DEVELOPMENT AND SOCIETY
Volume 47 | Number 3 | September 2018, 451-476
DOI 10.21588/dns/2018.47.3.006

**Article**

# The Heterogeneous Effects of Shadow Education on SAT Scores

Yool Choi | Korea National University of Education

*As many prior studies have pointed out, private shadow education, which includes commercial coaching and one-on-one tutoring, has important implications for educational opportunity and the process of social stratification. In this study, I analyze the heterogeneous effects of private shadow education on SAT scores by individual likelihood of participation using the Educational Longitudinal Study 2002. The key finding of this study is that the effects of shadow education are substantively different across propensity strata and that its effects increase as the propensity to participate in shadow education increases. That is, those who are more likely to use shadow education, who are socioeconomically advantaged and possess higher educational capital, benefit more from shadow education than those who are less likely to use it. Moreover, use of public resources neither alleviates the effects of shadow education nor changes the pattern of the heterogeneous effects of shadow education based on likelihood of use.*

**Keywords:** *shadow education, private supplementary education, heterogeneous effects, positive selection, negative selection*

## Introduction

This study aims to understand the impact of shadow education on educational inequality by examining the heterogeneous effects of shadow education on SAT scores. Stevenson and Baker defined shadow education as a "set of educational activities outside formal schooling that are designed to improve a student's chances of successfully moving through the allocation process" (1992, p. 2). The most prevalent forms of shadow education include commercial coaching, cram schools, and one-on-one tutoring. Extensive use of shadow education has been largely observed in a few countries, especially in several East Asian countries. However, recent research has shown that shadow education has become a worldwide phenomenon and is one of the fastest growing industries in many countries (Baker and LeTendre 2005; Bray 2001; Buchmann, Condron and Roscigno 2010; Byun 2014).

The prodigious growth of shadow education has caused much concern over educational quality and equality among policymakers (Bray 2009; Mori and Baker 2010). This is because in most cases shadow education completely relies upon private investment and is less accessible to families with limited socioeconomic capital (Baker and LeTendre 2005; Dang and Rogers 2008; Mori and Baker 2010; Stevenson and Baker 1992; Byun 2014). Given this, if shadow education does make a difference in academic achievement, it then carries important implications concerning educational opportunity and social stratification. Depending on its impacts, shadow education could serve as a mechanism for maintaining and increasing social stratification by allowing educational advantages on students who are already advantaged in terms of their economic, social, and cultural capital (Byun 2014). However, it is more difficult to control for the pervasiveness and repercussions of shadow education compared to other factors that affect educational consequences, since it takes place outside of formal schooling (Barry 2001; Grodsky 2010).

Given the importance of shadow education, this paper examines the effects of shadow education on SAT scores in the United States. This study aims to not only to understand how advantaged families utilize their economic resources to access the educational advantages that shadow education can provide, but also suggests how policymakers might approach the issue of shadow education to diminish educational stratification in the United States.

I build on and augment prior research on shadow education and on coaching effects on SAT scores in several important ways. First, using propensity score matching, I examine the causal relationship between shadow education and SAT achievement by carefully attending to issues of preexisting heterogeneity. Many prior studies have emphasized that selection bias is a major concern when examining the causal effect of shadow education on SAT scores, and only a few studies have tried to address this issue in the U.S. context by utilizing more advanced methods (Briggs 2001; Byun and Park 2012; Domingue and Briggs 2009; Hansen 2004; Powers and Rock 1999).

Second, in addition to examining the average effects of shadow education, I analyze variation in the effects on SAT scores based on the likelihood of receiving shadow education. Using semi- and non-parametric methods based on propensity scores, I summarize the systematic trend of heterogeneous effects across propensity scores. This approach could potentially reveal effects that have previously been masked. For example, although the average effects of shadow education are moderate, some part of the population might benefit significantly more than others from shadow education. If socioeconomically advantaged students who are also more likely to engage in shadow education benefit most from it, then shadow education is an important factor in widening the educational gap between social classes, as it advances advantaged students' educational achievement. However, if students less likely to participate in shadow education benefit more from it, the impact of shadow education on educational inequality is limited and it could in fact serve as a tool to diminish the educational gap between advantaged and disadvantaged students. In addition to socioeconomic background, students' propensity to engage in shadow education is an effective measure of students' educational environments. As I estimate students' likelihood of participating in shadow education according to demographic characteristics, educational motivation/aspiration, prior academic achievement, parents' expectation/involvement, and institutional characteristics, examining effect heterogeneity based on likelihood of participating in shadow education could tell us how the effects of shadow education vary by students' essential backgrounds that decide his or her academic success.

Third, considering various forms of shadow education, I use a multiple counter-factual approach to examine the impact of shadow education more precisely. While there is an ongoing debate on the definition of shadow education (Buchmann, Condron and Roscigno 2010; Grodsky 2010; Alon 2010), I conceptualize shadow education as taking two different forms: private (commercial coaching and one-on-one tutoring) and public (school prep courses, books, videos, and computer materials). In this study, I examine not only how private shadow education affects SAT scores differently depending on the alternative SAT preparation activities, but also whether public resources effectively allay the stratification effects of private shadow education. By attending to the effect heterogeneity of private shadow education with multiple counterfactual scenarios, this study mainly focuses on examining how the effects of shadow education differ based on individual social and educational position and how this phenomenon affects educational stratification in the United States.

## Contextual Background: *Shadow education in the United States*

Despite its important implications on social stratification, shadow education in the United States has received far less attention than other factors that are thought to contribute to educational inequality. This is due in part to the nature of the U.S. educational system. In general, shadow education is prevalent in countries that place heavy emphasis on formal examinations—particularly centrally administered examinations—and have tight linkages between educational achievement and later occupational and social status (Baker and LeTendre 2005; Dang and Rogers 2008; Stevenson and Baker 1992). For example, shadow education is extremely prevalent in most East Asian countries where great emphasis is placed on the importance of college entrance exams (Baker and LeTendre 2005; Bray 2001). In contrast, in the United States, in addition to formal exams, various non-cognitive factors such as extracurricular activities are also important in the college admissions process (Bray 2001; Buchmann, Condron and Roscigno 2010; Byun and Park 2012).

