College Admissions, Grading Policy, and Students' Time Allocation Responses :The Case of South Korea

Taehoon Kwon*

November 10, 2020

Abstract

College admissions competition is zero-sum game where students do care about their relative ranking resulting in the overinvestment of resources and an inefficient market outcome as the game participants are facing prisoner's dilemma. Can SAT and GPA grading policy reduce high school students' college prep activities? This paper estimates causal effects of the nationwide education policy in 2005 which introduces coarsening SAT grade structure and more competitive grading of high school GPA in Korea on students' time allocation behavior in a college admission game. I model a student's welfare maximization behavior to investigate the policy effects on her college admission preparation activities. Empirical estimation is done by difference-in-differences method using repeated cross-sectional data. The result shows that the policy did not achieve the goal as the SAT taking students reduced leisure compared to the CSAT non takers. Top GPA students respond to the policy more sensitively by decreasing leisure and increasing self study because of the heterogeneous marginal productivity and motivation.

Keywords: Allocation, Grading, Welfare, Coarsening, College Admission, Incentive

JEL Codes: D04, I25

*PhD Candidate, Department of Economics, University of Washington, Seattle, WA, United States

Teenagers in the modern society face a crucial trade-off between leisure and preparation of college. Leisure provides direct current utility to them. Also, quality leisure time for young people has been emphasized as not only a potential breeding ground for youth problems but also the endless possibilities for constructive development[1] in a broad range of positive outcomes (United Nations, 2006). At the same time, investment in college admission is believed to bring enormous returns in the labor market[2]. This implies that high school days are critical time for human capital investment (Bond et al., 2018). Zimmerman (2014) uses a regression discontinuity design to show the huge quarterly earnings difference by $1,593 between the two groups of students just above and below the admission cutoff for a university in Florida. Andrews et al. (2016) studies a case in Texas to find that the state's flagship college[3] graduation increases earnings significantly compared to the graduation from all other four-year public universities and public two-year colleges in Texas. In this context, estimating the high school student's lifetime welfare maximizing behavior who solves a problem of choosing her optimal leisure and admission investment may have political economic implications on how to achieve individually and socially desirable outcomes.

There exists a general concern that students and their families are investing a lot of time and money in college prep activities. If college admissions competition is zero-sum game where students do care about their relative ranking, the ongoing unproductive arms race may lead society to an inefficient outcome as the game participants are facing prisoner's dilemma in that the market outcome is not the welfare maximizing result. College admission policy may intervene to affect students and parents incentives to move away from the market failure. Would the implementation of the new grading systems for SAT and GPA enhance

[1]Furthermore, adolescent development has health implications throughout life (WHO).

[2]American Community Survey (2018) states that college graduates earned 75% more than high school graduates. Quality education in college improves labor productivity by providing access to information sources (Thomas et al., 1991), strengthening ability to decipher new information (Schultz, 1975) ability to learn (Rosenzweig, 1995), but also helps to join a more productive social network.

[3]University of Texas at Austin and Texas A&M University at College Station.

high school students' welfare or boost admission competition? How would student's optimal resources allocation change in response to the grading policy? To answer these questions, I focus on a Korean education policy which aims to reduce unnecessarily competitive admission race. Since most Korean colleges examine high school GPA and SAT score of applicants in determining admission decisions, a nationwide education policy coarsening the SAT grading scale and strengthening importance of high school GPA may affect students' welfare by affecting their time and financial resources allocation behavior.

Under this specific grading scheme I model that high school students decide how much time they would spend on each academic and leisure activity to maximize their expected lifetime utility during the three-year of race to college admission. I show that coarse grading SAT increases uncertainty in student's return to investment and results in larger expenditure on admission activities when students are loss aversion. Also, I shows that relative grading GPA makes it harder to achieve the highest status due to the increased competitiveness among cohort as it restrict the number of students in each grade category.

I empirically investigate the model's predictions about the policy effects on student's resource allocation behavior using Korean Education and Employment Panel Survey data. My difference-in-differences model compares SAT takers group and non-takers group in two academic cohorts before and after the policy changes in 2005. Acknowledging the possibility that the non-takers group might be exposed to the intervention in some extent, my empirical work finds a significantly negative effects on college applicants' current welfare. This implies that there was no policy achievement of reducing students' study burden due to the increased rank uncertainty in SAT and increased competitiveness in GPA.

The remainder of this chapter proceeds as follows. Section 1 provides literature review and policy background. Section 2 sets up the model. Section 3 describes data. Section 4 introduces the empirical strategy. Section 5 presents the empirical results. Section 6 checks robustness of empirical model and results. Section 7 concludes.

1 Background

"Adolescents, just like adults, need time every day to unwind and interact with their peers. Too much pressure in schools might mean that students feel compelled to spend more time studying, leaving less time for these non-academic activities, at the expense of students' quality of life." - OECD, 2017.

1.1 Literature and Contribution

A number of studies provides theoretical approach to college admissions. Hafalir et al. (2018) model college admissions contest as a Bayesian game to solve the equilibrium where heterogeneous *students* choose a college and a single dimensional effort level to maximize their utilities from the college admissions. Che and Koh (2016) develop a matching market model in the context of college admissions where *colleges* make admissions decisions to reach equilibria given the students' dominant strategies. Avery et al. (2014) design a non-price competition model to analyze strategic behaviors among colleges to attract highly qualified students, but their models also only focus on the supply side of admissions. Although there is an enormous literature on the education policy impact on students' performance or effort incentive, a relatively small literature examines policy effect on student's welfare and resource management behavior captured by their specific time use data. Han et al. (2016) show the incentive effects of coarsening the CSAT grades on the agents' effort levels. Dubey and Geanakoplos (2010) find that if students care primarily about their relative rank in a *games of status*, they are often best motivated by coarse grading to work harder.

I examine the nature of the fierce college competition in the framework of game of status (Hopkins and Kornienko (2004); Moldovanu et al. (2007)). In this view, the result of the competition can be socially undesirable in a sense that there exists over-expenditure in academic signaling activities such as GPA, SAT score, or resume to get better college admissions than the socially optimal level of investment. With limited household resource, time and money spending on those signaling efforts sacrifice students' leisure and consumption which would have enhanced directly teenagers' utility. Since the number of college admissions are limited, each applicant cares about her relative position, *status*, on the distributions of admission criteria such as SAT and GPA. According to the descending order of the applicants' academic performance (or SAT score and GPA), college admissions are assigned as the *status prizes* to those who are classified above the admission cutoff. A student's strategy is a mapping from time and financial resources to admission investment expenditure. Students want to achieve higher quality of admissions which allow better job market opportunities by signaling good quality of labor productivity. The game is described by a simultaneous move game under incomplete information. The equilibrium of this college admission game is similar to the symmetric Nash equilibrium in Hopkins and Kornienko (2004)[4] in the sense that the result is a Pareto-inferior outcome where students invest *too much* to the admission compared to the *socially* desirable outcome.

In the literature, however, study of teenagers' time use related to the college admission game is rare. Becker (1965) and Gronau (1977) introduce the time budget constraint of households to show the utility maximizing household's time allocation behavior. While the effect of non-working time use such as sleep on adult labor productivity has been recently studied (Gibson and Shrader, 2018), only little about young students' time allocation behavior related to adolescent welfare in response to the college admission policy change is investigated. My study models the lifetime welfare maximizing high school student's leisure allocation behavior in the context of Korean college admission competition. In this model, student's welfare is an increasing function of time use on *leisure*[5] and on college admission

[4]In this prisoners' dilemma in a non-cooperative one-shot admission game, the Nash equilibrium strategy of individual student (or her caregivers) is to invest her maximum amount of resource to get better admissions as long as incentive to marginal investment is greater than the marginal disutility or cost of effort or financial resource. Similar equilibrium concept in a *status* game where students care about their relative rank in the class is also described by Dubey and Geanakoplos (2010).

[5]In this study, leisure is defined as free time away from study excluding time use on sleeping at night

quality which takes time use on admission preparatory activities as input factors. As there is a trade-off between time use on current leisure and on study for the future, the optimal time allocation is strongly influenced by the college admission policy. This study also considers more choice variables such as students' time on leisure and admission preparatory activities to empirically capture the detailed changes in students' behavior connecting them to students' welfare. Most papers introduced above assume a simplified single dimensional effort choice by college applicants. My model expands student's choice variables to analyze the multi-dimensional policy effects.

My paper contributes to the empirical evidence of the demand side behaviors to the nationwide college admission policy changes by Korean government given the colleges' strategies. This study analyzes the college admissions game with a similar view to Bodoh-Creed and Hickman (2018). They model the college admissions market as a contest where students compete for admission by endogenously choosing the level of human capital to accrue prior to a rank-order admissions contest given the colleges ranking. Olszewski and Siegel (2016) also design a contest with many heterogeneous players and prizes to characterize equilibrium and Olszewski and Siegel (2019) show that a coarse performance-disclosure policy can be a Pareto improving policy which benefit all students. In contrast with their model prediction, this paper empirically evaluates the coarse performance-disclosure policy effect in Korea together with the relative grading policy in GPA to show that both rule changes may have negative effects on student's welfare.

Little empirical literature studies student time allocation. Grau (2018) builds a rank-order tournament structural model where high school students decide their level of effort and whether or not to take the college admissions test and empirically evaluates that an increase in GPA weight increases students' average effort. This paper finds the similar effect an increased GPA weight while my model assumes that SAT taking decision is not a choice

and having meals.

variable but exogenously determined by student's preference and family characteristics.

There exists a study which estimates the effect of the new CSAT grading policy that this paper is interested. Han et al. (2016) show mixed results that the students' efforts devoted to the CSAT Korean have decreased, but those devoted to the CSAT English have increased. However, they do not account for the simultaneous effect of GPA grading policy change on students' incentive which might have also affected the effort investment to CSAT. Also, they use different cohorts' labor market earnings from the sample cohorts as a outcome variable in the regression due to the data limitation. As a result, the estimates of effort incentives might have been biased. In this chapter, I estimate the comprehensive effect of the grading policy by considering changes of grading system in both CSAT and GPA to more precisely identify the aggregate effect on student's admission preparation behavior.

1.2 High School Education and College Admission Policy in Korea

In this seciton, I briefly introduce the Korean education system for understanding of the policy background. Korean government manages the highly standardized courses and evaluation system from elementary to high school education. Every Korean parent has a legal obligation to send her five to seven years old children to an elementary school and a middle school until they graduate therefrom[6]. After six years of elementary school and three years of middle school, most middle school graduates decide to go to private or public high schools.

[6]ELEMENTARY AND SECONDARY EDUCATION ACT (Act No. 16672, Dec. 3, 2019) CHAPTER II COMPULSORY EDUCATION Article 13 (Obligation of School Enrollment) (1) Every citizen shall send sons and daughters or children under his or her care to an elementary school from March 1 of the year following the year in which the date they reach six years of age falls and have them attend the elementary school until they graduate therefrom. (2) Notwithstanding paragraph (1), every citizen may send sons and daughters or children under his or her care to an elementary school in the year following the year in which the date they reach five years or seven years of age falls. In such cases, he or she shall have his or her sons and daughters or children attend an elementary school from March 1 of the year they are admitted to the elementary school until they graduate therefrom. (3) Every citizen shall send sons and daughters or children under his or her care to a middle school at the beginning of the school year following the school year in which they graduate from an elementary school and have them attend the middle school until they graduate therefrom.

In 2019, 463,130 middle school graduates out of 464,717 (99.7%) continued their education at high school. Seventy percent of middle school graduates choose to go to general high school while more than a quarter selects vocational high school. The rest of them go to the specialized high schools in natural science, foreign language, or art and music are special cases under this system. College entrance rate is also high. Public high school education cost is affordable. Annual tuition and admission fee of public high school in the most expensive region, Seoul, are only KRW 1,464,900 (USD 1,300) in 2020. Children in a household where its income is less than 50% of the median household income[7] are able to receive 100% subsidy from the government for the tuition, admission fee, and expenses for purchasing textbooks[8]. In 2008 academic year which is the sample year of this study, 82.8% of high school graduates entered university or college. These facts are summarized in Figure 1.

In this paper, I only choose regular (*jeongsi*) college applicants as my intervention group samples because rolling (*susi*) applicants have much less study pressure and little incentive to invest to College Scholastic Ability Test (CSAT) scores. Colleges may admit freshmen through two policies, regular (*Jeongsi*) and rolling (*Susi*). The main difference between *jeongsi* and *susi* is the significance and impact of CSAT scores on admission. Rolling admissions open in Spring and Autumn before *jeongsi* starts. Spring *susi* only requires GPA and resume while rolling in Autumn may set minimum cutoff CSAT grades which are typically very lower than the cutoffs in *jeongsi*[9]. Typically, *susi* applicants do better in GPA than in CSAT. Thus, *susi* applicants exert less efforts on CSAT to focus on other admission criteria. Indeed, some *susi* admitted students fail to satisfy the muinimum CSAT requirements[10].

[7]This is an annual income of KRW 23,223,468 which is equivalent to USD 20,609 for a household of three family members.

[8]Ministry of Education implemented the free high school education policy from 2020 as an intermediate step for the compulsory high school.

[9]For example, Seoul National University requires Autumn *susi* applicants to achieve just any two second-highest grades out of Korean, math, English, social science, and natural science CSAT subjects as the minimum admission requirement. In *jeongsi* admission, the cutoff student's average CSAT grade for social studies was 1.0575. This implies that if an applicant fails to get Grade 1 in all CSAT subjects, she might not get a *jeongsi* admission.

[10]For example, Korea University required the second-highest grades in two CSAT subjects as the minimum

The admission timeline for the 2008 academic year is illustrated in Figure 2.

Once a student is admitted from a college in the *susi* admission, she must register for the admitted school and cannot apply to other colleges in the same academic year. This prevents the admitted *Susi* students from competing in *jeongsi*. Regular admission typically starts after CSAT tested on November and requires both high school GPA and CSAT scores together with resume, essay, interview, or any combination of them. Nevertheless, CSAT is the primary source of evaluating *jeongsi* applicants for colleges as it is considered to be the most standardized test effectively testing students' ability to study in college. In 2005 academic year, 55.7% of college freshmen are admitted through the regular admission process while 46.9% of 2008 college freshmen achieved regular admission.

As in many countries, growing admission competition among college applicants is also pervasive in Korea. Bound et al. (2009) find data which shows that the US high school students in 2004 spent significantly more time on behavior associated with college preparation than did their counterparts from ten and twenty years before. The issue seems to be more serious in Korea. Korean college applicants do care about their relative ranking in CSAT and GPA to get a better college admission. With the limited number of schools relative to the number of students who want to be admitted, Korean students start taking extra lessons in their early ages and spend 6.4 years on the lessons to prepare college admission which is the longest among the twenty two high-income countries (OECD, 2017).

