Parental Investment, Child's Efforts, and Intergenerational Mobility

Hyunjae Kang*

October 15^{th} , 2022
Preliminary and incomplete. Please do not circulate

Abstract

This paper investigates the role of parents' investment and child's self-efforts made during adolescence on the intergenerational persistence of economic status. The competition for prestigious colleges between students is a channel where the heterogeneity of parental background comes in. While parental income strongly affects hours of private tutoring of a student, the level of a student's self-study is only weakly affected by parental income. To examine the role of parents' investment and the child's efforts on intergenerational mobility, I build and estimate a dynamic tournament model where each household chooses quality of private tutoring, hours of private tutoring, and hours of student self-study. Students compete for the finite seats of prestigious colleges, which are characterized by alumni's lifetime income profile. Using a unique longitudinal dataset on secondary school students, I find that heterogeneity in parental income in adolescent periods accounts for 46% of intergenerational persistence of earnings. Parental investment is responsible for a substantial portion of intergenerational persistence of earnings controlling for the child's efforts. Ignoring child's efforts from the mechanism leads to a significant increase in intergenerational persistence of earnings by 30%, which suggests the role of the self-effort as a mitigator of intergenerational persistence of earnings. Motivated by the recent policy of China banning private tutoring, I conduct a counterfactual experiment prohibiting private tutoring expenditure. A counterfactual simulation suggests that the introduction of the policy banning private tutoring would lead to an increase in intergenerational mobility at the expense of an increase in consumption inequality. Finally, in light of the advent of the extremely low fertility regimes, I assess the impact of the rapidly shrinking cohort size on parental investment. Based on the model projection, low income households spend money on private tutoring expenditure as cohort size decreases, while there is virtually no change in the parental investment spending of high income households.

*Email: jay.hj.kang@gmail.com. Mailing address: Department of Economics, Stony Brook University; 100 Nicolls Rd, Stony Brook, NY 11794. I am immensely grateful to my advisors, Steven Stern and Juan Pantano for invaluable guidance and support. I would also like to thank Meta Brown, David Wiczer, Gabriel Mihalache, Peter Arcidiacono, John Rust, Hugo Benitez-Silva, Samuele Centorrino, Arda Aktas, Miguel Poblete-Cazenave, Sunghun Cho, Marcos Fernandes, Jeremy Fox, Yujung Hwang, Fedor Iskhakov, Laura Karpuska, Sunham Kim, Chinhui Juhn, Alfonso Flores Lagunes, Bob Millard, Mark Montgomery, Victor Ronda, Greg Veramendi, Jiayi Wen, Sangha Yoon, Basit Zafar, Weilong Zhang and Yiyi Zhou. Special thanks to Warren Sanderson for initial encouragement and many helpful discussions. All errors are my own.

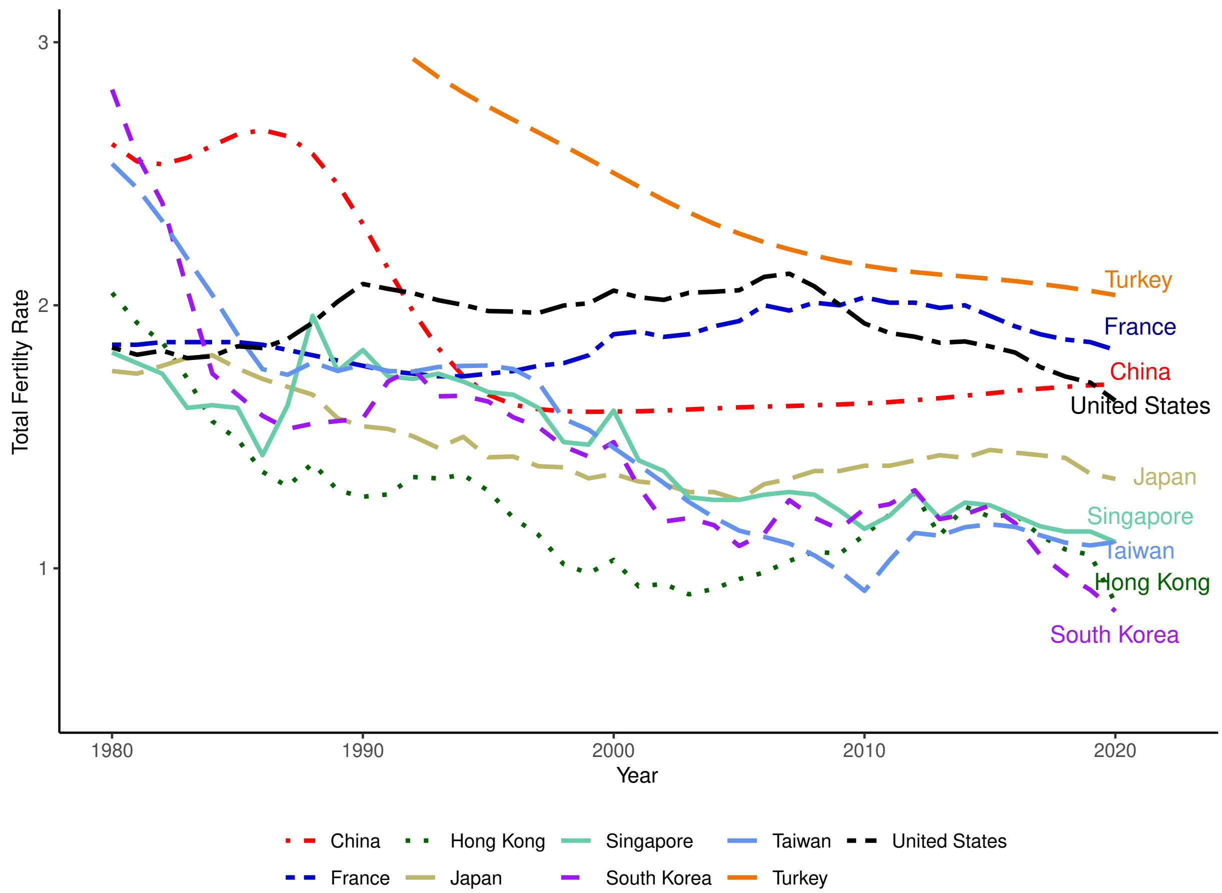
1 Introduction

Parental investment is a source of 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 (Guryan, Hurst and Kearney, 2008). Thus, parental investment potentially has an important consequence 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). The self-effort of the child is not responsive to parental background as much as parental investment is affected by parental background. Despite the potential relevance, few studies have attempted to disentangle and quantify the impact of parental investment on intergenerational mobility combined with effects of childrens' self-effort.

On the other hand, the competition motive is a driver of parental investment. 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), but seats for such prestigious colleges are limited, and evidence suggests that colleges do not adjust the seats to accommodate for increasing cohort size (Bound and Turner, 2007). The scarcity leads to competition with respect to getting into prestigious colleges, which drives more parental investment (Ramey and Ramey, 2010). In the United States, it is easy to see parents spend a lot of time helping their children with extracurricular activities. In East-Asian countries, parents spend a significant portion of their income on private tutoring expenditure whose main purpose is getting into better universities (Bray 1999; 2022). Competition can lead to a rat race among households, which might be the source of the increasing trend of parental investment (Ramey and Ramey, 2010). Most previous work on parental investment does not include this competition aspect into their framework.

This paper investigates these two aspects of parental investment. First, this paper seeks to shed light on the role of parental investment on intergenerational persistence or earnings. The inclusion of self-effort of the child, which is often ignored in the literature, might amplify or offset the link between the two generations. Second, the extent of household competition is affected by the size of the cohort for the limited number of colleges. If there are significant changes in the number of competitors, the decision of parental investment and child effort is likely to be affected. In fact, many developed countries face a drastic shift in demographic structure caused by a low fertility rate, as shown in Figure 1. Little is known about the consequence of the shift in the demographic structure on parental investment.

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.

To answer these questions, this paper builds and estimates a dynamic tournament model using a unique Korean 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 fit is quite good. 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 measure the effects of the shrinking cohort size on parental investment.

This study uses Korean datasets and is based on institutional features of the country. With a homogenous secondary school environment, the private expenditure of parents stands out as a primary contribution to the child's future outcomes. The importance of the final test score 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.

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 worker'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. 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). Second, the amount of parental investment significantly drops as students finish the college admission process. This shows 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.

I also report the empirical facts suggesting that parental investment and the child's self-efforts potentially have different implications for intergenerational mobility. As expected, data show that the 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. Omitting self-efforts of the child might result in an exaggeration of intergenerational persistence of earnings, which suggests the importance of modeling both parental investment and the child's self-efforts in studying the research question. Finally, I report empirical facts that hint at the importance of modeling the dynamic decisions of the household. Students' time allocations of effort choices change considerably as they proceed to the later periods in secondary school. Also, the exogenous characteristics of the household persistently affect the parental investment decisions throughout the secondary school periods. Each household self-selects the different levels of parental investment and child efforts over time based on their preconditions.

Motivated by the empirical evidence, I develop and estimate a 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 each period, a household makes decisions of parental investment and the level of the child's self-efforts, and these two are inputs of the test score. To capture the student's persistence in test-taking skills, I allow the previous test score to have its own direct effect on producing test score (Cunha and Heckman, 2007). The model structure repeats until the final test score is produced. Students are assigned to the college tiers based on their final test score, and the college tier is the sole determinant of the child's lifetime income.

I estimate the model using maximum simulated likelihood. The estimation results suggest that both the marginal effects of parental investments and self-study of the child decline over time. Also, the estimate of the substitution parameter of the production function suggests that the parental investments and hours of self-study are close to perfect substitutes. Compared to hours of tutoring and hours of self-study, there exists sizeable unobserved heterogeneity in the quality of tutoring. The estimated model fits the sample well.

Using the estimated structural model, I first quantify the role of heterogeneity in household income. Removing heterogeneity in the parental income during the adolescent period decreases the rank-rank slope, the slope between income percentiles of two generations, by 47.2%. Next, I quantify the role of parental investments and the self-efforts of the child on intergenerational persistence of earnings. I use the rank-rank slope as the measure of intergenerational persistence of earnings (Chetty *et al.*, 2014). I simulate the model by shutting down one of the choices. In particular, I compare the changes in the rank-rank slope by shutting down the choice of (i) self-study of the child or (ii) the parental investments for the child. Relative to the estimated model, the rank-rank slope increases by 30.2% when the channel of self-study is shut down. Also, the rank-rank slope decreases by 79.5% when the channel of parental investments is shut down. 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, motivated by the ban on private tutoring activities by the Chinese Community Party (CCP), I investigate the effects of private tutoring expenditure on consumption inequality. The CCP banned for-profit private tutoring of core subjects such as math, science, and history (Forbes, 2021). The purpose of the policy is to reduce the child-rearing costs caused by private tutoring, which in turn encourages couples to have a child. Such a policy is likely to increase consumption of households. As a result of the counterfactual experiment, I find that banning private tutoring increases consumption inequality, which implies that the presence of private tutoring expenditure decreases consumption inequality. This is because high-income households spend more income on private tutoring expenditure on average, and consumption inequality among households decreases.

Lastly, to understand the effects of the shrinking cohort size on the choices of households, I simulate the structural model using the projected number of high school graduates and the assumptions on changes in the number of seats in colleges and changes in the distribution of college quality. College admission competition is about winning the competition within the cohort. The tournament model enables studying the effects of changes in cohort size and the distribution of college quality. The counterfactual simulation shows that there are virtually no effects of the changing cohort size on parental investment unless there are significant changes in the distribution of college quality.

1.1 Related Literature and Contributions

This paper contributes to the burgeoning literature seeking to understand the source of intergenerational mobility. The literature of intergenerational mobility has been focusing on reporting estimates of intergenerational persistence of earnings (Solon, 1999; Mazumder, 2005; Chetty *et al.*, 2014). Only recently, there appeared a few papers investigating the mechanism that generates the intergenerational correlations in earnings. One approach in this literature is to build a model and use calibration (Lee and Seshadri, 2019; Caucutt and Lochner, 2020; Daruich, 2022; Yum, 2022). Another approach is to build a model and use estimation. The first group of papers tend to employ a general equilibrium framework, while the papers using estimation tend to build partial equilibrium models. Bolt *et al.* (2021b) estimate their dynastic structural model using the National Child Development Survey, a longitudinal dataset of British households. They find that 62% of the variation in lifetime wages can be explained by the characteristics of the individuals when they were 23. Gayle, Golan and Soytas (2022) build a dynastic model of household behavior incorporating parental investment and assortative mating. Estimating their model using the Panel Study of Income Dynamics, they find that parental traits account for between 58% and 68% of the intergenerational persistence in earnings and that the marginal impact of assortative mating on intergenerational in earnings is modest.

I contribute to this literature by building and estimating a dynamic model of a tournament that incorporates parental investment and the self-efforts of the child. Previous studies quantifying intergenerational mobility do not consider the self-efforts of the child in their framework. Also, the novel feature of my paper is to incorporate the competition among households into the dynamic structural model.

This paper relates to the large body of literature modeling post-birth parental choice.[1] Since Becker and Tomes (1979), economists have sought to understand how parents allocate resources to their children and how such decisions affect the child's outcomes such as cognitive development (Doepke *et al.*, 2019). Del Boca, Flinn and Wiswall (2014) build and estimate a dynamic model in which parents jointly choose the amount of time investment, amount of monetary investment, and decision of labor supply participation. Doepke and Zilibotti (2017) formulate a model of parenting styles. Agostinelli *et al.* (2020) extend their work by combining the choice of parenting with the child's peer formation. Papers in this literature have recently started incorporating externalities into parental choices. For example, a group of papers associates parental choices with social interactions (Agostinelli, 2018; Agostinelli *et al.*, 2020; Boucher *et al.*, 2022). While providing interesting implications, identification of social interaction is difficult, and the relevant data is often unavailable (Manski, 1993). Papers also model externalities by incorporating competition among students. 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. The theoretical model in this paper is different from theirs in several ways. First, I model the dynamic tournament allowing for uncertainties in the test score generation and the household choices. Second, I model the channel of monetary investment, which has direct implications for intergenerational transmission of earnings. Third, I also incorporate self-efforts of the child to the model.

