages in studying the research questions, a number of countries share these features. As I describe the characteristics of the system, I explain the possibility of generalization for other countries.

[12]As college competition in reality uses actual test scores rather than unobserved skills of the student, I do not apply the factor model techniques developed in the literature (see Cunha, Heckman and Schennach (2010); Agostinelli and Wiswall (2016)).

[13]As this paper employs private tutoring expenditure as a measure of parental investment, it also complements the literature of studies on private tutoring (Stevenson and Baker 1992; Cheo and Quah 2005; Tansel and Bircan Bodur 2005; Dang 2007; Ono 2007; Ryu and Kang 2013; Hof 2014; Kang and Park 2021).

3.1 High-Stakes College Entrance Exam

In Korea, the College Scholastic Ability Test (CSAT), the college entrance exam taking place at the end of 12th grade, is the single most important factor for college admission.[14] Students take Korean, Mathematics, English, and elective subjects. The exam starts at 8:40 am and finishes at 5:45 pm. For this exam, take-offs and landings of airplanes are suspended for 35 minutes during the English listening test. Firms and government offices are encouraged to delay their workday by an hour to help students avoid heavy traffic. All these suggest that the taking of the CSAT is a huge national event. After the exam, students receive a scoresheet that contains a standardized score and a stanine score for each subject.[15,16] Many educational consulting firms publish the "cutoff sheet" that contains the firm's prediction for the cutoffs for all colleges. The predictions are largely consistent across the firms and are close to the actual cutoffs. Based on the CSAT score and the predicted cutoffs, each student chooses up to three colleges in which to apply. Based on the CSAT score and the quota, colleges determine admission results for students. Several countries have their own high-stakes college entrance exam. *Gaokao* of China is a representative example in that the ranking in the exam is the most crucial factor in college admission. Baccalauréat of France is highly important for getting into grandes écoles, the group of elite colleges of the country.[17] The Scholastic Aptitude Test (SAT) of the United States is also utilized as an important factor in college admission, but other components such as high school grade-point-average and extra-curricular activities also matter.

3.2 Hierarchical College Structure and College-Tier

The institutional feature also prevalent in other countries is a hierarchical college structure. In many countries including Korea, college quality is unequal in terms of alumni outcomes. Empirical studies report that graduating from an elite college

[14]In South Korea, there has been a recent increase in the quota for the holistic review process, in which test score is not the only determinant for college admission. In 2019, 24.9% of total students were admitted through the holistic admission route (Bastedo 2021).

[15]There was one exception in 2007 in which only stanine scores were available for the college admission process. The original standardized score system was restored in 2008. Han, Kang and Lee (2016) estimate the changes in aggregate effort level of the students due to the grade scheme shift.

[16]A stanine score is nine discrete scales ranging from a low of 1 to a high of 9.

[17]Other examples include *Yükseköğretim Kurumları Sınavı* of Turkey, Exame Nacional do Ensino Médio of Brazil, Sijil Pelajaran Malaysia of Malayasia, and Ulttyq Biryńǵaı Testileý of Kazakhstan are highly similar in terms of their importance in the college admission process.

significantly affects a student's future labor market outcomes.[18] In South Korea, the college hierarchy has changed little (Kim and Lee 2006; Kim 2014). Starting from the top institution, Seoul National University, the applicants' preferences have been stable for decades, and "SKY" is a well-known acronym that refers to the top three universities in the country.[19,20]

Motivated by the college hierarchy, I categorize colleges in Korea into four ordered tiers based on the "cutoff sheet" published by Jinhak (2022), one of the major education consulting firms. Tier 1 includes the most prestigious universities. The cutoff of Tier 1 is around the top 1% of CSAT scores. Successively, the cutoffs of Tier 2 and 3 are approximately the top 5%, and top 15% of the CSAT score distribution, respectively.[21] Tier 4 is composed of graduates from 2 year colleges. Tier 5 is the residual tier that absorbs the rest of the students in the cohort. The member universities of each tier are specifically reported in Appendix B. I use this categorization of college tiers throughout this paper. In Section 4, I present empirical evidence suggesting the significant effects of the college tier on post-graduation labor market outcomes.

3.3 Homogeneous Secondary School and Private Tutoring Market

Secondary schools in South Korea are homogeneous, which provides a transparent environment where private expenditure translates into students' academic performance. First, the curriculum of secondary school is uniform and under the strict control of the Korean government. In addition to public schools, even private schools do not have autonomy in terms of the curriculum and tuition.[22] Second, as a result of the consecutive school-equalization policies, the quality of education provided by schools

[18]See, for example, Hoekstra (2009) for the United States, MacLeod *et al.* (2017) for Colombia, Zimmerman (2019) for Chile, Anelli (2020) for Italy, Sekhri (2020) for India, and Jia and Li (2021) for China.

[19]In the late 1990s, the term "In Seoul" has appeared, which refers to a group of all universities in Seoul. Anecdotally, Korean parents often say that they hope their children go to one of these "In Seoul" universities.

[20]Kim and Lee (2006) study this hierarchical market structure of universities in Korea and show that a strong university hierarchy is present in the country. They report that universities in the first three deciles strictly dominate the rest in terms of their measure of labor market outcomes, private donations, quality of faculties, and physical facilities.

[21]The top 15% score is the cutoff for the "In Seoul" universities previously mentioned.

[22]One of the few decisions of private secondary schools in Korea is that they can independently hire teachers. Park, Behrman and Choi (2013) provide evidence that the difference in the quality of teachers is not significant between private and public secondary schools in Korea.

is similar.[23] No schools are allowed to select students independently.[24] In fact, school assignments for middle school and high school are random within the residential district for most regions. After graduating from primary school, students are assigned to the middle schools within the residential education district by lottery.[25]

At the same time, 2.8% of GDP is spent on private tutoring activities for students by households in South Korea (Nam 2007). Parents spend 9% of their income on private tutoring activities for their children, which is a significant amount of expenditure.[26] The form of private tutoring varies. The most common form of private tutoring is *hagwon* (or cram school), the private academic institutions students go to after regular school hours. There are also one-on-one tutoring, group tutoring, and online classes. The country has an established private tutoring market. With the centralized school curriculum, private tutoring institutes are an effective substitute for parental time in teaching their kids. I use private tutoring as a measure of parental investment throughout the paper.

Two main features highlighting the education system of Korea are the homogeneous secondary schools and the fact that college admission relies heavily on the final exam. This feature provides a transparent environment in which the household income is translated into the educational outcome of the child.

4. Empirical Evidence

4.1 Data: Korean Educational Longitudinal Study 2005

I use the Korean Educational Longitudinal Study 2005 (KELS) for the main estimation procedure. The choice of data is motivated by the main goals of the paper: (i) to quantify the role of accumulated parental investment and student efforts on intergen-

[23]See Section II of Kim and Lee (2010) for a description of the history of school equalization policy. As of 2010, the high school equalization policy has been adopted for all major cities in South Korea.

[24]One exception is specialized high schools, which are not subject to the equalization policy. However, not like private schools in the United States, admission to specialized schools is mostly merit-based. The enrollment for the specialized schools accounts for only 3% of total enrollment. I expect the disparities due to the specialized high schools are captured by the household characteristics of the dataset.

[25]Papers in the literature exploit this random assignment feature to estimate the effects of various independent variables of interest on educational outcomes. See, for example, Kang (2007), Park, Behrman and Choi (2013), and Park, Behrman and Choi (2018). Park, Behrman and Choi (2013) show that the issue of non-compliers to the lottery policy is a minor concern.

[26]See Bray (1999, 2021) for a comprehensive cross-country comparison of private tutoring.

erational mobility, and (ii) to account for the dynamic selection of the effort choices of the household in the competition. Estimating the marginal effects of parental investment and hours of self-study in each period is necessary to achieve the goals. The Korean Educational Longitudinal Study 2005 includes information on private tutoring expenditure, hours spent for private tutoring, hours spent for self-study, income of the household, standardized test scores, and parental education, which is a rare combination for one dataset. Household income and private tutoring expenditure are collected each year. The hours spent in tutoring activities and the hours spent for self-study are collected as a weekly average. There are five different measures of academic performance available in the dataset. Academic performance in primary school is measured as an ordered discrete measure answered by the household. For 7^{th} to 9^{th} grades, the administrative test scores are of achievement tests standardized at the national level. For 12^{th} grade, the administrative College Scholastic Ability Test (CSAT) score is available. The actual scores are available for the three achievement tests and the CSAT, which I treat as continuous variables.

The nationally representative dataset tracks 6,908 students (1^{st} -year middle-school students) sampled from the country's 703,914 7^{th} grade students. The students are tracked starting from 2005 when they are 7th graders. In the first stage of the survey, the cohort is surveyed yearly up to 2012. In the second stage of the survey, namely the college and the labor market period, the cohort is surveyed semi-annually up to 2020, which is ten years after the cohort graduates from high school. The rules of selection and their effects are reported in Table C.1.[27] In addition, I include households missing one of the choice variables: tutoring expenditure, hours of tutoring, and hours of self-study. In the estimation section, I explain the rules to simulate the missing choice variables.

[27]The proportion of observations lost to missing the final test score is 0.48. Meanwhile, 99.9% of the students in the dataset report that they applied for the final exam, which suggests that the missing final exam score is not caused by the selection to take the final exam.

Table 1: Sample Moments

(a) Sample Moments: 7th - 9th grades

School grade	7th		8th		9th	
	Mean	Stdev	Mean	Stdev	Mean	Stdev
Tutoring Expenditure	25.8	20.0	25.1	19.6	36.1	31.0
Hours of Self-Study	5.48	5.04	5.97	5.13	6.45	5.27
Hours of Tutoring	11.37	8.50	9.69	7.22	11.29	9.90
Income	370.4	161.7	369.2	151.3	400.4	169.9
Test Scores	323.03	45.63	321.50	48.72	322.65	48.45
N	1792					

(b) Sample Moments: 10th - 12th grades

School grade	10th		11th		12th	
	Mean	Stdev	Mean	Stdev	Mean	Stdev
Tutoring Expenditure	38.3	36.5	47.9	48.6	29.5	41.7
Hours of Self-Study	7.65	5.68	8.45	6.00	14.42	9.14
Hours of Tutoring	7.40	6.74	9.16	9.45	5.69	7.89
Income	406.9	177.0	394.4	191.1	381.4	171.4
Test Scores	-	-	-	-	415.39	62.46
N	1792					

Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: The unit for measuring private tutoring expenditure is 10,000 KRW, approximately 8 USD, based on the average exchange rate during the data collection periods. Hours data are weekly measures.

Table 2: Sample Moments (Continued)

(a) Sample Moments: Other characteristics

	Mean	Stdev
Parental Education	13.27	2.01
6th grade Academic Performance	6.52	1.70
N	1792	

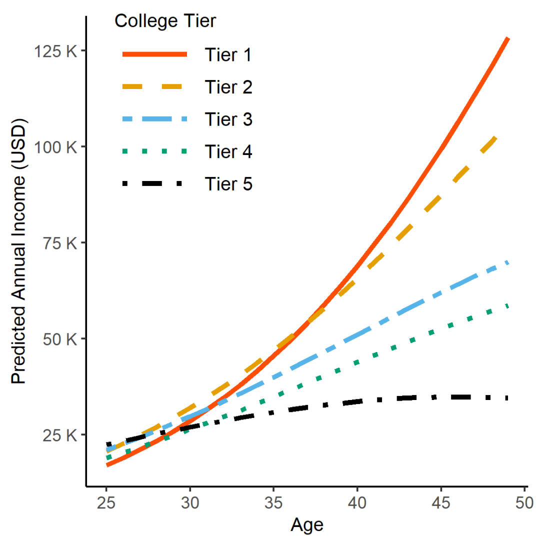
Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.

Table 1 and 2 present sample moments of KELS. While the average hours of self-study increase over time, the average hours of tutoring overall show a decreasing trend. I revisit the implications of such changes in hours allocation in Section 4.5. The moments of household income are stable over time. I use parental education data collected in the first year of the survey, and I assume that parental education does not change within the model period.[28]

To supplement the income information of KELS, I use the Korean Labor Income and Panel Study (KLIPS) to supplement the college tier-specific lifetime income. The description of KLIPS is in Appendix C.1.

4.2 The Lifetime Income Differential

Figure 2: Income Dynamics by College Tiers



Source: Korea Labor Income and Panel Study 1998-2012, Korea Labor Institute.
Note: The sample includes workers between 25 and 65 years old who work for wages or salary. I exclude workers who are born after 1992. The figure has units of 1,000 KRW, which is about 0.85 USD. Annual income is predicted using the Pooled-OLS estimates in column (1) of Table 3.

College ranking has a strong effect on the growth of alumni's income.[29] The effect is significant even after controlling for CSAT score. Columns (1), (2), and (3) in Table 3 provide the OLS estimates for the regression equations,

[28]This is a reasonable assumption given the relatively short period of time in the data. In fact, information on parental education is collected only in the first two years of the survey.

[29]Lee and Koh (2023) reports that the alumni of Tier 1 colleges in Korea earn 50.5% more compared to those from the bottom Tier group, based on their preferred specification and the tier definitions. The lifetime income empirical exercise in this section is consistent with their findings, but the specification differs to consistent with the structural model.

$$\ln y_{it} = \sum_{j=1}^{J} (\beta_j + \delta_j \cdot age_{it}) D_{i,j}^{Tier} + Z_{it}\gamma + \varepsilon_{it}^{y} \tag{1}$$

where $D_{i,j}^{Tier}$ is a dummy variable indicating that person i graduated from a tier j college, and Z_{it} is the set of explanatory variables including age, squared age, birth year, and gender of person i.[30]

Table 3: Log Income Regression

	(1) Pooled OLS	(2) Pooled OLS	(3) Pooled OLS	(4) RE	(5) RE	(6) RE
Top tier × age	0.067***	0.065***	0.111***	0.058***	0.058***	0.105***
	(0.009)	(0.009)	(0.014)	(0.009)	(0.009)	(0.024)
Second Tier × age	0.050***	0.052***	0.098	0.051***	0.055***	0.095***
	(0.011)	(0.010)	(0.054)	(0.010)	(0.011)	(0.031)
Third Tier × age	0.032***	0.037***	0.058**	0.034***	0.044***	0.011
	(0.008)	(0.007)	(0.019)	(0.012)	(0.013)	(0.050)
Fourth Tier × age	0.029***	0.027***	0.073***	0.022***	0.021***	0.038***
	(0.007)	(0.007)	(0.015)	(0.002)	(0.002)	(0.013)
N	29599	29599	685	29599	29599	752
Major	No	Yes	No	No	Yes	Yes
RE	No	No	No	Yes	Yes	Yes
CSAT	No	No	Yes	No	No	No

Source: Korea Labor Income and Panel Study 1998-2012, Korea Labor Institute.
Note: RE refers to "Random Effects." Explanatory variables used in the regressions such as squared age, birth year, and gender are excluded from the table for brevity. The sample includes workers between 25 and 65 years old who work for wages or salary. I exclude workers who are born after 1992.

Note that the regression equation captures both the effect of graduating from a tier j college on the level and the growth of an alumnus's income, respectively by β_j and δ_j. Columns (3) and (4) provide the estimates of the random effects model,

$$\ln y_{it} = \sum_{j=1}^{J} (\beta_j + \delta_j \cdot age_{it}) D_{i,j}^{Tier} + Z_{it}\gamma + \lambda_i^{y} + \eta_{it}^{y}$$

[30]A full set of estimates are available in Appendix C. The purpose of the birth year dummy variable is to capture the cohort difference in workers' income.

where λ_i^y and η_{it}^y are the individual-specific and the idiosyncratic errors respectively.[31] Columns (2) and (5) include the dummy variables of college-major, showing that the inclusion of major does not critically affect the main results of Columns (1) and (4), respectively. The Tier 1 dummy has the smallest estimate of intercept but the largest estimate of age differential. Figure 2 presents the predicted annual income of alumni using the estimates in Column (1) of Table 3. Before age 30, there is no economically significant difference in terms of annual income. On the other hand, the gap becomes significantly larger as people age. The effects are significant controlling for CSAT score, as can be seen in Columns (3) and (5) of Table 3.[32] The estimation results are consistent with the studies stressing the importance of using lifetime income in the returns to schooling literature (Haider 2001; Tamborini *et al.* 2015; Nybom 2017). The effects of parental investment on labor market outcomes through college reputation would be underestimated if researchers narrow their focus to the early labor market outcomes. The Pooled-OLS estimates in Column (1) of Table 1 are used in computing college-specific lifetime income, which is a component of the dynamic tournament model.

4.3 Competition Motives of Parental Investment

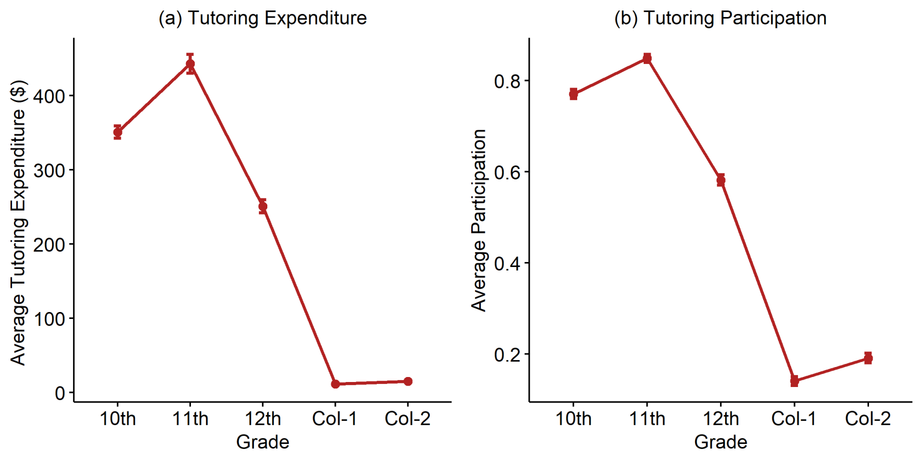
Competition with respect to getting into a more prestigious college is the primary motivation of parental investment. First, data suggest that the demand for private tutoring expenditure significantly drops as students finish the college admission process. Figure 3 presents the change of tutoring expenditure and participation rate over time for the sample cohort of KELS. Both expenditure and participation of private tutoring rapidly drop as soon as students graduate from high school, which suggests that the primary purpose of tutoring expenditure is associated with college admission. If the purpose of tutoring expenditure was for enhancing the student's human capital, it is unlikely that most students would completely stop private tutoring activities upon graduating from high school. Second, the number of seats at prestigious colleges is limited. Even with a very high final test score, students might not be able to go to a top-tier college if the seats are filled with students with higher test scores. The scarcity of seats at prestigious colleges and the fact that tutoring participation drops after

[31]Since the focus of the regression is the college tier, which is time-invariant, I do not consider the fixed effects model.