In *jeongsi* admissions, most of Korean colleges require students to take CSAT to submit the scores as one of the crucial admission criteria. Annual CSAT and quarterly mock exams are hosted by Korea Institute for Curriculum and Evaluation (KICE). CSAT can be taken only once a year at the end of the academic year, normally on the second or the third Thursday of November. As the only chance determines the quality of college admission for an applicant and the opportunity cost of failing the first take is another year of investment of

admission criteria in 2008 and the 35.5% of admitted students did not achieve them to lose their admissions.

financial and labor resources which cannot be underestimated by the twenty-year-old young people, high school seniors and their caregivers have huge incentives to exert their best efforts.

1.3 Policy Description and Timeline

"Cram schools like the one I taught in — known as "hagwon" in Korean — are a mainstay of the South Korean education system and a symbol of parental yearning to see their children succeed at all costs. "Hagwons" are soulless facilities, with room after room divided by thin walls, lit by long fluorescent bulbs, and stuffed with students memorizing English vocabulary, Korean grammar rules and math formulas. Students typically stay after regular school hours until 10 p.m. or later." — The New York Times, "How South Korea Enslaves Its Students." August 3, 2014.

Korean Ministry of Education decided to launch a new CSAT and GPA grading policy. The government aimed to strengthen public education, reduce unnecessary competition in CSAT, and promote diversity in college admissions[11]. The policy was implemented twofold. First, high school GPA grading regime changed from *absolute basis* to *relative basis* which emphasizes student's relative performance at school. Now, the GPA becomes more informative signal as it provides additional information such as percentile, mean, and standard deviation of each subject in the student's quarterly exams at school. This ranking information was not provided to students and colleges before the policy change. Furthermore, before the policy change, *easy-peasy* midterm and final exams at school were *pervasive*[12] as teachers have an incentive to boost their student's GPA to enhance their students' admission performance. Under the new policy, however, when there are many ties due to the easy test

[11] Ministry of Education, *Press Release*, August 26, 2004.
[12] Ibid.

their percentile is assigned by taking average of their position. That is, if there are twenty students with maximum score out of one hundred students, now the grades for them should be *Grade 2* for all twenty ties which is the second-highest grade instead of *Grade 1*.

Second, the CSAT results will be delivered only in the form of each subject's grade. No further information such as standard score and exact percentile was available as before. For example, a student who took 2008 CSAT and achieved 134 points in math which fell into 97 percentile of population and is assigned to be the first grade which is only given to the top four-percent students in the subject. Overall, the new grading policies give more weight on high school GPA and less weight on College Scholastic Ability Test (CSAT) in college admission process by adjusting informativeness of them. Table 1 shows the grade distributions of GPA and CSAT. Since a CSAT raw score is standardized to assign a grade over normal distribution, the median student's grade (50 percentile) in each subject will be Grade 5 which covers 40 $\sim$ 60 percentiles.

Figure 3 presents the policy summary and the implementation sequence. It was announced in 2002 and first applied to the 2005 high school freshmen. Before the policy, GPA was given by five grades in an *absolute basis*. For example, when a student gets 90 out of 100 which is the cutoff of the highest grade she would get grade 1. If she gets 75 out of 100 the grade 3 will be given regardless of difficulty of the test and the distribution of the test score. Further, there was no information about the distribution of test scores. Under previous regime, there existed an incentive for school teachers to make easier exams to give more generous GPA to their students. The result was inflation of GPA and colleges who know this situation distrusted high school GPA and more relied on CSAT which is the standardized nationwide exam.

The motivation of the policy can be found in a survey by the Organisation for Economic Co-operation and Development (OECD, 2017). Among OECD countries Korea ranked at the bottom in descending order of the percentage of 15-year-old students who reported being very satisfied with their life. Furthermore, the survey describes that more than 20% of student in Korea reported that they are not satisfied with their life while less than 4% of students in the Netherlands reported dissatisfaction. Another interesting result of the survey is that only 10% of Korean students cited the pleasure of studying as a motive for taking additional classes. Kahneman and Krueger (2006) show that those who report less satisfaction with their lives as a whole also spend a greater fraction of their time in an unpleasant state. Taking the fact that the 15-year-old Korean respondents of the survey have already taken 6.4 years, which is the longest among OECD countries, of extra instruction after school hours in average into account, Korean students would have been spending time on unpleasant learning activities throughout their childhood and adolescent.

The main purpose of those additional lessons in Korea is to prepare a specialized high school admission and a good college admission by learning English, math, and science in advance of the regular curriculum[13]. Meanwhile, percentage of Korean students who reported that they exercise or practice sports before or after school is only 46.3% which is the lowest not only among OECD countries but also among twenty one partner countries including China, Brazil, Russia, and Singapore. That is, Korean high school students spend their time at cram schools rather than at play fields. In my sample year of 2009, 62.8% of total high school students and 69.2% of GPA top 10% students have taken private lessons (Kostat, 2009). Also, Korean private tutoring market accounted for 1.2% of Korean GDP. The result is adverse welfare effect on teenagers' physical and mental health. An average Korean high school student slept only 5.4 hours a day while studied 11 hours a day. Their biggest concern was academic performance (69.1%). A half of high school students have a discriminatory experience due to their low academic performance. In addition, the main cause of parental conflict is academic achievement (50.1%) (Kostat, 2010).

[13]All types of prior learning has been strictly prohibited by the special law for normalization of public education from 2014.

2 Theoretical Framework

In this section, I model a welfare maximizing student's problem to show that, first, coarse grading increases incentive to invest more on admission preparation when agents are loss averters. Second, the model shows that the behavioral reactions to the admission weight change is theoretically ambiguous. Third, Coarse grading CSAT increases uncertainty in student's status to induce higher efforts. Forth, I show that the relative grading scheme for high school GPA raise the degree of competitiveness among students compared to the original absolute grading policy. Fifth, I show the policy can lead to inefficient learning. Sixth, testable implications are discussed.

2.1 Baseline Model: Students' Time Allocation

I model a student's two-period lifetime utility maximization problem to illustrate time allocation behavior to college admission related study and leisure activities. A high school student lives two periods; high school days ($t = 0$) and the rest of her life ($t = 1$). She wants to maximize expected total lifetime utility U. A representative student's optimization problem is modeled as follows:

$$\max_{l_0, l_q, k_q, x} u_0(l_0, x) + \beta u_1(Q) \tag{1}$$

$$s.t. \quad l_0, l_i, k_i, x > 0 \tag{2}$$

$$l_g + l_s + l_r + l_0 = l \tag{3}$$

$$p_g k_g + p_s k_s + p_r k_r + px \leq I \tag{4}$$

and

$$Q = w_g \cdot g(l_g, k_g) + w_s \cdot s(l_s, k_s) + w_r \cdot r(l_r, k_r) + \epsilon. \tag{5}$$

where u_0 is a current utility ($t = 0$) from current leisure l_0 in high school days and u_1 is a discounted future utility ($t = 1$) depending on a quality of college admission Q which is a function of high school GPA $g(l_g, k_g)$, SAT score or grade $s(l_s, k_s)$, resume $r(l_r, k_r)$, and unobservable variables ϵ. I consider myopic (or present bias) students who are assumed to enjoy future utility u_1 from their college admission directly, but the delayed benefit u_1 is discounted by $0 < \beta < 1$. I further assume that the utility function $u(\cdot)$ is increasing, concave, and continuously differentiable in its components.

The admission quality function Q is a non-decreasing function of each component which implies that the partial derivatives of Q with respect to g, s, and r are non-negative. Accordingly, getting a more competitive college admission Q brings higher lifetime utility after high school graduation as it potentially guarantees a better signal on labor market and a more productive social network. The determinants of admission quality g, s, and r are all non-decreasing functions of time and money. That is, the production functions of high school GPA (g), SAT scores (s), and quality of resume (r) are the increasing functions of time (or labor) and money (or capital) similar to a competitive firm in a goods market. Marginal productivity of time l_q and money k_q to each academic production $q = g, s, r$ are assumed to be diminishing in their production factors. For example, achieving higher grade in a CSAT subject is getting more difficult and requires more effort and financial investment.

I introduce weights in college admission w_q where

$$\sum_{q=g,s,r} w_q = 1 \tag{6}$$

to reflect the policy change interested. I assume that the admission quality Q is determined by the weighted average of three admission components g, s, and r. As an extreme case, when only CSAT score determines the admission result without considering high school GPA or resume at all, the weight on CSAT w_s in the function $Q(g, s, r)$ would only be positive and equal to one. Before 2007, most Korean colleges relied heavily on CSAT scores to screen their applicants. Further, some students from the specialized high schools such as natural science or foreign language high schools could apply to the CSAT-only admission channel of the major universities as they have adverse high school GPA due to the highly competitive school cohort. After the policy aiming to promote public education and reduce student's stress from preparation of CSAT, colleges had to enhance the weight on school records such as GPA and resume in their admission process following government policy. Ministry of Education was able to enforce this policy by controlling information of CSAT results to students and colleges and by implicit financial support threat to colleges as many colleges rely heavily on the central government subsidy.

Equation (5) implies that a student's admission quality is determined by her weighted average of CSAT score, high school GPA, resume, and unobservable factors. Colleges use Equation (5) in their admission process with different weights depending on their preference. By the definition of admission quality Q above, a partial effect of a marginal increase in GPA, CSAT, or resume on Q equals its weight. That is,

$$\frac{\partial Q}{\partial q} = w_q. \tag{7}$$

Equation (1) and (5) imply that student behavior and welfare would vary with the college admission policy. Noticing that most colleges in Korea assign weights on each component g, s, and r to look over students' applications, we can state the admission quality $Q(\cdot)$ as a function of the weighted average of three criteria and unobservables ϵ.

The representative intertemporal utility maximizing student faces a time constraint in Equation (2) and a household budget constraint in Equation (3) where l_q with $q \in \{g, s, r\}$ denotes input of time on each academic activity q, l_0 is time on non-academic activities, k_q is input of money on each academic activity q, p_q is the price of each academic activity q, and px denotes total money on other goods and services. This model assumes that students only use their time in studying or leisure activity. Current leisure l_0 provides direct utility at $t = 0$ while time spending on academic activities l_g, l_s, and l_r combined with capital investment k_q produces admission factors such as GPA, CSAT score, and resume which ultimately contribute to the admission quality $Q(g, s, r)$.

As almost all public and private high schools in Korea follow the common curriculum designed by the Ministry of Education, CSAT subjects including Korean, English, math, and social and natural science are almost identical to those being taught and examined quarterly in each high school. In fact, questions of CSAT should be set based on textbook written by the Ministry of Education or authorized publishers and workbook published by Korea Education Broadcasting System (EBS) which is the official CSAT preparation material designated by the Ministry of Education. In this context, the effect of one additional hour of study on CSAT math would also have positive impact on GPA. That is, it is natural to expect that for every level of l_q and k_q,

$$\frac{\partial q}{\partial l_q} > \frac{\partial q'}{\partial l_q} \geq 0 \tag{8}$$

$$\frac{\partial q}{\partial k_q} > \frac{\partial q'}{\partial k_q} \geq 0 \tag{9}$$

where $q, q' \in \{g, s, r\}$ and $q \neq q'$. For simplicity of analysis, instead, I only consider the net effect of l_q and k_q on q after normalizing the indirect effect to be zero. In other words, the complementary effect of allocating resource to one activity on other activity is considered null. This assumption yields

$$\frac{\partial q}{\partial l_q} > \frac{\partial q'}{\partial l_q} = 0 \tag{10}$$

$$\frac{\partial q}{\partial k_q} > \frac{\partial q'}{\partial k_q} = 0. \tag{11}$$

Equation (10) and (11) imply that even though the basic concepts and formula tested by CSAT and school exams for GPA are the same, when a student invests one additional hour to preparing CSAT math her marginal increase in math GPA would be zero since her math midterm exam tests the very specific chapter in the workbook or teacher-made test bank. To capture disparate students' inherent abilities to study or total factor productivity in producing an admission factor $q(\cdot)$, I introduce individual heterogeneity in the model. Each student has different production technology of GPA, CSAT, and resume denoted by i. Then, the individual admission quality Q_i can be written as

$$Q^i = w_g \cdot g^i(l_g^i, k_g^i) + w_s \cdot s^i(l_s^i, k_s^i) + w_r \cdot r^i(l_r^i, k_r^i) + \epsilon^i. \tag{12}$$

In other words, given the same amount of leisure l_q and money k_q invested in a certain admission activity q, the marginal productivity of time $\frac{\partial q^i}{\partial l_q}$ and capital $\frac{\partial q^i}{\partial k_q}$ are heterogeneous for every student. The individual ability or productivity for each activity g, s, and r is unobservable private information but assumed to be normally distributed over the population cohort. The equilibrium conditions imply that the lifetime utility maximizing student will balance her time so that the marginal utility gain from the last hour spent on each activity is identical. See Appendix for the derivation of equilibria.

2.2 Ambiguous Behavioral Reactions

How does the new grading policies affect students' behavior? This is a key question of this paper. Coarse grading of CSAT restricts the information quality of CSAT transcript. Thus, the policy forces colleges to adjust their weights on admission criteria. Depending on the individual productivity q^i and the predicted CSAT grades[14], the directions and magnitudes

[14]Korean students predict their CSAT rankings by referring to their scores of monthly private and quarterly public CSAT mock tests. A typical college applicant takes ten to thirty mock tests throughout the three years of high school period. It usually takes about eight hours to complete a set of CSAT mock test.

of students' resource allocation responses ∂l_q and ∂k_q to the policy change would vary. As Han et al. (2016) suggest, coarsening CSAT grades may have two offsetting effects. There can be an increase in time investment on CSAT when a student thinks her score is around the cutoff of a certain grade. This is because the marginal change of grades in any direction would make a critical change of her admission quality. Contrary to this, another student with a score far from upper or lower bound of a class or who does not need to enhance her grade would like to reduce resources allocated to CSAT. The equilibrium first-order conditions (21) - (24) also capture this ambiguous direction of effect. That is, the direction of alternative use of the reduced or increased leisure and money from CSAT preparation is uncertain. They can be consumed as l_0 and x to obtain more u_0 while keeping the level of Q constant or invested as another academic activities $l_{q\neq s}$ and $k_{q\neq s}$ to get better future rewards. The empirical exercise would be able to identify the average policy effect on the whole samples and also on the various sub-groups.