The two closest papers implementing student competition are Bodoh-Creed and Hickman (2019) and Grau (2018). 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. Also, Grau (2018) builds a static tournament model, estimates it's parameters and applies the estimated model to the college competition in Chile. Their static models abstract away from the measure of the resources and the efforts used for the college admission competition. I propose a dynamic tournament model and suggest plausible measures for the resources (household income) and the ef-

[1] See Chiappori, Salanié and Weiss (2017) for a model of joint decision of marriage and parental investment.

forts (private tutoring expenditure and hours of self-study) of the competition. By making the model dynamic, I can estimate the changing effects of parental investment and self-efforts of the child. Also, a dynamic model help capture how household self-select into the high and low level of investments. [2]

Another closely related paper is by Kim, Tertilt and Yum (2022), which studies the cause of the low fertility problem of South Korea. They propose a heterogeneous-agents 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. They also endogenize the fertility decisions of the parents. In their framework, the quantity-quality tradeoff governs the choice of the household where quality is where the status externality comes in. Using their calibrated model, they find that the absence of the status externality would increase fertility by 15%. The tournament model of this article complements their study by formally modeling the dynamic competition with respect to getting into prestigious colleges. The tournament structure can rationalize the underlying source of the status externality in their paper. Also, while I do not explicitly model fertility choice, the tournament model suffices to study the effects of the demographic structure shift on parental investment.

This paper contributes to the literature on childhood investments and skill development by estimating the effects of parental investment and the self-efforts of the child in the adolescent period. Most previous work focuses on estimating the effects of parental investment on child outcomes alone (e.g., Cunha and Heckman (2007), Cunha, Heckman and Schennach (2010), and Del Boca, Flinn and Wiswall (2014)). These studies find declining effects of parental time investment over time. Several studies estimate the effects of hours of self-study on academic achievements (e.g., Cooper *et al.* (2006); Stinebrickner and Stinebrickner (2008)), but they do not jointly estimate the effects of parental investments. Only recently, a few papers have estimated models incorporating both parental investment and self-efforts of the child. Del Boca, Monfardini and Nicoletti (2017) find that the effect of self-effort of the child is stronger than the effect of the mother's time investment during adolescence, and the effect of self-effort of the child increases over time. Del Boca *et al.* (2019) build a Stackelberg model of parent-child interaction and study the effects of conditional-cash-transfers on child outcomes. Such a line of research suggests the importance of modeling both parental investment and self-efforts of the child in studying the source of intergenerational mobility, which is the focus of this paper.[3][4]

As this paper employs private tutoring expenditure as a measure of parental investment, it also complements the literature of studies on private tutoring. Most previous studies exclusively focus on measuring the effects of private tutoring expenditure (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). Kang and Park (2021) take a step forward and report the heterogeneous effects of private tutoring expenditure among students with different pre-existing conditions. They find that there exists

[2] 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).

[3] Agostinelli and Sorrenti (2021) find that the trade-off between more household income from labor supply and parental time investment is significant for disadvantaged families (mothers).

[4] 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) and Agostinelli and Wiswall (2016)).

a significant selection into different amounts of private tutoring expenditure depending on household characteristics.[5] Almost all previous studies do not account for the dynamic selection. One exception is Choi (2013), who builds and estimates a dynamic model of hours of tutoring and hours of study. My paper contributes to the literature by providing "selection-corrected" estimates of the effects of private tutoring expenditure. Also, previous studies do not separately estimate the changing effects of private tutoring expenditure over time. The other important factor that is not considered in the previous studies is the cut-off effects of prestigious colleges (Eide, Brewer and Ehrenberg, 1998; Brewer, Eide and Ehrenberg, 1999; Zimmerman, 2019; Sekhri, 2020; Guo and Leung, 2021; Jia and Li, 2021) on private tutoring expenditure. The impact of private tutoring expenditure might be underestimated if it does not consider the "elite college premium," which is the key mechanism in the tournament model.

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

2 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.

2.1 High-Stakes College Entrance Exam

In Korea, the College Scholastic Ability Test (CSAT), the college entrance exam taking place at the end of 12$^{\text{th}}$ grade, is the single most important factor for college admission.[6] 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.[7]

[5] In Section 3.5, I present the evidence that households with favorable initial conditions tend to increase their tutoring expenditure, but households with disadvantaged initial conditions tend to decrease their average tutoring expenditure over time.

[6] 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).

[7] 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

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. 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. Baccalauréat of France is highly important for getting into grandes écoles, the group of elite colleges of the country. 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.

2.2 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 significantly affects a student's future labor market outcomes.[8] In South Korea, the college hierarchy has changed little (Kim and Lee, 2006). Starting from the top institution, Seoul National University, the applicants' preference has been stable for decades, and "SKY" is a well-known acronym that refers to the top three universities in the country. In 1980s, as the demand for elite college has increased, the SKY universities have become too far of a reach for many people. Then, the relatively new recent 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. 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.

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, 3 and 4 are approximately the top 5%, 10%, and 15% of the CSAT score distribution, respectively.[9] 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 3, I present empirical evidence suggesting the significant effects of the college tier on graduating labor market outcomes.

level of the students due to the grade scheme shift.

[8] 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.

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

2.3 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.[10] Second, as a result of the consecutive school-equalization policies, the quality of education provided by schools is similar.[11] No schools are allowed to select students independently.[12] 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.[13]

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.[14] 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.[15] 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.

3 Empirical Evidence

3.1 Data

I use the Korean Educational Longitudinal Study 2005 (KELS) for the main estimation procedure. To supplement the income information of KELS, I use the Korean Labor Income and Panel Study (KLIPS)

[10] One of the few autonomies 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.

[11] 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.

[12] 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.

[13] 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.

[14] See Bray (1999; 2021) for comprehensive cross-country comparisons of private tutoring.

[15] [Cite statistics on the proportion of parents helping their kids]

to supplement the college tier-specific lifetime income.

3.1.1 Korean Educational Longitudinal Study

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 intergenerational mobility, and (ii) to account for the dynamic selection of the effort choices of the household. 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 provides a rare combination of relevant data. The dataset 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. 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 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.

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

Table 1: Data Selection

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 1. 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. In Appendix 8, I show that the effects of the selection does not result in severe differences in the sample moments. The observations lost to income selection tend to have missing CSAT scores as well. Importantly, 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.

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					

(a) Sample Moments: 7th - 9th 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					

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

Table 2: Sample Moments

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

(a) Sample Moments: Other characteristics

Table 3: Sample Moments (Continued)

Table 2 and 3 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 3.5. 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. 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.

3.1.2 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 the workers graduated from, their 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 that of KLIPS. Employing KLIPS is more useful in predicting alumni's lifetime income as it contains data on workers of age between 20 and 65.[16]

[16] 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.

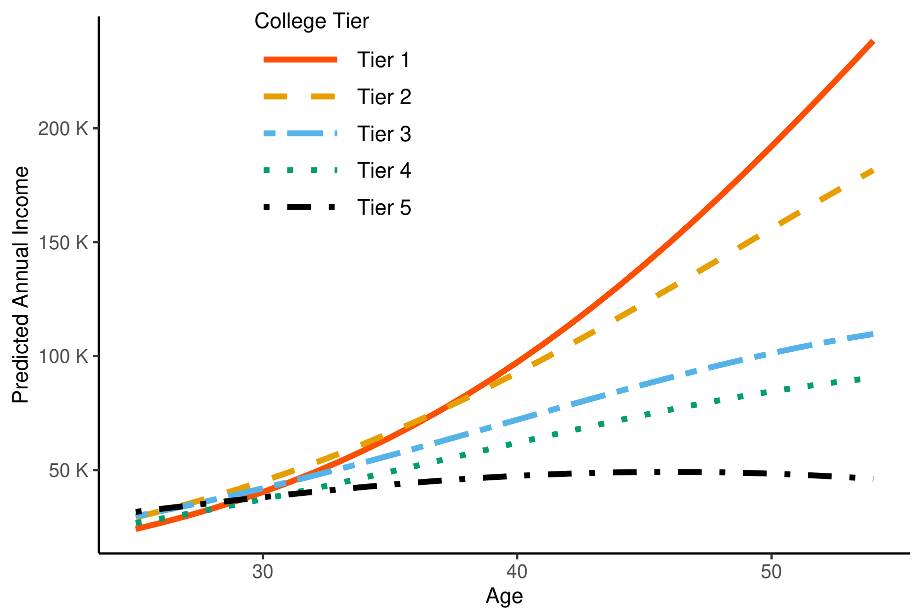
Table 4: Log Income Regression

	(1) Pooled OLS	(2) Pooled OLS	(3) RE	(4) RE
College Tier				
Top tier	-1.931***	-1.684***	-1.671***	-1.491***
	(0.347)	(0.325)	(0.279)	(0.288)
Second Tier	-1.332***	-1.195***	-1.409***	-1.364***
	(0.350)	(0.296)	(0.321)	(0.323)
Third Tier	-0.864**	-0.817***	-0.958***	-1.075***
	(0.269)	(0.196)	(0.363)	(0.364)
Fourth Tier	-0.895***	-0.618**	-0.727***	-0.524***
	(0.232)	(0.186)	(0.079)	(0.123)
age	0.092***	0.092***	0.095***	0.095***
	(0.001)	(0.001)	(0.002)	(0.002)
Interactions				
Top tier × age	0.067***	0.065***	0.058***	0.058***
	(0.009)	(0.009)	(0.009)	(0.009)
Second Tier × age	0.050***	0.052***	0.051***	0.055***
	(0.011)	(0.010)	(0.010)	(0.011)
Third Tier × age	0.032***	0.037***	0.034***	0.044***
	(0.008)	(0.007)	(0.012)	(0.013)
Fourth Tier × age	0.029***	0.027***	0.022***	0.021***
	(0.007)	(0.007)	(0.002)	(0.002)
N	29599	29599	29599	29599
Major	No	Yes	No	Yes
RE	No	No	Yes	Yes

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.

3.2 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 4.

College ranking has a strong effect on the growth of alumni's income. Columns (1) and (2) in Table 4 provide the OLS estimates for the regression equations,

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

where $D_{i,j}^{Tier}$ is the dummy variable indicating that person i graduated from a tier j college, and Z_i is the set of explanatory variables including age, squared age, birth year, and gender of person i.[17] Note that the regression equation captures both the effects of graduating from college tier j on the level and the growth of alumni's income, respectively by β_j and δ_j. Columns (3) and (4) provide the estimates of the random effects model,

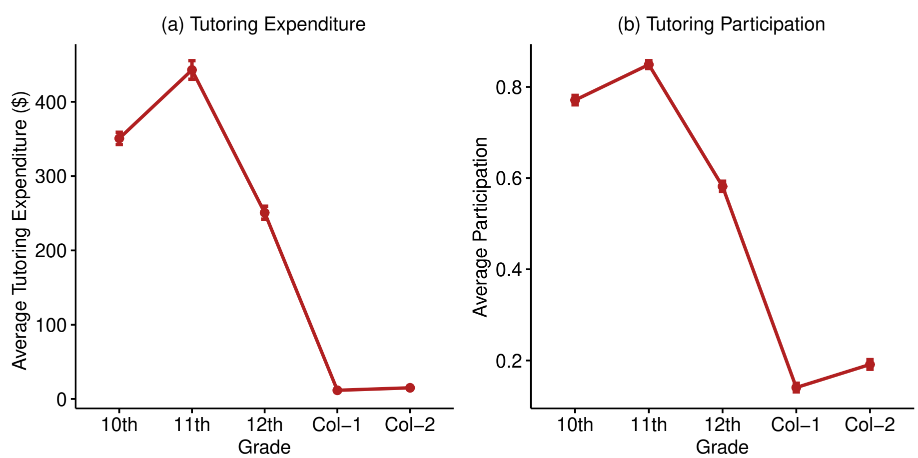
$$\ln y_{it} = \sum_{j=1}^{J} (\beta_j + \delta_j \cdot age_{it}) D_{i,j}^{Tier} + Z_i \gamma + \lambda_i + \eta_{it}$$

[17] The purpose of the birth year dummy variable is to capture the cohort difference in workers' income.

where λ_i and η_{it} are the individual-specific and the idiosyncratic errors respectively.[18] Columns (2) and (4) include the dummy variables of college-major, showing that the inclusion of major does not critically affect the main results of Columns (1) and (3), 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 4. 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 estimation results are consistent with the studies stressing the importance of using lifetime income in the returns to schooling literature (NEED SOME CITES). 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.