However, in accord with the recent expansion in secondary and post-secondary education in the United States, the importance of standardized tests has been growing (Alon 2009; Alon and Tienda 2007; Grodsky, Warren and Felts 2008). The growing importance of standardized testing in the U.S. admissions process is perhaps a major factor in pushing students to improve their scores on high-stakes tests, which fosters the development of shadow education (Buchmann, Condron and Roscigno 2010; Byun and Park 2012). The increased profitability of commercial coaching companies in the United States, such as the Princeton Review and Kaplan, is indicative of this growing trend (Buchmann, Condron and Roscigno 2010; Davies and Aurini 2006). For example, the Princeton Review earned $110.4 million in revenue for its test preparation services in 2009 (Princeton Review 2010) and is one of the fastest growing industries in the United States.

Despite its low profile, there are U.S. scholars who have called attention to the impact of shadow education on educational stratification (Buchmann, Condron and Roscigno 2010; Byun and Park 2012; Domingue and Briggs 2009). For example, Buchmann et al. (2010) referred to the variety of SAT/ACT test preparation services as American style shadow education. The most common form of shadow education in the United States are those activities that prepare students for SAT/ACT tests. In many other countries, preparing for college entrance exams is the most extensive and important form of shadow education (Bray 2001; Stevenson and Baker 1992). Therefore, examining the effects of SAT test preparation activities could lead to a thorough discussion of the implications of shadow education in the United States.

## Methodological Issues

### *Pretreatment heterogeneity issues in the study of shadow education*

Even though the effects of shadow education on educational achievement have been widely investigated in various countries, the empirical evidence is somewhat mixed. Generally, there is thought to be some positive relationship between shadow education and academic achievement; however, whereas some studies have found strong positive effects (Jacob and Lefgren 2004; Stevenson and Baker 1992; Buchman 2002; Dang 2007; Tansle and Bircan 2005), other studies have found modest positive or null effects (Domingue and Briggs 2009; Ha and Harpham 2005; Suryadarma et al. 2006; Kuan 2011; Cheo and Quah 2005; Kim 2010). One major factor contributing to these varied findings is the issue of selection bias in estimating the causal relationship between shadow education and its outcome (Byun 2014; Dang and Rogers 2008). That is, most of the research on this topic is limited when it comes to addressing pretreatment heterogeneity. Using OLS regression with various control variables is the most common approach to address this limitation; however, OLS regression cannot fully address the issue of pretreatment heterogeneity (Brand and Xie 2010; Rosenbaum and Rubin 1983). In this study, I use propensity score matching to address pretreatment heterogeneity in examining the causal effects of shadow education.

*Heterogeneous effects of shadow education*

Although some prior studies have recognized pretreatment heterogeneity in the relationship between shadow education and academic achievements, most studies have not attended to effect heterogeneity in examining the impact of shadow education on academic achievements. While most prior studies have focused on homogeneous shadow education effects, which assumes that the effects of shadow education are equal for each student, I believe it is reasonable to suppose that individual responses to shadow education can differ based on social background.

This study is particularly interested in examining treatment effect heterogeneity. Treatment effect heterogeneity refers to an estimation of the interaction between treatment and propensity for treatment (Brand and Davis 2011; Brand and Xie 2010; Brand and Simon-Thomas 2013; Xie). This approach addresses two sources of selection bias simultaneously; that is, by using propensity score matching, this study effectively attends to the pretreatment heterogeneity issue, and by examining how the effects of shadow education differ based on the propensity to use shadow education, the study can address systematic trends in the effect heterogeneity of shadow education.

In terms of shadow education, treatment effect heterogeneity can answer such questions as is shadow education particularly beneficial for already advantaged students who are also more likely participate in shadow education, or do students from less privileged backgrounds benefit enough to diminish the educational gap? If we understand these patterns of treatment effect heterogeneity, we can grasp the more profound implications of shadow education from a social stratification perspective. In so doing, we can also provide suggestions for effective policymaking geared toward assisting specific sub-populations.

Despite many prior studies on shadow education, only one study has shown some evidence of treatment effect heterogeneity in the United States. Domingue and Briggs (2009) estimated the effects of coaching on SAT scores using propensity score matching with the Educational Longitudinal Survey of 2002. They utilized one independent variable (commercial coaching) among various variables related to shadow education. In this analysis, they estimated propensity score stratum–specific coaching effects on SAT. While they did not explicitly summarize the trend in the variation of effects by propensity score strata, they found some evidence of effect heterogeneity and argued that effects are highest in higher subclasses (Domingue and Briggs 2009:19). In my analysis, I build on their research and expand on it by using more complex counterfactuals and by testing for systematic trends in the treatment effect heterogeneity using stratification-multilevel (SM) and smoothing-differencing (SD) models. I discuss these two methods in more detail in the analytical strategy section. Outside of the U.S. educational context, Choi and Park (2016) examined treatment effect heterogeneity in South Korea. They found that while the effects of shadow education do not differ by individual propensity to use shadow education when not controlling for hours per week spent on shadow education, intense shadow education (more than four hours per week) showed negative selection trends, which means that students less likely to participate in shadow education benefit most from it. In this study, I employ the same analytic strategy to examine the heterogeneous effects of shadow education in the U.S. context, which is greatly different from that of South Korea, and I further extend their study by differentiating between private and publish shadow education in this paper.

*What is shadow education?*

Along with methodological considerations, how one defines shadow education is a key factor in analyzing its effects. The operational definitions of shadow education are diverse and controversial. Buchmann et al. (2010) constructed a test preparation indicator, which includes four kinds of test preparation activities: (1) books, computer software, and/or videos; (2) high school prep courses; (3) private courses; and (4) private tutors. They included all four activities as forms of shadow education and did not distinguish between public and private resources. Grodsky (2010) criticized existing operational definitions of shadow education and emphasized that the distinction between public and private resources is particularly important in understanding the implications of shadow education. Thus, resources such as private courses and private tutoring should be seen as the only form of shadow education in the United States. Alon (2010) similarly emphasized the importance of the distinguishing between private and public resources in her comments and supplementary analysis to Buchman et al.'s work. Although she considered both private and public sources to be aspects of shadow education, she argued that distinguishing between the two was analytically important for understanding its economic aspects and effectiveness.