2.3 Coarse Grading and Uncertain Status of the Loss Averters

This subsection shows that the loss averse decision makers may want to avoid the situation of being in lower grades in the CSAT subjects than their *reference status* under the coarse grading scheme. Coarsening CSAT grades may cause two opposite effects on student's incentive to work hard depending on her *status*. More details will be discussed in 2.2. This subsection investigates the environmental change in student's status with respect to uncertainty of ranking under the new CSAT grade system. Harbaugh and Rasmusen (2018) shows that coarse grading scheme such as pass-fail grading reduces public uncertainty by minimizing expected and actual quality of sender because the extra participation outweighs the coarser reporting. As almost all college applicants take CSAT, however, college admission game ensures full participation already regardless of grading scheme. Thus, the relationship between the coarse grading and student's ranking uncertainty in the admission game might be different from the case in Harbaugh and Rasmusen (2018).

First, I show that there exists *ex-ante* uncertainty of a student's *rank* at the time of student's resource allocation decision before the score/grade distribution of CSAT is finalized after the actual tests. In my model as described in Section 2, a student exerts her best effort by choosing the optimal level of time use l_s^* and the optimal amount of money k_s^* to produce CSAT score (or grade) $s(l_s^*, k_s^*)$. Given a student's input level (l_s^*, k_s^*), standardized score method ensures almost continuously increasing rank function in CSAT score. That is, student i's CSAT score s_i would be distinguishable to student j's score s_j if there is a marginal difference between s_i and s_j. As long as $|s_i - s_j| > 0$, colleges can obtain an informative signal by comparing s_i and s_j. By contrast, coarse grading system may not generate an informative signal if $|s_i - s_j|$ is not sufficiently large enough to separate those two scores into different grades. A critical environment for students who make decisions on their resource allocation is the fact that no students, parents, and teachers can know the population distribution of CSAT score or grade prior to the actual test. Suppose the math in a CSAT is really easy so that more than 4% of the CSAT takers got the full credits, then the highest grade can be assigned to all those students even though the top 4% is the cutoff guideline for the first grade. In the opposite situation, if the math is too hard so only less than 1% of applicants got the full credits, the cutoff of the highest grade will be at around the top 4% and the cutoff score should be way lower. Suppose a student decides her optimal investment level for CSAT l_s^* and k_s^*. Ignoring any uncertainty in her SAT score without loss of generality, she will obtain the raw score of $s(l_s^*, k_s^*)$ in a CSAT subject. Her *status* or percentile, however, is *ex-ante* uncertain as the difficulty of the exam and the performance of her cohort may not be the same as previous years and previous mock exams. Thus, she does not know her exact status corresponding to her CSAT score $s(l_s^*, k_s^*)$ at the time she makes her time allocation decisions. Under the uncertain status, Equation (12) can be revisited with a random factor on the CSAT part to generate uncertainty of status affecting the admission quality.

$$Q^i = w_g \cdot g^i(l_g^i, k_g^i) + w_s \cdot R(s^i(l_s^i, k_s^i), s^{j \neq i}(l_s^j, k_s^j), d) + w_r \cdot r^i(l_r^i, k_r^i) + \epsilon^i. \tag{13}$$

The *status* (or ranking) function $R(\cdot)$ which converts individual CSAT score s^i to a rank or a grade over the population distribution is determined *after* the test by the degree of coarse grading d and the CSAT scores of all CSAT takers. The functional form of d is determined by the government. An increase in d means an increase of the number of students assigned into a grade. For example, the maximum d would be the *pass or fail* system where there exist only two partitions above and below a certain cutoff point. The minimum d would be the system with the finest grading where the raw scores are available for every player in the admission game. Note that this source of ex-ante uncertainty $R(\cdot)$ is different from the overall unobservable factors ϵ^i.

Second, I assume that Korean high school students have *loss aversion* preference on the future utility $u_1(\cdot)$. The key features of loss aversion is asymmetric reaction to gain and loss. That is, loss aversion describes the concavity of utility for gains and the convexity of utility for losses. The intuition is straightforward. A college applicant will be happy if she luckily gets a slightly better admission than she expected. By the way, if she gets a slightly lower admission than her reference college, her parents, teachers, friends, and herself will be very disappointed since everyone thinks that she could have done better if she performed as usual. As in Köbberling and Wakker (2005), I modify the future utility function $u_1(\cdot)$ to a value function

$$v_1(Q(\cdot)) = \begin{cases} \frac{1-e^{-\mu Q}}{\mu} & \text{for all} \quad Q \geq Q^* \\ \lambda(\frac{e^{vQ}-1}{v}) & \text{for all} \quad Q < Q^* \end{cases} \tag{14}$$

where Q^* is the reference level of admission quality determined by referring to the previous mock test results. λ is the degree of loss aversion and takes a value of $\lambda \geq 1$. For $0 < v < \mu$,

v_1 is convex for losses, concave for gains, and importantly, given the same deviation of admission quality ΔQ from a student's reference admission, the student losses more utility under the worse admission than the amount of utility gain from a higher admission relative to her reference admission. This implies $-v_1(Q^* - \Delta Q) > v_1(Q^* + \Delta Q)$. Adapting the loss aversion value function $v_1(\cdot)$ instead of $u_1(\cdot)$ yields

$$\max_{l_0, l_q, k_q, x} u_0(l_0, x) + \beta v_1(Q) \tag{15}$$

where

$$Q = w_g \cdot g^i(l_g^i, k_g^i) + w_s \cdot R(s^i(l_s^i, k_s^i), s^{j \neq i}(l_s^j, k_s^j), d) + w_r \cdot r^i(l_r^i, k_r^i) + \epsilon^i. \tag{16}$$

Equation (15) summarizes a student's lifetime welfare, consisting of the normal form of current utility u_0 and the loss averse utility v_1 discounted by β.

Finally, I investigate the effect of coarse grading on loss averse student's resource allocation behavior. What will happen if the degree of coarse grading d increases by adopting the nine grade-class system? I compare d_0 and d_1 where d_0 denotes the original finest grading scheme with the standardized CSAT scores and percentiles available. d_1 denotes the new coarse grading scheme with the nine grades proportionally assigned over the standardized normal distribution with no available percentile and raw scores. Accordingly R_0 denotes the student's CSAT rank under the finest grading scheme d_0 raging from the lowest CSAT score which is zero to the highest standardized score s_{max}. R_1 denotes the student's CSAT rank under the new coarse grading system.

$$\begin{cases} R_0 \in \{0, 1, 2, ..., 50, 51, ..., s_{max}\} \\ R_1 \in \{1, 2, ..., 9\} \end{cases} \tag{17}$$

Table 2 shows the population statistics of CSAT results from 2005 to 2009 school year. My sample cohorts are included in the 2005 and 2008 school year tests. Standardized scores are calculated under the assumption that the population scores are normally distributed[15]. When a CSAT subject is easier than the other years, the population mean is high and the standard deviation tends to low as students may get some baseline points from the easy questions. These two effects yield the lower Z-score. As a result, the cutoff standardized score for Grade 1 becomes lower. This implies that a harder exam yields the higher maximum standardized score s_{max}. For example, the cutoff standardized score of Grade 1 of CSAT Korean in 2006 school year is 125 which is the lowest among the other years because of its low degree of difficulty (and higher population mean). This means that if a smart student missed just only one single three-point question in the 2006 CSAT Korean by her trembling hand and got the raw score of 97 out of 100, then she will get Grade 2 which is *not* the highest grade. Under d_0, the same amount of gain or loss ds of raw score from a student's reference score s^* will bring a minimum change of the student's status dR_0. Under d_1, with the same amount of deviation (gain and loss) ds of raw score from one's reference score s^*, the utility change of the *loss averse* student will be greater in the case of loss rather than the case of gain. This implies that the new grading scheme d_1 increases variance of status $Var(R)$ given the same resource allocation s:

$$Var(R \mid d_0, s) < Var(R \mid d_1, s). \tag{18}$$

I assume that the individual variance of CSAT score is symmetric around the reference score. That is, the probability of getting n points higher score than the reference score s^* is the same as the probability of getting n points lower score than s^*. Repeated mock tests help

[15]Standardized score is a linear transformation of Z-score which is calculated by the formula

$$\text{Standardized Score} = (\text{Reference Mean}) + (\text{Z-score}) \times (\text{Reference Standard Deviation})$$

where Z-score = (Student's Raw Score - Population Mean) / Standard Deviation. Reference mean (Reference standard deviation) for Korean and math are 100 (20) and for social and natural sciences are 50 (10).

students to update their reference scores and to make the optimal exam day routine. As a result, the only factor that can affect her CSAT score other than her resource investment is *luck*. From Equation (18) a student's expected loss-averse utility under the new coarse grading scheme d_1 is now smaller than that under d_0 as d_1 gives more drastic changes in status. To avoid the critical loss of quality of admission under the new coarse grading policy, the incentive to work hard and to invest more money on cram schools would increase. If this is true, the policy which aimed to reduce competition in CSAT would have failed.

2.4 Relative Grading GPA and Harsh Competition

Another main change in the admission policy is the relative grading high school GPA which was implemented against the pervasive high school GPA inflation due to the *easy* midterm and final exams administered by each high school authority under the absolute grading scheme[16]. Under the new grading policy, students face more competitive environment at school as only top 4% of the school cohort can obtain the highest grade. When there are more than 4% of ties at the top, all students should get a grade corresponding to their average percentile. For example, if 20% of students in a school got full credits due to an easy exam, those 20% students' percentile becomes 10% from the top. This implies that no student can get the first grade as no one marked top 4% in this exam. Thus, teachers have now no incentive to make exams easy as low difficulty will harm all students in ties. As a result, it is natural to observe grade deflation under the new relative grading scheme. That is, now high school GPA provides more informative signal about individual student's status on the standardized normal distribution in her school cohort while students have strong incentive to invest more resources and compete with their friends than before to maintain their desired GPA given the limited portion of good grades.

[16]For example, Yonsei University reports in 2004 that among their applicants, 7 high schools gave grade A to *all* students in a subject. Also, 37 schools gave grade A to more than one hundred students in a subject by making very easy midterm and final exams.

2.5 Inefficient Learning

Efficient in learning is achieved when a student is able to complete high school courses at the minimized costs and obtain desired college admission. Contrary to efficient learning, when a student could get a high school degree and the admission only by sacrificing almost all of her household income or at least by investing more than the optimal level of resources, the learning is *inefficient* and over-invested. Suppose n applicants are applying to college A and only q qualified applicants per academic year will be admitted where $n > q$. Applicants are sorted by their admission quality function $Q^i(g^i, s^i, r^i, w_q, \varepsilon)$ where $i = 1, ..., n$. The q students from the top Q^1 to Q^q will be admitted and $(q+1)$th applicant cannot achieve the admission A. Admission A is a binary outcome

$$A = \begin{cases} 1 & \text{if} \quad Q^i \geq Q^q \\ 0 & \text{if} \quad Q^i < Q^q \end{cases}$$

where Q^q is the qth highest student's quality. For simplicity, I assume that there is no identical quality of applicant among any of two or more applicants $i \neq j$. That is, for all $i \neq j$,

$$Q^i \neq Q^j.$$

Assume that a student i updates her expected admission quality $E(Q^i|\Omega)$ and this year's expected admission threshold $E(Q^q|\Omega)$ by using information set Ω which includes previous CSAT mock test grades and score distribution, current HGPA, and honors and awards history. When she expects that her admission quality $E(Q^i|\Omega)$ is less than $E(Q^q|\Omega)$, she may want to invest more on studying CSAT, preparing midterm and final exam, and polishing resume to achieve higher Q^i than the threshold level of admission quality Q^q. Other applicants who know and observe their rivals' efforts may also have incentives to invest more on admission preparation. This implies that other student's performance would be reflected in Ω and then student i makes investment decision depending on her rival's effort level. Given individual productivity, students will compare an additional lifetime benefit of resource investment, which would be a better admission, and an additional loss of resources. With the limited number of admission, it is inevitable to compete with other students. This results marginal cost of producing one more unit of admission quality to be getting higher as more rivals join this competition and invest more. At some point of this investment war, students face extremely high marginal cost of Q due to diminishing marginal product of leisure and money invested to g, s, r and also due to the increased level of the expected threshold $E(Q^q|\Omega)$ by rivals' investment. That is, as the degree of admission competition increase, students should spend more resources to obtain the same level of admission A which would distort household decision and other markets leading to the over-investment in the college admission game.

3 Data

The analysis in this paper is based on Korean Education and Employment Panel Survey (KEEP) data surveyed by Korea Research Institute for Vocational Education & Training (KRIVET) which contain information on student and household demographics, income and wealth, school activities and test scores, and employment status. The survey is a longitudinal research study of two cohorts – young and old – that extracts nationally representative samples from the population group and traces them for over twelve years from 2004 to 2015. See Appendix for the explanation how the sample is constructed.

A total of 6,052 respondents were selected to participate in this survey, comprising the younger cohort consisting of 2,294 middle school seniors (3rd graders) and the older cohort consisting of 1,887 high school seniors (3rd graders) and 1,015 vocational and technical high school seniors (3rd graders) at the year of 2004. Supplemental data were provided by teachers and school administrators as well. The data tracks two cohorts. before and after the policy change at the period to investigate their behavior in high school. Thus, my sample from the different two cohorts can be regarded as repeated-cross sectional data.

3.1 CSAT Score Distribution

Table 2 shows the population statistics of CSAT grade and score distribution for four subjects, Korean, Math A, Math B, and English, from 2005 to 2009 including my two sample cohorts. Every year, the cutoff raw score for each grade in a subject varies depending on the exam difficulty. For the comparability across different cohorts, each CSAT subject's result is standardized and the grade cutoffs are assigned over the standard normal distribution. The around-median (40 $\sim$ 60 percentile) students in each CSAT subject receive grade 5. The top and the bottom 4% students will be assigned to grade 1 and grade 9, respectively. An average student in 2008 cohort has slightly higher grade ranging from 5.06 to 5.14 for three major subjects than the one in 2005 cohort whose grades ranging from 5.48 to 5.59, but those are located around the median grade of the population which is exactly 5.

Table 3 describes the available information of CSAT results before and after the policy change in my data. Some students took CSAT more than once over multiple years after their high school graduation to get a better college admission. In this paper, only the first-taken CSAT results of high school students in each cohort are summarized. 2005 admission CSAT (taken in 2004 Autumn) takers received all types of result available on their transcripts including standardized scores, grades, and percentile from the authority. College applicants provided all score information to the applying colleges, maximum three colleges for each student, combining with their high school GPA and resume. Students of 2008 admission CSAT (taken in 2007 Autumn) group were available only to obtain their grades for each CSAT subject. This implies that students do not know their exact score and the percentile on the population distribution. For example, the grade 1 which is the highest among nine grades is assigned to those who ranked within the highest four percentage in the population of the subject. That is, the top scorer and a student slightly above the four percentage cutoff would have the same grade. Colleges then are not able to distinguish among those two students by their CSAT scores. To screen proper applicants colleges need to more closely review other documents such as high school GPA and resume than the inexact CSAT grades.