3.3 Competition Motives of Parental Investment

Figure 3: Private Tutoring Expenditure and Participation in Tutoring



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

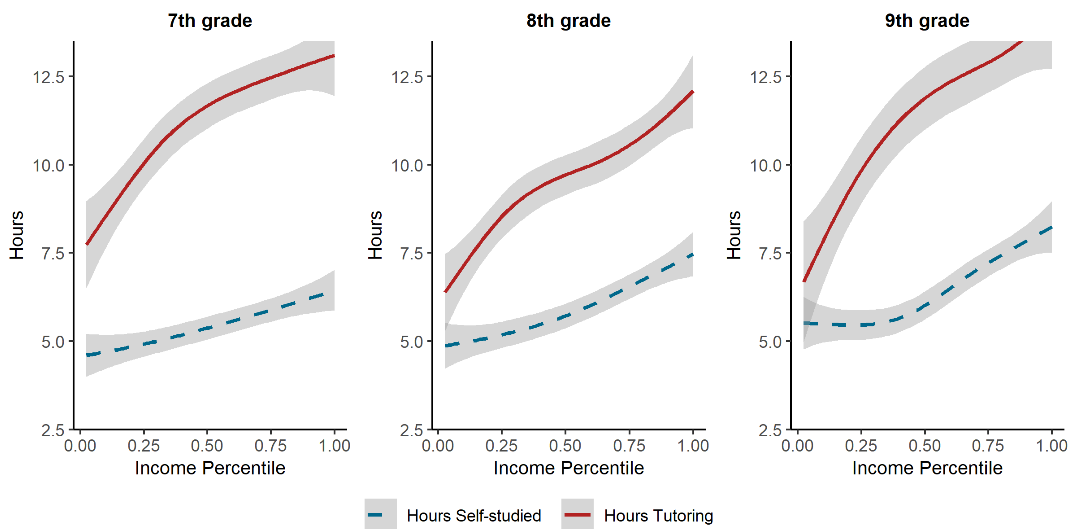
Competition with respect to getting into a more prestigious college is the primary motivation of parental investment. First, data suggest that the demand for private tutoring expenditure significantly drops as students finish the college admission process. Figure 3 presents the change of tutoring expenditure 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

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

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

3.4 Parental background and child's hours allocation

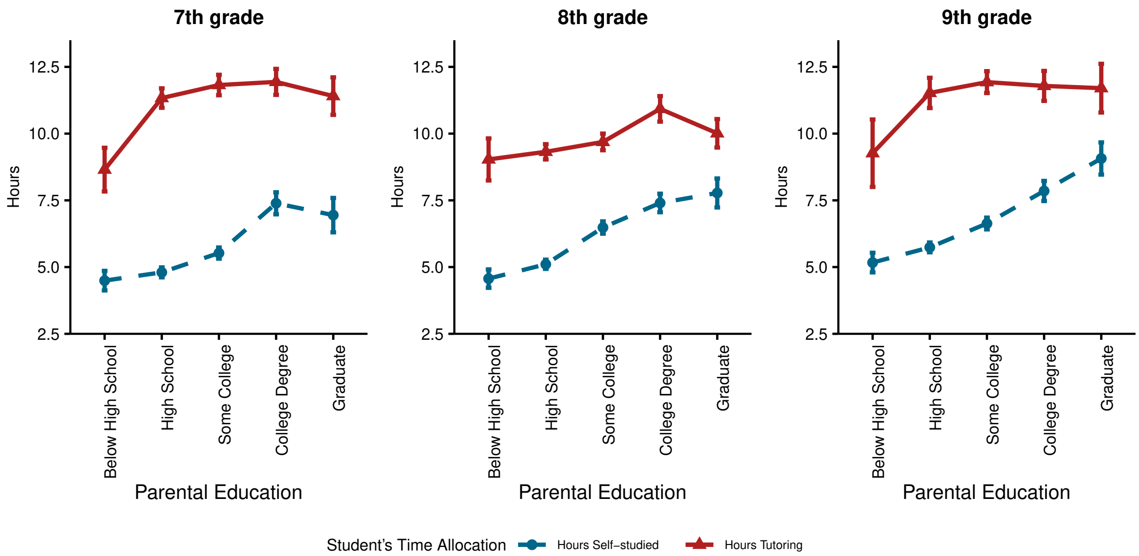
Figure 4: Income Gradient in Effort Decision



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 include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, and household income.

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 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 education m_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 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. Table 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) to (3) present the results where y_{it} is hours of self-study, and columns (4) to (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 5: 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***	0.152***	0.036
	(0.022)	(0.025)	(0.027)
Parental Edu		0.055***	
		(0.007)	
N	10454	10454	10454
Year	Yes	Yes	Yes
FE	No	No	Yes

	(1) log(1+Tutoring)	(2) log(1+Tutoring)	(3) log(1+Tutoring)
log(Income)	0.677***	0.616***	0.269***
	(0.027)	(0.030)	(0.037)
Parental Edu		0.038***	
		(0.008)	
N	9431	9431	9423
Year	Yes	Yes	Yes
FE	No	No	Yes

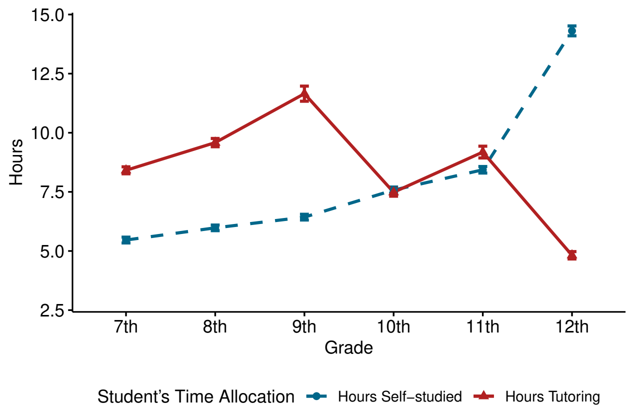
Source: Korea Educational Longitudinal Study 2005, Korean Educational Development Institute.
Note: log(1+Study) and log(1+Tutoring) refer to log of hours of self-study plus one and hours of tutoring plus one, respectively. I include households who do not have missing information on the following variables: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

3.5 Dynamic effort allocation of households

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.[19]

[19] 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 my best knowledge, there is no study reporting the changing effects of self-study over time.

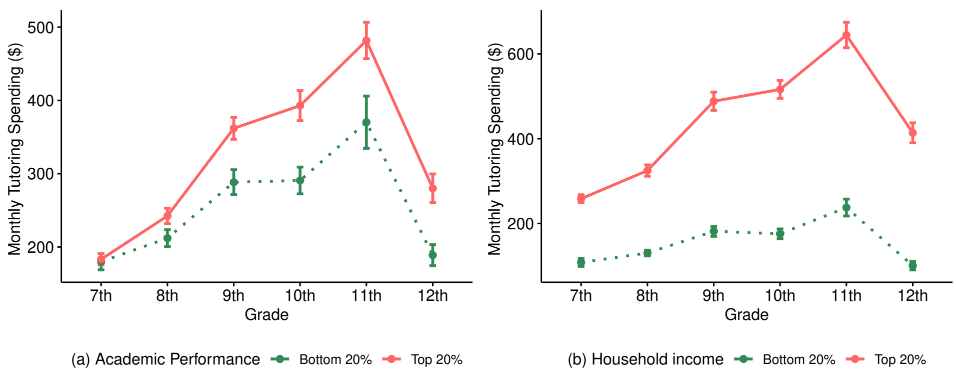
Figure 6: Dynamic allocation of efforts



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

The initial conditions of the household persistently affect the parental investment decisions throughout the secondary school periods. Figure 7 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 7 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 7 (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 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 7 (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 7: 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 include households who do not have missing information on the following variables: tutoring expenditure, CSAT scores, and household income.

The evidence suggests that households self-select into the different effort levels based on their preconditions, and the allocation of the two efforts changes over time. 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.

4 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 initiated by Lazear and Rosen (1981) and related to its applications in college admission competition (Han *et al.*, 2016; Grau, 2018; Tincani *et al.*, 2021).

4.1 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.

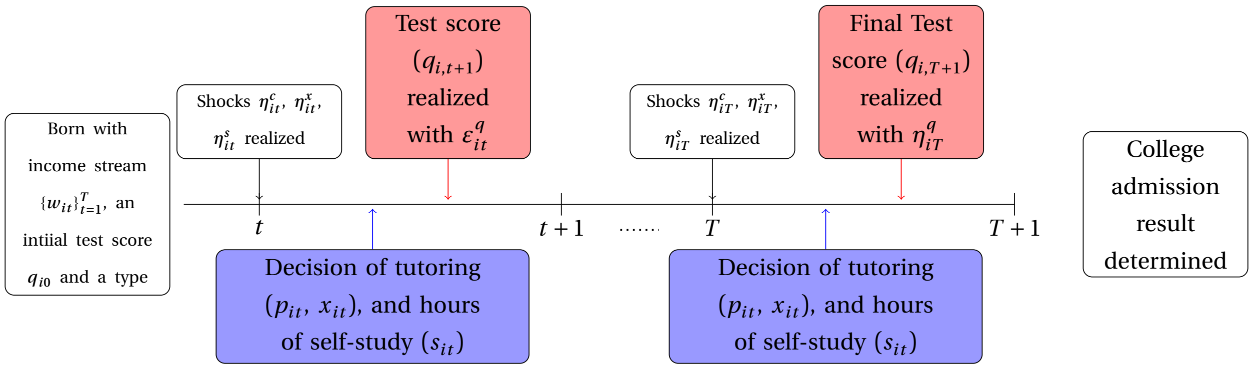


Figure 8: Model Timeline

Figure 8 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_{i0}. Also, each household has a specific type k. Different types of household have different person-specific characteristics that are unobserved by the econometrician. I define them as λ_k^c, λ_k^x, λ_k^s, and λ_k^q, which affect marginal utility from consumption, disutility from hours of tutoring, disutility from hours of self-study, and log of test score, respectively. Some households might value non-academic goods, such as travel, more than other households conditional on the observed characteristics (Lazear, 1977). Such unobserved taste for consumption would be captured by λ_k^c. Some households might prefer to encourage their child to study independently rather than send him/her to tutors, which would be 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.[20]

At each time t, as the household enters into period t, 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, 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 $(J-1)^{th}$ 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

[20] I introduce the joint distribution of the time-specific shocks (η_{it}^c, η_{it}^x, η_{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.

for the $(J-1)^{th}$ college tier and the students who do not go to college.[21] The assigned college tier is the sole determinant of ex-post lifetime income.

4.2 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.[22] 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 4.[23] As tier 1 is defined to be the top college tier, v decreases in j (i.e., $v_1 > v_2 > ... > v_{J-1} > v_J$).[24]

For student 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, $\{\tilde{Q}_j\}_{j=1}^J$ is where the competition comes in. In order for a student to be in tier j or better, she has to be above her 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.[25]

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

[21] 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%.

[22] The average interest rate is around 5% for South Korea in 2010. The corresponding discount rate is approximately 0.95.

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

[24] 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 empirical quantification of the consumption value of college.

[25] 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 *et al.* (2021) for the case that resulting cutoffs significantly deviate from the guessed cutoffs.

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 et al. (2022) as supporting evidence for Assumption 2. Using a dataset of one of the big 5 companies of Korea, they found that the effects of college reputation dominates the effects of college GPA on receiving an offer from the firm. 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. Also, other papers in the literature show that much of the variation in lifetime earnings happen before age 23 (Bolt *et al.* 2021b; Huggett, Ventura and Yaron, 2011).

Table 6: Job Offer Regression

	(1) probit
Tier=1	0.221***
	(0.039)
Tier=2	0.056***
	(0.005)
Tier=3	0.002
	(0.028)
ColGPA	0.006***
	(0.002)
N	9132

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.

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.[26] The total amount of tutoring

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

expenditure e_{it} is

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

The tutoring expenditure is constrained under two dimensions. A household cannot spend more tutoring expenditure than its income (i.e., $e_{it} \leq w_{it}$).[27] 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 [Support the claim using other datasets]. A couple of 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.

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.[28] 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

[27] I assume no borrowing.

[28] 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.

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 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 δ_{et} and δ_{st} determine the relative importance of the tutoring expenditure and hours of self-study, respectively. I define ε_{it}^q as a combined shock of λ_k^q and η_{it}^q, which is specified as $\ln \varepsilon_{it}^q = \lambda_k^q + \eta_{it}^q$.

4.3 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 ($z = c, x, s$), 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.[29] 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,

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

where $Prob(\ln \tilde{Q}_{j-1} \geq \ln q_{i,T+1} \geq \ln \tilde{Q}_j | \Gamma_{iT})$ 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 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_{it}^q. With

[29] 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.

the log-transformation, the terminal value is specified as

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

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. $\ln\tilde{g}_{ij} = \ln\tilde{Q}_{j-1} - \ln\widehat{q_{i,T+1}} - \lambda_k^q$), and F is the distribution of η_{it}^q. I assume F follows normal distribution in the spirit of rank-order tournament model (Lazear and Rosen 1981; Han *et al.* 2016; Grau 2018; Tincani *et al.* 2021).[30]

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 given by

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

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.[31]

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 $\{w_{it}\}_{t=1}^{T}$. The set of unobserved state variables Ψ_{it} includes the set of unobserved shocks and the type specific heterogeneity. Based on the timeline, the time-varying shock regarding test score is not an unobserved state variables. (i.e., $\Psi_{it} = \{\eta_{it}^c, \eta_{it}^x, \eta_{it}^s, \lambda_k^c, \lambda_k^x, \lambda_k^s, \lambda_k^q\}$.)

Information and Uncertainties: I assume a continuum of households and that the distribution of household is public information.

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

As stated in Assumption 1, each household correctly anticipates the set of college tier cutoffs $\{\tilde{Q}_j\}_{j=1}^{J}$.[32] They know the distribution of the final test scores in advance and make dynamic choices based upon the perfect guess.

[30] One can also adopt a functional form that η_{it}^q follows Generalized Extreme Value distribution which makes a Tullock (2001) contest.

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

[32] 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.

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

Also, 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 depicted in Figure 8, 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, they do not know about the test score shock η^q_{it} before they make a decision. Therefore, they make 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.

Choice	Notation	Incentives	Disincentives
Quality of Tutoring	p_{it}	Lifetime income of the child	Income
Hours of Tutoring	x_{it}	Lifetime income of the child	Income, disutility of tutoring
Hours of Self-study	s_{it}	Lifetime income of the child	disutiltiy of self-study

Table 7: Incentive structure of the model

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, household i solves

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

subject to (4), (7), and (8), where $\Gamma_{it} = \{Z_{it}, \Psi_{it}, \{\bar{Q}_j\}^J_{j=1}\}$ is the set of information upon making the decision. Each household faces a tradeoff between current flow utility and future payoffs. The incentive structure of the model is summarized in Table 7. Each choice variable incurs costs associated with the choice. In particular, investing more in parental investment (i.e., increasing p_{it} or x_{it}) requires suffering more from the disutility 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 of the final score occurs, the value function is specified as

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

where α_v is the altruism parameter. The altruism parameter implies the "exchange rate" between the current 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 a household thinks the child is "worth" sacrificing the current utility flow, it would exert more efforts using either parental investment or the child's self-efforts.