[32]As CSAT performance is collected as a discrete variable in KLIPS, the estimation is different with Regression Discontinuity Design.

the college entrance exam show that competition is the key feature determining the parental investment decision of the household.

Figure 3: Private Tutoring Expenditure and Participation in Tutoring

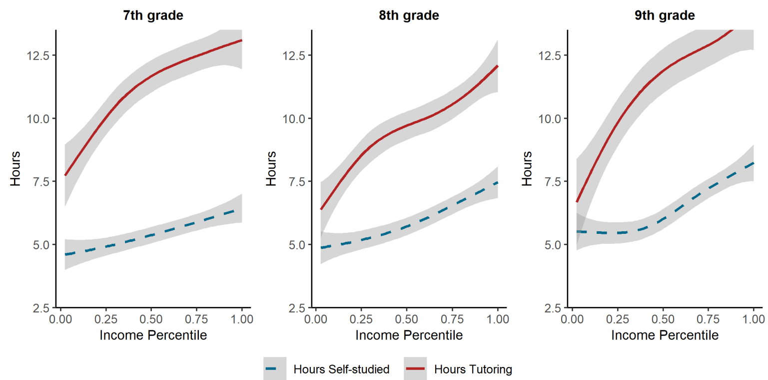


Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: I only include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, and household income.

4.4 Parental background and child's hours allocation

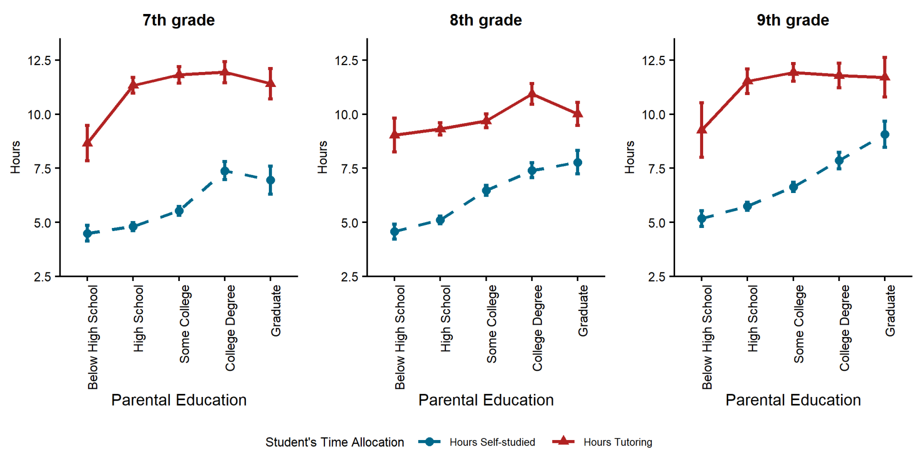
Compared to hours of tutoring, hours of self-study are less affected by parental income, which potentially has implications for intergenerational mobility. On the one hand, the income elasticity of hours of tutoring is higher than the income elasticity of hours of self-study. Figure 4 presents how hours of tutoring and hours of self-study vary with parental income when students are 7th, 8th, and 9th graders, using local linear regression. The slope of hours of tutoring is much steeper than the slope of hours of self-study, which shows that tutoring is an effort choice that is more responsive to parents' income. On the other hand, the covariation between hours of self-study and parental education is higher than the covariation between hours of tutoring and parental education, conditional on other household characteristics. Figure 5 presents how hours of tutoring and hours of self-study vary with parental education when students are 7th, 8th, and 9th graders. Unlike household income, the effect of parental education is higher on hours of self-study than the effect of parental education on hours of tutoring.

Figure 4: Income Gradient in Effort Decision



Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: The gray regions are confidence bands with a significance level of 0.05. I only include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, and household income.

Figure 5: Parental Education and Efforts Allocation



Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: In this graph, parental education is a categorical variable and based on the average years of parents educationm_i, which is defined as follows: Below High School if $m_i < 12$, High School if $m_i = 12$, Some College if $12 < m_i < 16$, College Degree if $m_i = 16$, and Graduate if $m_i > 16$. I only include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, household income, and parental education.

Parental education soaks up significant variation in hours of self-study, which leaves a relatively small variation with parental income. Tables 4 and 5 presents the pooled OLS estimates of the regression equation,

$$\ln(1 + y_{it}) = \beta_0 + \beta_1 \log(hhinc_{it}) + \beta_2 m_i + \epsilon_{it} \tag{2}$$

where $hhinc_{it}$ is the income and m_i is parental education of household i. Columns (1) through (3) present the results where y_{it} is hours of self-study, and columns (4) through (6) present the results where y_{it} is hours of tutoring. Columns (1) and (4) provide the estimates without including the average years of parents' education, and Columns (2) and (5) provide the estimates with including the average years of parents' education to equation (2). Overall, hours of tutoring are explained more by parents' income than hours of self-study. Moreover, much of the covariation between hours of self-study and income is absorbed after controlling for the average years of parents' education.

Such empirical relationships suggest that different household backgrounds can lead to different allocations of effort choice. Thus, omitting one of the effort choices (parental investment or child effort) might result in biased estimates of intergenerational mobility, which calls for including both effort choices in the theoretical framework.

Table 4: The Effects of Parental Background on the Hours Allocation

	(1) log(1+Study)	(2) log(1+Study)	(3) log(1+Study)
log(Income)	0.238*** (0.022)	0.152*** (0.025)	0.036 (0.027)
Parental Edu		0.055*** (0.007)	
N	10454	10454	10454
Year	Yes	Yes	Yes
FE	No	No	Yes

Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: log(1+Study) and log(1+Tutoring) refer to log of hours of self-study plus one and hours of tutoring plus one, respectively. I only include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, and household income. Parental Educ indicates average years of parents' education.

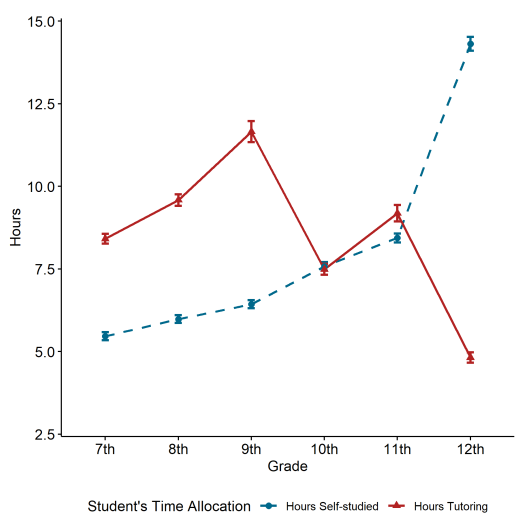
Table 5: The Effects of Parental Background on the Hours Allocation

	(4) log(1+Tutoring)	(5) log(1+Tutoring)	(6) log(1+Tutoring)
log(Income)	0.677*** (0.027)	0.616*** (0.030)	0.269*** (0.037)
Parental Edu		0.038*** (0.008)	
N	9431	9431	9423
Year	Yes	Yes	Yes
FE	No	No	Yes

Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: log(1+Study) refer to log of hours of self-study plus one and hours of tutoring plus one, respectively. I only include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, and household income. Parental Educ indicates average years of parents' education.

4.5 Dynamic effort allocation of households

Figure 6: Dynamic Allocation of Efforts



Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: I only include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, and household income.

Students' time allocation of effort choices considerably changes as students proceed to the later educational stages. Figure 6 presents how the average hours of self-study and the average hours of tutoring change with students' grade level. While the average hours of tutoring shows a decreasing trend, the average hours of self-study shows an increasing trend. In 12th grade, the average hours of self-study is almost three times the average hours spent for tutoring. Such changes in time allocation suggest that the marginal effects of hours of self-study and tutoring expenditures on academic outcomes might change over time.[33] In addition, I discuss how parental investments allocation differs based on their initial conditions in Appendix C.3.

The evidence suggests that the allocation of the two efforts changes over time and that households self-select into the different effort levels based on their preconditions. As suggested earlier, different effort choices might have different implications for intergenerational mobility. Capturing the changing behavior of the households is crucial to get the correct quantification of the statistics of interest.

5. A Dynamic Model of College Admission Tournament

Motivated by the empirical evidence, I build and estimate a dynamic model of competition where each household chooses the amount of parental investment and the level of child's efforts. The dynamic model is built upon the rank-order tournament first described by Lazear and Rosen (1981) and related to its applications in college admission competition (Han, Kang and Lee 2016; Grau 2018; Tincani, Kosse and Miglino 2021).

5.1 Timeline

There exist N households in the dynamic tournament. Each household is composed of one student and the parents. I assume the household makes a unitary decision. I abstract away from the intra-household decision-making process. The students compete for the final prize against other students in the same cohort.

Figure 7 illustrates the timeline of the model. The model begins as the student of the household enters into 7th grade, which is the first year of secondary school. Each household is born with the complete income stream $\{w_{it}\}_{t=1}^{T}$, parental education m_i, and initial test score q_{i1}. Also, each household has a specific type k. Different types of households have different type-specific characteristics that are unobserved by the econometrician. I define them as $\lambda_k^c, \lambda_k^x, \lambda_k^s$ and λ_k^q, which affect marginal utility from

[33] Several studies in the literature report that the effects of parental investment decrease with children's age (Cunha *et al.* 2010; Del Boca *et al.* 2017). To the best of my knowledge, there is no study reporting the changing effects of self-study over time.

consumption, disutility from hours of tutoring, disutility from hours of self-study, and log of test score, respectively. Some households value non-academic goods such as travel more than other households conditional on the observed characteristics (Lazear 1977). Such unobserved taste for consumption is captured by λ_k^c. Some households prefer to encourage their child to study independently rather than send her to tutors, which is captured by the relative size of λ_k^s to λ_k^x. Some students might be particularly good or bad in taking exams, which would be captured by λ_k^q.[34]

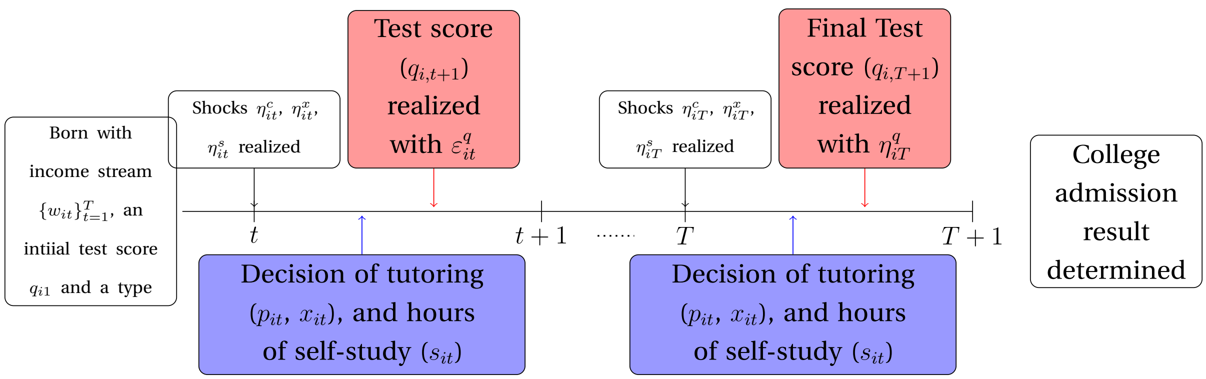


Figure 7: Model Timeline

At each time t, as the household enters into the period, the shock to the marginal utility of the consumption η_{it}^c, the shock to the marginal disutility from the tutoring activities η_{it}^x, and the shock to the marginal disutility from self-study η_{it}^s are realized. These shocks capture the unobserved time-varying components that are not accounted for by the deterministic components of the model. Based on those realized shocks and the observed state variables, each household chooses the quality of tutoring p_{it}, the hours spent on tutoring x_{it}, and the hours of self-study s_{it} to maximize its value function. The choices are subject to budget and time constraints. Subsequently, the test score $q_{i,t+1}$ is produced with the realization of the test score shock. This process repeats until the final test score $q_{i,T+1}$ is generated.

Each student is assigned to a college tier based on the ranking of the final test score and the fixed number of college seats in each tier. I denote n_j as the fixed number of seats for the j^{th} college-tier. In particular, denoting n_1 as the fixed number of seats for the first college tier, the first n_1 students are assigned to the top college tier, and the next n_2 students are assigned to the second tier. The process repeats until the $(J-1)^{th}$

[34] I introduce the joint distribution of the time-specific shocks ($\eta_{it}^c, \eta_{it}^x, \eta_{it}^s$, and η_{it}^q) and the specification of type-specific unobserved heterogeneity (λ_k^c,λ_k^x,λ_k^s and λ_k^q) when I explain the flow utility component of the model.

college tier is filled up with n_{J-1} students so that all seats for the college tiers bind. The bottom tier is a residual tier which is composed of students whose score is below the cutoff for the $(J-1)^{th}$ college tier and the students who do not go to college.[35] The assigned college tier is the sole determinant of ex-post lifetime income.

5.2 The Preliminaries of the Tournament

Prize: Lifetime Income. The prize for going to a more prestigious college tier is a higher expected lifetime income awarded to the student, which motivates the household to exert effort. There exist J college tiers that are characterized by expected lifetime income v_j. The tier-specific lifetime income v_j is the discounted sum of the predicted income of the graduates. In particular,

$$v_j = \sum_{t=T+1}^{T^*} \beta^{t-T} \hat{y}_{jt}$$

where $\hat{y}_{jt}$ is the estimated income of the alumni of college tier j in year t, T is the age when the student graduates from college, T^* is the retirement age, and β is the discount factor fixed to 0.95.[36] I define $\hat{y}_{jt}$ as the estimated tier-specific annual income at time t, which is predicted using Pooled-OLS estimates of Column (1) in Table 3.[37] As tier 1 is defined to be the top college tier, v decreases in j (i.e., $v_1 > v_2 > ... > v_{J-1} > v_J$).[38]

For the student of household i to obtain prize v_j, her final test score $q_{i,T+1}$ must be above the cutoff for tier j and below the cutoff for the tier $j-1$. In other words, student i is placed in college tier j iff

$$\tilde{Q}_{j-1} > q_{i,T+1} \geq \tilde{Q}_j$$

where $\tilde{Q}_j$ is the cutoff between college tier j and tier $j+1$. The cutoff $\tilde{Q}_j$ is the test score of the N_j^{th} highest student in the sample, where $N_j = \sum_{l=1}^{j} n_l$. Thus, $\{\tilde{Q}_j\}_{j=1}^{J}$

[35] The implicit assumption regarding the bottom tier is that everyone graduates high school. The high school drop-out rate in South Korea is less than 2%.

[36] The average interest rate is around 5% for South Korea in 2010.

[37] I assume no earnings in the college periods.

[38] I confine the prize to pecuniary rewards and rule out other benefits from the model. One might argue that the non-pecuniary value of attending an elite college should be considered part of the reward. However, it is difficult to separately measure the non-pecuniary value of attending better colleges due to data limitations. See Gong *et al.* (2019) for an empirical quantification of the consumption value of college.

is where the competition enters the model. In order for a student to be in tier j or better, she has to be above enough competitors by at least scoring the N_j^{th} highest final test score. As $q_{i,T+1}$ is a function of the effort choice of each household, $\{\tilde{Q}_j\}_{j=1}^J$ is endogenously determined by the competition across households. I assume that each household can correctly predict the final test score cutoffs.[39]

Assumption 1. *Each household correctly guesses the set of final test score cutoffs* $\{\tilde{Q}_j\}_{j=1}^J$.

The facts that (i) college-tier is assigned solely using the final test score $q_{i,T+1}$ and (ii) heterogeneity in college quality is the only variation of the lifetime income in this framework imply that the final test score of a student essentially determines the lifetime income of the student. That is, under the model environment, I assume that there is no extra opportunity to improve one's lifetime income once the college entrance exam is over.

Assumption 2. *The quality of the college one graduates from is the sole determinant of one's lifetime income.*

This is an arguably reasonable assumption under the institutional setting of the interest. I borrow the results of Kang, Kang, and Kim (2023) as supporting evidence for Assumption 2. Using a dataset of one of the big 5 companies of Korea, they find that the effect of college reputation dominates the effect of college GPA on receiving an final interview request from the company. Table 6 presents their estimates. Based on their probit estimates, increases in college GPA by 10 leads to a 6% increase in getting an offer from one of the subsidiary firms of the conglomerate. Meanwhile, graduating from one of the tier 1 colleges increases probability of getting an offer by 23% relative to graduating from a college below tier 3.

Parental Investment: One of the two modes of household effort is parental investment, which is embodied in private tutoring expenditure. Each household chooses the unit price (quality) of tutoring p_{it} and hours (quantity) of tutoring x_{it} to increase the child's test score.[40] The total amount of tutoring expenditure e_{it} is

$$e_{it} = p_{it} x_{it}.$$

[39] I assume away the inconsistency between the guessed cutoffs and the resulting cutoffs because the working sample did not go through significant policy shock that might cause the difference between the guessed and the resulting cutoffs. See Tincani, Kosse and Miglino (2021) for the case that resulting cutoffs significantly deviate from the guessed cutoffs.

[40] To the best of my knowledge, this is the first model to consider the quality and quantity of parental monetary investment simultaneously.

Table 6: Final Interview Request Regression (Probit)

	(1)	(2)	(3)
Tier=1	0.327***	0.737***	0.760***
	(0.038)	(0.103)	(0.113)
Tier=2	0.029	0.140***	0.161***
	(0.044)	(0.030)	(0.035)
Tier=3	-0.123	-0.018	-0.036
	(0.096)	(0.127)	(0.109)
ColGPA	0.006**	0.021***	0.022***
	(0.002)	(0.007)	(0.006)
N	9138	9132	9132
Major	No	Yes	Yes
Company	Yes	No	Yes
ClusterSE	No	Yes	Yes

Source: Confidential data of the conglomerate in late 2010s.
Note: The data are on the applicants to the subsidiary firms of the conglomerate for the latest three years. Other explanatory variables include the subsidiary firm's information and the applicants' information such as college major, age, and gender. The college GPA measured is scaled 0 to 100. ColGPA refers to the average of standardized college GPA.

The tutoring expenditure is constrained under two dimensions. A household cannot spend more tutoring expenditure than its income (i.e., $e_{it} \leq w_{it}$).[41] Also, hours of tutoring are bounded by the child's maximum available time, namely h. While the income constraint is unequal among households, available hours for the child are constant across all households.

Note that the time choice is solely about the time use of the child, which means I do not model the time allocation of parents. The data suggest that, in the secondary school periods, which the model concerns, the majority of parents do not teach their children themselves in middle school periods, and very few parents use their time to teach their child in the high school periods. A few potential explanations can be given for this empirical fact. As students grow, the test materials become more and more difficult to be taught by parents. Also, if there exists an established tutoring market, it would be a safer option for parents in terms of increasing student's test score. Note that the model concerns a regime with a high-stakes standardized test. Full-time tutors would have a comparative advantage in preparing students for exams over parents.