4 Estimation Strategy

Following the causal inference framework organized by Imbens and Rubin (2015), I use difference-in-differences estimation to investigate causal effects of the grading policy change. This can be said to be the differences between potential outcomes of each students' time allocation behavior with and without policy change. Due to the nature of nationwide policy I interested, all students who prepare CSAT should be affected by this policy. Given this policy environment, My identification strategy is to observe outcomes of CSAT takers and CSAT non-takers in two cohorts before and after policy change in 2005. As we do not observe counterfactual potential outcomes such as CSAT taking student's leisure time that she would have allocated when there was no policy change, I investigate appropriate comparison group to see the effect of policy on CSAT takers.

4.1 Difference-in-Differences Estimation

In this policy analysis, I use difference-in-differences estimation with repeated cross-sectional survey data of two different high school cohorts before and after the policy to identify causal effects on student's resource allocation behavior and welfare. As far as I know, no randomized controlled trial study for this education policy at the time of implementation was conducted. Thus, the policy effects can only be investigated by using observational data. Clearly, prior to the implementation of the policy in 2005 no student in my sample was exposed to the policy change. Also, the high school GPA policy change was applied to the 2008 CSAT cohort for the first time in 2005 when they were the high school freshmen. New CSAT grading policy was also applied to the CSAT taken in 2007 for 2008 academic year. Thus, only the younger cohort in my sample was affected as they studied at high school under the new grading regime[17]. Since the policy was announced in 2004 to apply to the 2008 CSAT cohort who are 2005 high school freshmen, there could be anticipation effects on them. Endogenous choice of CSAT take would be a possible threat to identifying the non-intervention group. This is discussed in Section 4.4.

Among those in the younger cohort, only students who want to go to college took CSAT. That is, CSAT taker group is more affected by the policy change. Interacting variables are *post* and *CSAT taker* dummy variables. Under the parallel trend assumption, the difference-in-differences (DD) estimator suggests the time changes over the policy in the average outcomes for the CSAT takers and non-takers. The validity of the assumption is discussed in Section 4.5.

4.2 Primary Outcome Variables: Weekday Leisure

The primary outcome variable of this study are student's *weekday leisure hours per day* during their *last year of high school*. As all Korean high school students are subject to the almost same daily class schedule planned by the central government, I expect that the quantity of time use is the key input for student's welfare and quality of college admission. Thus, in accordance with my model in the previous section, I take weekday leisure time as the primary outcome variable. Also, secondary outcome variables of interest are selected to evaluate students welfare. I consider time spending on private education and time of

[17] In fact, the 2008 CSAT cohort was the only one cohort who could not obtain their CSAT results in forms of percentile and standardized scores. Korean government regressed to the original CSAT grading policy from the 2009 CSAT cohort.

self study as the secondary outcome variables to measure the policy effect on students' welfare. Unfortunately, amount of financial resources spent to cram schools is not included in this study due to the inflation during the period and large numbers of missing data and inconsistency of survey questions across cohorts.

Figure 4 shows mean outcomes by group. I find that average weekday leisure is enjoyed more in the CSAT non-takers group than the CSAT takers group. Hours spent for self study and private lesson also differ between the groups at the baseline. Proportion of students who participate in academic club activity tends to be higher in CSAT takers group. CSAT takers are elected more as class or school leaders. These differences in baseline outcomes implies that there exists baseline level difference of outcome variables across groups. One way to interpret this difference is that CSAT takers typically spend more time to study than the non-takers who are relatively under less time pressure during their last year of high school.

4.3 Empirical Model

The econometric approach for analyzing the impact of the policy change on student's behavior is given by a difference-in-differences model where I compare outcomes for CSAT takers group to non-takers group. I estimate

$$y = \beta_0 + \delta_0 post + \beta_1 CSATtakers + \delta_1 post \times CSATtakers + \mathbf{x}\delta + u \tag{19}$$

where y is the outcome of interest such as time allocation. The DD estimator δ_1 captures the intervention effect of the policy on the outcome variables of CSAT takers group compared to the unexposed outcome without the policy. δ_1 is considered to be zero in the absence of the policy under the proper assumptions. I cluster standard errors by the sample students' high school.

The first estimation is done without covariates $\mathbf{x}$. The second estimation includes $\mathbf{x}$ which is a vector of controlled variables including baseline characteristics such as a vocational school dummy, parents' education level, household income and net asset, and number of siblings which have been fixed at the time of intervention. Also, I control metro-level school location to capture unobservable regional variation across the different school districts. In Korea, there exist huge regional differences in educational environment and fundamentals such as quality of school and teacher defined by historical college admission results, famous private academy lecturers, *hagwons*, and parents' information set. For example, the binary variable *Seoul* is 1 for students in schools located in Seoul which is the biggest city and has the most competitive private education market in Korea. Thus, considering for this school district fixed effects would be able to account for regional variation of educational environment. is a vector of controlled variables including baseline characteristics such as a vocational school dummy, parents' education level, household income and net asset, and number of siblings which have been fixed at the time of intervention.

4.4 Policy and Selection on CSAT Taking

An issue in the composition of the CSAT takers and non-takers groups is selection on CSAT taking due to the policy changes. A student's CSAT taking decision should not be affected by the policy change. In practice, it is hard to tell whether the CSAT taking decision by a student was made before or after the policy announcement or was made based on the consideration of the policy change. Willis and Rosen (1979) suggest that expected gains in life earnings, family-background, financial constraints, and tastes mainly influence the decision to attend college. Adopting this framework, it would be rare to find a student whose preference is almost indifferent between going to the college or working after high school graduation and the only critical factor for her to make college decision is the policy change.

Table 5 shows that there was no remarkable change of the CSAT taking trend among the high school graduates measured by the proportion of CSAT applicants before and after the policy change in 2008 out of population high school graduates. The proportion of CSAT takers maintained mid-seventy percent out of population graduates during 2005 to 2009 which covers my two sample cohorts of 2005 and 2008. Thus, I assume that the policy did not affect CSAT taking decision.

4.5 Balance in Potential Outcome Trends

My DD design requires that the trend of *outcome variables* observed in the non-takers group is the same trend that would have been observed in the CSAT takers group in the absence of the grading policy change. This so-called "*parallel trends*" assumption may be convincing if in Equation (19) the period fixed effect δ_0 and the average group fixed effect β_1 are invariant across the groups and time respectively, so that the *difference* in outcomes between groups is time invariant. In this study, parallel trends assumption implies that the changes or trends in resource allocation behavior of the CSAT taking group in the 2008 CSAT cohort would have been the same as that of the non-takers group in the 2008 CSAT cohort if there was no grading policy shock. Then, δ_1 in Equation (19) would be the valid DD estimator.

I provide evidence to support the hypothesis of parallel outcome trends between groups. There are several tests to *support* the assumption used in the empirical studies by investigating the data. First, the similarity in the pre-intervention outcome trends between the groups can make the assumption more plausible even though the assumption is about the trends in post-intervention counterfactual outcomes. It is useful to check the pre-intervention outcome data to predict the counterfactual post-intervention outcome trends. The more similar the pre-intervention outcome trends, we might be more easily convinced that the counterfactual outcome trend of CSAT takers would be similar to the post outcome trend of the non-takers group. Unfortunately, visual inspection of multiple periods of pre-intervention outcomes is not available in my observational data. It includes only two cohorts data and the first observed period of the post-intervention cohort is the year of the policy implementation. An alternative inspection is done by comparing the eight years of *post-intervention* outcome trends after the high school graduation between the groups. The key assumption of this investigation is that the policy only affect the students' high school leisure allocation behavior and does not have any effects after graduation. I take up this issue in Subsection 6.1. Second, the balance in descriptive pre-intervention characteristics between groups can be shown to check the similarity of two groups respect to the various baseline features in average. Subsection 6.2 discuss this.

5 Results

5.1 Average Policy Effect

Table 6 Column (1) shows that the policy reduces leisure by 30 minutes a day. This could suggest that the level of current period – high school days – utility u_0 would have been worsen off after the policy. Column (2) indicates that the policy encouraged students to take more private lessons. An average CSAT taker spent 0.8 hour more on cram schools after the policy while Column (3) shows a 1.3 hour increase in self study time. More private lesson and self study time imply that college applicants are worse off in terms of their current utility as a result of a decrease in current leisure l_0. They have invested more on college prep activities to enhance their GPA and CSAT. Column (4) indicates that 5% point more CSAT takers took the roles of class leader or school leader. Column (5) presents no significant impact on academic club activities.

Recalling the fact that the main policy goal is to reduce welfare-harming admission competition, the primary analysis above implies that there is no remarkable evidence of increase in students' leisure which is a key ingredient of current welfare u_0. Thus, it is hard to say that the policy had successfully achieved the goal. Indeed, students had to spend more time studying during their high school days.

5.2 Subgroup Analyses

I investigate the heterogeneous effects of the grading policy by high school GPA. Table 7 Panel A shows that the top 11 percentile GPA students (Grade 1 and Grade 2) reduced their leisure significantly by more than an hour per day. My model expects that the saved time from the shorter leisure would have been invested to college admission preparation activities. I find that the top GPA students invested more than an hour per day to study by themselves while there are no significant changes in private lesson, leadership and academic club activities. The estimation result confirms that the current leisure was substituted to additional investment in self studying.

Table 7 Panel B provides the restricted analysis results on the low GPA students. Contrary to the top students, I do not find any significant changes in all outcomes of the bottom 23 percentile GPA students (Grade 7 – 9) but a half-hour increase in self study. I find the heterogeneous effects depending on students' GPA. As GPA contains information about student's unobservable inherent productivity and motivation, students may respond differently to the incentive changes.

6 Robustness

6.1 Balance in Non-Intervention Outcome Trends

Unlike other DD studies, the grading policy is only applied to the college applicants. After the admission competition, all students become adults and are free from the policy except some students who decide to prepare their second or third CSAT. Since those retakers are not going back to the regular high school curriculum, their time consumption patterns are also considered to be one variation of early twenty's lifestyle. Thus, I compare the weekday leisure trends between the two groups in the two cohorts for eight years (from year $T+1$ to $T+8$) after the year of their first CSAT which were taken in the last year of high school (T).

Figure 9 displays box plots which show that all four groups have the similar trends of weekday leisure after the year of CSAT ($t=0$). The median weekday leisure hours per day in the first year after the CSAT ($t+1$) is about five hours in all four groups. This could possibly show that without the admission competition, The two groups have the similar time allocation patterns in average. From the second year after graduation, the trends are slightly going down for all groups as they are getting older to reach their late-twenty. Supplement time use trends of daily sleep hours in Figure 10 also shows similar patterns to Figure 9 across groups. Lack of sleep in adolescence may cause developmental problems such as learning, memory, attention, cognition, and emotion processing (Tarokh et al., 2016). Furthermore, insufficient sleep is significantly positively correlated with unsafe behaviors among youths such as risky driving, risky sexual activity, and drug-tobacco-alcohol use (Weaver et al., 2018).

6.2 Balance in Pre-Intervention Characteristics

The intervention which is the exogenous source of variation in this analysis is the time-specific grading policy changes on CSAT and GPA applied only to the younger cohort in my sample. To examine the impact of the policy, I focus on two high school graduate cohorts before (2005) and after (2008) the policy change. Total 435,538 college applicants out of 569,272 high school graduates in 2005 (76.51%) and 446,597 applicants out of 581,921 high school seniors in 2008 (76.75%) took CSAT for 2005 and 2008 college admissions, respectively. In this study, *CSAT non-takers groups* in each cohort consist of students who did not take CSAT ($CSATtaker = 0$) in their third-year of high school and were not affected by the policy changes. *CSAT takers groups* in each cohort are CSAT takers ($CSATtaker = 1$) who were planning to apply to college.

Figure 11 provides mean characteristics for all sample students by group and cohort. The female ratio of each group is more or less 50% which is not so far different from the population female ratio of 50%. Parents education level is suggested in row 2-3. The percentage of college graduate fathers is slightly higher in the CSAT takers groups for both cohorts. Except the mothers of non-takers in 2005 cohort, the ratio of college graduate mother is similar across groups. Household satisfaction is measured by a Likert scale from 1 to 5 to measure how students feel about their lives at home. Interesting fact is that CSAT taking students are more likely to have higher household satisfaction in both cohorts. Considering the fact that most of Korean parents want their children to go to college, this correlation seems to be plausible as parents and their child consented to the student's career. In summary, the two groups are balanced in terms of observable pre-intervention characteristics except some characteristics such as vocational school ratio and parents education.

7 Concluding Discussion and Policy Implications

I considered a college admission game in which high school students compete on the limited college admissions which are assigned in order of students' relative ranking. My model under the game of status suggests that the new policy of Coarse grading CSAT and relative grading GPA introduced uncertainty in expected return to CSAT investment and more competitiveness in GPA race which result in more investment in college prep activities.

This study finds empirical evidence that coarsening SAT grade and changing absolute to relative GPA grading resulted in unintended and negative impacts on student's leisure allocation. The results suggest that education policy makers need to understand the students' behavior under status uncertainty and highly competitive environment in designing college admission policies.

Enjoying life is important for all age groups. Especially, children who are in their physical and emotional development stages need to consume enough time with their family and friends to build up their healthy life style. When a market fails to bring a decent childhood for them, Government policy can change people's incentives to move away from an outcome failing to achieve the maximized social welfare. This study, however, shows that without comprehensive understanding of what makes people do what they do the government intervention may not work well.

References

Acemoglu, D. (2010, September). Theory, general equilibrium, and political economy in development economics. *Journal of Economic Perspectives 24*(3), 17–32.

Andrews, R. J., J. Li, and M. F. Lovenheim (2016). Quantile treatment effects of college quality on earnings. *Journal of Human Resources 51*(1), 200–238.

Athey, S. and G. W. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica 74*(2), 431–497.

Avery, C., S. Lee, and A. E. Roth (2014, December). College admissions as non-price competition: The case of south korea. Working Paper 20774, National Bureau of Economic Research.

Bailey, I. L., M. A. Bullimore, T. W. Raasch, and H. R. Taylor (1991). Clinical grading and the effects of scaling. *Investigative Ophthalmology & Visual Science 32(2)*, 422–32.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal 75*(299), 493–517.

Bertrand, M., E. Duflo, and S. Mullainathan (2004, 02). How Much Should We Trust Differences-In-Differences Estimates?*. *The Quarterly Journal of Economics 119*(1), 249–275.

Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review 19*(5), 547–556.

Bodoh-Creed, A. L. and B. R. Hickman (2018). College assignment as a large contest. *Journal of Economic Theory 175*(C), 88–126.

Bond, T. N., G. Bulman, X. Li, and J. Smith (2018). Updating human capital decisions: Evidence from sat score shocks and college applications. *Journal of Labor Economics 36*(3), 807–839.