4.4 Equilibrium of the Tournament

I first describe the definition of the dynamic equilibrium of the tournament model. Then I show the existence of the equilibrium using the Schauder Fixed-Point Theorem (Amir, 1996; Fey, 2008; Mertens and Judd, 2018; Engers, Hartmann and Stern, 2022). Finally, I show that the model equilibrium is unique with the provided assumptions.

Definition 1. Given the set of initial conditions and Assumptions 1, 2, and 3, a Markovian equilibrium of the model is a vector $\mathcal{K}^* = \{\{\Upsilon^*\}_{t=1}^T, \{\tilde{Q}_j^*\}_{j=1}^J\}$ such that:

1. $\{\Upsilon^*\}_{t=1}^T$ is a set of policy function that solves the household's maximization problems in equations (9) and (10) for every period t, subject to (4), (7), and (8).
2. The set of the final test score cutoffs $\{\tilde{Q}_j^*\}_{j=1}^J$ is produced with the test score function described in equation (4) and it is consistent with the policy function set $\{\Upsilon^*\}_{t=1}^T$.

I define a mapping $\aleph$ on a set $\mathcal{K} = \{\{\Upsilon\}_{t=1}^T, \{\tilde{Q}_j\}_{j=1}^J\}$. Using equations (9) and (10), and the constraints (4), (7), and (8), the mapping determines the value functions and the policy functions for each period t. Then, using the policy functions, it generates the distribution of test scores for each period t. As a result of the forward simulation, the set of final test score cutoffs $\{\tilde{Q}_j\}_{j=1}^J$ is produced. The equilibrium is a vector of policy functions and the test score cutoff $\mathcal{K}^* = \{\{\Upsilon^*\}_{t=1}^T, \{\tilde{Q}_j^*\}_{j=1}^J\}$.

Lemma 2. *Denoting $V_{it}^*(p_{it}, x_{it}, s_{it}; Z_{it}, \Psi_{it})$ as the choice specific value, the set of choice specific value $\{V_{it}^*(p_{it}, x_{it}, s_{it}; Z_{it}, \Psi_{it})\}_{t=1}^T$, the set of policy function $\{\Upsilon\}_{t=1}^T$, and the value function $\{V_{it}(Z_{it}, \Psi_{it})\}_{t=1}^T$ are compact.*

Proof. [In Appendix A.1] □

Lemma 3. *The mapping $\aleph$, which determines value functions $V_{it}(Z_{it}, \Psi_{it})$, the policy functions $\{\Upsilon\}_{t=1}^T$, and the set of test score cutoffs $\{\tilde{Q}_j^*\}_{j=1}^J$, 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. □

Assumption 5. *For the final period, the value function is concave.*

Theorem 5. *The equilibrium of the tournament model is unique*

Proof. Suppose one extreme equilibrium where all the households spend private tutoring expenditure as much as their income and all hours for either hours of self-study or hours of tutoring (i.e., $w_{it} = p_{it}x_{it}$ and $h = s_{it} + x_{it}$). In this extreme equilibrium, each household always has an incentive to cut down their investment and efforts to increase their flow utility. Suppose the other extreme equilibrium where all the households spend zero amount of private tutoring expenditure, hours of tutoring, and hours of self-study. In this extreme equilibrium, each household always has incentives to increase its investment and efforts to increase its future value. Therefore, both extreme equilibria are ruled out. The equilibrium is a set of policy function resulting from the household maximization problem. $\square$

4.5 Properties of the Model

The dynamic tournament model offers several features that help answer the research question of this paper. First, the rich heterogeneity of state variables and the choice set, and the specification of the test score production function help disentangle the source of intergenerational persistence of earnings. Each household can simultaneously choose the quality of tutoring, hours of tutoring, and hours of self-study in the model based on their state variables. The test score production function allows for a variety of inputs: previous test score q_{it}, quality of tutoring p_{it}, hours of tutoring x_{it}, and hours of self-study s_{it}. This enables me to separately quantify the impact of parental investment, self-efforts of the child, and other household characteristics on intergenerational persistence of earnings. In the model, each household can choose the other mode of investment even if they are not allowed to use one of the options. Thus, the model provides an opportunity to simulate the reaction of the household when one of the choices is restricted. In Section 7, I quantify the role of each choice by simulating the model with shutting down that particular choice from the model. Then I compare the simulation results with when all choices are allowed.

Second, the rank-order feature of the tournament model enables me to study the effects of the changes in cohort size and the role of disparity in college quality. Since the tournament model is about obtaining a limited number of seats in the better colleges, the model can be used to evaluate the exogenous changes in the number of competitors. As described in Section 1, countries with high private tutoring expenditure face a sharp decrease in total fertility rate. The reduction of cohort size due to declining fertility means there is a less number of competitors for higher college tiers, assuming there is no change in the number of seats for colleges.[33] The model provides an opportunity to simulate parental investment of households given the cohort size is reduced. The tournament model can also be used in assessing the role of disparity in college quality on parental investment. The well-known feature of the tournament model is that the size of the prize differential affects the effort choice of the agents (Lazear and Rosen, 1981). The prize differential in my model is the income differential of the higher college-tier. Such a feature of the tournament model captures the role of the distribution of college qualities on the investment decision of the household.

[33] In Section 8 I show changes in the number of seats in colleges and the number of high school graduates using administrative data of South Korea.

Third, as I allow for the time-varying effects of the choices, I can compare the effects of hours of self-study and hours of parental investment.

5 Estimation Strategy

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.

5.1 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

$$\mathscr{L}_i(\theta|q_{i0}, \{w_{it}\}_{t=1}^T, m_i) = \sum_{k=1}^{K} \left\{ \left(\Pi_{t=1}^T \mathscr{L}_{it}(\theta|Z_{it}, \Lambda_k) \right) \Pr(type = k) \right\} \tag{11}$$

which is conditional on the initial test score q_{i0}, the income stream $\{w_{it}\}_{t=1}^T$, and parental education m_i. The time-specific likelihood contribution $\mathscr{L}_{it}(\theta|Z_{it}, \Lambda_k)$ can be characterized in four different ways, which is determined by the combination of the tutoring-participation dummy variable d_{it}^x and self-study participation dummy variable d_{it}^s. In particular,

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

where $d_{it}^x = 1$ means that household participate in tutoring, 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 product of individual likelihood contributions,

$$\mathscr{L}_i(\theta) = \sum_{k=1}^{K} \left(\Pi_{t=1}^T \mathscr{L}_{it}(\theta|Z_{it}, \lambda_k) \right) \cdot \Pr(type = k).$$

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

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

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

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 the adoption of a good simulation method. To reduce the variance of simulation error, I use antithetic acceleration (Geweke, 1988; Stern, 1997; Stern and Zhou, 2018).

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 fill up 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}. In Appendix D, I show the derivation of the density and probability I use for computing the likelihood function, and I explain the simulation of unobserved variables.

5.2 Identification

The parameters of the structural model are jointly identified using the equilibrium conditions of the model. Specifically, the shock expression derived from the First Order Conditions of the structural model and other model definitions including the test score production function, jointly identify the parameters. The shock expressions are specified in Appendix D. What are exogenously given in the model are the academic performance in primary school q_{i1}, parental education m_i, and complete income stream of parents $\{w_{it}\}_{t=1}^T$.

Parameters of the model can be classified into the productivity parameters associated with the test score function and the parameters that directly affect value function. The productivity parameters in the test score production function are identified jointly using two conditions: (i) the covariation between the resulted test score $q_{i,t+1}$ and its inputs (q_{it}, p_{it}, x_{it}, and s_{it}) and (ii) the first order conditions govern the household choices, namely equations (12), (13), and (14). That is, the value function and the test score function altogether determines the productivity parameters. For example, δ_{et}, the parameter of the marginal effects of tutoring expenditure, is identified by the covariation between the effective tutoring expenditure term $\tilde{e}_{it}$, which is described in equation (3), and the first order conditions with respect to p_{it}, x_{it}, and s_{it}. A similar identification argument applies to the parameter of the marginal effects of self-study δ_{st}. As the data of the input variables are available for each period, I can separately identify the productivity parameters for each time t. The log of the total factor productivity term, $\ln A_{it}$, is the constant of the test score function. The constants are not independent from the equilibrium conditions of the tournament model. The tournament uses the final test score $q_{i,T+1}$ for assigning students to each college, and the previous test score, $q_{i,t}$ for $t < T$, are an argument of the expected value function, EV_{t+1}, for $t < T$. The substitution parameters ϕ and the effort parameters ν_t are also jointly identified using the test score functions and the equilibrium conditions.

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 by time. The element of the covariance matrix of the shocks are identified in maximizing the log-likelihood contribution of the associated shocks.

6 Estimation Results

6.1 Test score function parameters

Time-varying Parameters	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$
Previous	0.163	0.980	0.754	0.731	0.474	0.426
Test Score (δ_{qt})	(0.001)	(0.002)	(0.001)	(0.001)	(0.009)	(0.001)
Effort Parameters (ν_t)	0.556	0.659	0.420	0.360	0.150	0.020
	(0.001)	(0.014)	(0.009)	(0.001)	(0.002)	(0.001)
Share of tutoring	0.472	0.485	0.480	0.526	0.592	0.737
Expenditure (δ_{2t})	(0.007)	0.005)	(0.011)	(0.011)	(0.012)	(0.015)
Constants (δ_{0t})	4.030	-1.124	0.627	1.205	1.022	4.077
	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)
Time Invariant Parameters						
Substitution Parameter (ϕ)	0.880					
	(0.001)					
Intensity of Private tutoring	0.145					
Quality (κ)	(0.002)					

Table 8: Parameter Estimates: Test score production function

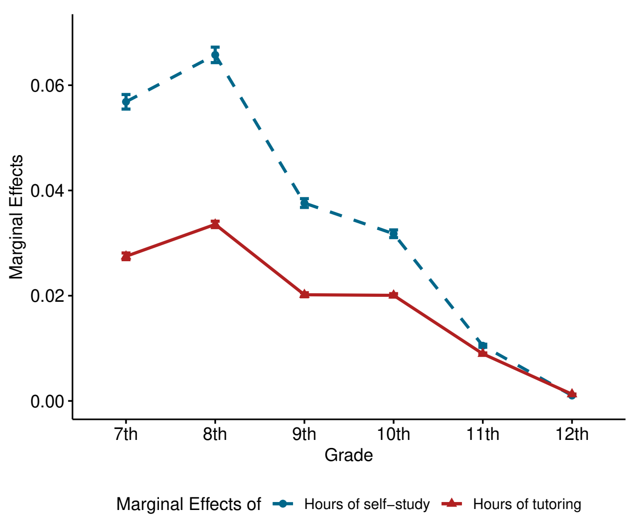
Table 8 presents the estimates of test score production function specified in equation (4). The interpretation of the previous test score parameter is the same as the log-log case of a linear regression equation. For example, for $t = 2$, an 1% increase in the previous test score leads to a 0.98% increase 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 relatively lower and it is more difficult to increase with the same amount of monetary or time investments.

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 be different by households because the calculation of marginal effects involves the data of p_{it}, x_{it}, and s_{it}. Figure 9 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 effects of hours of tutoring is slightly larger than the average marginal effects of hours of self-study. However, as it can be seen in 9, the difference is not statistically or economically significant.

The early period investments have delayed effects through evolving test scores. As it can be seen in Figure 9, the marginal effects of hours of self-study are already stronger in the earlier periods. Taking the delayed effect of hours of self-study through the evolving test scores ("self-productivity") into consideration, the hours devoted to self-study have strong effects on the final test score.

The hours of self-study and hours of tutoring are nearly perfect substitutes for each other based on the structural estimate. Table 8 includes the estimate of the substitution parameter ϕ, which is approximately 0.8798. The estimate of 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. Suppose hours of self-study is included as a choice variable. If a researcher conducts a counterfactual of quantifying intergenerational persistence of earnings using this one-choice model, a household cannot do anything against the restriction that they do not have an option of parental investment. Such an omission of mechanism might result in an exaggeration of the effects of parental investment. This substitutability plays an important role in computing intergenerational persistency because the channel of hours of self-study provides a household with a restrictive income constraint an opportunity to exert efforts in the tournament.

Figure 9: Average Marginal Effects of Hours Allocation



6.2 Preference and shock parameters

Table 9 presents the estimates of the preference parameters and the shock parameters. The preference parameters are parts of equations (5), (9), and (10). For the preference parameters, the estimates are the 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, $\exp(\tau_s m_i)$ 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. To esitmate the effect of higher parental education, I estimate τ_x and τ_s. Table 9 (a) presents 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, which captures that students with more parental education have less disutility from 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 heterogeneity of data. Regarding the unobserved heterogeneity, the unobserved heterogeneity from consumption shows the largest difference between the types of individuals. This could be due to the fact that I do not model the supply side of private tutoring, and the price difference across regions are not captured by the deterministic parts of the model.