In this study, I distinguish between public test preparation resources—which includes books, computer software, videos, and high school preparatory courses—and private resources, such as private courses and private tutors. Since private resources represent an apparent economic barrier for those with limited economic capital (Buchmann, Condron & Roscigno 2010; Byun & Park 2012), its implications for stratification are substantially different from those of other resources. Therefore, the dichotomous definition of shadow education between private and public aims to understand the degree to which shadow education contributes to educational inequality and to examine whether public resources are a comparable alternative to private resources.

## Hypothesis

The key interest of this study is understanding how the effects of shadow education differ according to likelihood of engaging in shadow education. There are two competing theoretical models for interpreting the patterns of treatment effect heterogeneity of shadow education. The first is the positive selection model, in which individuals who are most likely to participate in shadow education also benefit most from it. This is explained by the rational–behavioral model, an economic model that posits that individuals who think that they will have the highest returns from shadow education are most likely to participate in shadow education. Their utilization of shadow education is thus rational behavior based on a utility-maximizing strategy. Another possible explanation for the positive selection model is that since the quantity and quality of shadow education varies by cost, there is the possibility that those who are most likely to engage in shadow education do so for longer and that their shadow education is of higher quality than others based on their advantaged socioeconomic backgrounds (Choi and Park 2016).

The second theoretical model is negative selection, in which individuals who are less likely to use shadow education benefit more from shadow education. One of the possible explanations for negative selection is that students with lower socioeconomic status (SES) (and therefore less likely to participate in shadow education) who actually do participate in shadow education may be more strongly motivated than higher SES students, given that they would have had to overcome more barriers to participation. This indicates that the use of shadow education might be more culturally and socially driven, rather than by solely rational economic concerns (the aforementioned utility maximizing strategy). That is, it is possible that students from advantaged backgrounds might be engaged in shadow education because of their parents and are not strongly motivated. Also, we should consider the different counterfactuals between high propensity and low propensity participants. If shadow education is an effective tool to increase SAT scores, then extremely poor achievements by low propensity non-participants will reflect a larger difference than high-propensity participants. Since high propensity non-participants have diverse social and cultural capital and thus potentially alternative means for securing high test scores, they can diminish the outcomes gap between themselves and high propensity shadow education participants (Choi and Park 2016).

The implications of these two competing hypotheses for educational stratification are quite different. If the data indicate that positive selection is occurring, this means that shadow education plays an important role in widening the educational gap between advantaged and disadvantaged students and worsening educational inequality. In contrast, if the data support the negative selection hypothesis, then shadow education could serve as an effective tool to mitigate educational inequality between social classes, as those most disadvantaged students would benefit more from shadow education than do advantaged students. Therefore, in addition to providing a more accurate picture of the role of shadow education in educational stratification, examining treatment effect heterogeneity will also inform policymakers as to which populations benefit from shadow education and which populations require more support.

## Analytical Strategy: Methods, Data, and Measurement

### *Methods*

To estimate both the average and heterogeneous effects of shadow education, I first estimate individuals' propensity scores for receiving shadow education using probit regression as follows:

$$P = p(d_i = 1|X_i)$$

A propensity score is the conditional probability of treatment given the observed covariates $X$. In this form, $P$ is the propensity score, $d_i$ indicates whether student $i$ uses shadow education or not, and X is a vector of observed covariates. I estimate the propensity scores for three counter factual model *separately* (all models are described in greater detail in the measurement section below). Then, I invoke an ignorability assumption. The ignorability assumption means that after controlling for a given set of pretreatment covariates, there are no additional confounders between treated and untreated cases. The plausibility of the ignorability assumption depends on the richness of the observed covariates (Brand and Xie 2010). Therefore, I carefully include various covariates to predict the use of shadow education based on prior studies. However, it is important to note that there is always the possibility of unobserved causal factors.

Based on the estimated propensity scores, I examine the average treatment effect on the treated (ATT). ATT represents the average gain from shadow education for those who actually were treated:

$$\tau_{att} = E(y^{d=1} - y^{d=0} \mid d = 1)$$

I use nearest neighbor and kernel matching to estimate the ATT for each counterfactual model; unmatched differences are also calculated for comparison.

While the estimated results of ATT indicate the average effects of shadow education, I use SM and SD to estimate treatment effect heterogeneity. Both methods are based on propensity score matching approaches. Based on the estimated propensity scores for participating in shadow education, SM constructs a balanced propensity score strata, which means that those who do and those who do not participate in shadow education are not statistically different in terms of the mean value of every covariate and propensity score. Therefore, the only difference of those within the same balanced propensity score stratum is the treatment condition (i.e., whether or not the individual takes shadow education). Then, I estimate propensity score stratum–specific effects using OLS regression, and finally I summarize the systematic pattern of heterogeneous treatment effects in response to shadow education across propensity strata using a variance-weighted least squares regression (Brand, Pfeffer and Goldrick-Rab 2012). I repeat this analysis for three counterfactual scenarios (described in the next section).

In addition to the SM method, I conduct a sensitivity test using the SD method. The SM method has two key assumptions: everyone in each stratum has the same effects and it examines linear trend of heterogeneous effects across the strata. By using the SD method, I can test for sensitivity to the linearity and strata-specific homogeneity assumptions imposed in the SM method (Brand, Pfeffer and Goldrick-Rab 2012). Based on the estimated propensity scores, the SD method fits a separate nonparametric regress of the

SAT scores on the propensity scores and then takes the difference in the nonparametric curves between shadow education participants and non-participants (Brand, Pfeffer and Goldrick-Rab 2012:14). Therefore, the SD method capture curvilinear patterns of effect heterogeneity of shadow education

*Data*

In this analysis, I use the Educational Longitudinal Survey of 2002 (ELS), which contains extensive information about respondents' social and economic backgrounds and educational information such as grades, educational activities, and aspirations. ELS 2002 followed a nationally representative sample of U.S. high school sophomores in 2002 through their senior years in 2004 and beyond in 2006. I restrict my sample to those who have valid SAT or ACT scores to determine the effect of shadow education on SAT scores. I further restrict my sample to those who answered both the base and first follow-up surveys since only those samples have valid answers to questions about use of shadow education. Since I use only a portion of the total sample, it is difficult to generalize my findings to high school students in general. However, because my sample includes the majority of students in the sample who prepared for SATs in school and actually took the SAT, it has still important implications for understanding the role of shadow education in the college preparation process. The final sample consists of 6,900[1] high school seniors in the United States, but the actual sample varies according to each counterfactual model.