Bound, J., B. Hershbein, and B. T. Long (2009, December). Playing the admissions game: Student reactions to increasing college competition. *Journal of Economic Perspectives 23*(4), 119–46.

Brewer, M., T. F. Crossley, and F. Zilio (2019, May). What Do We Really Know about the Employment Effects of the UK's National Minimum Wage? IZA Discussion Papers 12369, Institute of Labor Economics (IZA).

Che, Y.-K. and Y. Koh (2016). Decentralized college admissions. *Journal of Political Economy 124*(5), 1295–1338.

Dubey, P. and J. Geanakoplos (2010). Grading exams: 100,99,98,... or a,b,c? *Games and Economic Behavior 69*(1), 72 – 94. Special Issue In Honor of Robert Aumann.

Gibson, M. and J. Shrader (2018). Time use and labor productivity: The returns to sleep. *The Review of Economics and Statistics 100*(5), 783–798.

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

Gronau, R. (1977). Leisure, home production, and work–the theory of the allocation of time revisited. *Journal of Political Economy 85*(6), 1099–1123.

Hafalir, I., R. Hakimov, D. Kübler, and M. Kurino (2018, 05). College admissions with entrance exams: Centralized versus decentralized. *Journal of Economic Theory 176*.

Han, C., C. Kang, and S.-H. Lee (2016). Measuring effort incentives in a tournament with many participants: Theory and application. *Economic Inquiry 54*(2), 1240–1250.

Harbaugh, R. and E. Rasmusen (2018, February). Coarse grades: Informing the public by withholding information. *American Economic Journal: Microeconomics 10*(1), 210–35.

Hopkins, E. and T. Kornienko (2004, September). Running to keep in the same place: Consumer choice as a game of status. *American Economic Review 94*(4), 1085–1107.

Imbens, G. W. and D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

Kahneman, D. and A. B. Krueger (2006, March). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives 20*(1), 3–24.

Köbberling, V. and P. P. Wakker (2005). An index of loss aversion. *Journal of Economic Theory 122*(1), 119 – 131.

Lu, X. and H. White (2014). Robustness checks and robustness tests in applied economics. *Journal of Econometrics 178*, 194 – 206. Annals Issue: Misspecification Test Methods in Econometrics.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. Number 9780195102680 in OUP Catalogue. Oxford University Press.

Moldovanu, B., A. Sela, and X. Shi (2007). Contests for status. *Journal of Political Economy 115*(2), 338–363.

OECD (2017). *PISA 2015 Results (Volume III)*.

Olszewski, W. and R. Siegel (2016). Large contests. *Econometrica 84*(2), 835–854.

Olszewski, W. and R. Siegel (2019). Pareto improvements in the contest for college admissions. *Manuscript, Department of Economics, Northwestern University*.

Rothschild, M. and J. E. Stiglitz (1970). Increasing risk: I. a definition. *Journal of Economic Theory 2*(3), 225 – 243.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services & Outcomes Research Methodology 2*, 169–188.

Sari, N. and M. Osman (2015). The effects of patient education programs on medication use among asthma and copd patients: a propensity score matching with a difference-in-difference regression approach. *BMC Health Services Research 15*(1), 1 – 9.

Schultz, T. W. (1975). The value of the ability to deal with disequilibria. *Journal of Economic Literature 13*(3), 827–846.

Stuart, E., H. Huskamp, K. Duckworth, J. Simmons, Z. Song, M. Chernew, and C. Barry (2014, 11). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services and Outcomes Research Methodology 14*(4), 166–182.

Su, D., Y.-c. Chen, H.-x. Gao, H.-m. Li, J.-j. Chang, D. Jiang, X.-m. Hu, S.-h. Lei, M. Tan, and Z.-f. Chen (2019). Effect of integrated urban and rural residents medical insurance on the utilisation of medical services by residents in china: a propensity score matching with difference-in-differences regression approach. *BMJ Open 9*(2).

Tarokh, L., J. M. Saletin, and M. A. Carskadon (2016). Sleep in adolescence: Physiology, cognition and mental health. *Neuroscience Biobehavioral Reviews 70*, 182 – 188. The Adolescent Brain.

Thomas, D., J. Strauss, and M.-H. Henriques (1991). How does mother's education affect child height? *The Journal of Human Resources 26*(2), 183–211.

United Nations (2006). *World Youth Report 2003 - Rethinking leisure time: Expanding opportunities for young people and communities.*

Weaver, M. D., L. K. Barger, S. K. Malone, L. S. Anderson, and E. B. Klerman (2018, 12). Dose-Dependent Associations Between Sleep Duration and Unsafe Behaviors Among US High School Students. *JAMA Pediatrics 172*(12), 1187–1189.

Willis, R. J. and S. Rosen (1979). Education and self-selection. *Journal of Political Economy 87*(5), S7–S36.

Wing, C., K. Simon, and R. A. Bello-Gomez (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health 39*(1), 453–469. PMID: 29328877.

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics 32*(4), 711–754.

A Appendix

A.1 Derivation of Equilibria

Equation (1) and (5) consist the objective function

$$\max_{l_0,l_q,k_q,x} u_0(l_0,x) + \beta u_1(w_g \cdot g(l_g,k_g) + w_s \cdot s(l_s,k_s) + w_r \cdot r(l_r,k_r) + \epsilon) \tag{20}$$

by expressing $Q(\cdot)$ as a function of a linear combination of CSAT, GPA, and resume. The equilibrium conditions of interior solution will be:

$$[l_0] \quad \frac{1}{\beta}\frac{\partial u_0}{\partial l_0} + w_g \frac{\partial u_1}{\partial g}\frac{\partial g}{\partial l_g}\frac{\partial l_g}{\partial l_0} + w_s \frac{\partial u_1}{\partial s}\frac{\partial s}{\partial l_s}\frac{\partial l_s}{\partial l_0} + w_r \frac{\partial u_1}{\partial r}\frac{\partial r}{\partial l_r}\frac{\partial l_r}{\partial l_0} = 0 \tag{21}$$

$$[x] \quad \frac{1}{\beta}\frac{\partial u_0}{\partial x} + w_g \frac{\partial u_1}{\partial g}\frac{\partial g}{\partial k_g}\frac{\partial k_g}{\partial x} + w_s \frac{\partial u_1}{\partial s}\frac{\partial s}{\partial k_s}\frac{\partial k_s}{\partial x} + w_r \frac{\partial u_1}{\partial r}\frac{\partial r}{\partial k_r}\frac{\partial k_r}{\partial x} = 0 \tag{22}$$

$$[l_q] \quad \frac{1}{\beta}\frac{\partial u_0}{\partial l_0}\frac{\partial l_0}{\partial l_q} + w_q \frac{\partial u_1}{\partial q}\frac{\partial q}{\partial l_q} + w_{q'} \frac{\partial u_1}{\partial q'}\frac{\partial q'}{\partial l_{q'}}\frac{\partial l_{q'}}{\partial l_q} + w_{q''} \frac{\partial u_1}{\partial q''}\frac{\partial q''}{\partial l_{q''}}\frac{\partial l_{q''}}{\partial l_q} = 0 \tag{23}$$

$$[k_q] \quad \frac{1}{\beta}\frac{\partial u_0}{\partial x}\frac{\partial x}{\partial k_q} + w_q \frac{\partial u_1}{\partial q}\frac{\partial q}{\partial k_q} + w_{q'} \frac{\partial u_1}{\partial q'}\frac{\partial q'}{\partial k_{q'}}\frac{\partial k_{q'}}{\partial k_q} + w_{q''} \frac{\partial u_1}{\partial q''}\frac{\partial q''}{\partial k_{q''}}\frac{\partial k_{q''}}{\partial k_q} = 0 \tag{24}$$

where $q, q', q'' \in \{g, s, r\}$ and $q \neq q' \neq q''$. The time constraint (3) implies that a marginal increase (or decrease) in current leisure l_0 sacrifices (or is allocated to) l_g, l_s, and l_r. This yields,

$$\frac{\partial l_g}{\partial l_0} + \frac{\partial l_s}{\partial l_0} + \frac{\partial l_r}{\partial l_0} = -1. \tag{25}$$

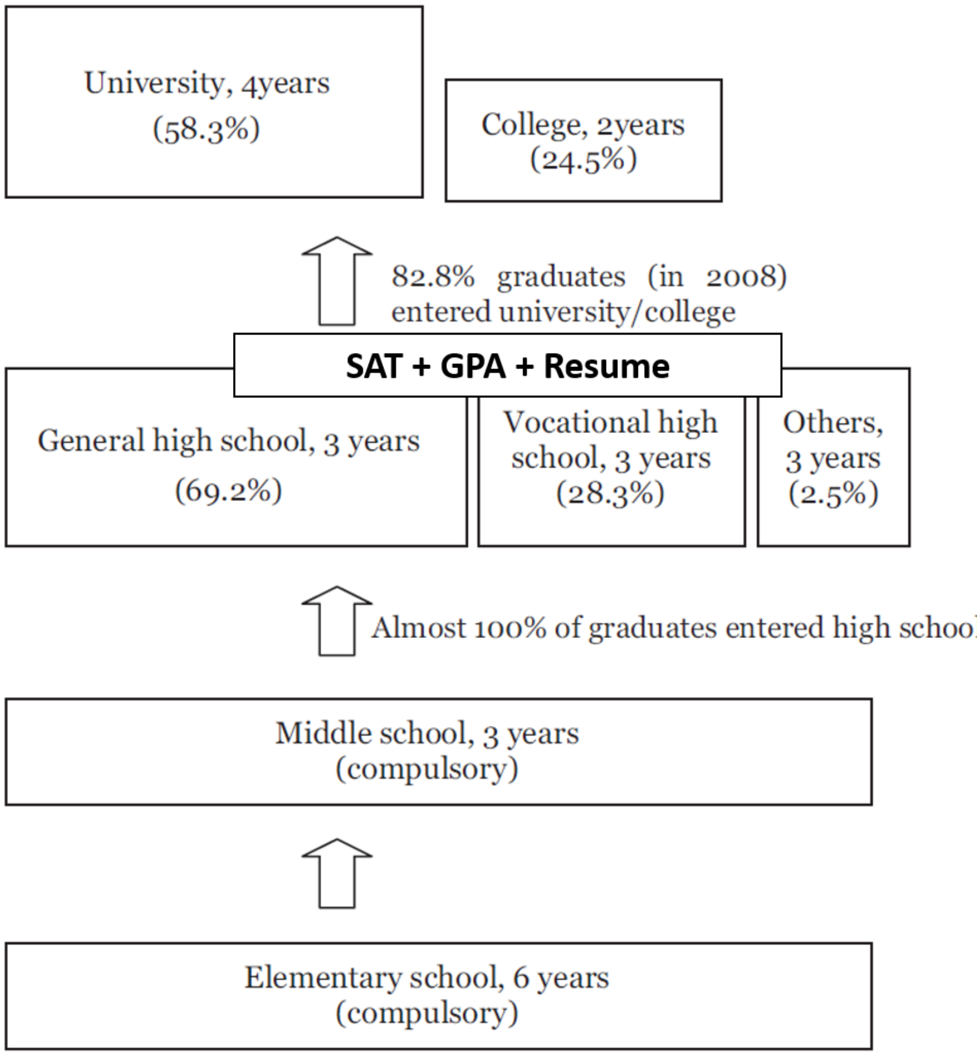


Figure 1: Korean Education System (Nam, 2014)