		Estimate	Standard error
Preference Parameters			
Taste for consumption	α_c	0.028	(0.000)
Altruism for the child's future	α_v	1.018	(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	-0.001	(0.002)
disutility from hours of self-study	τ_s	-0.005	(0.002)

(a) Preference parameters

Standard Deviation of		Estimate	Standard Error
Test score shock	$\sigma_{\eta q}$	0.230	(0.000)
Consumption shock	$\sigma_{\eta c}$	0.742	(0.014)
Study shock	$\sigma_{\eta c}$	0.549	(0.001)
Tutoring shock	$\sigma_{\eta l}$	0.472	(0.005)

(b) Shock parameters

$\frac{1}{N\times T\times 6}(\sum logL_i - Jacob) = -0.848$

Table 9: Parameter Estimates: Preference and Shock Parameters

6.3 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 $\hat{y}$ fits the actual data value y, for dependent variables $y = p, x, s, q$. Specifically, the expected data value conditional on the model predicted value is $E(y|\hat{y}) = \widehat{\kappa_0}(\hat{y})$,where

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

and

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

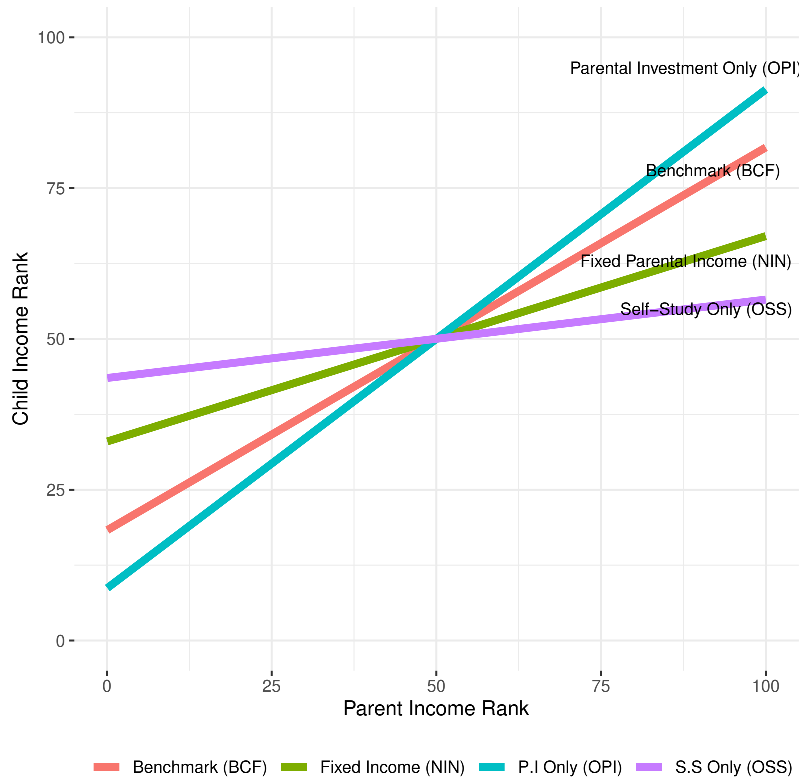
is the kernel function and b is the bandwidth. The farther the kernel curve deviates from the 45 degree line, the less the model is successful in fitting the data.

Figures E.1 to E.5 show the sample fit of the dependent variables of the likelihood function. Figures E.1 to E.5 present the sample fit of quality of tutoring, hours of tutoring, hours of self-study, and test scores, respectively. Overall, the fits are very good. While the quality of tutoring and the hours of self-study show excellent fits, hours of tutoring are somewhat a little bit overpredicted. The level of the final test score, as depicted in Figure E.5b, is somewhat overpredicted as well. 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 E.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.

7 Counterfactual Analyses

7.1 Quantification and Decomposition of Intergenerational Persistence of Earnings

Figure 10: Intergenerational Persistence of Earnings by Scenarios



The purpose of the quantification 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 role of the relevant channels. Each simulation produces a different distribution of the final test scores, which leads to a different distribution of the predicted income of the child. I define the predicted income of child as $childinc_i$. Two measures of the intergenerational persistence of earnings are intergenerational elasticity of earnings (IGE) and the rank-rank slope. In particular, IGE is the coefficient of the regression equation,

$$\ln childinc_i = \delta_{00} + \delta_{IGE} \ln hhinc_i + \varepsilon_{it},$$

where $hhinc_i$ is the average of household i's income over the periods. The estimate of IGE is $\widehat{\delta_{IGE}}$, and (1-$\widehat{\delta_{IGE}}$) is the measure of intergenerational mobility (Black and Devereux, 2010). On the other hand, the rank-rank slope (Chetty *et al.*, 2014) is the estimate of the regression equation,

$$R_i = \delta_{01} + \delta_{RR} P_i + \upsilon_{it},$$

where R_i is the rank of the child income within the generation, and P_i is the rank of the parental income within the generation. 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.[34] Different counterfactual simulations might result in different inequalities in $childinc_i$, which can make the interpretation difficult. 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.

Table 10: Definitions of Counterfactual Simulations

	Household Inc	6th grade Score	Parental Educ	Parental Investments	Child's Self-Study
Benchmark Counterfactual (BCF)	O	O	O	O	O
Parental Investments (OPI)	O	O	O	O	X
Self-Study (OSS)	O	O	O	X	O
Without 6th grade Test (NST)	O	X	O	O	O
Without Household Inc (NIN)	X	O	O	O	O

Table 10 provides definitions of the counterfactual simulations. In particular,

- BCF is the baseline of the counterfactuals in which 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.
- In NIN, I fix all monthly net household income to 4,000,000 KRW (approximately 2800 USD).

Table 11 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 0.635. The estimated IGE with the benchmark model (BCF) is 0.269. The OSS simulation is interesting on its own because it provides implications for the tutoring ban policy of China. I discuss this in Section (7.2).

[34] In particular, $\delta_{IGE} = \rho_{ch} \frac{\sigma_{\ln childinc}}{\sigma_{\ln hhinc}}$, where σ_x is a standard deviation of data x and ρ_{ch} is a correlation between $\ln childinc_i$ and $\ln hhinc_i$.

Table 11: Intergenerational Persistence of Earnings under the Counterfactual Simulations

(a) Rank-rank Slope Estimates

	(1) BCF	(2) OPI	(3) OSS	(4) NIN
pincprctile	0.635***	0.827***	0.130***	0.341***
	(0.018)	(0.013)	(0.023)	(0.022)
R-squared	0.403	0.684	0.017	0.116

(b) Intergenerational Elasticity of Earnings Estimates

	(1) BCF	(2) OPI	(3) OSS	(4) NIN
log(hhinc)	0.269***	0.313***	0.068***	0.173***
	(0.008)	(0.005)	(0.009)	(0.011)
R-squared	0.398	0.708	0.031	0.119

The quantification exercise highlights several findings. First, removing heterogeneity in parental income decreases the rank-rank slope by 46.2%, which can be found in the result of the NIN simulation in Column (5) in Table 11a. 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 leads to significant increases in the intergenerational persistence of earnings. The estimated rank-rank slope under the OPI simulation is 0.827, which is greater than the benchmark simulation by 30.2%, as shown in Column (2) in Table 11a. At the same time, when the channel of parental investment is removed, the rank-rank slope decreases by 79.5%, as shown in Column (3) in Table 11a. Such results suggest that while parental investment reinforces the intergeneration persistence of earnings, the self-study of the child mitigates it. Third, 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 12. For example, BCF' is the same simulation as BCF, but the only difference is 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.

Table 12: Intergenerational Persistence of Earnings Fixing Initial Conditions

(a) Rank-rank Slope Estimates

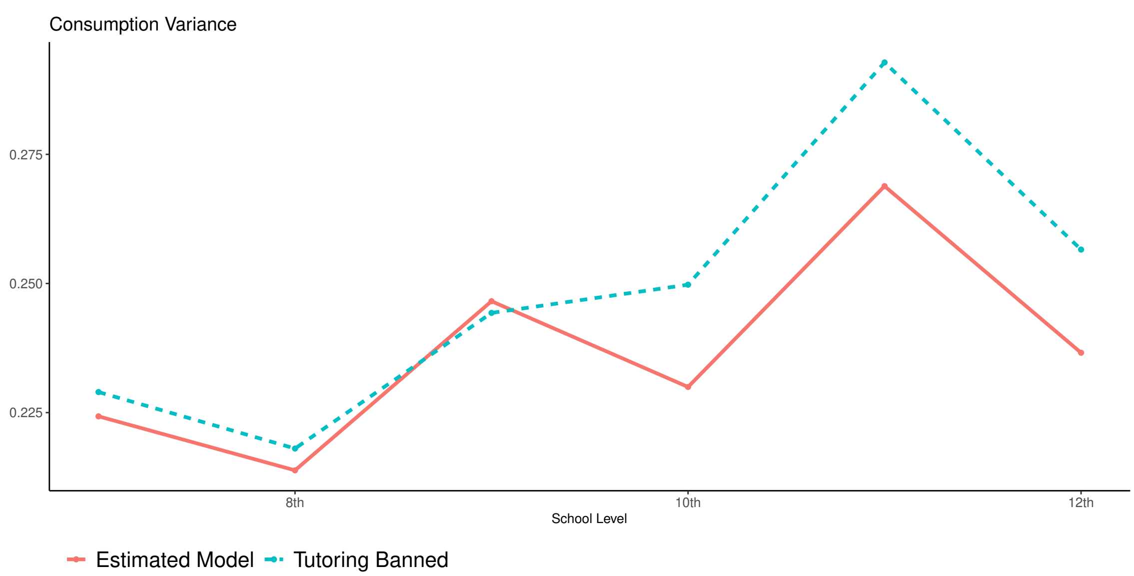
	(1) BCF'	(2) OPI	(3) OSS'
pincprctile	0.641***	0.834***	0.323***
	(0.018)	(0.013)	(0.022)
R-squared	0.411	0.695	0.104

(b) Intergenerational Elasticity of Earnings Estimates

	(1) BCF'	(2) OPI'	(3) OSS'
logpinc	0.393***	0.312***	0.100***
	(0.011)	(0.005)	(0.007)
R-squared	0.401	0.719	0.097

7.2 The Effects of China's Private Tutoring Ban Policy

Figure 11: Counterfactual: Consumption Inequality

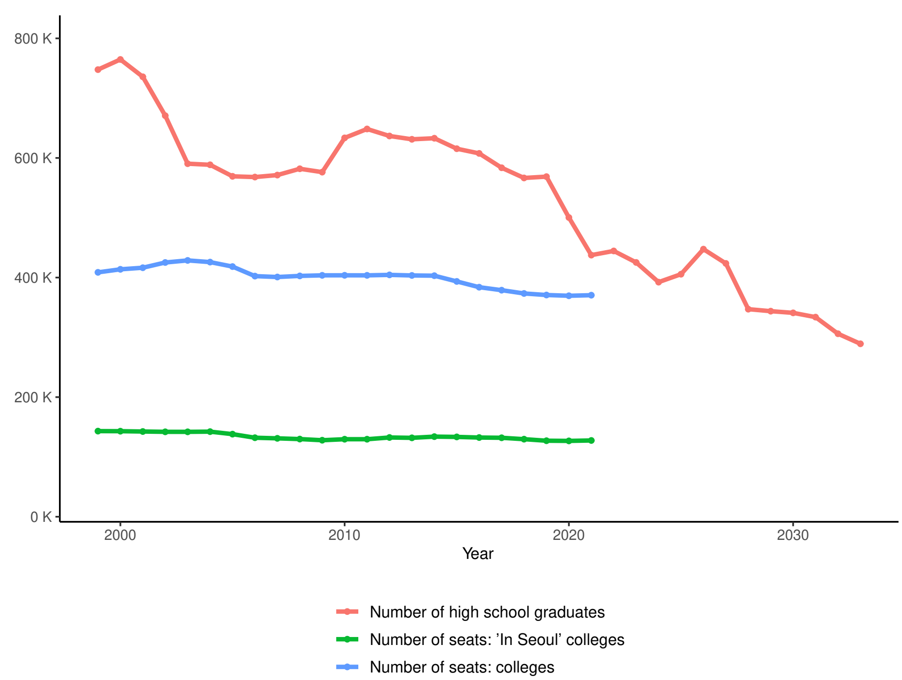


I evaluate the private tutoring ban policy of China based on two criteria: (i) intergenerational mobility and (ii) consumption inequality. First, Intergenerational mobility is assessed using rank-rank slope quantification in the Section 7.1. The OSS simulation is used because it is where private tutoring expenditure is prohibited. Second, I measure the changes in consumption inequality by the changes in the variance of consumption. Overall, prohibiting private tutoring expenditure increases consumption inequality. Figure 11 presents the changes in consumption inequality of the dynamic model in each period. Removing private tutoring expenditure increases consumption inequality because high income households were already spending a lot of income on private tutoring expenditure. Meanwhile, low income households' consumption does not increase as much relative to the benchmark case. On the aggregate level, this suggests that the presence of private tutoring expenditure decreases consumption inequality.

Based on the simulation results of Section 7.1, which can be found in Table 11a, the private tutoring ban policy would decrease intergenerational persistence of earnings. In other words, such a policy would increase intergenerational mobility. However, the increase in mobility would come with the expense of an increase in consumption inequality.

7.3 Parental Investment with declining fertility

Figure 12: Number of seats

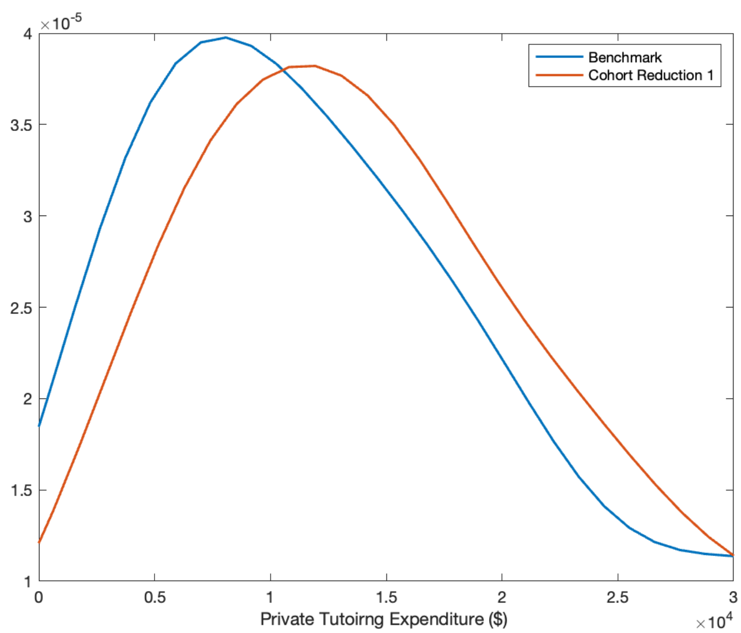


Cohort Reduction Simulation The purpose of this counterfactual experiment is to simulate parental investment decisions of households when there is a drastic reduction in the size of the cohort. Countries with a high amount of average parental monetary investments tend to experience extremely low fertility rates. South Korea, China, Turkey, Singapore, Taiwan are the countries where the demand for private tutoring is high (Bray, 1999, 2021), and they are experiencing a drastic reduction in the size of the cohort as can be seen in Figure 1. The reduction in the cohort size is equivalent to the reduction in the number of competitors in the college admission tournament. If there is a radical reduction in the number of competitors, the degree of competition would be less fiercer given the number of seats for colleges do not change.