*Measurements*

First, in terms of the treatment variable, I generate three categories for the shadow education variables: (1) those who did not participate in any form of shadow education, (2) those who did so using only public resources, and (3) those who either used only private resources or used private resources in addition to public resources.[2] Since I use these three categories for the shadow education variable, I generate multiple contrasts. The effect of private

[1] Following the Institute of Education Sciences' restricted-use data security procedures, I rounded the unweighted Ns to the nearest 10.

[2] I do this because many private shadow education users also utilize public resources. I also conduct additional analysis with those who used only private resources as a key treatment variable and the results do not greatly differ from this study's findings.

test preparation is my key interest and thus I examine the following three counterfactuals: (1) the effect of private shadow education vs. anything else (including both public and no test preparation), (2) private shadow education vs. public shadow education, and (3) private shadow education vs. no test preparation.

Second, I utilize various independent variables to estimate a student's propensity to participate in private shadow education compared to the three counterfactuals. These include demographic characteristics, family background, prior educational achievement, educational motivation/aspiration, parents' expectation/involvement, and high school characteristics. All measures of the independent variables are from base year data (2002). Most variables have only a small amount of missing data; however, five variables do have a relatively greater amount of missing data (remedial course enrollment, PSAT plan, college info-seeking activities, discussing SAT prep with parents, and discussing school courses with parents). Instead of removing those missing cases, I create an additional level that denotes missing for each of the five variables, following prior studies on shadow education (Byun and Park 2012; Domingue and Briggs 2009; Hansen 2004). I also conduct the same analyses with a non-missing sample (listwise deletion), which yields similar results.

1) Demographic characteristics: Gender is a dummy variable (male = 1) and race/ethnicity includes white [reference category], Black, Asian, Hispanic, and other races (Native Hawaiian, Pacific Islander, American Indian, and Alaska native)
2) Family background: I use a dichotomous variable for family composition such as whether or not a student lives with both parents (both parents = 1). The SES index is provided by ELS and is a standardized composite score based on five variables, including father/mother's education, family income, and father/mother's occupation.
3) Prior educational achievement: I utilize four variables to measure prior educational achievement. Tenth grade GPA, math and reading composite test scores provided by ELS, and number of AP and remedial courses. Except for the remedial course variable, the other three variables are continuous. The remedial course variable consists of three categories (yes, no [reference category], and missing).
4) Educational motivation/aspiration: Educational motivation includes five variables. These are students' educational expectations (1–8),

importance of good grades (1–4), homework hours per week (more than 10 hours = 1, less than 10 hours = 0), plans to take the PSAT (yes, no [reference category], and missing), and seeking college information (yes, no [reference category], and missing).

5) Parents' expectations and involvement: Three variables are used for measuring parental characteristics. These are parent's educational expectations (1–7), discussing SAT preparations with parents (never [reference category], sometimes, often, and missing), and discussing school courses with parents (never [reference category], sometimes, often, and missing).
6) High school characteristics: Students' school features include three variables, which are private school (yes = 1), school region (south [reference category], northeast, midwest, and west), and urbanity (urban [reference category], suburban, and rural)

Finally, the SAT score is my outcome variable. For those with only ACT scores, ELS 2002 converted the ACT scores to the SAT scale, which ranges from 400–1600.

## Results

### *Descriptive statistics*

Table 1 shows the results of descriptive statistics by shadow education use. My key interest group, those who used private shadow education, is quite different from both the group that used only public shadow education and the group that did not use any shadow education.

First, it is evident that private shadow education participants have much better economic backgrounds than non-participants, which confirms our prior assertion that participation in private shadow education greatly depends on individual economic background. High school characteristics also shows a clear distinction between private shadow education participants and others. A greater percentage of private shadow education participants attend private schools and schools in urban areas.

In terms of demographic characteristics, there are more women who participate in private shadow education, and Black and Asian students are more likely to engage in private shadow education than White students.

Prior educational achievement shows somewhat mixed results. For

**TABLE 1**
**Descriptive Statistics by Shadow Education Use: Educational Longitudinal Study 2002 (N=6,900)**

| | Full sample | Private shadow education | Public shadow education | No shadow education |
|---|---|---|---|---|
| ***De mographic Characteristics*** | | | | |
| Male | .46(.007) | .42(.016) | .41(.010) | .57(.014) |
| *Race* | | | | |
| White | .72(.007) | .63(.015) | .73(.009) | .79(.011) |
| Black | .11(.005) | .17(.012) | .11(.006) | .07(.007) |
| Asian | .05(.003) | .08(.006) | .05(.004) | .03(.004) |
| Hispanic | .09(.004) | .09(.009) | .09(.006) | .09(.008) |
| Other race | .03(.003) | .04(.007) | .03(.004) | .03(.004) |
| ***Family backgrounds*** | | | | |
| Parents' SES | .23(.011) | .42(.023) | .20(.014) | .14(.020) |
| Both parents(0/1) | .67(.007) | .67(.015) | .67(.010) | .67(.014) |
| ***Educational Achievements*** | | | | |
| $10^{th}$ grade GPA | 3.06(.010) | 3.05(.022) | 3.11(.013) | 2.99(.020) |
| AP courses | 1.09(.025) | 1.35(.058) | 1.07(.034) | .921(.049) |
| Remedial class(0/1) | | | | |
| Yes | .08(.004) | .09(.011) | .08(.005) | .08(.008) |
| No | .87(.005) | .85(.011) | .87(.007) | .87(.010) |
| Missing | .06(.003) | .06(.008) | .05(.005) | .05(.007) |
| Test score($10^{th}$) | 54.74(.13) | 54.17(.30) | 54.71(.166) | 55.20(.278) |
| ***Educational Motivation/Aspiration*** | | | | |
| Educational expectation (1-8) | 6.43(.016) | 6.68(.033) | 6.47(.022) | 6.18(.033) |
| HW hours per week (10h) (0/1) | .29(.007) | .37(.016) | .29(.010) | .22(.012) |
| Importance of good grades (1/4) | 3.55(.009) | 3.64(.018) | 3.59(.013) | 3.40(.021) |