[41] I assume no borrowing.

Child's hours of self-study: Hours of self-study is the other household's mode of effort in the tournament. Each household chooses how much time to allocate for hours of self-study s_{it} which is constrained by h. Unlike parental investment, the resource of self-study does not vary over households as time is equally granted to everyone. The taste for self-study, however, can be considerably heterogeneous across students. For example, some students might prefer studying independently rather than re-learning the same materials from the tutors. Others may prefer reviewing materials with tutors rather than studying alone. I allow the taste for hours of self-study to vary by parental education and the associated shock.

Test Score Production Function: The final test score is the result of accumulated dynamic choices of the household along with its given initial conditions. The initial academic performance q_{i1} is exogenously given and proxied by academic performance in primary school.[42] The three choices affecting test scores are quality of tutoring p_{it}, hours of tutoring x_{it}, and hours of self-study s_{it}. I allow that the quantity (hours) and quality (unit price) of the tutoring activity have different intensities in contributing to the test score production. Denoting κ as intensity of quality of tutoring, the transformed tutoring input is specified as

$$\tilde{e}_{it} = p_{it}^{\kappa} x_{it}^{1-\kappa} \tag{3}$$

where $\kappa < 0.5$ and follows decreasing returns to scale (DRS). The DRS restriction is necessary to prevent the household from choosing an infinitesimal quantity of tutoring hours. If $\kappa \geq 0.5$, the household always has an incentive to make p_{it} greater and x_{it} smaller. The opposite case of a household choosing extremely large hours of tutoring does not occur as available time is restricted by h.

For each time $t = 1, 2, ..., T$, the test score $q_{i,t+1}$ is produced following

$$q_{i,t+1} = g(\theta_t^q, q_{it}, p_{it}, x_{it}, s_{it}, \eta_{it}^q, \lambda_k^q)$$

where η_{it}^q is the test score shock, λ_k^q is the type-specific error, and θ^q is the set of relevant parameters for the test score production. The inclusion of the test score produced in the previous period, q_{it}, allows that the previous test score has its own effects in

[42] Although an earlier measure of the initial child's ability would be more desirable, this is the earliest time period that the academic performance data are available.

generating subsequent test score (Cunha and Heckman 2007). Furthermore, I allow the subset of production parameters to change across periods. The effect of the combined efforts of the household is likely to change over time. As students grow older, the materials taught become more advanced, which makes it harder for students with insufficient background to catch up. Thus, private tutoring expenditure and hours of self-study can be less effective in the later stages of education. In addition, the relative importance of each investment might change over time. For example, the marginal effects of parental investment might increase (decrease) while the effects of self-study decrease (increase) over time. To reflect such changing effects, I let the marginal effects parameters ν_t, δ_{pt}, and δ_{st} be different for each period $t = 1, 2, ...T$.

For estimation, the production function g is a Constant Elasticity of Substitution (CES) production function and is specified as

$$q_{i,t+1} = A_t q_{it}^{\delta_{qt}} \left[\delta_{et}(1 + \tilde{e}_{it})^{\phi} + \delta_{st}(1 + s_{it})^{\phi} \right]^{\frac{\nu_t}{\phi}} \varepsilon_{it}^{q} \tag{4}$$

where A_t is total factor productivity, ν_t is the parameter of marginal effect of the combined effort choices, and ϕ is the parameter governing substitution between tutoring and self-study. The marginal effect of the total effort decision is captured by ν_t, while the relative importance of the tutoring expenditure and hours of self-study are captured by δ_{et} and δ_{st}, respectively. I define ε_{it}^q as a combined shock of λ_k^q and η_{it}^q, which is specified as $\ln \varepsilon_{it}^q = \lambda_k^q + \eta_{it}^q$.

5.3 Household

Flow Utility: The utility function of the unitary household is comprised of three parts: (i) the marginal utility from the household consumption c_{it}, (ii) the marginal disutility from hours spent on tutoring x_{it}, and (iii) the marginal disutility from hours of self-study s_{it}. I denote α_c, α_x, and α_s as taste parameters for household consumption, hours of tutoring, and hours of self-study, respectively. The taste parameters may depend on the fixed characteristics of the household. I assume additive and separable log utility, which is specified as

$$u(c_{it}, x_{it}, s_{it}, \varepsilon_{it}) = \alpha_c \varepsilon_{it}^c \log(c_{it}) + \alpha_x \varepsilon_{it}^x \log(1 + x_{it}) + \alpha_s \varepsilon_{it}^s \log(1 + s_{it}) \tag{5}$$

where ε_{it}^{c} is the shock to the marginal utility from consumption, ε_{it}^{x} is the shock to the disutility from hours of tutoring, ε_{it}^{s} is the shock to the disutility from hours of self-study, and $\varepsilon_{it} = \{\varepsilon_{it}^{c}, \varepsilon_{it}^{x}, \varepsilon_{it}^{s}\}$. The shocks are distributed joint normal and separated into the type-specific and the time-varying components. In particular, I denote λ_k^z and η_{it}^z as type-specific and time-varying components of ε_{it}^z $(z = c, x, s, q)$, respectively. The shocks are decomposed as

$$\begin{pmatrix} \ln \varepsilon_{it}^{c} \\ \ln \varepsilon_{it}^{x} \\ \ln \varepsilon_{it}^{s} \\ \ln \varepsilon_{it}^{q} \end{pmatrix} = \begin{pmatrix} \eta_{it}^{c} \\ \eta_{it}^{x} \\ \eta_{it}^{s} \\ \eta_{it}^{q} \end{pmatrix} + \begin{pmatrix} \lambda_{k}^{c} \\ \lambda_{k}^{x} \\ \lambda_{k}^{s} \\ \lambda_{k}^{q} \end{pmatrix}, \text{ and } \begin{pmatrix} \eta_{it}^{c} \\ \eta_{it}^{x} \\ \eta_{it}^{s} \\ \eta_{it}^{q} \end{pmatrix} \sim N(0, \Omega^{\eta})$$

where Ω^{η} is the covariance matrix for the time-varying shocks.[43] I assume that the correlations between the time-varying shocks η_{it}^z $(z = c, x, s, q)$ are 0.

Note that I do not specify the utility flow from the current test score. Each household is concerned solely about the final outcome, and the role of the current test score is limited to the stepping stone for the final test score. That is, the current test score affects the decision of the household only through the value of the future. The specification of future value is introduced with the recursive formulation at the end of the subsection.

Terminal Value: Expected lifetime income is the terminal value of the model, which drives the dynamic choices of the tournament model. With the tier-specific lifetime income v_j, the expected lifetime income is a weighted sum,

$$\sum_{j=1}^{J} \left\{ \ln(v_j) * Prob(\ln \tilde{Q}_{j-1} \geq \ln q_{i,T+1} \geq \ln \tilde{Q}_j \Big| \Gamma_{iT}) \right\} \tag{6}$$

where $Prob(\ln \tilde{Q}_{j-1} \geq \ln q_{i,T+1} \geq \ln \tilde{Q}_j \Big| \Gamma_{iT})$ is the probability of getting into college tier j. The randomness of the admission probability comes from the test score shock η_{it}^q. Each student would have a different probability of going to a college tier j as they have different characteristics affecting the evolution of the test scores. The disparity among students in terms of going to each college tier leads to the discrepancies in expected lifetime income, which generates the heterogenous incentives among households. The

[43] In modeling the self-study shock, an alternative specification involves assuming that there exists unobserved heterogeneity in terms of the productivity of hours of self-study. Such an assumption, however, is computationally burdensome if the test score production function is CES.

higher expected lifetime income leads to bigger the terminal value of the household, which makes it more appealing for the parents to invest in the child.

The functional form of the expected lifetime income is determined by the test score shock ε^q_{it}. With the log-transformation, the terminal value is specified as

$$\sum_{j=1}^{J}\left\{ \ln(v_j) * Prob(\ln \tilde{Q}_{j-1} \geq \ln q_{i,T+1} \geq \ln \tilde{Q}_j \middle| \Gamma_{iT})\right\}$$
$$= \sum_{j=1}^{J}\left\{ \ln(v_j) * \left\{ F_q(\frac{\ln \tilde{g}_{i,j-1}}{\sigma_q}\middle|\Gamma_{iT}) - F_q(\frac{\ln \tilde{g}_{ij-1}}{\sigma_q}\middle|\Gamma_{iT})\right\}\right\}$$

where $\ln \tilde{g}_{ij}$ is the distance between the deterministic components of log final test score of student i and the log cutoff of the college tier j (i.e. $\ln \bar{g}_{ij} = \ln \tilde{Q}_{j-1} - \ln \widehat{q_{i,T+1}} - \lambda^q_k$), and F is the distribution of η^q_{it}. I assume F follows normal distribution in the spirit of rank-order tournament model (Lazear and Rosen 1981; Han, Kang and Lee 2016; Grau 2018; Tincani, Kosse and Miglino 2021).[44]

Budget and Time Constraints: The choices of the household are restricted by the budget and the time constraints. The budget constraint is given by

$$c_{it} + p_{it}x_{it} \leq w_{it} \tag{7}$$

where w_{it} is household income, and the time constraint is

$$x_{it} + s_{it} \leq h \tag{8}$$

where h is student's disposable time. I define h as the maximum time each student can use every week, which is assumed to be 63.[45]

State Variables: There are observed and unobserved state variables in the dynamic model. The set of observed state variables Z_{it} includes the previous test score q_{it}, parental education m_i, and the complete income stream $\{w_{it}\}^T_{t=1}$. The set of unobserved state variables Ψ_{it} includes the set of unobserved shocks and the type specific

[44] One can also adopt a functional form that η^q_{it} follows Generalized Extreme Value distribution which results in a Tullock (2001) contest.

[45] I assume each student can use 9 hours everyday for non-leisure activities other than hours spent in regular school

heterogeneity. Based on the timeline, the time-varying shock regarding test score is not an unobserved state variables. (i.e., $\Psi_{it} = \{\eta^c_{it}, \eta^x_{it}, \eta^s_{it}, \lambda^c_k, \lambda^x_k, \lambda^s_k, \lambda^q_k\}$).

Information and Uncertainty: I assume a continuum of households. The continuum assumption is useful in that the information of other households can be summed up as a distribution of households.

Assumption 3. *The distribution of household is common knowledge.*

As stated in Assumption 1, each household correctly anticipates the set of college tier cutoffs $\{\tilde{Q}_j\}_{j=1}^J$.[46] They know the distribution of the final test scores in advance and make dynamic choices based upon the perfect guess.

Assumption 4. *Each household knows its complete wage stream.*

There is no uncertainty in the income process. In fact, each household is assumed to know its complete wage stream as the model begins. As this is a markov model, past wages are irrelevant after conditioning on the remaining state variables. As depicted in Figure 7, each household learns about the realization of the consumption shock η^c_{it}, the disutility shock to hours of tutoring η^x_{it}, and disutility shock to hours of self-study η^s_{it} at the beginning of each period. However, it does not know about the test score shock η^q_{it} before it makes a decision. Therefore, it makes a set of choices based on the expectation over η^q_{it}, $\eta^c_{i,t+1}$, $\eta^x_{i,t+1}$, and $\eta^s_{i,t+1}$, conditional on observed state variables and type-specific unobserved heterogeneity.

Household Value Function: Building upon the model components, I describe the value function of the household. As stated earlier, each household chooses the unit price (quality) of tutoring p_{it}, hours of tutoring x_{it}, and hours of self-study s_{it} based on the anticipation of future values. In particular, at each time t, the household i solves

$$V_{it}(Z_{it}, \Psi_{it}) = \max_{p_{it}, x_{it}, s_{it}} \left\{ u(c_{it}, x_{it}, s_{it}, \varepsilon_{it}) + \beta \mathop{E}_{\eta^q_{it}, \eta_{it}} \left[V_{i,t+1}(Z_{i,t+1}, \Psi_{i,t+1} \middle| \Gamma_{it}) \right] \right\}, \tag{9}$$

subject to equation (4) and constraints (7) and (8), where $\Gamma_{it} = \{Z_{it}, \Psi_{it}, \{\bar{Q}_j\}_{j=1}^J\}$ is the set of information before making the decision and $\eta_{it} = \{\eta^c_{it}, \eta^x_{it}, \eta^s_{it}\}$ is the set of unobserved time-varying shocks. Each household faces a tradeoff between current flow utility and future payoffs. Each choice variable incurs costs associated with the

[46] In the static model of Grau (2018), Assumption 3 implies that the tournament participants can correctly guess the cutoffs. In my dynamic model, however, Assumption 3 does not guarantee the perfect foresight due to the presence of future shocks that each individual cannot predict.

choice. In particular, investing more in parental investment (i.e., increasing p_{it} or x_{it}) requires suffering more from the disutility from hours of tutoring and sacrificing current consumption. Spending more time on hours of self-study leads to an increase in the disutility from hours of self-study. This dynamic incentive structure governs the decision of the household.

At the final test stage ($t = T$), where the tournament term appears, the value function is

$$V_{iT}(Z_{iT}, \Psi_{iT}) = \max_{p_{iT}, x_{iT}, s_{iT}} \Big\{ u(c_{iT}, x_{iT}, s_{iT}, \varepsilon_{iT}) \\ + \alpha_v \sum_{j=1}^{J} \ln(v_j) \times Prob(\ln \tilde{Q}_{j-1} \geq \ln q_{i,T+1} \geq \ln \tilde{Q}_j \Big| \Gamma_{iT}) \Big\} \tag{10}$$

where α_v is an altruism parameter. The altruism parameter measures the "exchange rate" between the current household utility and the child's future lifetime income. All-in-all, each household makes a choice between the child's lifetime income and its flow utility. If the marginal value to the household is greater than the marginal loss of flow utility of the household, it exerts more efforts using either parental investment, the child's self-efforts, or both.

5.4 Equilibrium of the Tournament

In this section, I define the dynamic equilibrium of the tournament model. Then I prove the existence of the equilibrium using the Schauder Fixed-Point Theorem (Amir 1996; Fey 2008; Mertens and Judd 2018; Engers, Hartmann and Stern 2022). I define a set $k = \{\{V_t(p_t, x_t, s_t; Z_t, \Psi_t)\}_{t=1}^T, \{\tilde{Q}_j\}_{j=1}^J\}$, where $\{V_t(p_t, x_t, s_t; Z_t, \Psi_t)\}_{t=1}^T$ is a set of value functions that are specified in equations (9) and (10), and $\{\tilde{Q}_j\}_{j=1}^J$ is the set of college-tier cutoffs. I define $\mathcal{K}$ as a set of all possible k.

Definition 1. *Given the set of initial conditions and Assumptions 1, 2, and 3, a Markovian equilibrium of the model is a vector* $k^* = \left\{ \{V_t^*(p_t, x_t, s_t; Z_t, \Psi_t)\}_{t=1}^T, \{\tilde{Q}_j^*\}_{j=1}^J \right\}$*, which is generated by the following process:*

1. I define $\mathcal{K}_a$ as a set of all possible combinations of choice variables $\{p_t, x_t, s_t\}_{t=1}^T$. Given the set of initial conditions $\{q_{i1}, \{w_{it}\}_{t=1}^T, m_i\}$, a mapping $\aleph_a$ maps $\mathcal{K}$ into

$\mathcal{K}_a$ ($\aleph_a : \mathcal{K} \to \mathcal{K}_a$), based on the value functions specified in equations (9) and (10). [47]

2. I define $\mathcal{K}_b$ as possible distributions of the final test score q_{T+1}. A mapping $\aleph_b$ maps $\mathcal{K}_a$ into $\mathcal{K}_b$ ($\aleph_b : \mathcal{K}_a \to \mathcal{K}_b$), based on the test score production function specified in equation (4) .
3. I define $\mathcal{K}_c$ as possible sets of resulting cutoffs $\{\check{Q}_j\}_{j=1}^J$. Given the number of seats for each college tier $\{n_j\}_{j=1}^J$, a mapping $\aleph_c$ maps the distribution of the final test score q_{T+1} and $\{n_j\}_{j=1}^J$ into the set of cutoffs, $\{\check{Q}_j\}_{j=1}^J$ ($\aleph_c : \mathcal{K}_b \to \mathcal{K}_c$). The mapping $\aleph_c$ is based on the rules of college admission.
4. A mapping $\aleph_d$ maps $\{\check{Q}_j\}_{j=1}^J$ into $\mathcal{K}$ ($\aleph_d : \mathcal{K}_c \to \mathcal{K}$).
5. In equilibrium, the set of guessed cutoffs $\{\tilde{Q}_j\}_{j=1}^J$ match the set of realized cutoffs $\{\check{Q}_j\}_{j=1}^J$.

Finally, I define a mapping $\aleph : \mathcal{K} \to \mathcal{K}$. The mapping $\aleph$ is the composition of submappings. In particular,

$$\begin{aligned}\aleph =& \aleph_a \circ \aleph_b \circ \aleph_c \circ \aleph_d \\ =& \aleph_a(\aleph_b(\aleph_c(\aleph_d(k)))).\end{aligned}$$

Lemma 2. *The mapping $\aleph$ is compact*

Proof. [In Appendix A.1] □

Lemma 3. *The mapping $\aleph$ is continuous*

Proof. [In Appendix A.2] □

Theorem 4. *A Markovian equilibrium exists.*

Proof. Previous results establish that $\mathcal{K}$ is a nonempty, compact, and closed subset of a locally convex Hausdorff space. The map $\aleph$ is continuous. Therefore, the set of fixed points of $\aleph$ is nonempty and compact. The mapping satisfies all the requirements of Schauder Fixed-Point Theorem. Hence a fixed point exists.

□

[47] The mapping $\aleph_a$ involves backward recursion.

6. Estimation Strategy

The touranment term at the final period T generates heterogenous incentives for households with different state variabels. These features make the policy function highly non-linear. The choice variables of the structural model involves the pairs of interior-interior, interior-corner, and corner-corner. In addition, the structural model involves endogenous regressors. To this end, I estimate the parameters of the model using Maximum Simulated Likelihood. I describe the likelihood function and discuss the sources of identification underlying the estimation procedure.