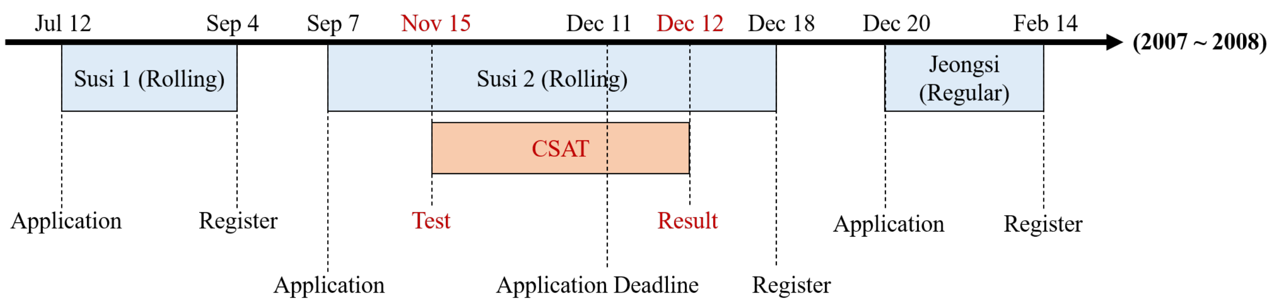


Figure 2: 2008 Academic Year College Admission Timeline

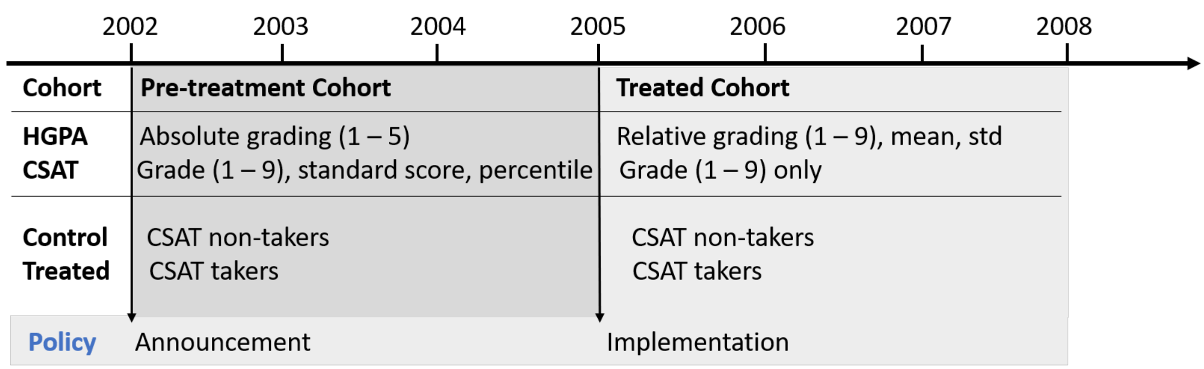


Figure 3: Policy Implementation Timeline

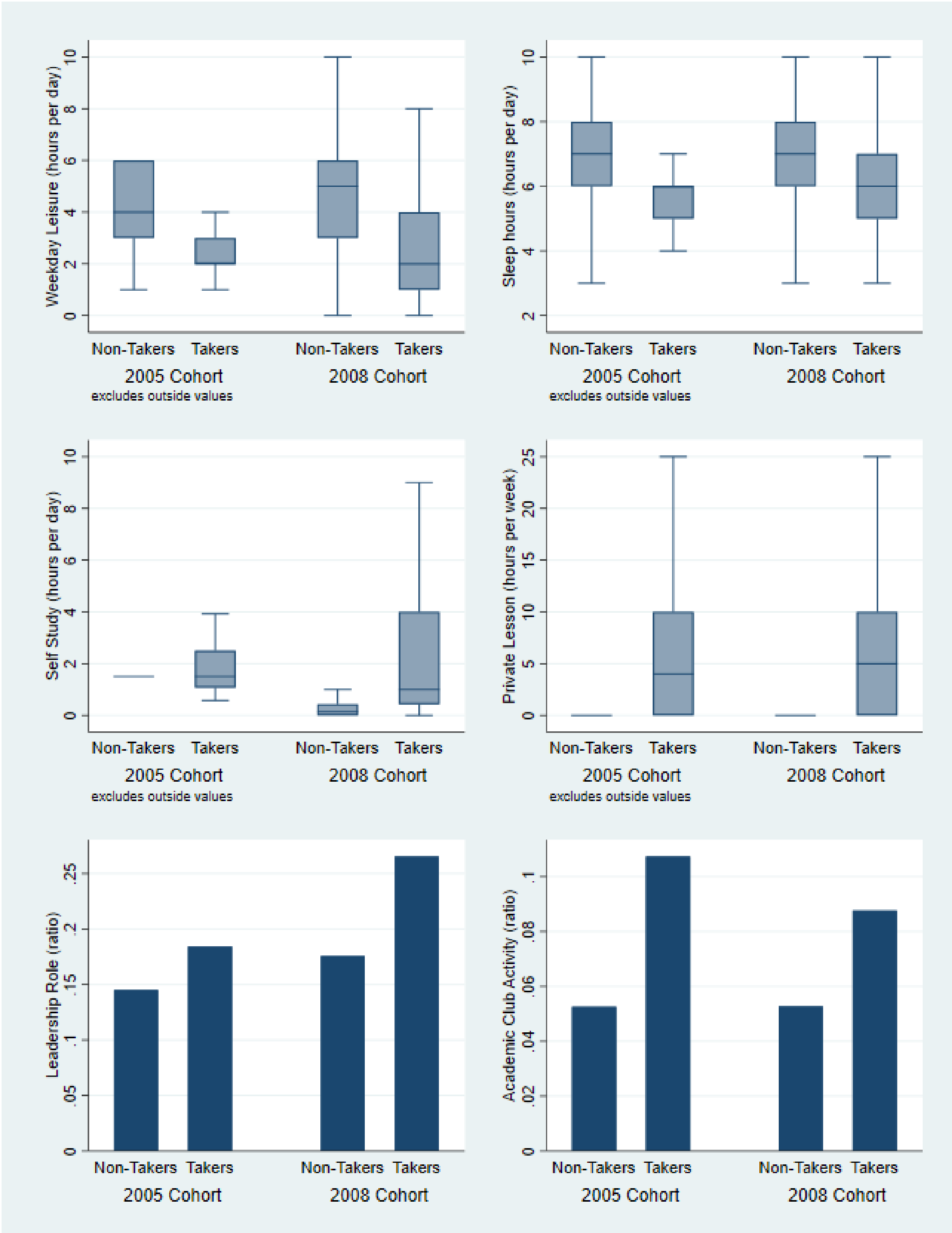
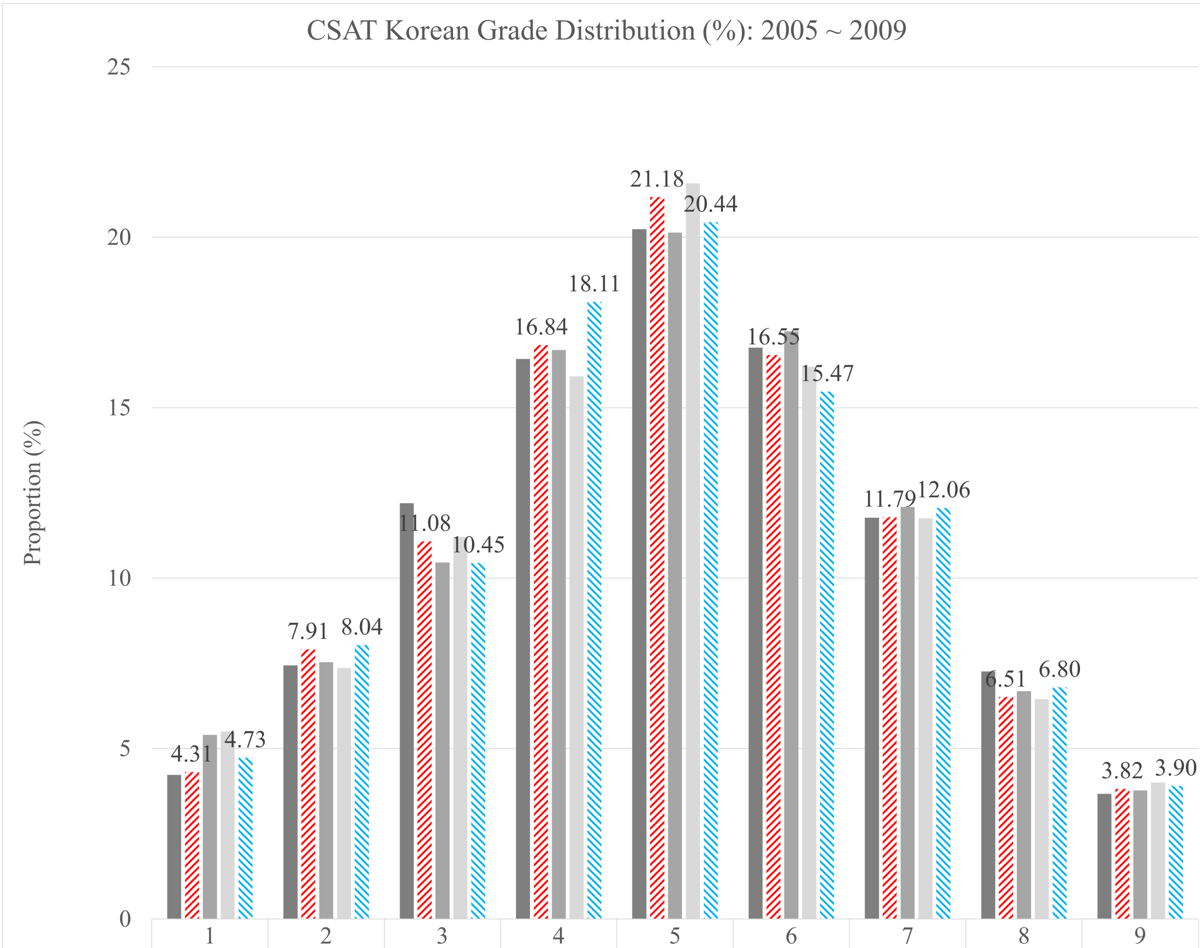


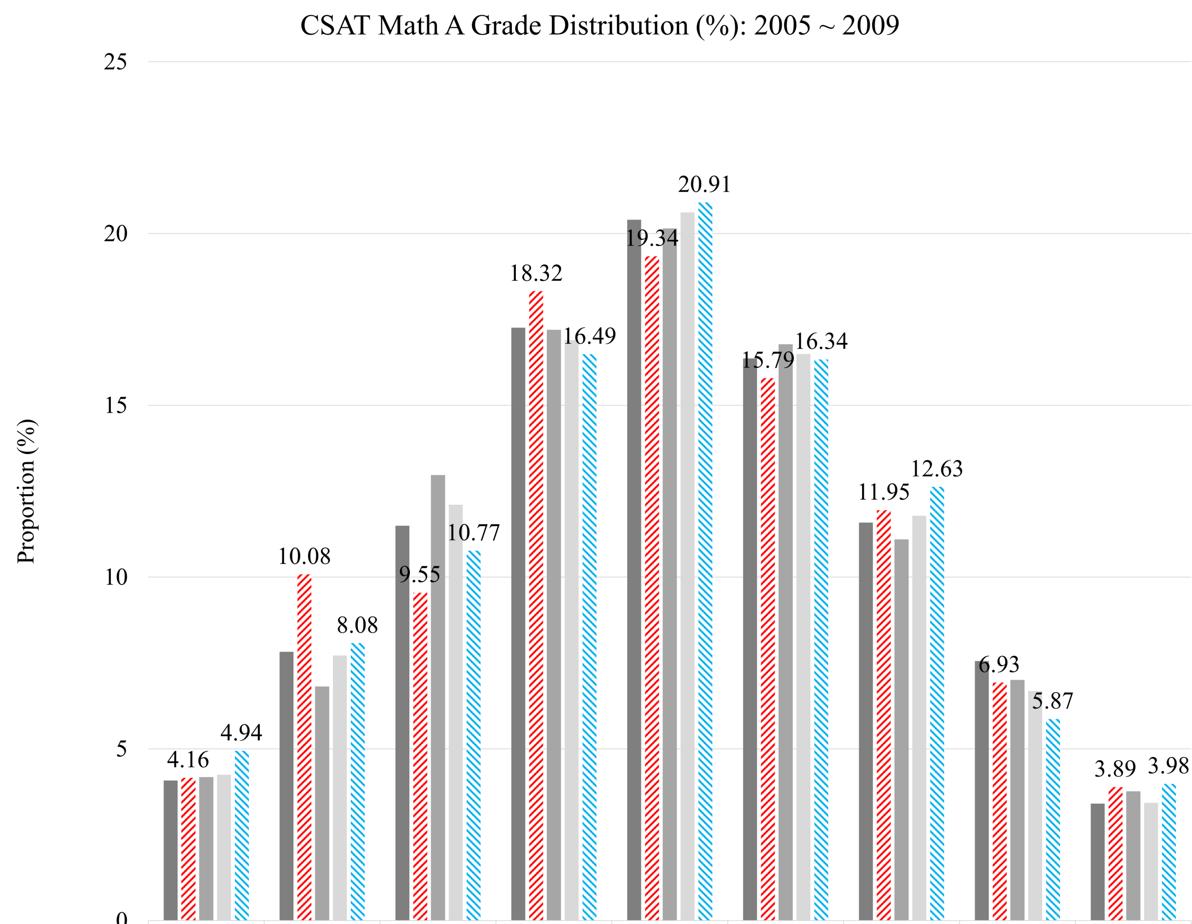
Figure 4: Mean Outcomes



	1	2	3	4	5	6	7	8	9
09 Korean	4.23	7.44	12.2	16.43	20.24	16.76	11.77	7.26	3.67
08 Korean	4.31	7.91	11.08	16.84	21.18	16.55	11.79	6.51	3.82
07 Korean	5.4	7.53	10.46	16.69	20.14	17.24	12.09	6.68	3.77
06 Korean	5.5	7.36	11.21	15.92	21.58	16.21	11.76	6.45	4
05 Korean	4.73	8.04	10.45	18.11	20.44	15.47	12.06	6.80	3.90

a. Source: Ministry of Education.
b. Grade 1 is the highest. Students raw CSAT scores are normalized. Each student's percentile is matched to one of nine preset grades in Table 1. When there are ties around the grade cutoff, the higher grade will be given to all.

Figure 5: CSAT Korean Grade Distribution



	1	2	3	4	5	6	7	8	9
09 Math A	4.08	7.83	11.5	17.26	20.4	16.37	11.59	7.56	3.41
08 Math A	4.16	10.08	9.55	18.32	19.34	15.79	11.95	6.93	3.89
07 Math A	4.18	6.82	12.97	17.2	20.15	16.78	11.1	7.01	3.77
06 Math A	4.25	7.72	12.11	16.9	20.62	16.5	11.79	6.69	3.43
05 Math A	4.94	8.08	10.77	16.49	20.91	16.34	12.63	5.87	3.98

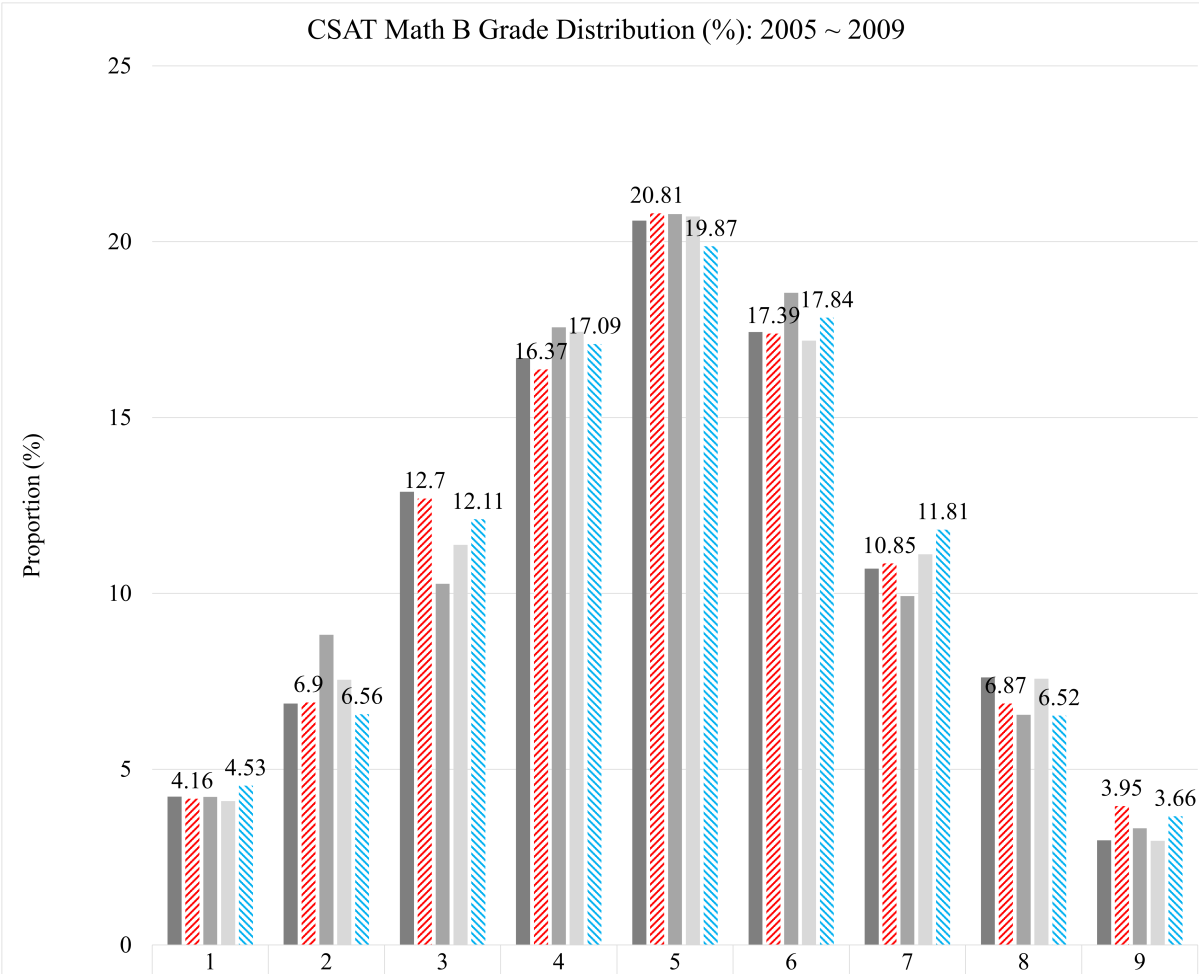
CSAT Math A Grade (1: Highest - 9: Lowest)

09 Math A 08 Math A 07 Math A 06 Math A 05 Math A

a. Source: Ministry of Education.

b. Grade 1 is the highest. Students raw CSAT scores are normalized. Each student's percentile is matched to one of nine preset grades in Table 1. When there are ties around the grade cutoff, the higher grade will be given to all.