Colleges have not decreased the number of seats amid the demographic shift. Figure 12 presents the number of high school graduates up to 2033, the number of seats of colleges in Korea, and colleges in Seoul up to 2022. The number of high school graduates after 2022 is projected using the average dropout rates and the number of graduates from younger cohorts in 2022. The number of projected high school graduates in 2033 is 289,216, which is only 44.5% of the number of high school graduates in 2011. On the other hand, colleges do not adjust the number of seats. Cohort size already was already decreasing around 2000, but colleges do not change the number of seats, as can be found in Figure 12. Colleges in Seoul, which is the equivalent of Tier 1 to Tier 3 colleges of the earlier classification, do not decrease the number of seats either.

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 50%. 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 increasing by twice. For now, I assume that the inequality of the college qualities remain same.

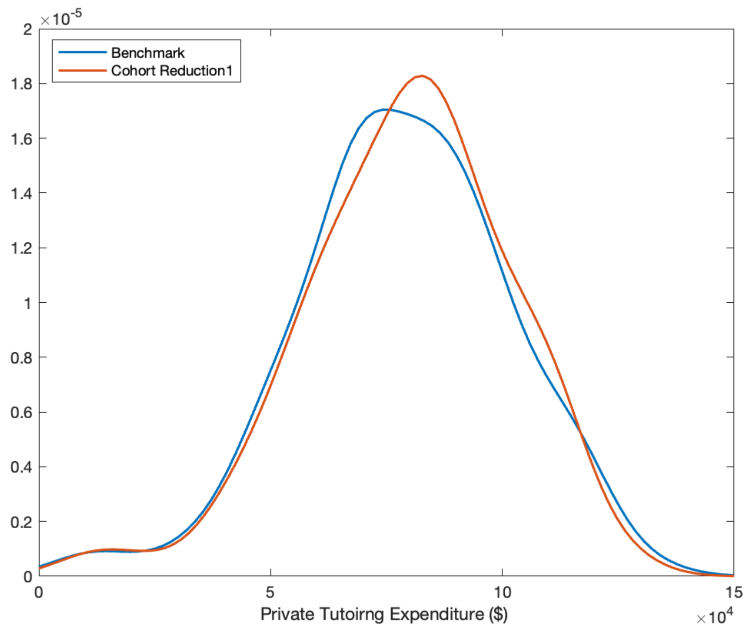
Figure 13: Cohort Simulation: Low Income Households



According to the simulation results, the amount of private tutoring expenditure slightly increases assuming the disparities in college quality do not change. Figures 14 and 13 present the density of private tutoring expenditures of low income households (lowest 5%) and high income households (highest

5%) respectively. High income households increase their total spending by 1.06% whereas low income households increase their spending by 12.57%. The increase of parental investment might be driven by the assumption that college inequality does not change over time. As a result of cohort reduction, students can access to the good income prospects with less fiercer competition. In turn, people who previously had no chance of getting into an upper college tiers would have better chance, For low income households, the increased number of seats increases the probability of going to higher tier college, which makes low income households spend more.

Figure 14: Cohort Simulation: High Income Households

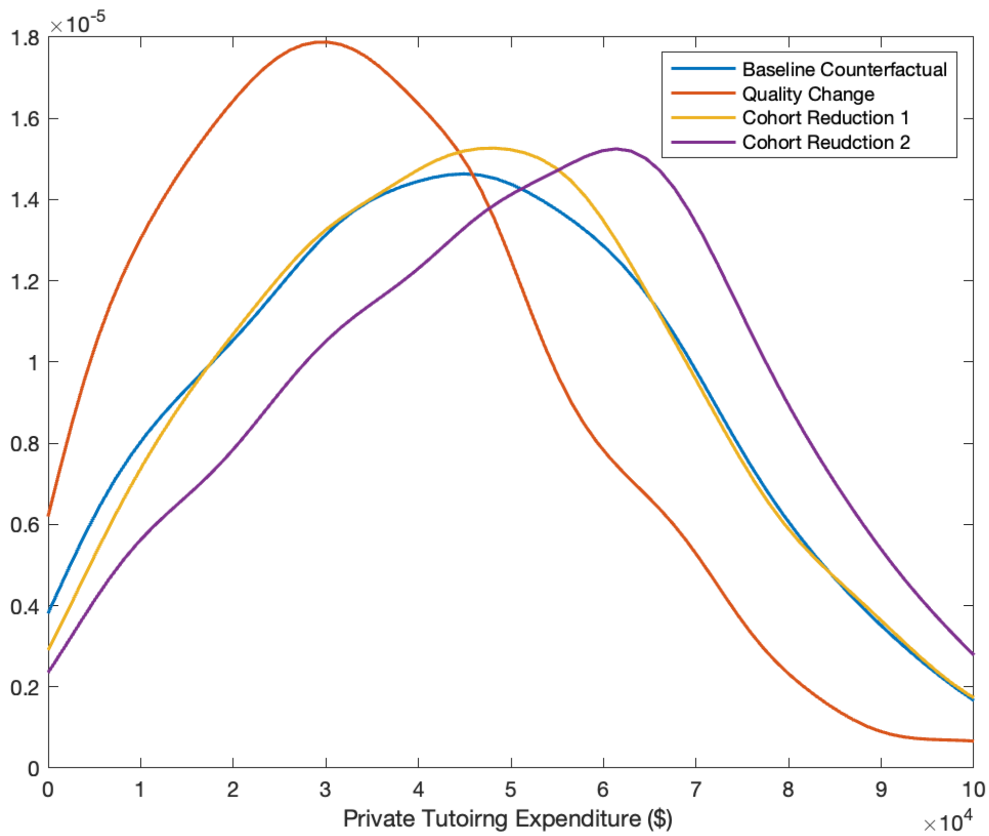


College Inequality Simulation To evaluate the effects of college inequality on parental investment, I change the values of $\{v_j\}_{j=1}^{J}$, the earning prospects of alumni for each tier. The relative size of v_1, v_2, v_3 and v_4 relative to v_5 are 2.22, 1.95, 1.45, 1.25 respectively. In this simulation, I change them to 1.5, 1.4, 1.3, and 1.2 respectively. I set the percentiles that students need to make for each tier as the same with the benchmark simulation. Figure 15 presents the resulting changes on parental investment for the households. The decrease in college inequality leads to a substantial decrease in private tutoring expenditure of the entire households. Table 13 presents the quantification of the changes in private tutoring expenditure. I report the changes amount of total private tutoring expenditure compared to the benchmark simulation where I normalize private tutoring expenditure as 100. It can be seen that the changes in college hierarchy results in a significant decrease in total private tutoring expenditure.

Table 13: Simulated Private Tutoring Expenditure

	Benchmark	Cohort Reduction 1	Cohort Reduction 2	College Hierarchy
Entire Household	100	101.57	114.06	76.13
Low Income Household	100	114.38	137.44	68.63
High Income Household	100	101.07	102.77	83.98

Figure 15: College Inequality Simulation



8 Conclusion

- [Motivation]
- [Empirical findings]
- [Model]

- [Counterfactual findings]
- [Contribution]

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. (2020) It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects,.

Agostinelli, F. and Sorrenti, G. (2021) Money vs. Time: Family Income, Maternal Labor Supply, and Child Development.

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.

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.

Black, S. E. and Devereux, P. J. (2010) Recent developments in intergenerational mobility, Tech. rep., National Bureau of Economic Research.

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.

Brewer, D. J., Eide, E. R. and Ehrenberg, R. G. (1999) Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings, *The Journal of Human Resources*, **34**, 104.

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.

Chiappori, P.-A., Salanié, B. and Weiss, Y. (2017) Partner choice, investment in children, and the marital college premium, *American Economic Review*, **107**, 2109–67.

Choi, J. (2013) Essays on private tutoring and single-sex schooling, *Dissertations available from ProQuest. AAI3565039.*

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.

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. (2019) Actors in the Child Development Process, Working Paper 25596, National Bureau of Economic Research.

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., Sorrenti, G. and Zilibotti, F. (2019) The Economics of Parenting, *Annual Review of Economics*, **11**, 55–84.

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

Eide, E., Brewer, D. J. and Ehrenberg, R. G. (1998) Does it pay to attend an elite private college? Evidence on the effects of undergraduate college quality on graduate school attendance, *Economics of Education Review*, **17**, 371–376.

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

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

Forbes (2021) China Bans For-Profit Tutoring In Reforms Aimed At Boosting The Birth Rate, section: Healthcare.

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.

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.

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.

Huggett, M., Ventura, G. and Yaron, A. (2011) Sources of Lifetime Inequality, *The American Economic Review*, **101**, 2923–2954, publisher: American Economic Association.

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. 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. (2022) Status Externalities in Education and Low Birth Rates in Korea.

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, 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.

Manski, C. F. (1993) Identification of Endogenous Social Effects: The Reflection Problem, *The Review of Economic Studies*, **60**, 531.

Mazumder, B. (2005) Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data, *Review of Economics and Statistics*, **87**, 235–255.

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.

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.

Solon, G. (1999) Intergenerational Mobility in the Labor Market, in *Handbook of Labor Economics*, Elsevier, vol. 3, pp. 1761–1800.

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

Stern, S. N. and Zhou, Y. (2018) Antithetic acceleration and estimation of unobserved heterogeneity parameters, *Available at SSRN 3297698*.

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.

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.

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

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

Proof. Each period, the flow utility term $u(c_{it}, x_{it}, s_{it}, \varepsilon_{it})$ is closed and bounded. Consumption is bounded by household income and the hours variables are bounded by the maximum available hours. For the final period, the tournament term $\sum_{j=1}^{J} \ln(v_j) \times Prob(\ln \tilde{Q}_{j-1} \geq \ln q_{i,T+1} \geq \ln \tilde{Q}_j \Big| \Gamma_{iT})\Big\}$ is closed and bounded as the v_j term is finite and greater than 0. Following the backward recursion, the choice specific value for the earlier period ($t < T$) is closed and bounded. It follows that the policy function Υ is closed and bounded for each period. By the Heine-Borel theorem, the choice specific value and the policy function are compact. □

Appendix A.2

Proof of Lemma 3

Proof. I begin by showing the continuity of value function. Recall that the value function consists of the flow utility term and the expected value of $t+1$. The flow utility term is continuous in its arguments because all the arguments are continuous variables. One way to show the continuity of the expected value function is show that it is sequentially continuous. that For any sequence of the arguments of the value function,

$$\{Z_t^n, \Psi_t^n\} \to \{Z_t^0, \Psi_t^0\},$$

we have

$$\int\limits_{\Psi} V_{t+1}(Z_{t+1}^n, \Psi_{t+1}^n) d\Psi \to \int\limits_{\Psi} V_{t+1}(Z_{t+1}^0, \Psi_{t+1}^0) d\Psi.$$

Note that Z_t is a set of continuous variables, which consists of household income w_{it}, previous test score q_{it}, and parental education m_i. Recall that $\Psi_{t+1} = \{\eta_{i,t+1}^c, \eta_{i,t+1}^x, \eta_{i,t+1}^s, \eta_{it}^q, \lambda_k^c, \lambda_k^x, \lambda_k^s, \lambda_k^q\}$. As I assume the expectation of the unobserved shocks has finite expectation, the expected value term has finite expectation as well. As all the inputs of the expected value term are continuous, $\int_{\Psi} V_{t+1}(Z_{t+1}^0, \Psi_{t+1}^0) d\Psi$ is continuous by the Dominated Convergence Theorem.

Now I am ready to show the continuity of the mapping $\aleph$. The equations defining the model can be thought of as mappings which map its inputs on the RHS to the term on LHS. For example, in equation (9), the flow utility term $u(c_{it}, x_{it}, s_{it}, \varepsilon_{it})$ and the expected value term $\underset{\eta_{it}^q, \eta_{i,t+1}}{E}\left[V_{i,t+1}(Z_{i,t+1}, \Psi_{i,t+1} \Big| \Gamma_{it})\right]$ are the inputs of the mapping, and the resulted value function $V_{i,t+1}(Z_{i,t+1}, \Psi_{i,t+1})$ is the outcome of the mapping. For the policy functions, the First Order Conditions with respect to the choice variables can be understood as a mapping that determines the optimum choice of the model. I denote each of such mapping as $\Xi_a : A_a \to B_a$ for each rule a. As value functions for each period is sequentially continuous, the mapping Ξ_a is sequentially continuous. As the composition of two sequentially continuous mappings is sequentially continuous, the composition of any number of sequentially continuous mappings is sequentially continuous. As the mapping $\aleph$ is a composition of Ξ_a, its continuity follows from the continuity of Ξ_a. □

Appendix B

	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 graduate

Appendix C

Define the first order conditions as

$$
\begin{aligned}
V_p &= \alpha_c \varepsilon_{it}^c u_p^c(c_{it}) + \beta\left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial p_{it}} \\
V_x &= \alpha_c \varepsilon_{it}^c u_x^c(c_{it}) + \alpha_x \varepsilon_{it}^x u_x^x(x_{it}) + \beta\left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial x_{it}} \\
V_s &= \alpha_s \varepsilon_{it}^s u_s^s(s_{it}) + \beta\left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial s_{it}}
\end{aligned}
$$

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

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

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

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 D

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

$$
\begin{aligned}
\frac{\partial}{\partial p_{it}} &: \alpha_c\varepsilon^c_{it}u^c_p(c_{it}) + \beta\left[\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial p_{it}} = 0;\\
\frac{\partial}{\partial x_{it}} &: \alpha_c\varepsilon^c_{it}u^c_x(c_{it}) + \alpha_x\varepsilon^x_{it}u^x_x(x_{it}) + \beta\left[\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial x_{it}} = 0;\\
\frac{\partial}{\partial s_{it}} &: \alpha_s\varepsilon^s_{it}u^s_s(s_{it}) + \beta\left[\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial s_{it}} = 0.
\end{aligned}
$$

For the households who do not participate in tutoring expenditure, I assume that the choice variable regarding tutoring activity is the total amount of tutoring expenditure, e_{it}, not like the participants who separately choose the quantity (hours of tutoring) and the quality (the unit payment of tutoring) of tutoring expenditure. The non-participants' first order conditions are

$$
\begin{aligned}
\frac{\partial}{\partial e_{it}} &: \alpha_c\varepsilon^c_{it}u^c_e(c_{it}) + \beta\left[\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial e_{it}} = 0;\\
\frac{\partial}{\partial s_{it}} &: \alpha_s\varepsilon^s_{it}u^s_s(s_{it}) + \beta\left[\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial s_{it}} = 0,
\end{aligned}
$$

where u^c_e denotes the partial derivative of u^c with respect to e, private tutoring expenditure.