6.1 The likelihood function

I denote θ as the set of parameters, Z_{it} as the set of observed state variables, and λ_k as the set of unobserved type-specific characteristics. The individual likelihood contribution of household i is

$$\mathcal{L}_i(\theta|q_{i1}, \{w_{it}\}_{t=1}^{T}, m_i) = \sum_{k=1}^{K} \left\{ \left(\Pi_{t=1}^{T} \mathcal{L}_{it}(\theta|Z_{it}, \lambda_k) \right) \Pr(type = k) \right\} \tag{11}$$

which is conditional on the initial test score q_{i1}, the income stream $\{w_{it}\}_{t=1}^{T}$, and parental education m_i. The time-specific likelihood contribution $\mathcal{L}_{it}(\theta|Z_{it}, \lambda_k)$ can be characterized in four different ways depending on the combination of the tutoring-participation dummy variable d_{it}^x and self-study participation dummy variable d_{it}^s. In particular,

$$\begin{aligned}
\mathcal{L}_{it}(\theta|Z_{it}, \lambda_k) = & \Big[f(p_{it}, x_{it}, s_{it}, q_{it}) \Big]^{d_{it}^x d_{it}^s} \\
& \times \Big[\Pr(p_{it}, x_{it}, s_{it} = 0) \cdot f_{q_{it}}(q_{it}|x_{it}, s_{it} = 0) \Big]^{d_{it}^x (1-d_{it}^s)} \\
& \times \Big[\Pr(x_{it} = 0, s_{it}) \cdot f_{q_{it}}(q_{it}|x_{it} = 0, s_{it}) \Big]^{(1-d_{it}^x) d_{it}^s} \\
& \times \Big[\Pr(x_{it} = 0, s_{it} = 0) \cdot f_{q_{it}}(q_{it}|x_{it} = 0, s_{it} = 0) \Big]^{(1-d_{it}^x)(1-d_{it}^s)}
\end{aligned}$$

where $d_{it}^x = 1$ means that household participates in tutoring at time t, and $d_{it}^s = 1$ means that the student of the household i has non-zero hours of self-study at time t.

The final form of the likelihood function is a sum of the log likelihood contributions,

$$\log \mathcal{L}(\theta) = \sum_{i=1}^{N} \log \mathcal{L}_i(\theta | q_{i1}, \{w_{it}\}_{t=1}^{T}, m_i).$$

6.2 Identification

Parameters of the model can be classified into the productivity parameters associated with the test score function and the taste parameters that directly affect value function. The productivity parameters in the test score production function are identified by the covariation between the subsequent test score $q_{i,t+1}$ and the inputs (q_{it}, p_{it}, x_{it}, and s_{it}). As data of the inputs are available for each period, I can separately identify the productivity parameters for each time t.

The taste parameters α_c, α_x, α_s, and the altruism parameter α_v affect the value function, and do not directly affect the test score function. These parameters are the constants for the likelihood contribution of the corresponding choice variables. I do not differentiate the taste parameters for each period. The element of the covariance matrix of the shocks are identified in maximizing the log-likelihood contribution of the associated shocks.

The exogenous variables in the model are the academic performance in primary school q_{i1}, the parental education m_i, and the complete income stream of parents $\{w_{it}\}_{t=1}^{T}$. The identifying assumption is the time varying shocks are orthogonal to the initial conditions. In particular,

$$\{\eta_{it}^c, \eta_{it}^x, \eta_{it}^s, \eta_{it}^q\}_{t=1}^{T} \perp \left\{ q_{i1}, m_i, \{w_{it}\}_{t=1}^{T} \right\}.$$

7. Estimation Results

7.1 Test score function parameters

Table 7 presents the estimates of test score production function specified in equation (4). The interpretation of the previous test score parameter is the same as the log-log case of a linear regression equation. For example, for $t = 2$, a 1% increase in the previous test score leads to a 0.98% increase in the subsequent test score controlling for other inputs. The marginal effects of the effort parameters ν_t largely decline over time. Especially in the final period, the effect plummets to 0.02. This estimate suggests that, in the final period, the marginal effects of both the parental investment and hours of self-study are significantly lower and it is more difficult to increase a test score with the same amount of monetary or time investments.

Table 7: Parameter Estimates: Test score production function

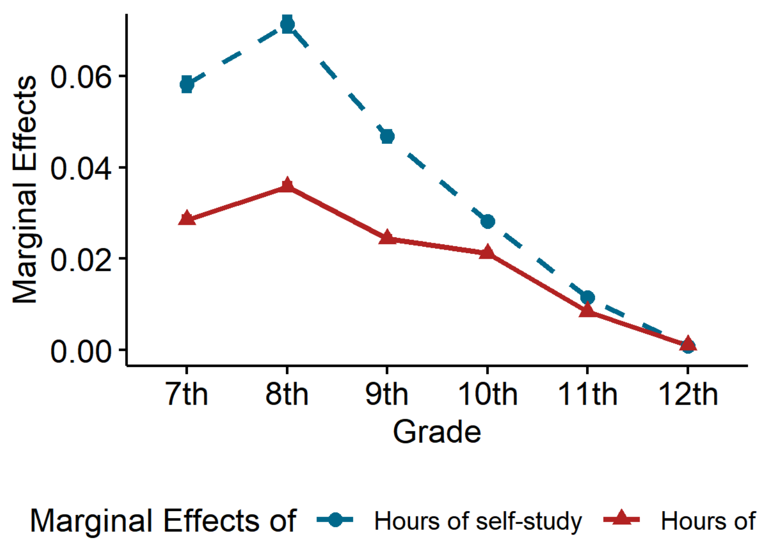
Time-varying Parameters	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$
Previous	0.161	0.972	0.741	0.701	0.445	0.415
Test Score (δ_{qt})	(0.001)	(0.002)	(0.001)	(0.001)	(0.009)	(0.001)
Effort Parameters (ν_t)	0.575	0.710	0.515	0.350	0.152	0.017
	(0.001)	(0.014)	(0.009)	(0.001)	(0.002)	(0.001)
Share of tutoring	0.481	0.485	0.477	0.575	0.555	0.716
Expenditure (δ_{et})	(0.007)	(0.005)	(0.011)	(0.011)	(0.012)	(0.015)
Constants (δ_{0t})	4.031	-1.125	0.540	0.936	0.863	4.484
	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)
Time Invariant Parameters						
Substitution Parameter (ϕ)	0.869					
	(0.001)					
Intensity of Private tutoring	0.154					
Quality (κ)	(0.002)					

Note: Standard errors are computed using delta method and are in parentheses below estimates. Based on the CES test score function, share of hours of self-study is implied by share of tutoring expenditure. (i.e., $\delta_{st} = 1 - \delta_{et}$).

Hours of self-study have stronger average marginal effects on the subsequent test score than hours of tutoring. The marginal effects are computed by the partial derivative of test score $q_{i,t+1}$ with respect to either hours of tutoring (x_{it}) or hours of self-study (s_{it}). I present the average marginal effects, which could vary by households because the calculation of marginal effects involves the data of p_{it}, x_{it}, and s_{it}. Figure 8 presents the comparison of the average marginal effects of hours of self-study and hours of tutoring. In almost all periods, hours of self-study have greater marginal effects than hours of tutoring. Only in the final period, the average marginal effect of hours of tutoring is slightly larger than the average marginal effects of hours of self-study. However, as seen in Figure 8, the difference of the marginal effects in the final period is neither statistically nor economically significant.

The early period investments have delayed effects through evolving test scores. As seen in Figure 8, the marginal effect of hours of self-study are already stronger in the earlier periods. Considering the delayed effects of hours of self-study through the evolving test scores, hours of self-study have a strong effect on the final test score.

Figure 8: Average Marginal Effects of Hours Allocation



Note: This figure presents the average of marginal effects of hours of self-study and hours of tutoring over time. Due to the functional form of the test score function, the marginal effects differ by each individual. The marginal effects are computed using the first order derivative with respect to hours of self-study (s_{it}) or hours of tutoring (x_{it}) and the estimated parameters. The vertical interval at each point indicates the standard deviation of the marginal effects.

The estimate of the substitution parameter shows that the impact of parental investments could be exaggerated in simulating the model if the child's self-study is not incorporated into the mechanism. The hours of self-study and hours of tutoring are nearly perfect substitutes for each other based on the structural estimate. Table 7 includes the estimate of the substitution parameter ϕ, which is approximately 0.88. Suppose a researcher wants to conduct a counterfactual experiment of restricting parental investment using a structural model without other options of investing in the child. Each household must accept the restriction as given. Such an omission of mechanism might result in an exaggeration of the effects of parental investment on outcomes such as intergenerational persistence of earnings. This substitutability plays an important role because the channel of hours of self-study provides a household with a restrictive income constraint an opportunity to exert efforts in the tournament.

7.2 Preference and shock parameters

Table 8: Parameter Estimates: Preference and Shock Parameters

		Estimate	Standard error
Preference Parameters			
Taste for consumption	α_c	0.028	(0.000)
Altruism for the child's future	α_ν	1.040	(0.001)
Disutility from hours of tutoring	α_x	-0.006	(0.000)
Disutility from hours of self-study	α_s	-0.005	(0.006)
Parental education parameters			
disutility from hours of tutoring	τ_x	-1e-05	(2e-05)
disutility from hours of self-study	τ_s	-0.001	(0.002)

(a) Preference parameters

Standard Deviation of		Estimate	Standard Error
Test score shock	$\sigma_{\eta q}$	0.248	(0.000)
Consumption shock	$\sigma_{\eta c}$	0.745	(0.014)
Study disutility shock	$\sigma_{\eta s}$	0.534	(0.001)
Tutoring disutility shock	$\sigma_{\eta x}$	0.521	(0.005)

(b) Shock parameters

Note: Standard errors are computed using delta method and are in parentheses below estimates. $\frac{1}{N \times T \times 6}(\sum logL_i - Jacob) = -0.848$.

Table 8 presents the estimates of the preference parameters and the shock parameters. The preference parameters are components of equations (5), (9), and (10). For the preference parameters, the estimates are relative estimates of the other preference parameters. The altruism parameter is estimated as 1.018. To capture the observed heterogeneity of the household, I allow the preference parameters to vary by parental education. In particular, $\exp(\tau_x D_i^{pedu})$ is multiplied to the disutility from hours of tutoring α_x and $\exp(\tau_s D_i^{pedu})$ is multiplied to the disutility from hours of self-study, where D_i^{pedu} is 1 for household whose average years of parental education is strictly greater than 12. Table 8 (a) includes the estimates of the effects of parental education on the preference parameters. Based on the estimates, parental education alleviates the disutility to hours of self-study. Specifically, a child of a household whose average education of parents is greater than 12 years feels more disutility of study by 0.005. In contrast, the effect of parental education on mitigating disutility from hours of tutoring is not statisticially different from 0.

The estimated standard deviations of unobserved shocks are overall modest, which suggests that the observed characteristics and the structural model capture a considerable proportion of heterogeneity in the data. The unobserved heterogeneity from consumption is considerably different among the different types of households. This could be due to the fact that I do not model the supply side of private tutoring, and the price differences across regions are not captured by the deterministic parts of the model.

7.3 Model Fit

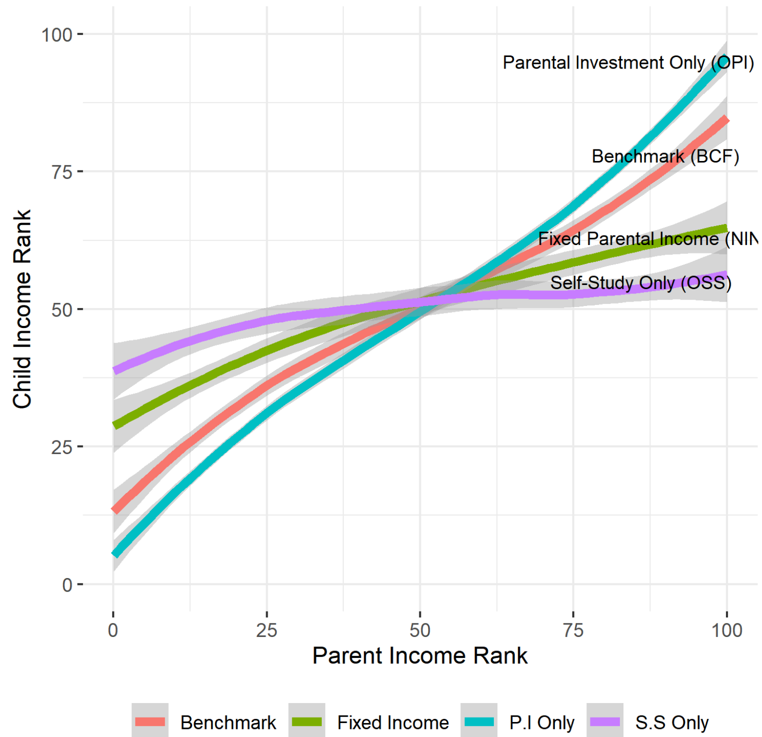
To examine the goodness-of-fit of the structural model, I use a local linear regression estimator to see how well the model prediction $\widehat{y}$ fits the actual data value y, for dependent variables $y = e, p, x, s, q$. [48] Figures F.1 to F.5 present the sample fit of tutoring expenditure, hours of tutoring, quality of tutoring, hours of self-study, and test scores, respectively. Overall, the fits are very good. While the hours of tutoring and the hours of self-study show excellent fits, the quality of tutoring is somewhat overpredicted. Also, while the level of the final test score fits reasonably well, the level of the earlier test scores are also somewhat underpredicted. This is due to the fact that the constants of the test scores are part of the tournament model as specified in the tournament component of equation (10). Thus, the constant of the test scores cannot be separated from the constants of all the other dependent variables' likelihood contributions. Nevertheless, the model fits the distribution of the test scores very well, as shown in Figure F.5a. As the tournament model is about the ranking of the final test score, the distributions of the test scores are the major concern in simulating the model, which is captured by the tournament model.

[48] Appendix G includes the specification of local linear regression used to judge the fit.

8. Counterfactual Analyses

8.1 Decomposition of Intergenerational Persistence of Earnings

Figure 9: Intergenerational Persistence of Earnings by Scenarios



Note: This graph presents local linear curves that fit child income rank and parental income rank under different counterfactual scenarios. Parent Income Rank is the average income over six years. Child Income Rank is computed based on the simulated results of each scenario. BCF is the benchmark counterfactual; NIN is a simulation where each household income is fixed to the mean; OPI is a simulation where each household can use only parental investment; and OSS is a simulation where each household can use only hours of self-study.

The purpose of the quantification exercise is to decompose the role of channels affecting intergenerational persistence of earnings. Using the structural estimates, I simulate the model under the counterfactual environments that help quantify the impact of the relevant channels. As the cutoffs are unknown in this counterfactual case, I simulate the household behavior until their guess of the cutoffs becomes identical with the resulting cutoffs.[49] Each simulation produces a different distribution of the

[49] I use this algorithm again in the college constraint counterfactual analyses described in the subsequent subsection.

final test scores, which leads to a different distribution of the predicted income of the child. I define the predicted income of the child of household i as $childinc_i$.

The rank-rank slope (Chetty *et al.* 2014) is the estimate of the regression equation,

$$R_i = \delta_{01} + \delta_{RR} P_i + \upsilon_i \tag{12}$$

where R_i is the percentile rank of the child income within the generation, and P_i is the rank of the parental income within the generation. [50]

Each counterfactual scenario is defined in the following way.

- BCF is the status quo where the model is simulated without a counterfactual modification.
- OPI is the counterfactual where only parental investment is the means of the tournament model, and hours of self-study are excluded from the choice of the household and fixed to 0.
- OSS is the counterfactual where only child's self-study is the means of the tournament, and parental investment is excluded from the choice and fixed to 0.[52]
- NIN is the counterfactual where all monthly net household income is fixed to the average household income, which is about 4,000,000 KRW (approximately 3000 USD).

Table 9 presents the estimates of the rank-rank slope and the intergenerational elasticity of earnings under five different simulations. The estimated rank-rank slope for the benchmark counterfactual (BCF) is 0.572.[53]

The quantification exercise highlights several findings. First, removing heterogeneity in parental income decreases the rank-rank slope by 43.5%, which can be found in the result of the NIN simulation in Column (5) in Table 9a. The result suggests that heterogeneity in parental income is responsible for a substantial part of the intergenerational persistence of earnings. Second, omitting hours of self-study

[50]Although I present both IGE and the rank-rank slope for each counterfactual simulation, the preferred estimate of intergenerational persistence of earnings is the rank-rank slope. The IGE is sensitive to the ratio of the income inequalities of the two generations.[51] To minimize this issue, the discussion is based on the results of estimates of the rank-rank slope, which is more robust to the difference in the income variance across the generations.

[52]The OSS simulation is equivalent to China's tutoring ban policy in that it completely prohibits private tutoring activities. Gu and Zhang (2024) evaluate the tutoring ban policy using their macroeconomic model.

[53]The estimated IGE with the benchmark model (BCF) is 0.231, which aligns with the intergenerational elasticity of earnings estimates in the literature (Ueda 2013).

leads to significant increases in the intergenerational persistence of earnings. The estimated rank-rank slope under the OPI simulation is 0.782, which is greater than the benchmark simulation by 36.7%, as shown in Column (2) in Table 9a. At the same time, when the channel of parental investment is removed, the rank-rank slope decreases by 66.4%, as shown in Column (3) in Table 9a. Such results suggest that while parental investment reinforces the intergenerational persistence of earnings, the self-study of the child mitigates it. This finding is consistent with the substitution parameter estimate in Table 7, which suggests that self-study serves as a substitute for parental investment.[54]

Table 9: Intergenerational Persistence of Earnings under the Counterfactual Simulations

(a) Rank-Rank Slope Estimates

	(1) BCF	(2) OPI	(3) OSS	(4) NST	(5) NIN	(6) NED
pincprctile	0.572***	0.782***	0.192***	0.572***	0.323***	0.573***
	(0.019)	(0.015)	(0.023)	(0.019)	(0.022)	(0.019)
R-squared	0.327	0.611	0.037	0.327	0.105	0.328

(b) Intergenerational Elasticity of Earnings Estimates

	(1) BCF	(2) OPI	(3) OSS	(4) NST	(5) NIN	(6) NED
log(hhinc)	0.231***	0.254***	0.074***	0.231***	0.166***	0.233***
	(0.008)	(0.005)	(0.008)	(0.008)	(0.011)	(0.008)
R-squared	0.302	0.624	0.051	0.303	0.110	0.305

Note: Table (a) and (b) provide the estimates of rank-rank slope and intergenerational elasticity of earnings, respectively. BCF is a benchmark counterfactual. NIN is a simulation where household income is fixed to the mean. OPI is a simulation where household can only use parental investment. OSS is a simulation where household can only use hours of self-study.

8.2 Relaxing College Constraints

Leveraging the number of seats and the tier-specific returns in the tournament model, I simulate the demand for parental investment under two counterfactual scenarios: (i)

[54] Appendix G provides results fixing other initial conditions.

a 50% increase in the number of seats in Tier 1 colleges, and (ii) a 50% reduction in cohort size.