Figure 6: CSAT Math A Grade Distribution



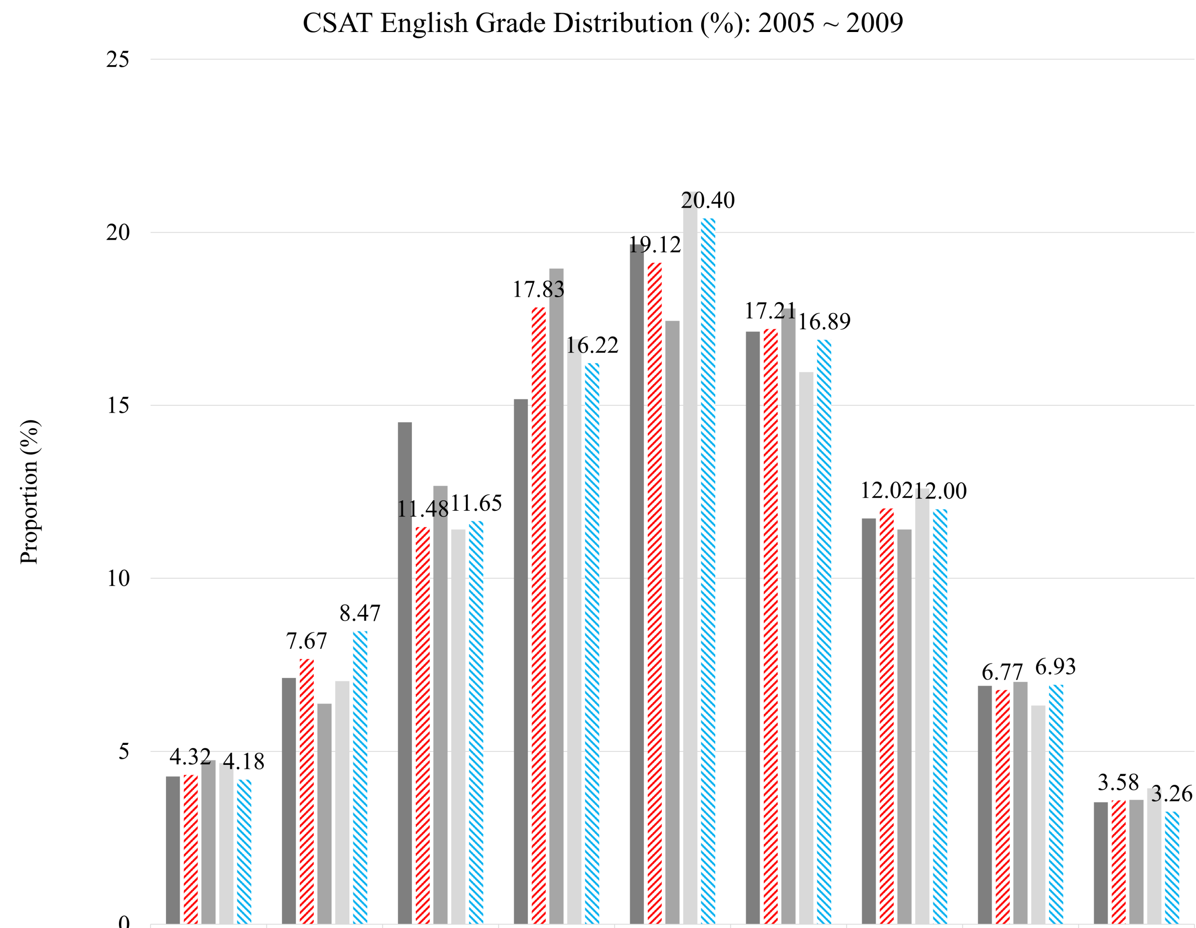
	1	2	3	4	5	6	7	8	9
09 Math B	4.22	6.87	12.89	16.69	20.6	17.43	10.7	7.61	2.98
08 Math B	4.16	6.9	12.7	16.37	20.81	17.39	10.85	6.87	3.95
07 Math B	4.21	8.82	10.27	17.57	20.79	18.55	9.92	6.55	3.32
06 Math B	4.09	7.54	11.38	17.44	20.72	17.19	11.11	7.57	2.96
05 Math B	4.53	6.56	12.11	17.09	19.87	17.84	11.81	6.52	3.66

CSAT Math B Grade (1: Highest - 9: Lowest)

09 Math B 08 Math B 07 Math B 06 Math B 05 Math B

a. Source: Ministry of Education.
b. Grade 1 is the highest. Students raw CSAT scores are normalized. Each student's percentile is matched to one of nine preset grades in Table 1. When there are ties around the grade cutoff, the higher grade will be given to all.

Figure 7: CSAT Math B Grade Distribution



	1	2	3	4	5	6	7	8	9
09 English	4.27	7.12	14.51	15.18	19.65	17.13	11.73	6.89	3.53
08 English	4.32	7.67	11.48	17.83	19.12	17.21	12.02	6.77	3.58
07 English	4.74	6.38	12.67	18.95	17.44	17.8	11.41	7.01	3.6
06 English	4.66	7.03	11.41	16.91	21.18	15.96	12.6	6.32	3.93
05 English	4.18	8.47	11.65	16.22	20.40	16.89	12.00	6.93	3.26

CSAT English Grade (1: Highest - 9: Lowest)

09 English 08 English 07 English 06 English 05 English

a. Source: Ministry of Education.

b. Grade 1 is the highest. Students raw CSAT scores are normalized. Each student's percentile is matched to one of nine preset grades in Table 1. When there are ties around the grade cutoff, the higher grade will be given to all.

Figure 8: CSAT English Grade Distribution

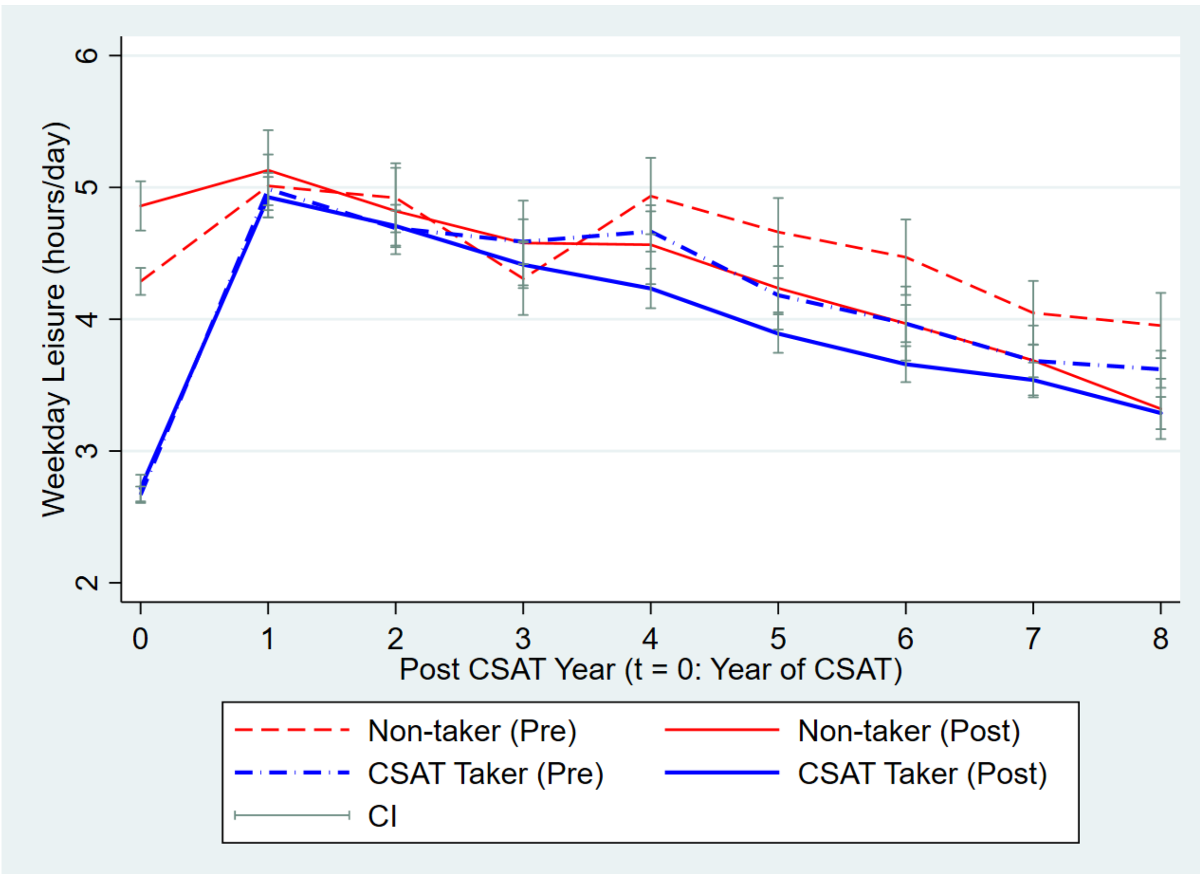


Figure 9: Non-Intervention Outcome Trends after CSAT (hours/day)

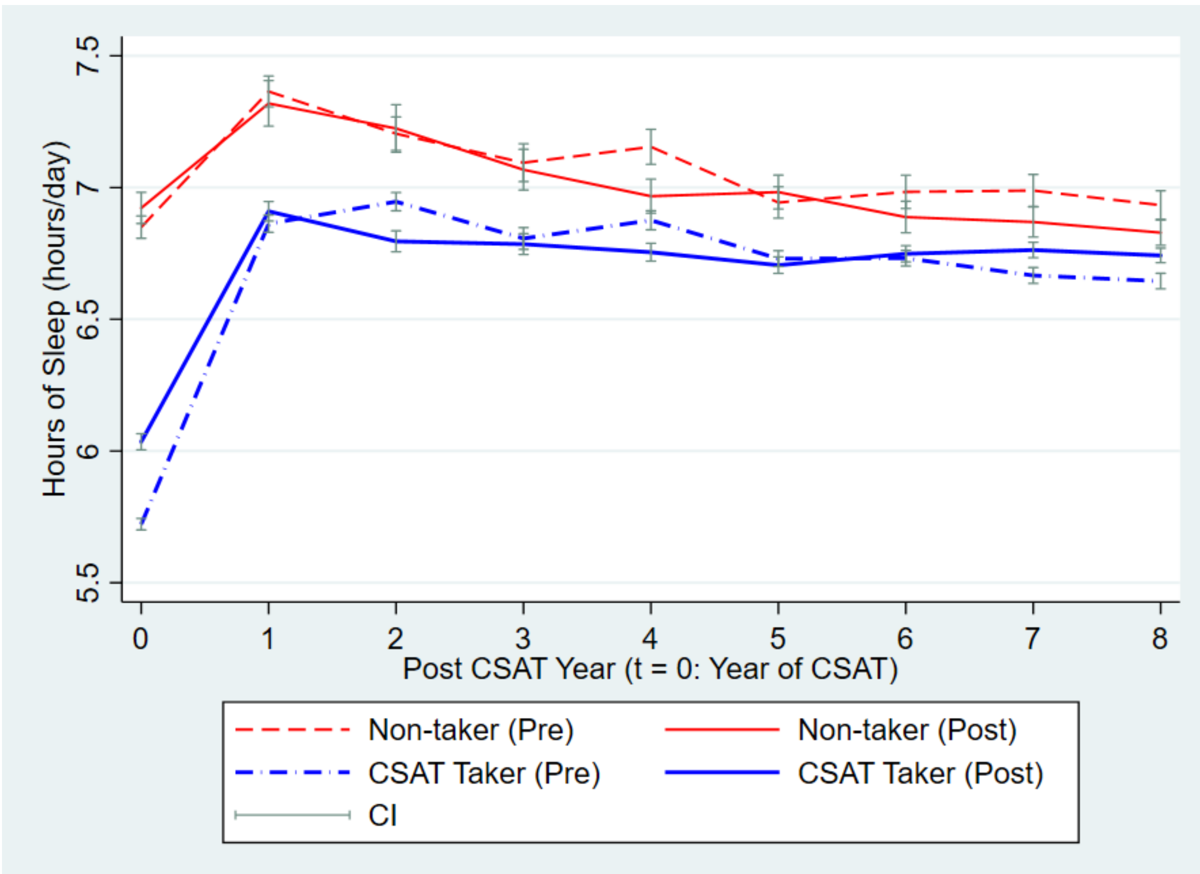


Figure 10: Non-Intervention Sleep Hours Trends after CSAT (hours/day)