**TABLE 1**
**(CONTINUED)**

| | Full sample | Private shadow education | Public shadow education | No shadow education |
|---|---|---|---|---|
| Plans to take the PSAT | | | | |
| Yes | .75(.007) | .81(.013) | .76(.009) | .68(.014) |
| No | .20(.006) | .13(.011) | .19(.008) | .26(.013) |
| Missing | .06(.003) | .06(.008) | .05(.004) | .06(.007) |
| Seeking college info | | | | |
| Yes | .82(.006) | .87(.011) | .84(.007) | .75(.013) |
| No | .10(.005) | .06(.007) | .09(.006) | .16(.011) |
| Missing | .08(.004) | .08(.009) | .07(.005) | .09(.008) |
| ***Parents' Expectation /Involvement*** | | | | |
| Educational expectation (1-7) | 5.61(.016) | 5.82(.032) | 5.62(.021) | 5.46(.032) |
| Discuss SAT prep with parents | | | | |
| Never | .30(.007) | .18(.012) | .28(.009) | .41(.014) |
| Sometimes | .41(.007) | .41(.016) | .44(.010) | .35(.014) |
| Often | .17(.006) | .28(.014) | .16(.007) | .12(.010) |
| Missing | .12(.005) | .13(.011) | .12(.006) | .12(.009) |
| Discuss courses with parents(1-3) | | | | |
| Never | .10(.004) | .07(.008) | .09(.006) | .15(.010) |
| Sometimes | .47(.007) | .43(.016) | .48(.010) | .49(.015) |
| Often | .31(.007) | .37(.016) | .32(.010) | .25(.013) |
| Missing | .12(.005) | .13(.010) | .11(.006) | .11(.009) |
| ***High school Characteristics*** | | | | |
| Private (0/1) | .11(.003) | .18(.009) | .11(.004) | .08(.005) |
| Region | | | | |
| Northeast | .19(.006) | .23(.014) | .19(.008) | .18(.012) |
| Midwest | .28(.007) | .21(.013) | .28(.009) | .34(.014) |

TABLE 1
(CONTINUED)

| | Full sample | Private shadow education | Public shadow education | No shadow education |
|---|---|---|---|---|
| West | .17(.006) | .19(.014) | .17(.009) | .17(.013) |
| South | .35(.007) | .38(.015) | .37(.010) | .31(.013) |
| Urbanity | | | | |
| Urban | .26(.007) | .33(.015) | .25(.009) | .24(.012) |
| Suburban | .53(.007) | .53(.016) | .54(.010) | .54(.015) |
| Rural | .20(.006) | .14(.011) | .22(.009) | .23(.012) |
| ***Outcome*** | | | | |
| SAT scores | 1011.171 (2.99) | 1020.52 (6.82) | 1006.68 (3.92) | 1012.79 (6.15) |
| N | 6,900 | 1610 | 3530 | 1770 |

Note: Descriptive statistics are weighted, with unweighted sample sizes reported in the last row. Standard deviations in parentheses.

example, average 10th grade test scores are highest in non-participants and lowest for private shadow education participants, though the gap is very small. However, 10th grade GPA is highest for public shadow education participants and lowest for the non-participant group. In terms of number of AP courses taken, private shadow education participants took a greater number of AP courses than other groups. These results suggest that shadow education in the United States seems to be used for both remediation and enrichment.

In terms of educational motivation and aspiration, the gap between private shadow education participants and public shadow education participants is relatively small, but there is a substantive difference between non-shadow education participants and others. This result indicates that motivation/aspiration is an important factor in determining participation in any SAT preparation activities, but that this does not seem to be a critical factor in distinguishing between private and public resources. The results show that private education participants' parents have higher educational expectations and are more involved in their children's education than other groups. In particular, they are much more likely to discuss SAT preparation with their children, which indicates that the use of private shadow education seems to be strongly determined by parents' expectations or level of involvement.

*Average effects of shadow education*

As a first step, I estimate the propensity score for each individual who participated in private shadow education using a probit regression model. Appendix table A shows the results of the probit regression model according to three counterfactual scenarios. The three models show very similar results and confirm the patterns observed in the descriptive statistics that suggest that socioeconomic resources and both parents' and children's educational motivation and aspirations positively affect the likelihood of participating in shadow education. Also, the mixed findings for prior academic achievements seem to point to the dual function of private shadow education in the United States as both remediation and enrichment.

Based on the estimated propensity scores, I estimate the average treatment effect on the treated of private shadow education under the three counterfactual models. Table 2 shows the results of the matching estimates and unmatched differences of private shadow education. First of all, private shadow education has a statistically significant effect on SAT scores in every model and with every matching method. For example, private shadow education participants scored about 18 points higher on the SAT than those who did not use private shadow education (private shadow education vs. anything else). Although the estimates of the second model (private shadow

**TABLE 2**
**Matching Estimates of Private Shadow Education on SAT Scores**

| | Private shadow education vs. | | |
|---|---|---|---|
| | Anything else | Public resources | None |
| Unmatched Differences | 29.480***[a]<br>(5.743)[b] | 30.170***<br>(6.075) | 28.105***<br>(7.166) |
| Nearest Neighbor Matching (k = 5) | 17.431*<br>(6.968) | 19.868**<br>(7.285) | 22.294*<br>(10.370) |
| Kernel Matching | 18.667**<br>(6.433) | 19.299**<br>(6.705) | 21.358*<br>(9.713) |
| N | 6900 | 5130 | 3380 |

Note: a. †$p < .10$ *$p < .05$ **$p < .01$ ***$p < .001$, [b] Standard errors in parentheses;

education vs. public shadow education) are slightly lower than those of the third model (private shadow education vs. none), the differences are relatively small. For example, the effect of private shadow education compared to public shadow education is about 19 points higher and the effects of private shadow education compared to none is about 22 points higher. This result suggests that public shadow education does not effectively substitute for private shadow education. In short, under the assumption of effect homogeneity, private shadow education had clearly positive effects on SAT scores regardless of the counterfactual and increased SAT scores by about 17 to 22 points.