Table 10: Description of College Constraint Simulation

	Tier 1		Tier 2		Tier 3		Tier 4		Tier 5	
	# Seats	Prize	# Seats	Prize	# Seats	Prize	# Seats	Prize	# Seats	Prize
Status Quo	n_1	v_1	n_2	v_2	n_3	v_3	n_4	v_4	n_5	v_5
Simulation I	$1.5n_1$	$v_1(1-R)$	n_2	$v_2 - \frac{O_1}{2n_2}$	n_3	$v_3 - \frac{O_1}{4n_3}$	n_4	$v_4 - \frac{O_1}{4n_4}$	n_5	v_5
Simulation II	$2n_1$	$v_1(1-R)$	$2n_2$	$v_2(1-R)$	$2n_3$	$v_3(1-R)$	$2n_4$	$v_4(1-R)$	n_5	$v_5 - \frac{O_2}{n_5}$

Note: For each simulation, the overflow of lifetime income is defined as $O_1 = v_1(1-r_1)(n_1+n_1') + v_1 n_1$ and $O_2 = 2(1-R)\sum_{j=1}^{4} n_j v_j - \sum_{j=1}^{4} n_j v_j$

An increase in the number of seats across college tiers can either raise or lower the overall demand for private tutoring. Expanding the number of seats makes private tutoring more valuable for students who previously had a lower probability of entering a higher tier, as their chances of admission increase. On the other hand, if the increase in seats sufficiently boosts a student's chances of entering into the upper tier, the value of private tutoring may decline. The average expenditure on private tutoring will decrease if the second effect outweighs the first, and increase if the first effect is stronger.

To account for the zero-sum game nature of the college admission competition, it is necessary to make assumptions about (i) the trade-offs between an additional number of tier-specific seats and the marginal decrease in tier-specific returns, and (ii) how to deal with the increased total lifetime income relative to the status quo. I define R_j as the loss of return for Tier j. Specifically, if there is an n_j' increase for the seat numbers of Tier j, the sum of future lifetime income of Tier j alumni is $(n_j + n_j')(v_j(1-R_j))$. The "overflow" of the aggregated lifetime income caused by changes in all tiers, is defined as

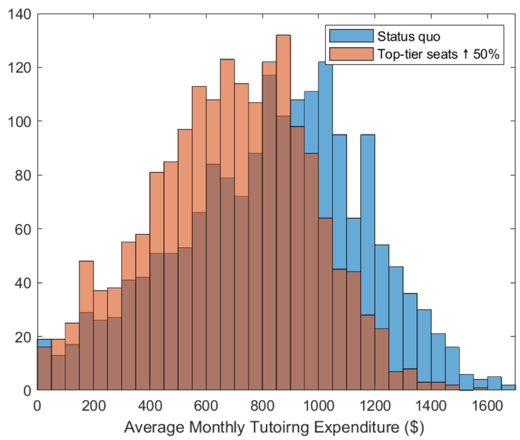
$$O = \sum_{j=1}^{J} v_j(1-R_j)(n_j+n_j') - \sum_{j=1}^{J} v_j n_j.$$

Two counterfactual scenarios are characterized by different combinations of n_j', R_j, and O, which are described in Table 10.

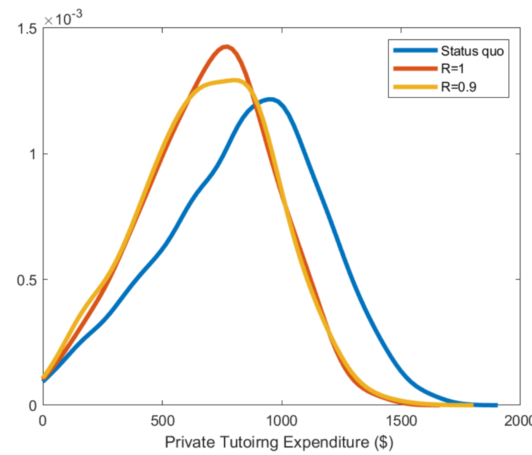
Simulation I: Relaxation of the Elite College Constraint In this simulation, I increase the number of seats in Tier 1 by 50% fixing the number of seats of other tiers ($n_1' = 1.5$ and $n_j = 0$ for all other js). Only the Tier 1 burdens the cost of increasing seats, thus the tier-specific lifetime income is decreased to $v_1(1 - R_1)$. Then, the overflow of Simulation I is defined as $O_1 = v_1(1 - r_1)(n_1 + n_1') - v_1 n_1$. Half of the overflow of the future lifetime income is subtracted from the future lifetime income of Tier 2, and the rest is equally burdened by Tier 3 and 4. That is, v_2 is changed to $v_2 - \frac{O_1}{2n_2}$, and v_3 and v_4 are changed to $v_3 - \frac{O_1}{4n_3}$ and $v_4 - \frac{O_1}{4n_4}$, respectively. Tier 5 stays the same.

Figure 10: Simulation I: Elite College Constraints

(a) Simulation with $R_1 = 0.9$



(b) Simulation with changing R_1



Note: Graph (a) depicts the distribution of the simulated monthly private tutoring expenditure when the assumed return R_1 is 0.9, compared to the status quo. Graph (b) shows the distribution of the simulated monthly private tutoring expenditure for $R_1 = 1, 0.9$ and the status quo.

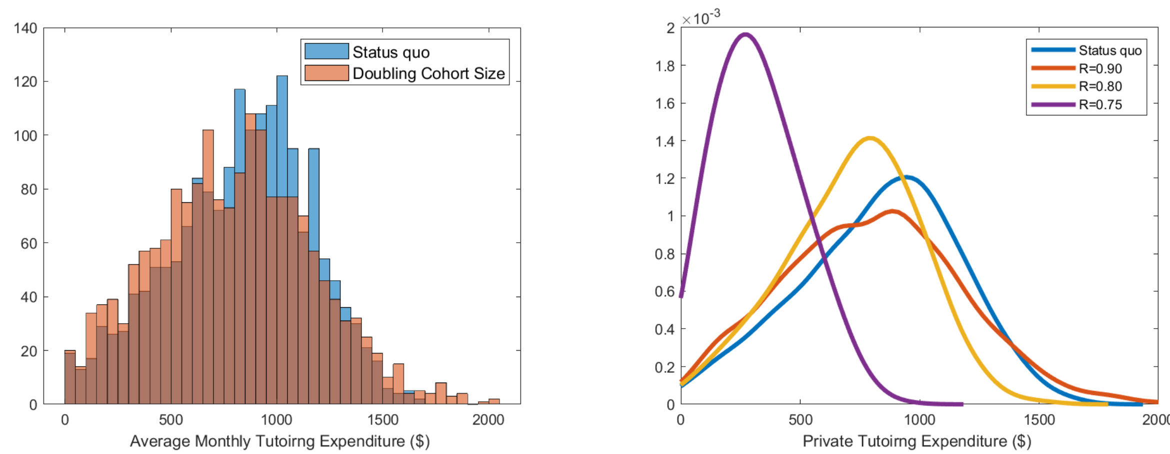
Table 11 presents the results of the college constraint simulation. An increase in the number of Tier 1 seats by 50% leads to a decrease in the average private tutoring expenditure by 15.5% to 16.3% when R_j values are assumed to be 1 and 0.9, respectively. Figure 10a shows the distribution of the average monthly private tutoring expenditure for the expansion scenario and the status quo. Compared to the status quo, the distribution shifts to the left. These results suggest that expanding the number of seats in the elite college overall leads to a decrease in the average private tutoring expenditure.

Simulation II: Cohort Size Reduction Motivated by the cohort reduction in South Korea, I simulate changes in the amount of parental investment when the size of the cohort decreases by half. The effects of the cohort changes are reflected through the increased seats of college tiers. As there are half of the competitors relative to the unchanged number of college seats, it is equivalent to that the number of seats for each tier doubles.

In this simulation, I increase the number of seats in Tier 1 to 4 by 100% ($n_j' = n_j$ and $n_5 = 0$). Tier 1 to 4 burdens the cost of increasing seats, thus the tier-specific lifetime income is decreased to $v_j(1-R)$ for $j = 1, 2, 3, 4$. Then, the overflow of Simulation II is defined as $O_2 = 2(1-R)\sum_{j=1}^{4} n_j v_j - \sum_{j=1}^{4} n_j v_j$. The overflow of the future lifetime income is subtracted from the future lifetime income of Tier 5. That is, v_5 is changed to $v_5 - \frac{O_2}{n_5}$.

Figure 11: Simulation II: Shrinking Cohort

(a) Simulation with $R = 0.9$ (b) Simulation with changing R



Note: Graph (a) depicts the distribution of the simulated monthly private tutoring expenditure when the assumed return R_1 is 0.75, compared to the status quo. Graph (b) shows the distribution of the simulated monthly private tutoring expenditure for $R_1 = 0.9, 0.8, 0.75$ and the status quo.

A decrease in cohort size does not necessarily lead to a reduction in average private tutoring expenditure unless there is a significant decrease in the returns from graduating from a better-tier college. Figure 11b presents the average private tutoring expenditure when the seat-to-cohort ratio doubles for $R = 0.90, 0.80, 0.75$. When there is only a 5% decrease in R, there is a substantial increase in average private tutoring expenditure. Figure 11a presents the simulation results for $R = 0.90$. Compared to the status quo, total private tutoring decreases only by 4%. Private tutoring expenditure significantly decreases only when there are significant decrease in alumni for all tiers. When there are 25% decrease in all college tiers, private tutoring significantly drops.

Table 11: Changes in choice variables under the college-constraint simulation

		Tutoring Expenditure	Hours of Self-study
Status quo		100	100
Simulation I	Increasing Elite College Seats		
	R=1	84.5	85.1
	R=0.9	83.7	99.5
Simulation II	Shrinking Cohort Size		
	R=0.9	96.0	141.1
	R=0.8	86.4	90.3
	R=0.75	39.1	61.6

Note: Both tutoring expenditure and hours of self-study are the aggregated values through all time periods. I standardize the value by setting the status quo values as 100.

9. Conclusion

I develop and estimate a dynamic tournament model of college admissions in which each household utilizes both private tutoring expenditures and the child's hours of self-study. [**State the key mechanism of the model**]

Using the estimated model, I quantify the roles of private tutoring expenditures, hours of self-study, and other household characteristics. I find that heterogeneity in parental income during adolescence accounts for 46

I also use the model to evaluate the impact of relaxing college admission constraints. Expanding the number of seats in elite colleges leads to a 15% decrease in average private tutoring expenditures. The estimated model shows that a sharp decline in cohort size does not necessarily reduce the demand for private tutoring, unless accompanied by a significant decline in the returns to graduating from an upper-tier college.

The findings of this paper suggest two avenues for future research. First, this paper does not allow the possibility of wealth transmission within the household. As Becker and Tomes (1979) suggest, the transmission of capital can be an alternative way of inheriting the income of the parents, especially when the child does not perform well academically. Incorporating the channel of capital transmission within a family requires at least decent data on wealth for more than one generation, which is not an easy data requirement. Second, this paper does not allow for unobserved heterogeneity of labor income conditional on college quality, mainly due to the limitations of micro-data. An efficiency analysis on the rat-race nature of the college admission competition would be feasible with the addition of the channel. I leave this for future research.

References

Agostinelli, F. (2018) Investing in Children's Skills: An Equilibrium Analysis of Social Interactions and Parental Investments, *Unpublished Manuscript*.

Agostinelli, F., Doepke, M., Sorrenti, G. and Zilibotti, F. (2023) It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects, *Journal of Political Economy*.

Agostinelli, F. and Wiswall, M. (2016) Estimating the Technology of Children's Skill Formation, Working Paper 22442, National Bureau of Economic Research.

Amir, R. (1996) Continuous stochastic games of capital accumulation with convex transitions, *Games and Economic Behavior*, **15**, 111–131, publisher: Elsevier.

Anelli, M. (2020) The Returns to Elite University Education: a Quasi-Experimental Analysis, *Journal of the European Economic Association*, **18**, 2824–2868, publisher: Oxford Academic.

Arcidiacono, P., Aucejo, E. M. and Hotz, V. J. (2016) University differences in the graduation of minorities in STEM fields: Evidence from California, *American Economic Review*, **106**, 525–62.

Bastedo, M. (2021) Holistic admissions as a global phenomenon, in *Higher Education in the Next Decade*, Brill, pp. 91–114.

Becker, G. S. and Tomes, N. (1979) An equilibrium theory of the distribution of income and intergenerational mobility, *Journal of political Economy*, **87**, 1153–1189.

Bodoh-Creed, A. and Hickman, B. R. (2019) Pre-College Human Capital Investments and Affirmative Action: A Structural Policy Analysis of US College Admissions, *Working Paper*, publisher: Elsevier BV.

Bolt, U., French, E., Maccuish, J. H. and O'Dea, C. (2021a) The intergenerational elasticity of earnings: exploring the mechanisms.

Bolt, U., French, E., Maccuish, J. H. and O'Dea, C. (2021b) Intergenerational altruism and transfers of time and money: A life-cycle perspective, *Unpublished Manuscript*.

Borsch-Supan, A., Hajivassiliou, V. and Kotlikoff, L. J. (1992) Health, Children, and Elderly Living Arrangements: A Multiperiod-Multinomial Probit Model with Unobserved Heterogeneity and Autocorrelated Errors, in *Topics in the Economics of Aging*, University of Chicago Press, pp. 79–108.

Boucher, V., Bello, C. L. D., Panebianco, F., Verdier, T. and Zenou, Y. (2022) Education Transmission and Network Formation, *Journal of Labor Economics*, p. 718981.

Bound, J. and Turner, S. (2007) Cohort crowding: How resources affect collegiate attainment, *Journal of public Economics*, **91**, 877–899.

Bray, M. (2021) Shadow Education in Europe: Growing Prevalence, Underlying Forces, and Policy Implications, *ECNU Review of Education*, **4**, 442–475.

Bray, M. (2022) Shadow Education in Asia and the Pacific: Features and Implications of Private Supplementary Tutoring, in *International Handbook on Education Development in Asia-Pacific* (Eds.) W. O. Lee, P. Brown, A. L. Goodwin and A. Green, Springer Nature Singapore, Singapore, pp. 1–23.

Bray, T. (1999) *The shadow education system: Private tutoring and its implications for planners*, UNESCO International Institute for Educational Planning.

Caucutt, E. M. and Lochner, L. (2020) Early and Late Human Capital Investments, Borrowing Constraints, and the Family, *Journal of Political Economy*, **128**, 1065–1147.

Chen, J. and Shum, M. (2010) Estimating a tournament model of intra-firm wage differentials, *Journal of Econometrics*, **155**, 39–55.

Cheo, R. and Quah, E. (2005) Mothers, Maids and Tutors: An Empirical Evaluation of their Effect on Children's Academic Grades in Singapore, *Education Economics*, **13**, 269–285.

Chetty, R., Hendren, N., Kline, P. and Saez, E. (2014) Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *, *The Quarterly Journal of Economics*, **129**, 1553–1623.

Cooper, H., Robinson, J. C. and Patall, E. A. (2006) Does Homework Improve Academic Achievement? A Synthesis of Research, 1987–2003, *Review of Educational Research*, **76**, 1–62.

Cunha, F. and Heckman, J. (2007) The Technology of Skill Formation, *American Economic Review*, **97**, 31–47.

Cunha, F., Heckman, J. J. and Schennach, S. M. (2010) Estimating the technology of cognitive and noncognitive skill formation, *Econometrica*, **78**, 883–931.

Dang, H.-A. (2007) The determinants and impact of private tutoring classes in Vietnam, *Economics of Education Review*, **26**, 683–698.

Daruich, D. (2022) The Macroeconomic Consequences of Early Childhood Development Policies.

De Groote, O. (2023) Dynamic effort choice in high school: Costs and benefits of an academic track, *Journal of Labor Economics*.

Del Boca, D., Flinn, C. and Wiswall, M. (2014) Household choices and child development, *Review of Economic Studies*, **81**, 137–185.

Del Boca, D., Flinn, C. J., Verriest, E. and Wiswall, M. J. (2023) Parenting with patience: Parental incentives and child development, *Unpublished Manuscript*.

Del Boca, D., Monfardini, C. and Nicoletti, C. (2017) Parental and child time investments and the cognitive development of adolescents, *Journal of Labor Economics*, **35**, 565–608.

Doepke, M. and Zilibotti, F. (2017) Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission, *Econometrica*, **85**, 1331–1371.

Engers, M., Hartmann, M. and Stern, S. (2022) A Dynamic Model of Equilibrium with Private Information, *Unpublished Manuscript*.

Epple, D., Figlio, D. and Romano, R. (2004) Competition between private and public schools: testing stratification and pricing predictions, *Journal of public Economics*, **88**, 1215–1245, publisher: Elsevier.

Epple, D. and Romano, R. (2008) Educational vouchers and cream skimming, *International Economic Review*, **49**, 1395–1435, publisher: Wiley Online Library.

Epple, D. and Romano, R. E. (1998) Competition between private and public schools, vouchers, and peer-group effects, *American Economic Review*, pp. 33–62, publisher: JSTOR.

Epple, D., Romano, R. E. and Urquiola, M. (2017) School vouchers: A survey of the economics literature, *Journal of Economic Literature*, **55**, 441–92.

Epple, D., Romano, R. E. and Urquiola, M. (2021) Is Education Different? A Review of the Voucher Literature and Lessons for Implementation, *The Routledge Handbook of the Economics of Education*, pp. 150–168.

Fey, M. (2008) Rent-seeking contests with incomplete information, *Public Choice*, **135**, 225–236, publisher: Springer.

Fu, C. and Mehta, N. (2018) Ability tracking, school and parental effort, and student achievement: A structural model and estimation, *Journal of Labor Economics*, **36**, 923–979.

Gayle, G.-L., Golan, L. and Soytas, M. A. (2022) What is the source of the intergenerational correlation in earnings?, *Journal of Monetary Economics*.

Geweke, J. (1988) Antithetic acceleration of Monte Carlo integration in Bayesian inference, *Journal of Econometrics*, **38**, 73–89.

Gong, Y., Lochner, L., Stinebrickner, R. and Stinebrickner, T. (2019) The Consumption Value of College, *NBER Working Paper Series*.

Grau, N. (2018) The impact of college admissions policies on the academic effort of high school students, *Economics of Education Review*, **65**, 58–92.

Gu, S. and Zhang, L. (2024) The macroeconomic consequences of competition for college admissions, *Working Paper*.

Guo, N. and Leung, C. K. Y. (2021) Do elite colleges matter? The impact on entrepreneurship decisions and career dynamics, *Quantitative Economics*, **12**, 1347–1397.

Guryan, J., Hurst, E. and Kearney, M. (2008) Parental Education and Parental Time with Children, *Journal of Economic Perspectives*, **22**, 23–46.

Haider, S. J. (2001) Earnings instability and earnings inequality of males in the United States: 1967–1991, *Journal of labor Economics*, **19**, 799–836, publisher: The University of Chicago Press.

Han, C., Kang, C. and Lee, S. H. (2016) Measuring Effort Incentives in a Tournament with many Participants: Theory and Application, *Economic Inquiry*, **54**, 1240–1250.

Hoekstra, M. (2009) The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach, *The Review of Economics and Statistics*, **91**, 717–724.