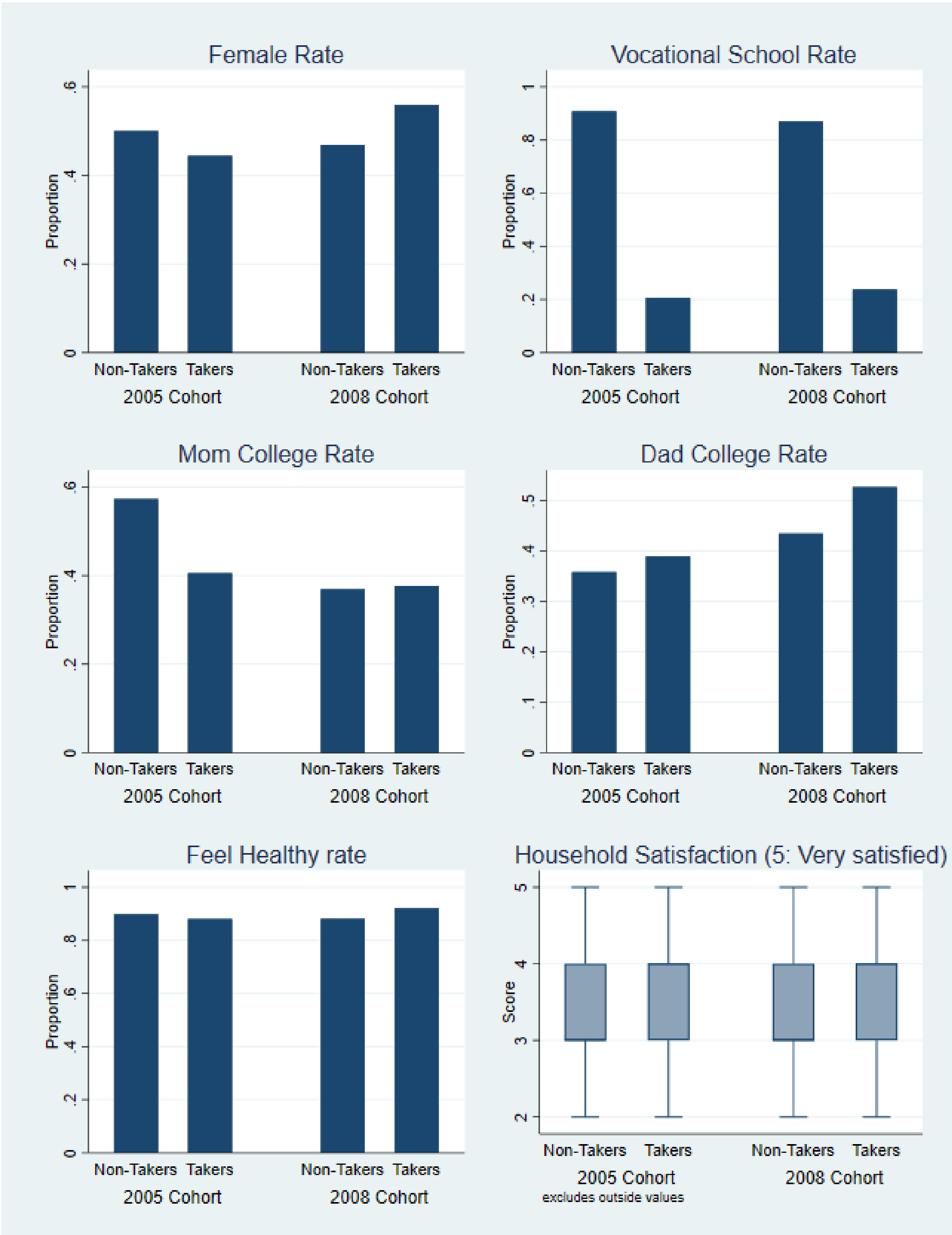


Figure 11: Mean Characteristics

Grade	Cumulative Percentage
1	4% or Under
2	4% $\sim$ 11%
3	11% $\sim$ 23%
4	23% $\sim$ 40%
5	40% $\sim$ 60%
6	60% $\sim$ 77%
7	77% $\sim$ 89%
8	89% $\sim$ 96%
9	Over 96%

[a] Source: Ministry of Education.

[b] Grade 1 is the highest. Students raw CSAT scores are normalized. Each student's percentile is matched to one of nine preset grades.

Table 1: GPA and CSAT Preset Grade Scale

2009 School Year

	Korean				Math A				Math B				English			
Grade	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%
1	92	131	23,615	4.23	81	135	4,965	4.08	79	138	16,795	4.22	95	131	23,590	4.27
2	86	125	41,589	7.44	72	125	9,542	7.83	68	128	27,346	6.87	89	126	39,343	7.12
3	79	117	68,183	12.2	63	116	14,010	11.50	55	116	51,255	12.89	79	117	80,173	14.51
4	71	108	91,839	16.43	53	106	21,027	17.26	42	104	66,402	16.69	69	108	83,865	15.18
5	61	96	113,095	20.24	42	95	24,847	20.40	28	90	81,935	20.60	55	95	108,563	19.65
6	51	85	93,689	16.76	32	84	19,947	16.37	19	82	69,325	17.43	41	82	94,631	17.13
7	41	74	65,794	11.77	22	74	14,117	11.59	14	77	42,566	10.70	30	72	64,785	11.73
8	30	62	40,553	7.26	14	65	9,213	7.56	9	73	30,279	7.61	22	65	38,046	6.89
9	Under	Under	20,495	3.67	Under	Under	4,160	3.41	Under	Under	11,869	2.98	Under	Under	19,507	3.53

2008 School Year

	Korean				Math A				Math B				English			
Grade	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%
1	90	131	23,693	4.31	99	128	5,103	4.16	93	135	15,952	4.16	96	130	23,502	4.32
2	83	124	43,478	7.91	93	123	12,346	10.08	85	129	26,484	6.90	90	125	41,729	7.67
3	76	116	60,919	11.08	86	118	11,699	9.55	71	118	48,724	12.70	82	118	62,432	11.48
4	67	107	92,557	16.84	74	108	22,447	18.32	54	105	62,811	16.37	69	107	96,995	17.83
5	56	95	116,450	21.18	59	95	23,696	19.34	36	91	79,851	20.81	54	94	103,988	19.12
6	46	85	91,002	16.55	45	84	19,347	15.79	24	81	66,732	17.39	39	81	93,620	17.21
7	37	75	64,820	11.79	31	72	14,642	11.95	17	76	41,623	10.85	30	74	65,381	12.02
8	27	65	35,812	6.51	20	63	8,489	6.93	12	72	26,356	6.87	22	67	36,843	6.77
9	Under	Under	21,000	3.82	Under	Under	4,764	3.89	Under	Under	15,167	3.95	Under	Under	19,455	3.58

2007 School Year

	Korean				Math A				Math B				English			
Grade	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%
1	95	127	29,708	5.40	89	134	4,906	4.18	96	137	16,142	4.21	96	130	25,899	4.74
2	91	123	41,408	7.53	81	126	8,000	6.82	84	128	33,812	8.82	91	126	34,834	6.38
3	86	118	57,503	10.46	71	116	15,216	12.97	72	119	39,342	10.27	82	118	69,201	12.67
4	78	109	91,778	16.69	61	106	20,171	17.20	53	104	67,339	17.57	69	107	103,520	18.95
5	67	97	110,757	20.14	50	95	23,627	20.15	34	90	79,649	20.79	56	95	95,258	17.44
6	55	85	94,825	17.24	39	84	19,684	16.78	22	81	71,064	18.55	41	82	97,263	17.80
7	43	72	66,465	12.09	27	73	13,021	11.10	16	76	38,029	9.92	30	72	62,342	11.41
8	32	60	36,756	6.68	18	64	8,224	7.01	11	72	25,103	6.55	22	65	38,311	7.01
9	Under	Under	20,760	3.77	Under	Under	4,424	3.77	Under	Under	12,716	3.32	Under	Under	19,693	3.60

2006 School Year

	Korean				Math A				Math B				English			
Grade	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%
1	98	125	30,362	5.50	84	134	5,595	4.25	84	139	14,995	4.09	92	133	25,615	4.66
2	95	121	40,617	7.36	76	125	10,170	7.72	73	128	27,684	7.54	83	127	38,608	7.03
3	90	117	61,835	11.21	67	116	15,963	12.11	60	118	41,758	11.38	74	118	62,724	11.41
4	83	110	87,830	15.92	56	106	22,267	16.90	42	103	64,019	17.44	63	107	92,923	16.91
5	74	99	119,028	21.58	41	94	27,167	20.62	30	89	76,028	20.72	50	93	116,413	21.18
6		86	89,420	16.21		83	21,736	16.50		82	63,095	17.19		82	87,700	15.96
7		72	64,862	11.76		73	15,531	11.79		78	40,793	11.11		74	69,267	12.60
8		57	35,550	6.45		65	8,819	6.69		74	27,795	7.57		68	34,718	6.32
9		Under	22,050	4.00		Under	4,521	3.43		Under	10,849	2.96		Under	21,576	3.93

2005 School Year

	Korean				Math A				Math B				English			
Grade	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%	Raw Score	Std. Score	N	%
1	94	128	26,870	4.73	88	131	7,206	4.94	88	140	16,231	4.53	96	132	23,863	4.18
2	88	123	45,647	8.04	81	125	11,776	8.08	77	131	23,528	6.56	89	125	48,316	8.47
3	84	117	59,328	10.45	73	118	15,702	10.77	60	117	43,391	12.11	76	117	66,462	11.65
4	76	108	102,838	18.11	62	108	24,043	16.49	40	101	61,270	17.09	60	108	92,500	16.22
5	66	97	116,112	20.44	45	94	30,487	20.91	27	90	71,218	19.87	40	95	116,385	20.40
6		86	87,880	15.47		82	23,825	16.34		83	63,941	17.84		82	96,374	16.89
7		73	68,507	12.06		72	18,414	12.63		79	42,343	11.81		72	68,452	12.00
8		59	38,601	6.80		67	8,560	5.87		76	23,387	6.52		66	39,509	6.93
9		Under	22,167	3.90		Under	5,810	3.98		Under	13,126	3.66		Under	18,570	3.26

[a] Source: KICE, Megastudy, Kim01 Edu Consulting.
[b] Raw scores are not announced by KICE. They are estimated by private institutes, *Megastudy* and *Kim01 Edu Consulting*. Students get higher raw scores under an easy exam which results in higher Grade 1 cutoff.

Table 2: Population Statistics of CSAT Results (2005 – 2009)

		2005 CSAT Takers				2008 CSAT Takers			
Result	Subject	Obs	Mean	SD	Min / Max	Obs	Mean	SD	Min / Max
Grade	Korean	2,184	5.58	1.91	9 / 1	1,262	5.12	2.26	9 / 1
	Math	1,943	5.48	1.81	9 / 1	1,181	5.05	2.15	9 / 1
	English	2,172	5.59	1.91	9 / 1	1,246	5.11	2.30	9 / 1
Percentile	Korean	2,199	40.67	27.77	0 / 100				
	Math	2,199	37.20	28.81	0 / 100				
	English	2,199	39.93	28.24	0 / 100				
Score	Korean	2,199	92.95	22.00	0 / 135				
	Math	2,199	83.42	34.56	0 / 150				
	English	2,199	92.11	22.15	0 / 139				

[a] No percentile and standardized score data of 2008 CSAT is available due to the grading policy. The population median grade is 5 for all subjects. Difference in observations between categories in 2005 CSAT is due to missing data. In 2008 CSAT, some students selectively took exams depending on the admission criteria of applying colleges.

Table 3: Sample statistics of CSAT results

	2005 Non-takers (N = 951)	2005 CSAT Takers (N = 2199)	2008 Non-takers (N = 569)	2008 CSAT Takers (N = 1266)
Female	50%	44%	47%	56%
Mother College	57%	41%	37%	38%
Father College	36%	38%	43%	53%
Health Condition	3.54	3.52	3.66	3.74
HH Satisfactory	3.25	3.70	3.43	3.79
Vocational School	91%	21%	87%	24%
Punishment	17%	8%	13%	5%
Prize	36%	40%	22%	45%
Smoking	27%	8%	23%	6%
Drinking	63%	42%	69%	49%
Daily Gaming	28%	10%	29%	11%

[a] Health condition (5: Very healthy) and Household satisfactory level (5: Very satisfied) range from 1 to 5. Punishment at school, prize, smoking, and drinking are binary variables where they take 1 with any experience.

Table 4: Mean Characteristics by Group and Cohort

Year	Number of Graduates	Number of CSAT Applicants
2001	736,171	603,238 (81.94%)
2002	670,713	541,662 (80.76%)
2003	590,413	482,089 (81.65%)
2004	588,550	476,129 (80.90%)
2005	569,272	435,538 (76.51%)
2006	568,055	422,310 (74.34%)
2007	571,357	425,396 (74.45%)
2008	581,921	446,597 (76.75%)
2009	576,298	448,472 (77.82%)
2010	633,539	532,436 (84.04%)

[a] Data source: KOSIS, KICE

[b] There was no propotional changes in CSAT applicants out of population high school graduates around the policy changes from 2005 to 2009 implying that the average CSAT taking decision was not affected by the policy.

Table 5: High School Graduates and CSAT Applicants

	(1) Weekday Leisure	(2) Private Lesson	(3) Self Study	(4) Leader Role	(5) Academic Club
Post	0.549***	0.085	-0.965***	0.031	0.000
	(0.134)	(0.248)	(0.073)	(0.024)	(0.014)
CSAT Taker	-1.619***	5.126***	0.474***	0.039**	0.055***
	(0.079)	(0.300)	(0.046)	(0.015)	(0.012)
Post × CSAT	-0.502***	0.814*	1.294***	0.051*	-0.020
	(0.163)	(0.469)	(0.155)	(0.029)	(0.020)
Observations	4985	4985	4985	4985	4985
Mean Outcome	3.236	4.652	1.718	0.196	0.086

[a] Cluster standard errors by school in parentheses.

[b] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[c] Units are hours per day for weekday leisure and self study, hours per week for private lesson, and binary for leader role and academic club activities.

[d] Mean outcome is the average of total observations including both cohorts and groups.

Table 6: Average Policy Effects on Time Allocation

A. TOP GPA STUDENTS

	(1) Weekday Leisure	(2) Private Lesson	(3) Self Study	(4) Leader Role	(5) Academic Club
Post	0.976**	0.153	-0.893***	0.048	0.035
	(0.437)	(0.715)	(0.117)	(0.103)	(0.061)
CSAT Taker	-1.420***	6.347***	0.871***	0.115	0.097**
	(0.264)	(0.918)	(0.152)	(0.089)	(0.048)
Post × CSAT	-1.062**	-0.319	1.155***	-0.032	-0.084
	(0.463)	(1.194)	(0.274)	(0.117)	(0.069)
Observations	480	480	480	480	480
Mean Outcome	3.116	5.944	2.074	0.392	0.119

B. BOTTOM GPA STUDENTS

	(1) Weekday Leisure	(2) Private Lesson	(3) Self Study	(4) Leader Role	(5) Academic Club
Post	1.348***	-0.284	-1.160***	-0.000	-0.025
	(0.363)	(0.340)	(0.083)	(0.030)	(0.020)
CSAT Taker	-1.272***	4.091***	0.189**	0.037	0.005
	(0.169)	(0.541)	(0.077)	(0.028)	(0.021)
Post × CSAT	0.180	-0.182	0.520*	0.104	0.026
	(0.590)	(1.277)	(0.301)	(0.076)	(0.043)
Observations	627	627	627	627	627
Mean Outcome	3.891	2.818	1.220	0.097	0.040

[a] Cluster standard errors by school in parentheses.
[b] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[c] Units are hours per day for weekday leisure and self study, hours per week for private lesson, and binary for leader role and academic club activities.
[d] Mean outcome is the average of total observations including both cohorts and groups.

Table 7: Policy Effects on Time Allocation by High School GPA