*Heterogeneous effects of shadow education*

Next, to estimate treatment effect heterogeneity, I construct balanced propensity score strata based on estimated propensity scores of using private shadow education under the three counterfactual models. The mean values of every covariate and propensity score between treated and untreated students in the same propensity strata are not statistically different. Appendix table B shows the mean value of every covariate and treatment variable for the "private vs. anything else" model. It presents a very clear pattern of students' family backgrounds, educational motivation/aspirations, parents' educational expectations/involvement, and the percentage of private and

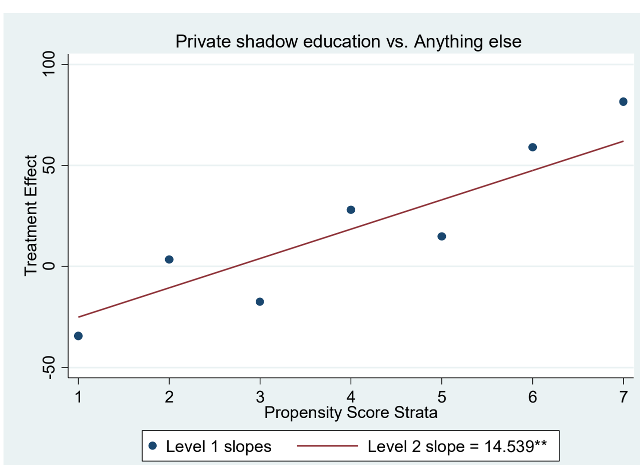



**Fig. 1.**—Stratification Multilevel (SM) Method for Heterogeneous Treatment Effects

urban schools increasing as propensity score strata increases. This indicates that individuals in higher propensity strata tend to come from more advantaged social and educational positions than those in lower propensity strata.

Next, based on propensity score strata, I estimate stratum-specific treatment effects (level-1) and then I examine the systematic pattern of heterogeneous treatment effects across propensity score strata with variance weighted least squares regression (level-2). Figure 1 shows the linear trend of treatment effects across the strata for the "private vs. anything else" model (Appendix C presents every coefficient and significance for the level-1 and level-2 slopes for all three counterfactual models). The level-2 slope is 14.539 and is statistically significant at the 0.01 alpha level. This means that as one strata increases, the effect of private shadow education on SAT scores increases by 14.539 points. This pattern clearly supports the positive selection hypothesis, which indicates that those who are most likely to participate in shadow education benefit most from it. This positive selection hypothesis is particularly important in terms of educational stratification in the United States. This is because, as seen in Appendix Table B, high propensity participants are already advantaged in terms of their social and educational backgrounds and thus, positive selection could further exacerbate educational inequality between advantaged and disadvantaged students. In sum, although the average effects of shadow education seem to be moderate, the effect of private shadow education is substantially large for the more advantaged, high-propensity students.

Next, I examine the pattern of effect heterogeneity of private shadow education based on alternative counterfactuals. I repeat the SM method for

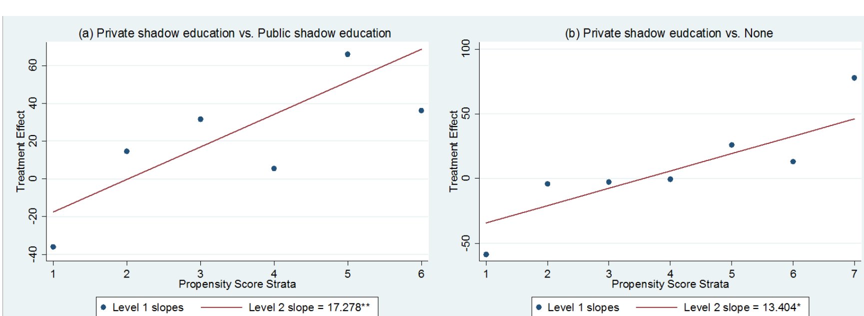


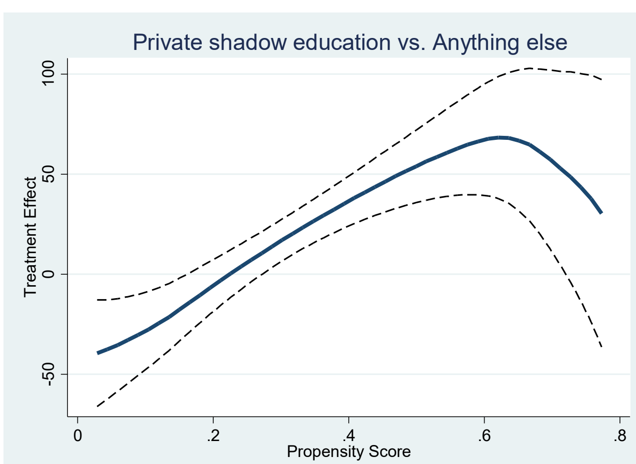

FIG. 2.—Stratification Multilevel (SM) Method for Heterogeneous Treatment Effects

the “private vs. public” and “private vs. none” models. Since public shadow education is a more accessible resource for the majority of students, understanding the effectiveness of public shadow education is particularly important when seeking to alleviate educational inequality.

Figures 2-a and 2-b show the results of two alternative counterfactual models. Both counterfactuals also show similar upward linear slopes for “private vs. anything else.” The level-2 slope for “private vs. public” is 17.278 and it is statistically significant (P-value = .008) and the level-2 slope for “private vs. none” is 13.075, which is also statistically significant (P-value = .013). Both counterfactuals clearly indicate a positive selection pattern. These results suggest that use of public shadow education does not change the pattern of positive selection for private shadow education.