Hof, S. (2014) Does private tutoring work? The effectiveness of private tutoring: A nonparametric bounds analysis, *Education Economics*, **22**, 347–366.

Jia, R. and Li, H. (2021) Just above the exam cutoff score: Elite college admission and wages in China, *Journal of Public Economics*, **196**, 104371, publisher: Elsevier.

Jinhak (2022) www.jinhak.com.

Kang, C. (2007) Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea, *Journal of Urban Economics*, **61**, 458–495.

Kang, C. and Park, Y. (2021) Private Tutoring and Distribution of Student Academic Outcomes: An Implication of the Presence of Private Tutoring for Educational Inequality, *The Korean Economic Review*, **37**, 287–326.

Kim, S. (2014) Effects of academic cliques on the first job offers of college graduates, *Journal of Educational Studies*, **45**.

Kim, S. and Lee, J. (2010) Private Tutoring and Demand for Education in South Korea, *Economic Development and Cultural Change*, **58**, 259–296.

Kim, S. and Lee, J.-H. (2006) Changing Facets of Korean Higher Education: Market Competition and the Role of the State, *Higher Education*, **52**, 557–587.

Kim, S., Tertilt, M. and Yum, M. (2024) Status Externalities in Education and Low Birth Rates in Korea, *American Economic Review*, **Forthcoming**.

Lavy, V., Palumbo, M. and Stern, S. (1998) Simulation of multinomial probit probabilities and imputation of missing data, *Advances in Econometrics*, **13**, 145–180, publisher: JAI PRESS INC.

Lazear, E. (1977) Education: consumption or production?, *Journal of Political Economy*, **85**, 569–597.

Lazear, E. P. and Rosen, S. (1981) Rank-Order Tournaments as Optimum Labor Contracts, *Journal of Political Economy*, **89**, 841–864, publisher: University of Chicago Press.

Lee, J. and Koh, Y. (2023) University ranking and the lifetime wage gap, *The Korean Journal of Economic Studies*, **71.2**.

Lee, S. Y. T. and Seshadri, A. (2019) On the Intergenerational Transmission of Economic Status, *Journal of Political Economy*, **127**, 855–921.

MacLeod, W. B., Riehl, E., Saavedra, J. E. and Urquiola, M. (2017) The Big Sort: College Reputation and Labor Market Outcomes, *American Economic Journal: Applied Economics*, **9**, 223–261, publisher: American Economic Association.

Mertens, T. M. and Judd, K. L. (2018) Solving an incomplete markets model with a large cross-section of agents, *Journal of Economic Dynamics and Control*, **91**, 349–368.

Nam, K. (2007) Time Series Trend of the Scale of Private Supplementary Tutoring, *The Journal of Economics and Finance of Education*, **16**, 57–79.

Nybom, M. (2017) The distribution of lifetime earnings returns to college, *Journal of Labor Economics*, **35**, 903–952, publisher: University of Chicago Press Chicago, IL.

Ono, H. (2007) Does examination hell pay off ? A cost–benefit analysis of "ronin" and college education in Japan, *Economics of Education Review*, **26**, 271–284.

Park, H., Behrman, J. R. and Choi, J. (2013) Causal Effects of Single-Sex Schools on College Entrance Exams and College Attendance: Random Assignment in Seoul High Schools, *Demography*, **50**, 447–469.

Park, H., Behrman, J. R. and Choi, J. (2018) Do single-sex schools enhance students' STEM (science, technology, engineering, and mathematics) outcomes?, *Economics of Education Review*, **62**, 35–47.

Ramey, G. and Ramey, V. A. (2010) The Rug Rat Race, *Brookings Papers on Economic Activity*, **41**, 129–199.

Ryu, D. and Kang, C. (2013) Do Private Tutoring Expenditures Raise Academic Performance? Evidence from Middle School Students in South Korea: Do Tutoring Expenditures Raise Performance?, *Asian Economic Journal*, **27**, 59–83.

Sekhri, S. (2020) Prestige Matters: Wage Premium and Value Addition in Elite Colleges, *American Economic Journal: Applied Economics*, **12**, 207–225.

Stern, S. (1997) Simulation-Based Estimation, *Journal of Economic Literature*, **35**, 2006–2039.

Stevenson, D. L. and Baker, D. P. (1992) Shadow Education and Allocation in Formal Schooling: Transition to University in Japan, *American Journal of Sociology*, **97**, 1639–1657.

Stinebrickner, R. and Stinebrickner, T. R. (2004) Time-use and college outcomes, *Journal of Econometrics*, **121**, 243–269.

Stinebrickner, R. and Stinebrickner, T. R. (2008) The causal effect of studying on academic performance, *The BE Journal of Economic Analysis & Policy*, **8**.

Stinebrickner, T. R. (1999) Estimation of a duration model in the presence of missing data, *Review of Economics and Statistics*, **81**, 529–542.

Sullivan, P. (2009) Estimation of an Occupational Choice Model when Occupations are Misclassified, *Journal of Human Resources*, **44**, 495–535.

Tamborini, C. R., Kim, C. and Sakamoto, A. (2015) Education and Lifetime Earnings in the United States, *Demography*, **52**, 1383–1407.

Tansel, A. and Bircan Bodur, F. (2005) Effect of private tutoring on university entrance examination performance in Turkey.

Tincani, M., Kosse, F. and Miglino, E. (2021) Subjective beliefs and inclusion policies: Evidence from college admissions.

Todd, P. and Wolpin, K. I. (2018) Accounting for mathematics performance of high school students in mexico: Estimating a coordination game in the classroom, *Journal of Political Economy*, **126**, 2608–2650.

Tullock, G. (2001) Efficient rent seeking, in *Efficient rent-seeking*, Springer, pp. 3–16.

Ueda, A. (2013) Intergenerational mobility of earnings in South Korea, *Journal of Asian Economics*, **27**, 33–41.

Vukina, T. and Zheng, X. (2007) Structural Estimation of Rank-Order Tournament Games with Private Information, *American Journal of Agricultural Economics*, **89**, 651–664.

Vukina, T. and Zheng, X. (2011) Homogenous and Heterogenous Contestants in Piece Rate Tournaments: Theory and Empirical Analysis, *Journal of Business & Economic Statistics*, **29**, 506–517.

Yum, M. (2022) Parental time investment and intergenerational mobility, *Unpublished Manuscript*.

Zimmerman, S. D. (2019) Elite Colleges and Upward Mobility to Top Jobs and Top Incomes, *American Economic Review*, **109**, 1–47.

Appendix A

Appendix A.1

Proof of Lemma 2: Compactness

Proof. A value function is the sum of flow utility and the discounted future value. The flow utility term $u(c_{it}, x_{it}, s_{it}, \varepsilon_{it})$ is monotone in its arguments. Also, u is defined at the lower and upper bounds of c_{it}, x_{it}, s_{it}. Thus, $u(c_{it}, x_{it}, s_{it}, \varepsilon_{it})$ is closed and bounded. The expected future value EV_{t+1} is closed and bounded. For the final period, the tournament term described in equation (10) is closed and bounded because (i) the v_j term is finite and greater than 0, and (ii) $Prob(\ln \tilde{Q}_{j-1} \geq \ln q_{i,T+1} \geq \ln \tilde{Q}_j \Big| \Gamma_{iT}) \in [0, 1]$. Therefore, the choice-specific value function of the final period, $V_{it}(Z_{it}, \Psi_{it})$ is closed and bounded for $t = T$. Following the backward recursion, $V_{it}(Z_{it}, \Psi_{it})$ is closed and bounded. □

Appendix A.2

Proof of Lemma 3: Continuity

Proof. I start by showing that the value function V_{it} is continuous. To show V_{it} is continuous, It suffices to show that both $u(c_{it}, x_{it}, s_{it}, \varepsilon_{it})$ and $\int_\eta V_{t+1}(Z_{t+1}, \mathit{\Psi}_{t+1}) f(\eta) d\eta$ are continuous. □

- Start from the final period and show that the final term is continous: Bounded right hand side. Left hand side is continous in its arguemnt. Use Dominated Convergence Theorem.
- Previous period same
- Then move on to the continuity of the mapping

Proof. One way to show the continuity of the expected value function is show that it is sequentially continuous. For any sequence of the arguments of the value function,

$$\{Z_t^n, \mathit{\Psi}_t^n\} \to \{Z_t^0, \mathit{\Psi}_t^0\},$$

we have

$$\int_{\Psi} V_{t+1}(Z^n_{t+1}, \Psi^n_{t+1}) d\Psi \to \int_{\Psi} V_{t+1}(Z^0_{t+1}, \Psi^0_{t+1}) d\Psi.$$

Recall that $\Psi_{t+1} = \{\eta^c_{i,t+1}, \eta^x_{i,t+1}, \eta^s_{i,t+1}, \eta^q_{it}\}$. As the expectation of the unobserved shocks has finite expectation, the expected value term has finite expectation as well. $\int_{\Psi} V_{t+1}(Z^0_{t+1}, \Psi^0_{t+1}) d\Psi$ is continuous by the Dominated Convergence. Each supmapping is continuous as its elements are continuous. As each submapping is continuous. By induction, the composition of mapping is continuous. Therefore, $\aleph$ is continuous. $\square$

Appendix B

Colleges for each college tier

	List of the member colleges
First Tier	Seoul National, Yonsei, Korea , Sogang, SKKU, Hanyang, KAIST, Pusan, Ewha, Postech
Second Tier	Choongang, Kyunghee, HUFS, University of Seoul, KU, Dongguk, Kyongpook, Sookmyung, Ajou, Honggik, Inha, Hangkong, Kookmin, Soongsil, Sejong, Dankook, Kwangwoon, Cheonnam, Seoul Industrial University
Third Tier	Myongji, Sangmyeong, Catholic, Choongam, Choongbook, Seongshin, Kyeongki Kyongwon, Deoksong women, Dongdeok women, Dong-A, Bookyeong
Fourth Tier	The rest of the 2 year colleges
Fifth Tier	High school graduates

Appendix C

Korean Labor Income and Panel Study

The college tier-specific lifetime income is inferred from the Korean Labor Income and Panel Study (KLIPS). KLIPS is a panel dataset of representative Korean households from 1998 to 2021. The dataset provides information on which college each worker graduated from, her major, income history, and other demographic characteristics. Using KLIPS, I generate the average lifetime income of the alumni for each college tier and complement the labor market information of KELS. In fact, KELS also provides individual information on the early labor market outcomes of the sample. Still, both the income data and the participation data have a substantial proportion of missing data compared to KLIPS. Employing KLIPS is more useful in predicting alumni's lifetime income as it contains data on workers of age between 20 and 65.[55]

Selection Rules and Effects

Table C.1: Data Selection

Original Sample Size	6,908
Cause of Exclusion	
Missing CSAT	3,310
Missing at least one period of Income	1,576
Zero Income	16
Missing Initial Test Score	40
Missing one of the parental education	59
Tutoring Expenditure greater than income	6
All choice variables missing	62
Implausible unit price of tutoring	47
Remaining Sample Size	1,792

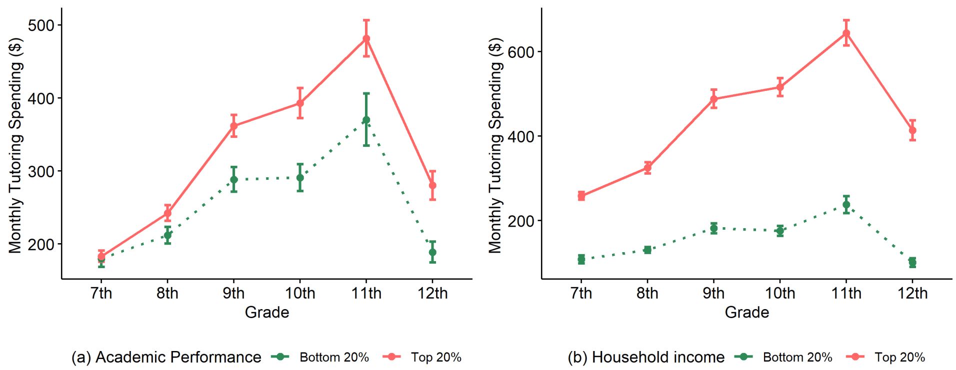
Additional Descriptive Evidence: Dynamic Allocation of Effort choices

The initial conditions of the household persistently affect the parental investment decisions throughout the secondary school periods. Figure A.1 presents changes in the average hours of tutoring expenditure over time differentiated by two of households' pre-conditions: the initial academic performance and the initial parents' income. To see how these initial conditions affect the investment decision of households, I present the changes in average tutoring expenditure of two sub-groups: the top 20% and the bottom 20% of the ordered initial conditions. In particular, the solid lines of Figure A.1 connect the average tutoring expenditure of the highest 20% of households classified by the two initial conditions. In the same manner, the dotted lines connect the average tutoring expenditure of the bottom 20% of households. Figure A.1 (a) shows the increasing gap in tutoring expenditure between those who were in the top 20% of the test score in 6th grade and who were in the bottom 20% of the test score in 6th grade over time. In 7th grade, there is no significant difference between the

[55]The Lifelong Career Survey (LCS) by the Korea Research Institute for Vocational Education & Training (KRIVET) is an alternative dataset that could be used to generate the proxy of the prize of the tournament (Han, Kang and Lee 2016). For the purpose of this paper, KLIPS is preferred because it can recover the age-specific income profile.

two groups in terms of tutoring expenditure. From 8th grade on, there is an evident gap in tutoring expenditure between these two groups. Based on the average tutoring expenditure in 12th grade, students who were in the top 20% of the test score in 7th grade increased their tutoring expenditure compared to when they were in 7th grade. In comparison, the students who were in the lowest 20% of the test score in 7th grade decreased their tutoring expenditure compared to when they were in 7th grade. Figure A.1 (b) presents the average tutoring expenditure of high-income and low-income groups. The gap is significant in 7th grade and becomes greater over time. On average, high-income households' tutoring expenditure increases in 12th grade compared to when the students were in 7th grade. On the other hand, low-income households' tutoring expenditure decreases on average compared to when the students were in 7th grade.

Figure A.1: Dynamic of Parental Investment by Initial Conditions



Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: In this figure, academic performance is measured in 6th grade and used for subsequent years. I only include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, and household income.

Liftime Income Estimates

	(1) Pooled OLS	(2) Pooled OLS	(3) Pooled OLS	(4) RE	(5) RE	(6) RE
College Tier						
Top tier	-1.931***	-1.684***	-2.958***	-1.671***	-1.491***	-2.194***
	(0.347)	(0.325)	(0.494)	(0.279)	(0.288)	(0.723)
Second Tier	-1.332***	-1.195***	-2.460	-1.409***	-1.364***	-1.908**
	(0.350)	(0.296)	(1.683)	(0.321)	(0.323)	(0.820)
Third Tier	-0.864**	-0.817***	-1.549**	-0.958***	-1.075***	0.190
	(0.269)	(0.196)	(0.507)	(0.363)	(0.364)	(1.249)
Fourth Tier	-0.895***	-0.618**	-1.954***	-0.727***	-0.524***	-0.425
	(0.232)	(0.186)	(0.361)	(0.079)	(0.123)	(0.436)
age	0.092***	0.092***	0.167***	0.095***	0.095***	0.168***
	(0.001)	(0.001)	(0.018)	(0.002)	(0.002)	(0.040)
Interactions						
Top tier $\times$ age	0.067***	0.065***	0.111***	0.058***	0.058***	0.105***
	(0.009)	(0.009)	(0.014)	(0.009)	(0.009)	(0.024)
Second Tier $\times$ age	0.050***	0.052***	0.098	0.051***	0.055***	0.095***
	(0.011)	(0.010)	(0.054)	(0.010)	(0.011)	(0.031)
Third Tier $\times$ age	0.032***	0.037***	0.058**	0.034***	0.044***	0.011
	(0.008)	(0.007)	(0.019)	(0.012)	(0.013)	(0.050)
Fourth Tier $\times$ age	0.029***	0.027***	0.073***	0.022***	0.021***	0.038***
	(0.007)	(0.007)	(0.015)	(0.002)	(0.002)	(0.013)
N	29599	29599	685	29599	29599	752
Major	No	Yes	No	No	Yes	Yes
RE	No	No	No	Yes	Yes	Yes
CSAT	No	No	Yes	No	No	No

Appendix D

Define the first order conditions as

$$
\begin{aligned}
V_p &= \alpha_c \varepsilon^c_{it} u^c_p(c_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial p_{it}} \\
V_x &= \alpha_c \varepsilon^c_{it} u^c_x(c_{it}) + \alpha_x \varepsilon^x_{it} u^x_x(x_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial x_{it}} \\
V_s &= \alpha_s \varepsilon^s_{it} u^s_s(s_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial s_{it}}
\end{aligned}
$$

$$
\frac{\partial p}{\partial w} = -\frac{\frac{\partial V_p}{\partial w}}{\frac{\partial V_p}{\partial p}}
$$

$$
\frac{\partial V_p}{\partial w} = \alpha_c \varepsilon^c_{it} \frac{x_{it}}{(w_{it} - p_{it} x_{it})^2}
$$

$$
\begin{aligned}
\frac{\partial V_p}{\partial p} =& \frac{\partial}{\partial p} \varepsilon^c_{it} \frac{-x_{it}}{(w_{it} - p_{it} x_{it})} + \beta \frac{\partial}{\partial p} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \left(\nu_t \frac{\delta_{2t}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{2t}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{3t}(1 + s_{it})^{\phi}]} (\kappa {p_{it}}^{\kappa-1} {x_{it}}^{1-\kappa}) \right) \\
=& -2\alpha_c \varepsilon^c_{it} \frac{x^3_{it}}{(w_{it} - p_{it} x_{it})^3} + \beta \left[\frac{\partial^2}{\partial^2 \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \left(\nu_t \frac{\delta_{2t}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{2t}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{3t}(1 + s_{it})^{\phi}]} (\kappa {p_{it}}^{\kappa-1} {x_{it}}^{1-\kappa}) \right)^2 \\
& + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \nu_t \Bigg(- \frac{\delta^2_{2t} \kappa^2 p^{(2\kappa-2)}_{it} \phi x^{(2-2\kappa)}_{it} (1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{2\phi-2}}{[\delta_{2t}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{3t}(1 + s_{it})^{\phi}]^2} (\kappa {p_{it}}^{\kappa-1} {x_{it}}^{1-\kappa}) \\
& + \frac{\delta_{2t} \kappa^2 p^{(2\kappa-2)}_{it} (\phi - 1) x^{(2-2\kappa)} (1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi-2}}{[\delta_{2t}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{3t}(1 + s_{it})^{\phi}]} \\
& + \frac{\delta_{2t}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{2t}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{3t}(1 + s_{it})^{\phi}]} ((\kappa - 1)\kappa p^{\kappa-2}_{it} x^{1-\kappa}_{it}) \Bigg)
\end{aligned}
$$

As $\phi < 1$, $\kappa < 0.5$, and $\frac{\partial^2}{\partial^2 \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) < 0$, $\frac{\partial V_p}{\partial p} < 0$ and $\frac{\partial V_p}{\partial w} > 0$, $\frac{\partial p}{\partial w} > 0$.