Simplifying the first order conditions further, we can get three first order conditions with respect to each choice variable:

$$\alpha_c \varepsilon_{it}^c + \beta \frac{1}{u_p^c(c_{it})} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial p_{it}} = 0; \tag{12}$$

$$\alpha_c \varepsilon_{it}^c u_x^c(c_{it}) + \alpha_x \varepsilon_{it}^x u_x^x(x_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial x_{it}} = 0; \tag{13}$$

$$\alpha_s \varepsilon_{it}^s u_s^s(s_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial s_{it}} = 0. \tag{14}$$

Equation (12) provides the analytical expression of ε_{it}^c. To obtain the analytical expressions of ε_{it}^x, plug in the result of (12) to (13):

$$\alpha_c \varepsilon_{it}^c u_x^c(c_{it}) + \alpha_x \varepsilon_{it}^x u_x^x(x_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial x_{it}} = 0; \tag{15}$$

$$\underbrace{[-\beta \frac{1}{u_p^c(c_{it})} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial p_{it}}]}_{\alpha_c \varepsilon_{it}^c} u_x^c(c_{it}) + \alpha_x \varepsilon_{it}^x u_x^x(x_{it}) + \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial x_{it}} = 0 \tag{16}$$

The expression gives

$$\alpha_x \varepsilon_{it}^x u_x^x(x_{it}) = \underbrace{[\beta \frac{1}{u_p^c(c_{it})} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial p_{it}}]}_{\alpha_c \varepsilon_{it}^c} u_x^c(c_{it}) - \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial x_{it}}$$

$$\begin{aligned}
\alpha_x \varepsilon_{it}^x u_x^x(x_{it}) &= \beta \frac{u_x^c(c_{it})}{u_p^c(c_{it})} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial p_{it}}] - \beta \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial x_{it}} \\
\varepsilon_{it}^x &= \frac{\beta}{\alpha_x u_x^x(x_{it})} \frac{u_x^c(c_{it})}{u_p^c(c_{it})} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial p_{it}}] - \frac{\beta}{\alpha_x u_x^x(x_{it})} \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \frac{\partial \ln q_{i,t+1}}{\partial x_{it}} \\
\varepsilon_{it}^x &= \frac{\beta}{\alpha_x u_x^x(x_{it})} \left\{ \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \left(\frac{u_x^c(c_{it})}{u_p^c(c_{it})} \frac{\partial \ln q_{i,t+1}}{\partial p_{it}} - \frac{\partial \ln q_{i,t+1}}{\partial x_{it}} \right) \right\} \\
\varepsilon_{it}^x &= -\frac{\beta}{\alpha_x u_x^x(x_{it})} \left\{ \left[\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\cdot,\cdot,\cdot) \right] \left(\frac{\partial \ln q_{i,t+1}}{\partial x_{it}} - \frac{u_x^c(c_{it})}{u_p^c(c_{it})} \frac{\partial \ln q_{i,t+1}}{\partial p_{it}} \right) \right\}
\end{aligned}$$

All together the shocks can be expressed in terms of the exogenous variables, and I denote such optimizing points of the shocks as $\widetilde{\varepsilon_{it}^c}$, $\widetilde{\varepsilon_{it}^l}$, and $\widetilde{\varepsilon_{it}^s}$. They are specified as

$$\widetilde{\varepsilon}^c_{it} = -\frac{1}{\alpha_c u^c_p(c_{it})}\left[\beta\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\left(\frac{\partial \ln q_{i,t+1}}{\partial p_{it}}\right);$$
$$\widetilde{\varepsilon}^x_{it} = -\frac{1}{\alpha_x u^x_x(x_{it})}\left\{\beta\left[\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\left(\frac{\partial \ln q_{i,t+1}}{\partial x_{it}} - \frac{u^c_x(c_{it})}{u^c_p(c_{it})}\frac{\partial \ln q_{i,t+1}}{\partial p_{it}}\right)\right\}$$
$$\widetilde{\varepsilon}^s_{it} = -\frac{1}{\alpha_s u^s_s(s_{it}))}\beta\left[\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right]\frac{\partial \ln q_{i,t+1}}{\partial s_{it}}$$

where $u^c_x(c_{it})$ and $u^c_p(c_{it})$ are first order derivatives of $u^c(c_{it})$ with respect to x_{it} and p_{it} respectively, and $u^l_x(l_{it})$ and $u^l_s(l_{it})$ are the first order derivatives with respect to x_{it} and s_{it} respectively.

$$\widetilde{\ln\varepsilon^c_{it}} = \tilde{\eta}^c_{it} + \lambda^c_i = \ln(-\frac{1}{\alpha_c u^c_p(c_{it})}) + \ln\left[\beta\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right] + \ln\left(\frac{\partial \ln q_{i,t+1}}{\partial p_{it}}\right);$$
$$\widetilde{\ln\varepsilon^x_{it}} = \tilde{\eta}^x_{it} + \lambda^x_i = \ln(-\frac{1}{\alpha_x u^x_x(x_{it})}) + \ln\left[\beta\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right] + \ln\left(\frac{\partial \ln q_{i,t+1}}{\partial x_{it}} - \frac{u^c_x(c_{it})}{u^c_p(c_{it})}\frac{\partial \ln q_{i,t+1}}{\partial p_{it}}\right)$$
$$\widetilde{\ln\varepsilon^s_{it}} = \tilde{\eta}^s_{it} + \lambda^s_i = \ln(-\frac{1}{\alpha_s u^s_s(s_{it}))}) + \ln\left[\beta\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right] + \ln\frac{\partial \ln q_{i,t+1}}{\partial s_{it}}$$

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

$$EV_{i,T+1} = v_1 - \sum_{j=1}^{J}\Big(\ln(v_j) - \ln(v_{j+1})\Big)\Phi(\frac{\ln\bar{q}_j - \ln q_{iT+1} - \lambda^q_i}{\sigma_q});$$
$$\frac{\partial}{\partial \ln q_{i,T+1}}EV_{i,T+1}(\cdot,\cdot,\cdot) = \underbrace{\sum_{j=1}^{J}\Big(\ln(v_j) - \ln(v_{j+1})\Big)\frac{1}{\sigma_q}\phi(\frac{\ln\bar{q}_j - \ln q_{iT+1} - \lambda^q_i}{\sigma_q})}_{devt/dq}.$$

Parameterization with CES (Case when $t < T$)

With the functional form assumptions of log utility,

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

The derivative of $\ln q_{i,t+1}$ with respect to the choice variables are

$$
\begin{aligned}
\frac{\partial \ln q_{i,t+1}}{\partial p_{it}} &= \nu_t \frac{\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{3t}(1+s_{it})^{\phi}]}(\kappa {p_{it}}^{\kappa-1}{x_{it}}^{1-\kappa});\\
\frac{\partial \ln q_{i,t+1}}{\partial x_{it}} &= \nu_t \frac{\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{3t}(1+s_{it})^{\phi}]}((1-\kappa) {p_{it}}^{\kappa}{x_{it}}^{-\kappa});\\
\frac{\partial \ln q_{i,t+1}}{\partial s_{it}} &= \nu_t \frac{\delta_{3t}(1+s_{it})^{\phi-1}}{[\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{3t}(1+s_{it})^{\phi}]}.
\end{aligned}
$$

Also,

$$
\begin{aligned}
&\frac{\partial \ln q_{i,t+1}}{\partial x_{it}} - \frac{u_x^c(c_{it})}{u_p^c(c_{it})}\frac{\partial \ln q_{i,t+1}}{\partial p_{it}}\\
&\qquad= \frac{\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{3t}(1+s_{it})^{\phi}]}\left\{(1-\kappa){p_{it}}^{\kappa}{x_{it}}^{-\kappa} - \kappa\frac{p_{it}}{x_{it}}({p_{it}}^{\kappa-1}{x_{it}}^{1-\kappa})\right\}\\
&\qquad= \frac{\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{3t}(1+s_{it})^{\phi}]}\left\{(1-\kappa){p_{it}}^{\kappa}{x_{it}}^{-\kappa} - \kappa({p_{it}}^{\kappa}{x_{it}}^{-\kappa})\right\}\\
&\qquad= \frac{\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{2t}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{3t}(1+s_{it})^{\phi}]}({p_{it}}^{\kappa}{x_{it}}^{-\kappa})(1-2\kappa)
\end{aligned}
$$

Finally, the shocks are specified

$$
\tilde{\eta}_{it}^c = \ln(\frac{1}{\alpha_c}\frac{(w_{it}-p_{it}x_{it})}{x_{it}}) + \ln\left[\beta\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right] + \ln\left(\nu_t\frac{\delta_{et}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{et}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{st}(1+s_{it})^{\phi}]}(\kappa {p_{it}}^{\kappa-1}{x_{it}}^{1-\kappa})\right) - \lambda_i^c \tag{17}
$$

$$
\tilde{\eta}_{it}^x = \ln(-\frac{1}{\alpha_x}(1+x_{it})) + \ln\left[\beta\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right] + \ln\left(\nu_t\frac{\delta_{et}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi-1}}{[\delta_{et}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{st}(1+s_{it})^{\phi}]}({p_{it}}^{\kappa}{x_{it}}^{-\kappa})(1-2\kappa)\right) - \lambda_i^x \tag{18}
$$

$$
\tilde{\eta}_{it}^s = \ln(-\frac{1}{\alpha_s}(1+s_{it})) + \ln\left[\beta\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}(\cdot,\cdot,\cdot)\right] + \ln\left(\nu_t\frac{\delta_{st}(1+s_{it})^{\phi-1}}{[\delta_{et}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{st}(1+s_{it})^{\phi}]}\right) - \lambda_i^s \tag{19}
$$

$$
\tilde{\eta}_{it}^q = \ln q_{i,t+1} - \delta_{0t} - \delta_{1t}\ln q_{it} - \frac{\nu_t}{\phi}\ln[\delta_{et}(1+{p_{it}}^{\kappa}{x_{it}}^{1-\kappa})^{\phi}+\delta_{st}(1+s_{it})^{\phi}] - \lambda_i^q \tag{20}
$$

where $\frac{\partial}{\partial \ln q_{i,t+1}}EV_{i,t+1}$ is simulated with $\sum_{r=1}^{R} F\hat{V}_{i,t+1}$, where $F\hat{V}_{i,t+1}$ is the slope of interpolating function $\hat{V}_{i,t+1}$ with respect to $\ln q_{i,t+1}$.

With the presence of error correlation, the computation should be done numerically. Based on the model timeline, household use the expectation over the test score error $\ln \varepsilon_{it}^q$ in making their decision of tutoring expenditure. Note that this is after the realization of the shocks relevant to the choice variables. Therefore, the expectation is a conditional expectation over $\ln \varepsilon_{it}^q$ conditional on the realized values of η_{it}^c, η_{it}^l and η_{it}^s. In particular, for the case of $\tilde{\eta}_{it}^c$, the equation

$$\tilde{\eta}^c_{it} = \ln(\frac{1}{\alpha_c}\frac{(w_{it}-p_{it}x_{it})}{x_{it}}) + \ln\left[\beta\frac{\partial}{\partial \ln q_{i,t+1}} EV_{i,t+1}(\ln\widehat{q_{i,g+1}} + \lambda^q_i + \eta^q_{it} + \Sigma_{12}\Sigma_{22}^{-1}\begin{pmatrix}\tilde{\eta}^c_{it}\\ \tilde{\eta}^l_{it}\\ \tilde{\eta}^s_{it}\end{pmatrix}, \cdot, \cdot)\right]$$
$$+\ln\left(\nu_t\frac{\delta_{et}(1+p_{it}{}^{\kappa}x_{it}{}^{1-\kappa})^{\phi-1}}{[\delta_{et}(1+p_{it}{}^{\kappa}x_{it}{}^{1-\kappa})^{\phi}+\delta_{st}(1+s_{it})^{\phi}]}(\kappa p_{it}{}^{\kappa-1}x_{it}{}^{1-\kappa})\right) - \lambda^c_i$$

needs to be solved. Due to the form of $V_{i,t+1}$, the equation does not have a closed form solution. Thus, with the presence of the error correlation, I need to write a routine which minimizes the three first order conditions jointly.