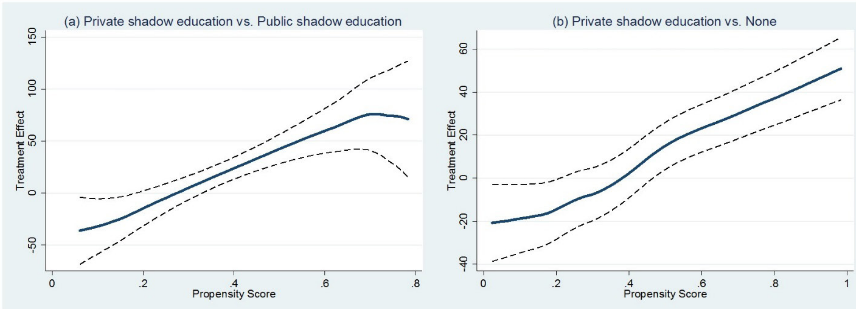
*Auxiliary analysis*

Lastly, using an SD method, I conduct a sensitivity test for the linearity assumption imposed by the SM method. Figure 3 presents the results of the “private vs. anything else” model. The x-axis represents the continuous propensity score and the y-axis represents the differences in nonparametric regressions between the treated and un-treated groups. The effects of private




Note: Solid line indicates local polynomial fit. Dashed lines indicate 95 percent confidence interval

**Fig. 3.**—Smoothing-Differencing (SD) Method for Heterogeneous Treatment Effects




Note: Solid line indicates local polynomial fit. Dashed lines indicate 95 percent confidence interval

**FIG. 4.**—Smoothing-Differencing (SD) Method for Heterogeneous Treatment Effects

shadow education increase as propensity score increases; however, after a .6 propensity score, the effects of private shadow education decrease until the highest propensity score. Thus, the effects of shadow education are strongest around a propensity score of .6 and weakest in the lowest propensity distribution.

Figure 4 presents the results of the SD method for two alternative counterfactuals. In both figures, the effects of private shadow education increase as propensity score increases, which clearly indicates positive selection. While the effects of private shadow education compared to public resources in Figure 4-a slightly decrease after the .7 propensity score, the effect of private shadow education to none in Figure 4-b continuously increases across the propensity scores. In conclusion, the overall trend found using the SD method for three counterfactual scenarios also support the positive selection hypotheses. Although the effects of shadow education slightly decrease at the highest propensity scores in some models, the key implications of this remain unchanged. That is, advantaged students tend to benefit more from private shadow education and the most disadvantaged students benefit least from it.

## Conclusion

The overall goal of this study was to examine the effects of shadow education on SAT scores and thus to investigate the impact of shadow education on educational inequality in the United States. The key questions that asked in this study are as follows: (1) What is the average effect of private shadow education on SAT scores? (2) How do the effects of private shadow education differ based on individual likelihood of participating in shadow education? and (3) How do the average and heterogeneous effects of private shadow education differ according to multiple counterfactual scenarios? By examining variation in SAT scores by likelihood of participating in private shadow education, this study challenges the effect homogeneity assumption and sheds light on the exact role of shadow education in educational stratification. Moreover, by considering multiple counterfactual conditions, I investigated the possibility that public resources could serve as a substitute for private shadow education.

Using rich data from the Educational Longitudinal Study of 2002, we found significant variation in SAT scores by likelihood of participation in private shadow education. Those who were most likely to use private shadow education benefitted the most from it. Since higher propensity participants are already socioeconomically advantaged, this pattern suggests that private shadow education widens the education gap between social classes. That is, although the average effects of private shadow education are moderate, which limits the relevance of private shadow education to educational inequality, the positive selection pattern suggests that the most advantaged students benefit considerably more from private shadow education and thus it functions as a strong mechanism of educational stratification in the United States. Moreover, I find that using public resources in lieu of private shadow education neither ameliorates the impact of private shadow education on academic outcomes nor changes the positive selection pattern.

This study broadens our understanding of how private shadow education is a critical factor in exacerbating educational stratification in the United States. However, the policy implications of this study are quite complex. Intervening in shadow education would be extremely difficult due to the fact that shadow education operates in the private sector. Some countries such as South Korea, Uganda, and Mauritius have tried to ban shadow education, but these bans have been ineffective (Bray 2006). If direct intervention in shadow education is difficult, then, this study raises questions for public education in terms of its system and quality in the United States. In this sense, it would be useful to examine what current educational environments boost the prevalence and impact of shadow education. If the impact of shadow education centers on SAT preparation activities and the college entrance process, what kind of alternative systems could be put in place? And more specifically, given that public shadow education does not help to mitigate the disproportionate impact of private shadow education, what differences exist between those two educational activities? What kind of additional learning opportunities within formal schooling could be provided to disadvantaged students, and how can these activities be made effective compared to private shadow education? These questions are only part of what policymakers and education scholars need to examine in order to ameliorate the educational inequality that is caused by private shadow education.

In conclusion, given the increasing prevalence of private shadow education in the United States, this study emphasizes that the acquisition of educational advantages greatly depends on private investment and that the pathway to social mobility through education appears to cross the border of public education. Therefore, extensive efforts to enhance educational equality is necessary, taking into account both public and private educational contexts.

(Submitted: June 24, 2018; Revised: July 30, 2018; Accepted: September 4, 2018)

## Reference

Alon, Sigal. 2009. "The evolution of class inequality in higher education competition, exclusion, and adaptation." *American Sociological Review* 74: 731-755.

______. 2010. "Racial differences in test preparation strategies: A commentary on shadow education, American style: Test preparation, the SAT and college enrollment." *Social Forces* 89: 463-474.

Alon, Sigal and Marta Tienda. 2007. "Diversity, opportunity, and the shifting meritocracy in higher education." *American Sociological Review* 72: 487-511.

Baker, David and Gerald K LeTendre. 2005. *National differences, global similarities: World culture and the future of schooling*: Stanford University Press.