Appendix E

Details of the Likelihood Function

The likelihood contributions of the choice variables are computed by transforming the characterized expression of the shocks, using the Jacobian-transformation. In particular, the time-specific likelihood contribution can be expressed as

$$
\begin{aligned}
\mathcal{L}_{it}(\theta|S_{it},\lambda_k) = & \left[f_{\eta^c_{it}}(\tilde{\eta}^c_{it}) \cdot f_{\eta^x_{it}}(\tilde{\eta}^x_{it}) \cdot f_{\eta^s_{it}}(\tilde{\eta}^s_{it}) \cdot f_{\eta^q_{it}}(\tilde{\eta}^q_{it}) |det\left(\frac{\partial(\tilde{\eta}^c_{it},\tilde{\eta}^x_{it},\tilde{\eta}^s_{it},\tilde{\eta}^q_{it})}{\partial(p_{it},x_{it},s_{it},q_{it})}|\right]^{d^x_{it}} \\
& \times \left[\int_{\tilde{\eta}^s_{it}} \left(f_{\eta^c_{it}}(\tilde{\eta}^c_{it}) \cdot f_{\eta^x_{it}}(\tilde{\eta}^x_{it}) \cdot f_{\eta^s_{it}}(\eta^s_{it}) \cdot f_{\eta^q_{it}}(\tilde{\eta}^q_{it}) \right) d\eta^s_{it} | \det \frac{\partial(\tilde{\eta}^c_{it},\tilde{\eta}^x_{it},\tilde{\eta}^q_{it})}{\partial(p_{it},x_{it},q_{it})}| \right]^{d^x_{it}(1-d^s_{it})} \\
& \times \left[\int_{-\infty}^{\infty} \int_{\tilde{\eta}^x_{it}(\eta^c_{it})} \left[f_{\eta^c_{it}}(\eta^c_{it}) \cdot f_{\eta^l_{it}}(\eta^x_{it}|\eta^c_{it}) \cdot f_{\eta^s_{it}}(\tilde{\eta}^s_{it}) \cdot f_{\eta^q_{it}}(\tilde{\eta}^q_{it}) \right] d\eta^x_{it} d\eta^c_{it} | \det \frac{\partial(\tilde{\eta}^s_{it},\tilde{\eta}^q_{it})}{\partial(s_{it},q_{it})}| \right]^{(1-d^x_{it})d^s_{it}} \\
& \times \Bigg[\Bigg(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Big[\Pr\{V_{00}(\eta^c_{it},\eta^x_{it},\eta^s_{it}) > V_{x0}(\eta^s_{it}), V_{00}(\eta^c_{it},\eta^x_{it},\eta^s_{it}) > V_{0s}(\eta^c_{it},\eta^x_{it}), V_{00}(\eta^c_{it},\eta^x_{it},\eta^s_{it}) > V_{xs})\} \\
& \quad \times d\eta^c_{it} d\eta^x_{it} d\eta^s_{it} \Big] f(\tilde{\eta}^q_{it}) | \det \frac{\partial \tilde{\eta}^q_{it}}{\partial q_{i,t+1}}| \Bigg]^{(1-d^x_{it})(1-d^s_{it})}
\end{aligned}
$$

where V_{00} is the value when $x_{it} = s_{it} = 0$, V_{x0} is the value when $x_{it} > 0$ and $s_{it} = 0$, and V_{0s} is the value when $x_{it} = 0$ and $s_{it} > 0$.[56]

To evaluate the integrals in the likelihood function, I use the Montecarlo simulation. Borsch-Supan, Hajivassiliou and Kotlikoff (1992) show that the MSL estimates perform well under a moderate number of draws, such as 20, with an adoption of a good simulation method. To reduce the variance of simulation error, I use antithetic acceleration (Geweke 1988; Stern 1997).

About 8.3% of the household-year observations are missing, creating "holes" in the household data. I simulate the unobserved choice variables using the value function of the model (Lavy, Palumbo and Stern 1998; Stinebrickner 1999; Sullivan 2009). In particular, for each draw of the set of errors, I replace the unobserved choice variables with the optimized choices that maximize the value function of the model. Also, for periods 4 and 5, the test score data are unobserved. I simulate the unobserved test scores for each draw of test score error η^q_{it} using equation (4). In the next subsection, I show the derivation of the density and probability I use for computing the likelihood function, and I explain the simulation of unobserved variables.

First-order conditions used for likelihood contribution

The goal of this section is to get a closed form expression of the shocks, which are the building blocks of the likelihood function. I denote $u^c_p(c_{it})$ and $u^c_x(c_{it})$ as the first order derivatives of $u^c(c_{it})$ with respect to x_{it} and p_{it} respectively, and $u^l_x(l_{it})$ and $u^l_s(l_{it})$ as the

[56]For the case $d^x_{it} = d^s_{it} = 0$, I am working on a G.H.K type of simulation to reduce the variance of simulation error.

first order derivatives with respect to x_{it} and s_{it} respectively. The first order conditions of the value function in equation (9) are

$$
\begin{aligned}
\frac{\partial}{\partial p_{it}} &: \alpha_c \exp(\eta^c_{it} + \lambda^c_k) + \beta \frac{1}{u^c_p(c_{it})} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(Z_{i,t+1}(\ln q_{i,t+1}(p_{it}, x_{it}, s_{it}), \Psi_{i,t+1}) \right] \frac{\partial \ln q_{i,t+1}}{\partial p_{it}} = 0; \\
\frac{\partial}{\partial x_{it}} &: \alpha_c \exp(\eta^c_{it} + \lambda^c_k) u^c_x(c_{it}) + \alpha_x \exp(\eta^k_{it} + \lambda^c_k) u^x_x(x_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(Z_{i,t+1}(\ln q_{i,t+1}(p_{it}, x_{it}, s_{it}), \Psi_{i,t+1}) \right] \frac{\partial \ln q_{i,t+1}}{\partial x_{it}} = 0; \\
\frac{\partial}{\partial s_{it}} &: \exp(\eta^s_{it} + \lambda^s_k) u^s_s(s_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(Z_{i,t+1}(\ln q_{i,t+1}(p_{it}, x_{it}, s_{it}), \Psi_{i,t+1}) \right] \frac{\partial \ln q_{i,t+1}}{\partial s_{it}} = 0.
\end{aligned}
$$

With the functional form assumptions of log utility,

$$
\begin{aligned}
u^c_x(c_{it}) =& -\frac{p_{it}}{w_{it} - p_{it} x_{it}}; \\
u^c_p(c_{it}) =& -\frac{x_{it}}{w_{it} - p_{it} x_{it}}; \\
u^s_s(s_{it}) =& \frac{1}{1 + s_{it}}; \\
u^x_x(x_{it}) =& \frac{1}{1 + x_{it}}
\end{aligned}
$$

And with the functional form of the test score function,

$$
\begin{aligned}
q_{i,t+1} &= A_{it} q^{\delta_{1t}}_{it} \left[\delta_{et}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{st}(1 + s_{it})^{\phi} \right]^{\frac{\nu_t}{\phi}} \exp(\lambda^q_k + \eta^q_{it}) \\
\ln q_{i,t+1} &= \ln A_{it} + \delta_{1t} \ln q_{it} + \frac{\nu}{\phi} \ln[\delta_{et}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{st}(1 + s_{it})^{\phi}] + \lambda^q_k + \eta^q_{it}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \ln q_{i,t+1}}{\partial p_{it}} =& \nu_t \frac{\delta_{et}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{et}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{st}(1 + s_{it})^{\phi}]} (\kappa {p_{it}}^{\kappa-1} {x_{it}}^{1-\kappa}); \\
\frac{\partial \ln q_{i,t+1}}{\partial x_{it}} =& \nu_t \frac{\delta_{et}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{et}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{st}(1 + s_{it})^{\phi}]} ((1-\kappa) {p_{it}}^{\kappa} {x_{it}}^{-\kappa}); \\
\frac{\partial \ln q_{i,t+1}}{\partial s_{it}} =& \nu_t \frac{\delta_{st}(1 + s_{it})^{\phi-1}}{[\delta_{et}(1 + {p_{it}}^{\kappa} {x_{it}}^{1-\kappa})^{\phi} + \delta_{st}(1 + s_{it})^{\phi}]}.
\end{aligned}
$$

The first order conditions with respect to p_{it} is characterized as

$$\alpha_c \exp(\eta_{it}^c + \lambda_k^c) - \beta \frac{w_{it} - p_{it}x_{it}}{x_{it}} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(Z_{i,t+1}(\ln q_{i,t+1}(p_{it}, x_{it}, s_{it})), \Psi_{i,t+1}) \right]$$
$$\times \nu_t \frac{\delta_{et}(1 + {p_{it}}^\kappa {x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{et}(1 + {p_{it}}^\kappa {x_{it}}^{1-\kappa})^\phi + \delta_{st}(1 + s_{it})^\phi]} \times (\kappa {p_{it}}^{\kappa-1} {x_{it}}^{1-\kappa}) = 0. \tag{13}$$

The first order conditions with respect to x_{it} is characterized as

$$- \alpha_c \exp(\eta_{it}^c + \lambda_k^c) \frac{p_{it}}{w_{it} - p_{it}x_{it}} + \alpha_x \exp(\eta_{it}^x + \lambda_k^x) \frac{1}{1 + x_{it}} + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\ln q_{i,t+1}(p_{it}, x_{it}, s_{it}), \Psi_{i,t+1}) \right]$$
$$\times \nu_t \frac{\delta_{et}(1 + {p_{it}}^\kappa {x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{et}(1 + {p_{it}}^\kappa {x_{it}}^{1-\kappa})^\phi + \delta_{st}(1 + s_{it})^\phi]} \times (1 - \kappa) {p_{it}}^\kappa {x_{it}}^{-\kappa} = 0. \tag{14}$$

The first order conditions with respect to s_{it} is characterized as

$$\alpha_s \exp(\eta_{it}^s + \lambda_k^s) \frac{1}{1 + s_{it}} + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(Z_{i,t+1}(\ln q_{i,t+1}(p_{it}, x_{it}, s_{it})), \Psi_{i,t+1}) \right]$$
$$\times \nu_t \frac{\delta_{st}(1 + s_{it})^{\phi-1}}{[\delta_{et}(1 + {p_{it}}^\kappa {x_{it}}^{1-\kappa})^\phi + \delta_{st}(1 + s_{it})^\phi]} = 0. \tag{15}$$

This difference between the previous period and the final period can be confusing. For the final period,

$$\begin{aligned} EV_{i,T+1} =& v_1 - \sum_{j=1}^{J} \Big(\ln(v_j) - \ln(v_{j+1}) \Big) \Phi(\frac{\ln \bar{q}_j - \ln q_{iT+1} - \lambda_i^q}{\sigma_q}); \\ \frac{\partial}{\partial \ln q_{i,T+1}} EV_{i,T+1}(\cdot, \cdot, \cdot) =& \sum_{j=1}^{J} \Big(\ln(v_j) - \ln(v_{j+1}) \Big) \frac{1}{\sigma_q} \phi(\frac{\ln \bar{q}_j - \ln q_{iT+1} - \lambda_i^q}{\sigma_q}), \end{aligned}$$

while for $t < T$, EV_{it} is an interpolated value function.

Computation of Likelihood Contribution

(Case 1) ($x_{it} > 0$ and $s_{it} > 0$)

I define $\tilde{\eta}_{it}^z$ for $z = c, x, s$ as the particular realization of η_{it}^z that satisfies the first order conditions. The likelihood contribution for all-positive case is

$$
\begin{aligned}
f(p_{it}, x_{it}, s_{it}, q_{it}) =& f(\tilde{\eta}^c_{it}, \tilde{\eta}^x_{it}, \tilde{\eta}^s_{it}, \tilde{\eta}^q_{it}) \cdot |det\left(\frac{\partial(\widetilde{\eta}^c_{it}, \widetilde{\eta}^x_{it}, \widetilde{\eta}^s_{it}, \widetilde{\eta}^q_{it})}{\partial(p_{it}, x_{it,} s_{it}, q_{it})}\right)| \\
=& \phi(\tilde{\eta}^c_{it}, \tilde{\eta}^x_{it}, \tilde{\eta}^s_{it}, \tilde{\eta}^q_{it}) \cdot |det\left(\frac{\partial(\widetilde{\eta}^c_{it}, \widetilde{\eta}^x_{it}, \widetilde{\eta}^s_{it}, \widetilde{\eta}^q_{it})}{\partial(p_{it}, x_{it,} s_{it}, q_{it})}\right)| \\
=& (2\pi)^{-4/2} |det(\Omega)|^{-1/2} \exp\Bigg[-0.5 \underbrace{\begin{pmatrix} \widetilde{\eta}^c_{it} \\ \widetilde{\eta}^x_{it} \\ \widetilde{\eta}^s_{it} \\ \widetilde{\eta}^q_{it} \end{pmatrix}'}_{1\times 4} \underset{4\times 4}{\Omega^{-1}} \underbrace{\begin{pmatrix} \widetilde{\eta}^c_{it} \\ \widetilde{\eta}^x_{it} \\ \widetilde{\eta}^s_{it} \\ \widetilde{\eta}^q_{it} \end{pmatrix}}_{4\times 1} \Bigg] \Bigg| det\left(\frac{\partial(\widetilde{\eta}^c_{it}, \widetilde{\eta}^x_{it}, \widetilde{\eta}^s_{it}, \widetilde{\eta}^q_{it})}{\partial(p_{it}, x_{it,} s_{it}, q_{it})}\right) \Bigg|
\end{aligned}
$$

(Case 2) ($x_{it} > 0$ and $s_{it} = 0$)

This is the case where household participate in tutoring, but have zero hours of self-study. First, I define the joint probability of such case, and separate the density of η^c_{it} and η^x_{it} out using Bayes' theorem. I denote $A_{x_{it}, s_{it}=0}$ as the corresponding region that the joint integration of η^c_{it}, η^x_{it}, and η^s_{it} needs to be made.

$$
\begin{aligned}
\Pr(p_{it}, x_{it}, s_{it} = 0) =& \Pr(s_{it} = 0 | p_{it}, x_{it}) f(p_{it}, x_{it}) \\
=& \Pr(\eta^s_{it} > \underline{\eta}^s_{it} | \tilde{\eta}^c_{it}, \tilde{\eta}^x_{it}) f_\eta(\tilde{\eta}^x_{it}, \tilde{\eta}^c_{it}) | \det \frac{\partial(\tilde{\eta}^c_{it}, \tilde{\eta}^x_{it})}{\partial(p_{it,} x_{it})}|,
\end{aligned}
$$

where $\underline{\eta}^s_{it}$ is the minimum value of η^s_{it} that leads to zero hours of self-study. I use the first order condition with respect to s_{it}, equation (15), in computing the critical value.

With the zero correlation assumption between eta,

$$
\begin{aligned}
& \Pr(\eta^s_{it} > \underline{\eta}^s_{it} | \tilde{\eta}^c_{it}, \tilde{\eta}^x_{it}) f_\eta(\tilde{\eta}^x_{it}, \tilde{\eta}^c_{it}) | \det \frac{\partial(\tilde{\eta}^c_{it}, \tilde{\eta}^x_{it})}{\partial(p_{it,} x_{it})}| \\
=& \Pr(\eta^s_{it} > \underline{\eta}^s_{it}) f(\tilde{\eta}^x_{it}) f(\tilde{\eta}^c_{it}) | \det \frac{\partial(\tilde{\eta}^c_{it}, \tilde{\eta}^x_{it})}{\partial(p_{it,} x_{it})}| \\
=& \left(1 - \Phi(\underline{\eta}^s_{it})\right) \frac{1}{\sigma_x} \phi(\frac{\tilde{\eta}^x_{it}}{\sigma_x}) \frac{1}{\sigma_c} \phi(\frac{\tilde{\eta}^c_{it}}{\sigma_x}) | \det \frac{\partial(\tilde{\eta}^c_{it}, \tilde{\eta}^x_{it})}{\partial(p_{it,} x_{it})}|,
\end{aligned}
$$

which is what I use for computing the likelihood contribution for (Case 2).

(Case 3) ($x_{it} = 0$ and $s_{it} > 0$)

This is the case where household do not participate in tutoring, but do positive hours of self-study. Since $p_{it} > 0$ for all households, $p_{it}x_{it} = 0$ is equivalent to $x_{it} = 0$. For people who have $x_{it} = 0$, I let them consider minimum quality of tutoring, $\bar{p}$, which is equivalent the minimum market price.