Simulation of unobserved variables

Appendix E

Figure E.1: Sample Fit: Quality of Tutoring

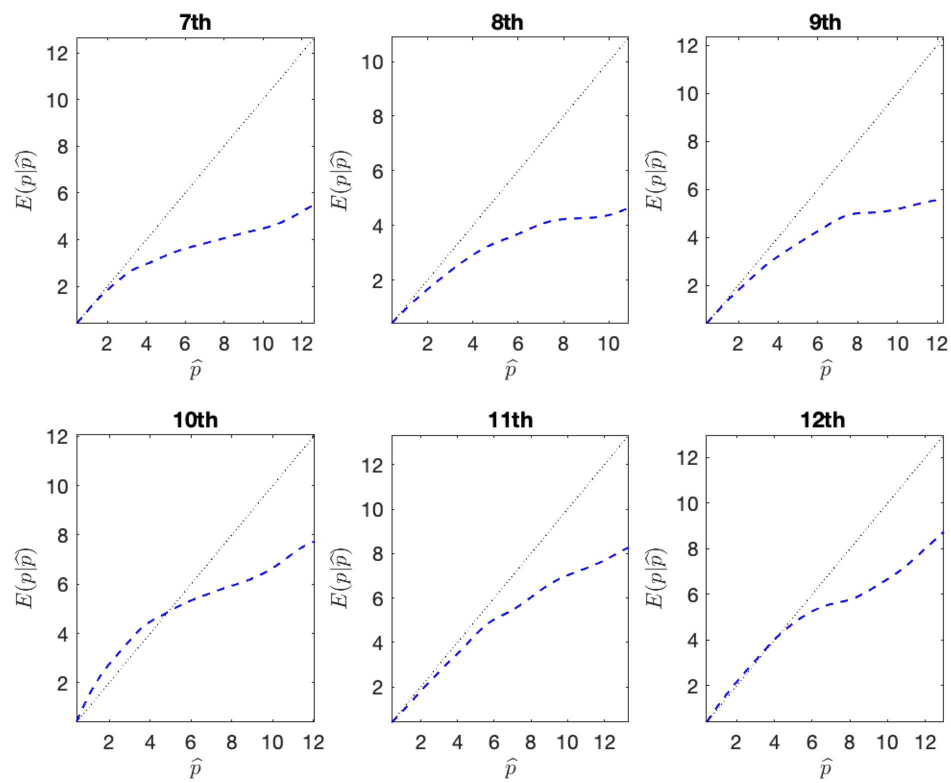


Figure E.2: Hours of Tutoring

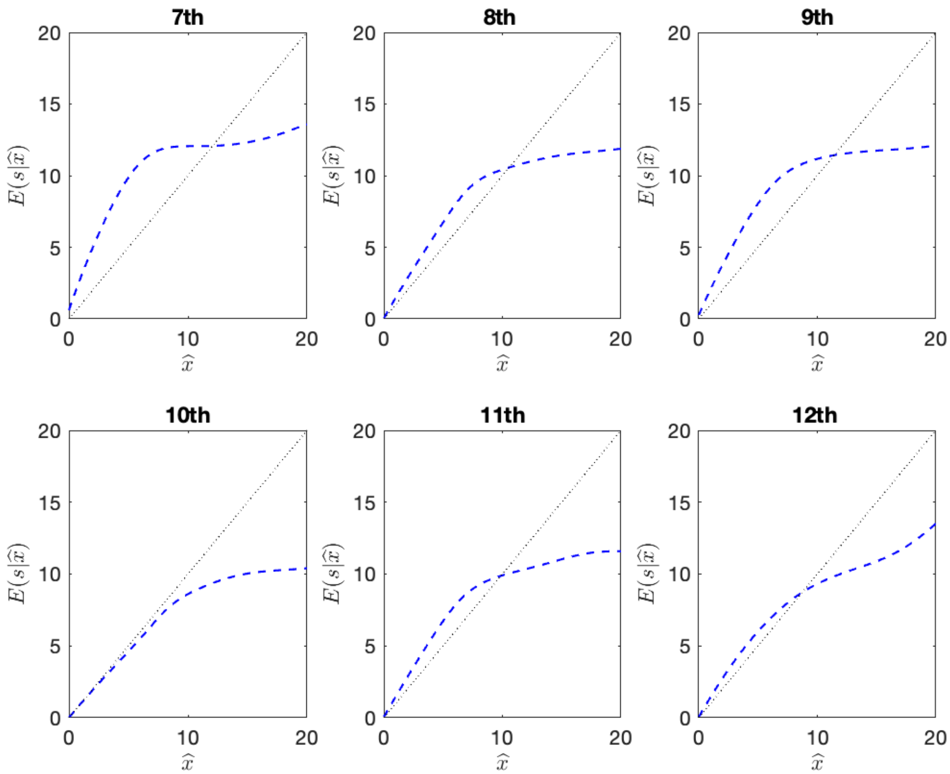


Figure E.3: Sample Fit: Private Tutoring Expenditure

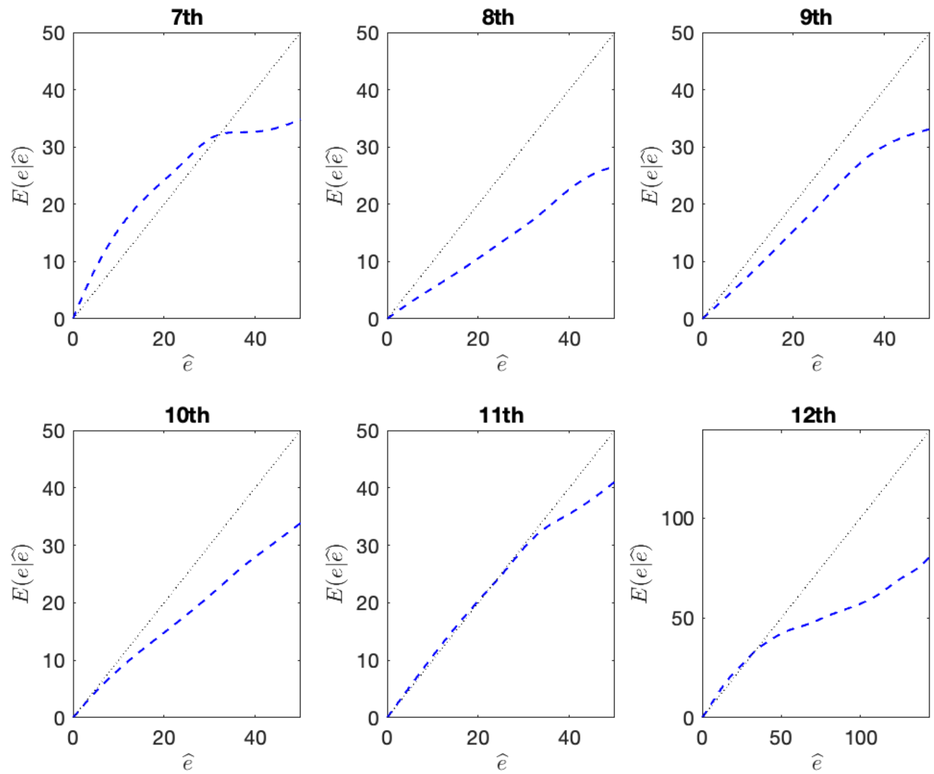


Figure E.4: Sample Fit: Hours of self-study

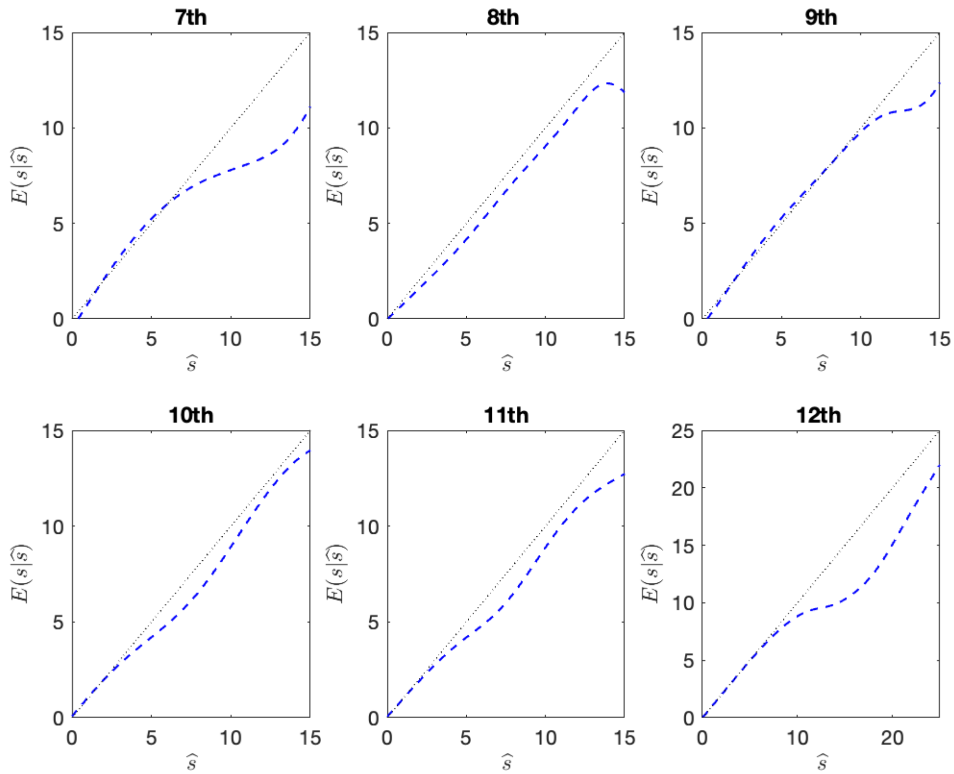
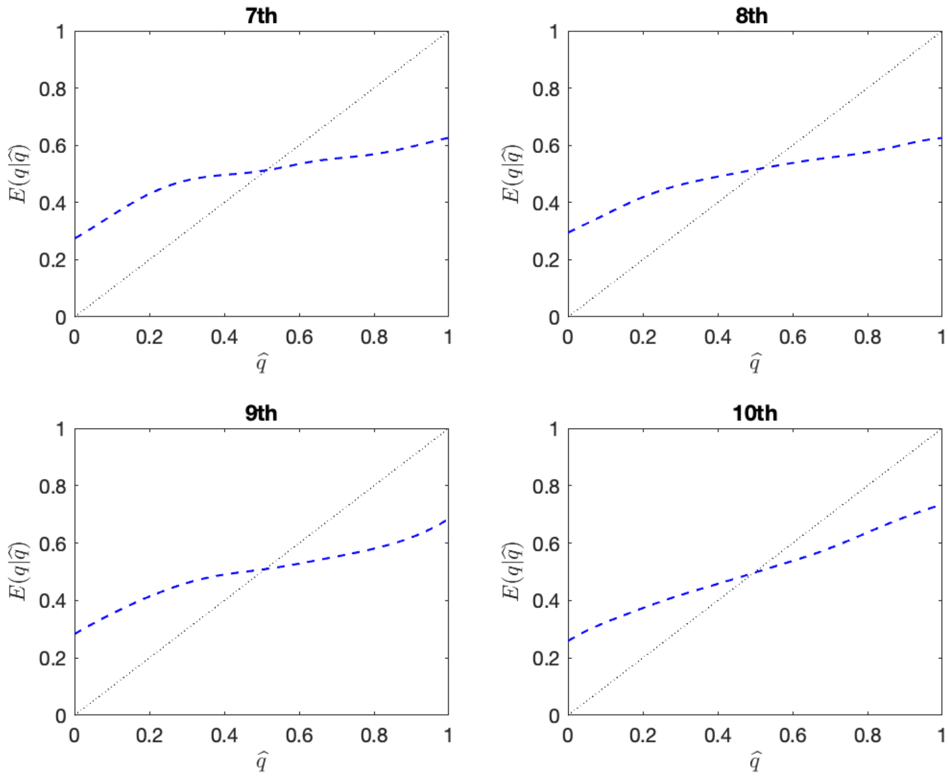
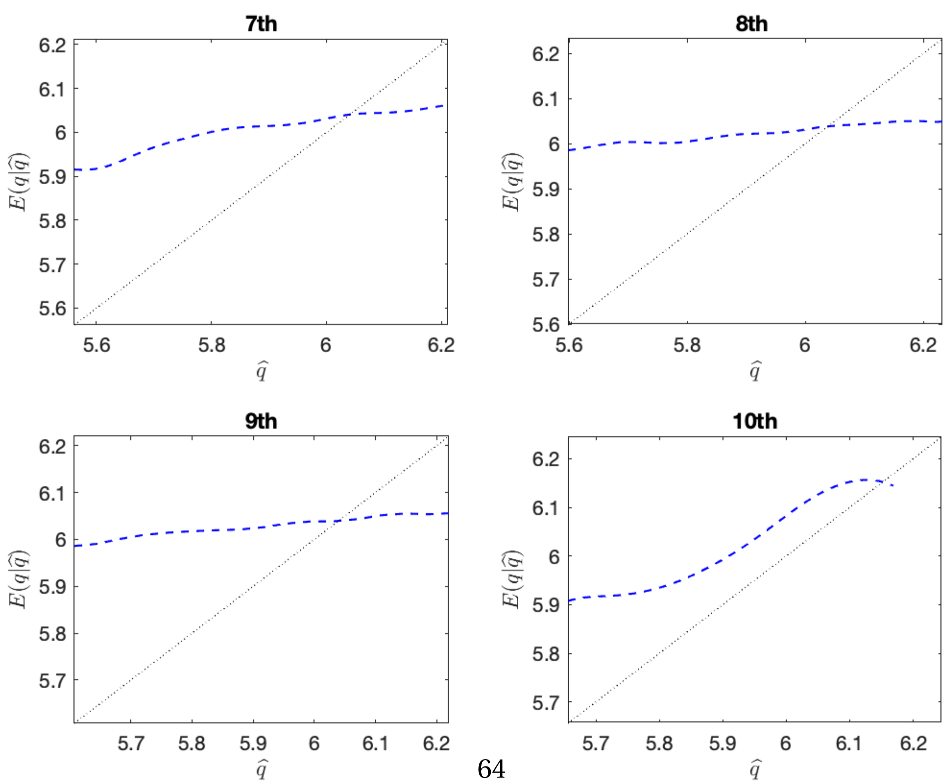


Figure E.5: Sample Fit: Log Test Scores

(a) Fit by distribution



(b) Fit by level



Appendix F: Additional Figures

Counterfactual Analyses