Brand, Jennie E. and Dwight Davis. 2011. "The impact of college education on fertility: Evidence for heterogeneous effects." *Demography* 48: 863-887.

Brand, Jennie E., Fabian T. Pfeffer, and Sara Goldrick-Rab. 2012. "Interpreting community college effects in the presence of heterogeneity and complex counterfactuals." *Unpublished Working Paper. WISCAPE.*

Brand, Jennie E. and Juli Simon Thomas. 2013. "Causal Effect Heterogeneity." Pp. 189-213 in *Handbook of Causal Analysis for Social Research*: Springer.

Brand, Jennie E. and Yu Xie. 2010. "Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education."

*American Sociological Review* 75: 273-302.
Bray, Mark. 2001. “The Shadow Education System: private tutoring and its implications for planners.” *Assessment in Education* 8.
______. 2006. “Private supplementary tutoring: Comparative perspectives on patterns and implications.” *Compare* 36: 515-530.
______. 2009. *Confronting the shadow education system: What government policies for what private tutoring?*: United Nations Educational, Scientific and Cultural Organization; International Institute for Educational Planning.
Briggs, Derek C. 2001. “The effect of admissions test preparation: Evidence from NELS: 88.” *Chance* 14: 10-18.
Buchmann, Claudia. 2002. “Getting ahead in Kenya: social capital, shadow education, and achievement.” *Research in the Sociology of Education* 13: 133-159.
Buchmann, Claudia, Dennis J Condron, and Vincent J Roscigno. 2010. “Shadow education, American style: Test preparation, the SAT and college enrollment.” *Social Forces* 89: 435-461.
Byun, Soo-yong. 2014. “Shadow Education and Academic Success in Republic of Korea.” Pp. 39-58 in *Korean education in changing economic and demographic contexts*: Springer.
Byun, Soo-yong and Hyunjoon Park. 2012. “The Academic Success of East Asian American Youth The Role of Shadow Education.” *Sociology of Education* 85: 40-60.
Choi, Yool and Hyunjoon Park. 2016 “Shadow education and educational inequality in South Korea: Examining effect heterogeneity of shadow education on middle school seniors’ achievement test scores.” *Research in Social Stratification and Mobility* 44: 22-32.
Cheo, Roland and Euston Quah. 2005. “Mothers, maids and tutors: An empirical evaluation of their effect on children’s academic grades in Singapore.” *Education Economics* 13: 269-285.
Dang, Hai-Anh. 2007. “The determinants and impact of private tutoring classes in Vietnam.” *Economics of Education review* 26: 683-698.
Dang, Hai-Anh and F. Halsey Rogers. 2008. “The growing phenomenon of private tutoring: Does it deepen human capital, widen inequalities, or waste resources?” *The World Bank Research Observer* 23: 161-200.
Davies, Scott and Janice Aurini. 2006. “The franchising of private tutoring: a view from Canada.” *Phi Delta Kappan* 88: 123.
Domingue, Ben and Derek C. Briggs. 2009. “Using linear regression and propensity score matching to estimate the effect of coaching on the SAT.” *Multiple Linear Regression Viewpoints* 35: 12-29.
Grodsky, Eric. 2010. “Learning in the Shadows and in the Light of Day: A commentary on Shadow Education, American Style: Test Preparation, the SAT and College Enrollment.” *Social Forces* 89: 475-481.
Grodsky, Eric, John Robert Warren, and Erika Felts. 2008. “Testing and social

stratification in American education." *Annu. Rev. Sociol* 34: 385-404.

Ha, Tran Thu and Trudy Harpham. 2005. "Primary Education in Vietnam: Extra Classes and Outcomes." *International Education Journal* 6: 626-634.

Hansen, Ben B. 2004. "Full matching in an observational study of coaching for the SAT." *Journal of the American Statistical Association* 99: 609-618.

Jacob, Brian A. and Lars Lefgren. 2004. "Remedial education and student achievement: A regression-discontinuity analysis." *Review of economics and statistics* 86: 226-244.

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

Kuan, Ping-Yin. 2011. "Effects of cram schooling on mathematics performance: Evidence from junior high students in Taiwan." *Comparative education review* 55: 342-368.

Lavy, Victor and Analia Schlosser. 2004. "Targeted remedial education for under-performing teenagers: costs and benefits." National Bureau of Economic Research.

Mori, Izumi and David Baker. 2010. "The origin of universal shadow education: What the supplemental education phenomenon tells us about the postmodern institution of education." *Asia Pacific Education Review* 11: 36-48.

Powers, Donald E. and Donald A Rock. 1999. "Effects of coaching on SAT I: Reasoning test scores." *Journal of Educational Measurement* 36: 93-118.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70: 41-55.

Princeton Review. 2010. "The Princeton Review Reports Fourth Quarter and Full Year 2009 Financial Results." Available at: http://www.princetonreview.com/releases.cfm.

Stevenson, David Lee and David P Baker. 1992. "Shadow education and allocation in formal schooling: Transition to university in Japan." *American journal of sociology*: 1639-1657.

Suryadarma, Daniel, Asep Suryahadi, Sudarno Sumarto, and F. Halsey Rogers. 2006. "Improving student performance in public primary schools in developing countries: Evidence from Indonesia." *Education Economics* 14: 401-429.

Tansel, Aysıt and Fatma Bircan. 2006. "Demand for education in Turkey: A tobit analysis of private tutoring expenditures." *Economics of Education review* 25: 303-313.

***YOOL CHOI*** is assistant professor in the department of social studies education at Korea National University of Education. He received his PhD in sociology at the University of California, Los Angeles in 2015. His areas of interest include social stratification, social demography, and sociology of education. His recent publications appeared in Research in Higher Education, Journal of Aging and Health, Research in Social Stratification and Mobility, and Korean Journal of Sociology. *Address*: 511, Education Hall, Korea National University of Education, 250, Taeseong tapyeonro, Kangnae myeon, Heungdeok Gu, Cheongju, Chungbuk, Korea. [*E-mail*: ychoi@knue.ac.kr]