Denote $A_{x_{it}=0}$ as the corresponding region that the joint integration of η^c_{it} and η^x_{it} needs to be made. First, I separate out the marginal density of η^s_{it} using Bayes' theorem, which gives me

$$\begin{aligned}\Pr(x_{it}=0, s_{it}) =& \Pr(x_{it}=0|s_{it})f(s_{it})\\ =& \Pr(\eta^c_{it}, \eta^x_{it} \in A_{x_{it}=0,s_{it}}|\tilde{\eta}^s_{it}) \cdot \frac{1}{\sigma_s}\phi(\frac{\tilde{\eta}^s_{it}}{\sigma_s})|\frac{\partial\tilde{\eta}^s_{it}}{\partial(s_{it})}|.\end{aligned}$$

As I assume there is no correlation between η_{it},

$$\begin{aligned}&\Pr(\eta^c_{it}, \eta^x_{it} \in A_{x_{it}=0,s_{it}}|\tilde{\eta}^s_{it}) \cdot \frac{1}{\sigma_s}\phi(\frac{\tilde{\eta}^s_{it}}{\sigma_s})|\frac{\partial\tilde{\eta}^s_{it}}{\partial(s_{it})}|\\ &= \Pr(\eta^c_{it}, \eta^x_{it} \in A_{x_{it}=0,s_{it}}) \cdot \frac{1}{\sigma_s}\phi(\frac{\tilde{\eta}^s_{it}}{\sigma_s})|\frac{\partial\tilde{\eta}^s_{it}}{\partial(s_{it})}|.\end{aligned}$$

Here, I use the first order condition, equation (15), in characterizing the unique values of $\tilde{\eta}^s_{it}$. Define $\underline{\eta}^x_{it}$ as a minimum amount of shock that makes individual start doing zero hours of tutoring. Again, with the assumption of no correlation between η_{it},

$$\begin{aligned}&\Pr(\eta^c_{it}, \eta^x_{it} \in A_{x_{it}=0,s_{it}}) \cdot \frac{1}{\sigma_s}\phi(\frac{\tilde{\eta}^s_{it}}{\sigma_s})|\frac{\partial\tilde{\eta}^s_{it}}{\partial(s_{it})}|\\ =& \Pr(\eta^c_{it}, \eta^x_{it} > \underline{\eta}^s_{it}|\tilde{\eta}^s_{it})d\eta^x_{it}d\eta^c_{it}\frac{1}{\sigma_s}\phi(\frac{\tilde{\eta}^s_{it}}{\sigma_s})|\frac{\partial\tilde{\eta}^s_{it}}{\partial(s_{it})}|\\ =& \Bigg[\int_{-\infty}^{\infty}\Bigg\{\int_{\underline{\eta}^x_{it}(\eta^c_{it})}^{\infty}\frac{1}{\sigma_x}\phi(\frac{\eta^x_{it}}{\sigma_x})d\eta^x_{it}\Bigg\}\frac{1}{\sigma_c}\phi(\frac{\eta^c_{it}}{\sigma_c})d\eta^c_{it}\Bigg]\frac{1}{\sigma_s}\phi(\frac{\tilde{\eta}^s_{it}}{\sigma_s})|\frac{\partial\tilde{\eta}^s_{it}}{\partial(s_{it})}|\\ =& \Bigg[\int_{-\infty}^{\infty}\Bigg\{1-\Phi(\frac{\underline{\eta}^x_{it}(\eta^c_{it})}{\sigma_x})\Bigg\}\frac{1}{\sigma_c}\phi(\frac{\eta^c_{it}}{\sigma_c})d\eta^c_{it}\Bigg]\frac{1}{\sigma_s}\phi(\frac{\tilde{\eta}^s_{it}}{\sigma_s})|\frac{\partial\tilde{\eta}^s_{it}}{\partial(s_{it})}|\end{aligned}$$

(Case 4) ($x_{it} = 0$ and $s_{it} = 0$)

This is the case where $x_{it} = 0$ and $s_{it} = 0$. To make the notation concise, I denote V_{00} as the value when $x_{it} = s_{it} = 0$. V_{x0} denotes the case $x > 0$ and $s = 0$. V_{0s} denotes the case $x = 0$ and $s > 0$.

$$\begin{aligned}\Pr(x_{it} = 0, s_{it} = 0) &= \Pr(\eta^c_{it}, \eta^x_{it}, \eta^s_{it} \in A_{x_{it}=0, s_{it}=0}) \\ &= \Pr(V_{00}(\eta^c_{it}, \eta^x_{it}, \eta^s_{it}) > V_{x0}(\eta^s_{it}), V_{00}(\eta^c_{it}, \eta^x_{it}, \eta^s_{it}) > V_{0s}(\eta^c_{it}, \eta^x_{it}), V_{00}(\eta^c_{it}, \eta^x_{it}, \eta^s_{it}) > V_{xs}).\end{aligned}$$

$$\begin{aligned}\Pr(x_{it} = 0, s_{it} = 0) =& \Pr(\eta^c_{it}, \eta^x_{it}, \eta^s_{it} \in A_{x_{it}=0, s_{it}=0}) \\ =& \Bigg(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1\{V_{00}(\eta^c_{it}, \eta^x_{it}, \eta^s_{it}) > V_{x0}(\eta^s_{it}), V_{00}(\eta^c_{it}, \eta^x_{it}, \eta^s_{it}) > V_{0s}(\eta^c_{it}, \eta^x_{it}), V_{00}(\eta^c_{it}, \eta^x_{it}, \eta^s_{it}) > V_{xs})\} \\ =& f(\eta^s_{it}) f(\eta^x_{it}) f(\eta^c_{it}) d\eta^s_{it} d\eta^x_{it} d\eta^c_{it} \Bigg)\end{aligned}$$

The integral does not have an analytical solution and needs to be simulated.
Simulation algorithm is
(1) I draw an unconditional set of $\eta^r_{it} = \{\eta^{cr}_{it}, \eta^{xr}_{it}, \eta^{sr}_{it}\}$
(2) Let household optimize their choices.
(3) Count the proportion of cases that household chooses $x_{it} = 0$ and $s_{it} = 0$
In particular define x^r and s^r such that

$$(x^r, s^r) = \arg\max_{x_{it}, s_{it}} V_{it}(\eta^c_{it}, \eta^x_{it}, \eta^s_{it})$$

Compute

$$\frac{1}{R} \sum_{r=1}^{R} \mathbb{1}(x*, s* = 0).$$

So

$$\Pr(x_{it} = 0, s_{it} = 0) \approx \frac{1}{R} \sum_{r=1}^{R} \mathbb{1}(x*, s* = 0).$$

Simulation of unobserved variables

For each missing choice variables, I draw a set of corresponding error. For example, if x_{it} is missing for person i, the simulation algorithm is

(1) I draw a simulation for the corresponding error. In this example, it is η_{it}^{xr}

(2) Let household optimize their choice

$$x^r = \frac{1}{R} \sum_{r=1}^{R} \left\{ \arg\max_{x_{it}, s_{it}} V_{it}(\eta_{it}^{c}, \eta_{it}^{xr}, \eta_{it}^{s}) \right\}.$$

The optimized choice is used for computing likelihood function.

For missing test score, I draw a set of errors for η_{it}^{q}. Then the unobserved test score is simulated using equation (4).

Appendix F

The expected value of data y conditional on the model predicted value $\widehat{y}$ is $E(y|\widehat{y}) = \widehat{\kappa_0}(\widehat{y})$, where

$$\begin{pmatrix} \widehat{\kappa_0}(y) \\ \widehat{\kappa_1}(y) \end{pmatrix} = \sum_i \left[K(\frac{y_i - \widehat{y}_i}{b}) \begin{pmatrix} 1 \\ y_i - \widehat{y}_i \end{pmatrix} \begin{pmatrix} 1 & y_i - \widehat{y}_i \end{pmatrix} \right]^{-1} \cdot \left[K(\frac{y_i - \widehat{y}_i}{b}) \begin{pmatrix} 1 \\ y_i - \widehat{y}_i \end{pmatrix} y_i \right],$$

and

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp(-0.5x^2)$$

is the kernel function with bandwidth b. The farther the kernel curve deviates from the 45 degree line, the less the model is successful in fitting the data.

Figure F.1: Sample Fit: Private Tutoring Expenditure

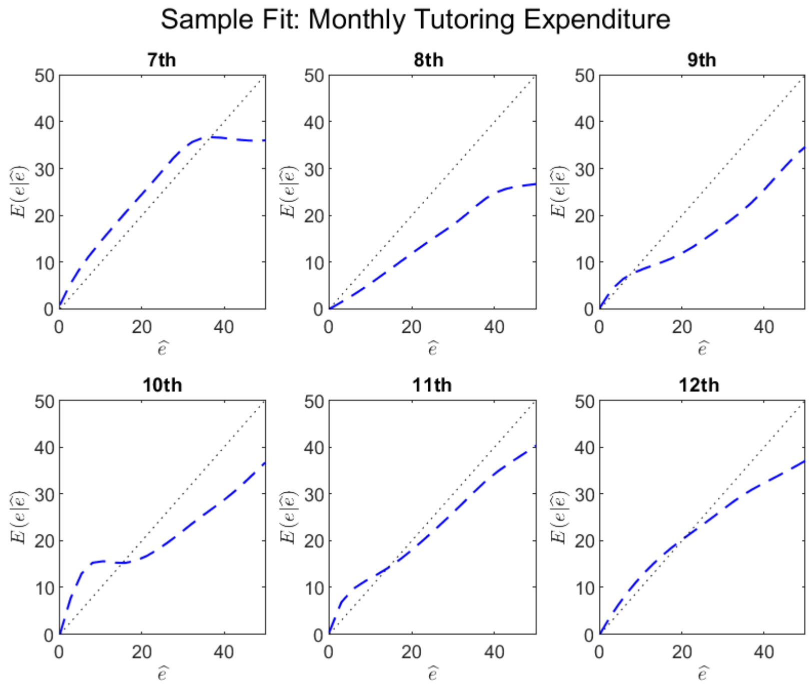


Figure F.2: Hours of Tutoring

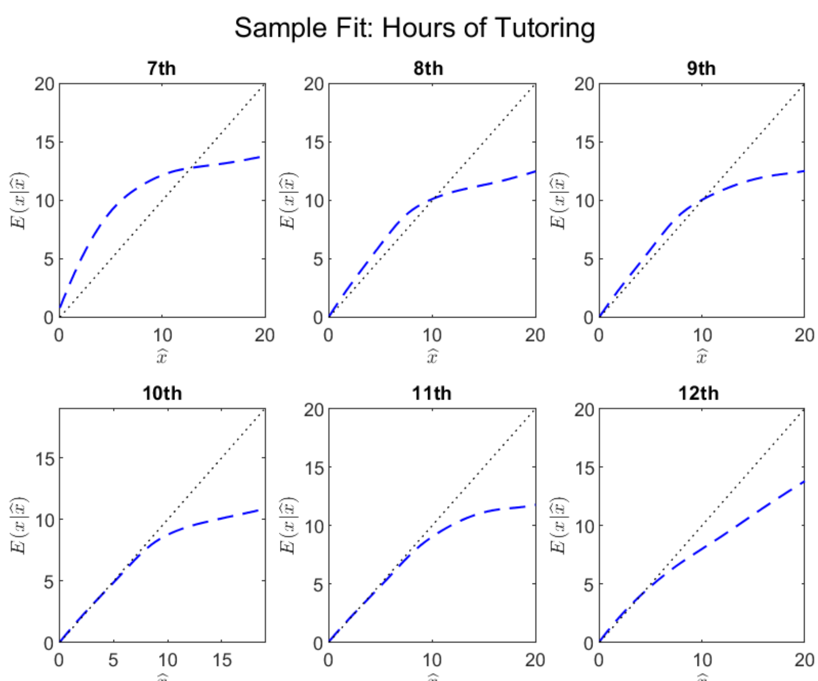


Figure F.3: Sample Fit: Quality of Tutoring

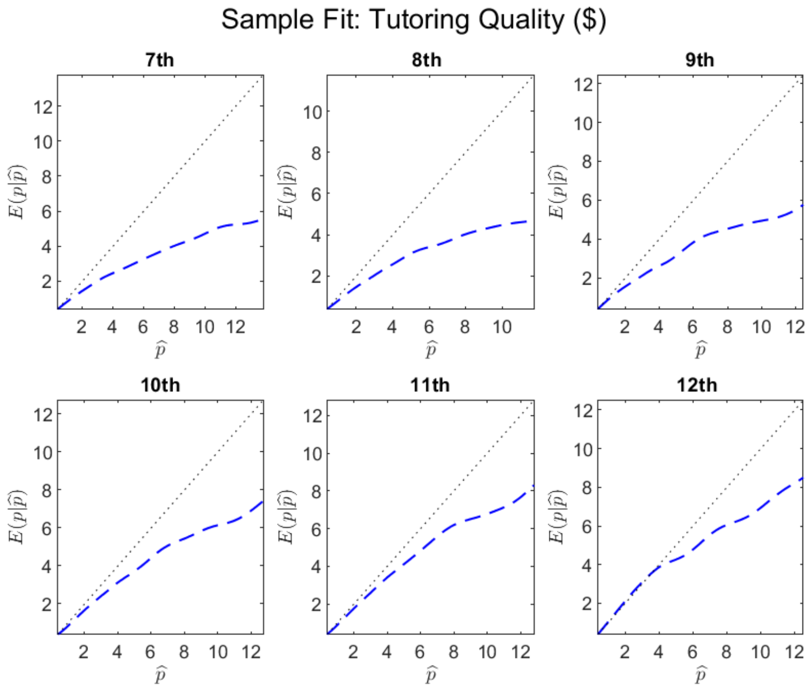


Figure F.4: Sample Fit: Hours of self-study

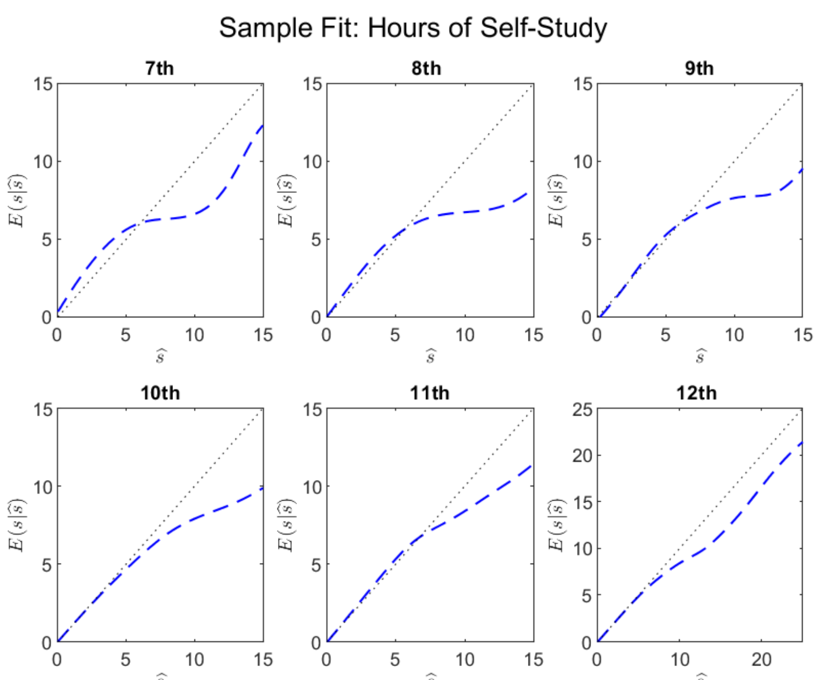
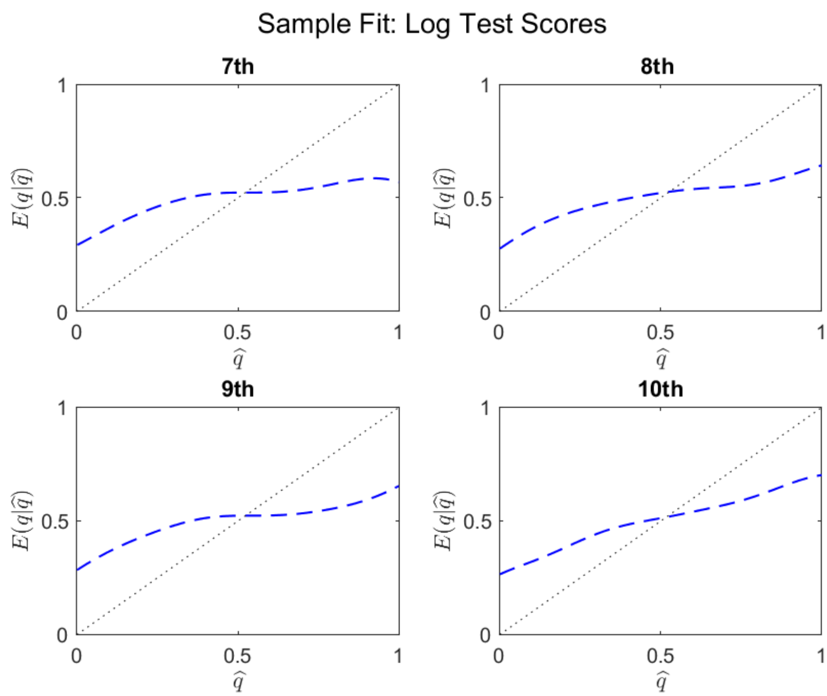
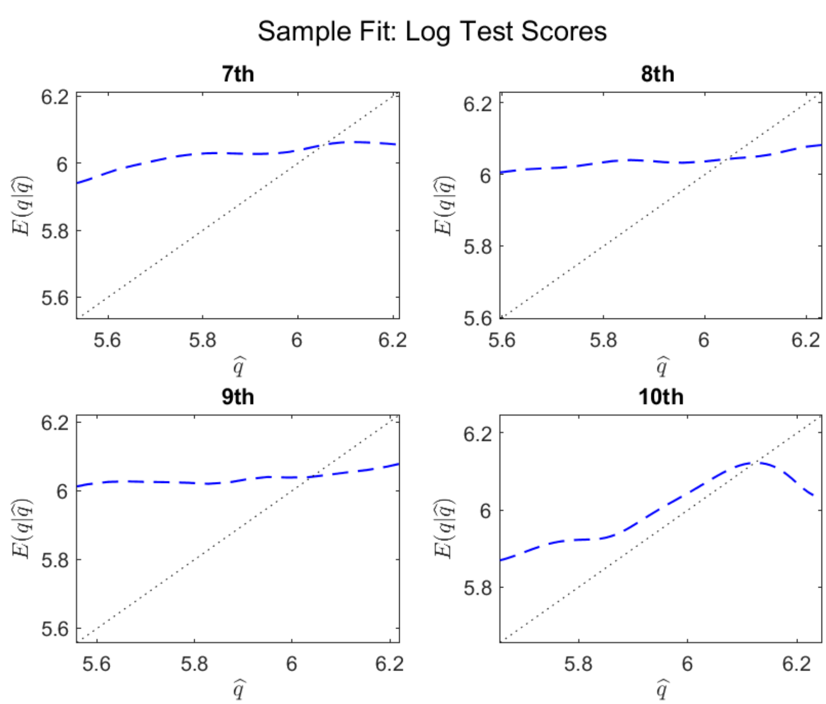


Figure F.5: Sample Fit: Log Test Scores

(a) Fit by distribution



(b) Fit by level



Appendix G

Intergenerational persistence of earnings

Table G.5: Intergenerational Persistence of Earnings Fixing Initial Conditions

The effect of heterogeneity in the academic performance in primary school on the intergenerational persistence of earnings is modest. To control for the difference among students before 7th grade, I run the counterfactual simulations with fixing the academic performance in primary school and parental education, which can be found in Table G.5. For example, BCF' is the same simulation as BCF except that 6th-grade academic performance and parental education are fixed across households. The results are consistent with the original counterfactual simulations that are conducted without fixing the household characteristics.

(a) Rank-rank Slope Estimates

	(1) BCF'	(2) OPI	(3) OSS'
pincprctile	0.573*** (0.019)	0.715*** (0.017)	0.380*** (0.022)
R-squared	0.329	0.511	0.144

(b) Intergenerational Elasticity of Earnings Estimates

	(1) BCF'	(2) OPI'	(3) OSS'
logpinc	0.232*** (0.008)	0.230*** (0.005)	0.139*** (0.008)
R-squared	0.304	0.539	0.137

Note: Table (a) and (b) provide the estimates of rank-rank slope and intergenerational elasticty of earnings, respectively. In order to assess the importance of initial conditions, all simulations are conducted fixing the initial test score and parental education. Parental income is not fixed except for the NIN' simualtion. BCF' is the benchmark counterfactual; NIN' is a simulation where each household income is fixed to the mean; OPI' is a simulation where each household can use only parental investment; and OSS' is a simulation where each household can use only hours of self-study. The linearity of the lines follows from the linearity of the rank-rank equation.