

# **Who benefits most from shadow education?: The heterogeneous effects of SAT preparation activities on SAT scores**

Yool Choi  
*University of California – Los Angeles*

## **Introduction**

In this study, I want to examine the effect of shadow education on SAT scores. Stevenson and Baker (1992) 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” (p. 2). In the past, the extensive use of shadow education was only observed in a few countries, especially in several East Asian countries. However, recent research shows that the expansion and prevalence of shadow education is a worldwide phenomenon, and shadow education is one of the fastest growing industries in many countries around the world (Baker and LeTendre 2005; Bray 1999; Buchmann, Condron and Roscigno 2010; Byun [Forthcoming]).

The prodigious growth of shadow education has caused much concern over educational quality and equality among policymakers (Bray, l'éducation and UNESCO 2009; Mori and Baker 2010). This is because in most cases shadow education accompanies significant private capital investment on the part of certain families and, at the same time, it is less feasible for many families with limited economic capital (Baker and LeTendre 2005; Dang and Rogers 2008; Mori and Baker 2010; Stevenson and Baker 1992). Also, as Grodsky points out, “Shadow education is different from other mechanisms that contribute to

educational inequality such as school segregation and tracking or ability grouping within schools that fall under the control of educational organizations.” (Grodsky 2010:476). That is, it is more difficult to control the pervasiveness and repercussions of shadow education compared to other factors since it takes place outside of formal schooling. Some countries such as South Korea, Uganda and Mauritius have tried to ban shadow education, but the ban was ineffective (Bray 2006).

Along these lines, if shadow education does make a difference in academic achievement, it carries important implications concerning educational opportunity and stratification in society. It is obviously a mechanism for maintaining and increasing social stratification by conferring educational advantages on students who are already advantaged in terms of their rich economic, social and cultural capital. In this study, I propose to examine the effects of shadow education in the context of the United States. Given the increasing trend of shadow education in the U.S, I pose two key questions as to its effects: does shadow education have a causal effect on academic performance, and how do the effects of shadow education differ depending on an individual’s social and economic backgrounds? By examining these two questions, this research has important implications, not only for understanding how more advantaged families utilize their economic resources to acquire the educational advantages of shadow education, but also for suggesting how policymakers might approach the issue of shadow education to diminish educational stratification in the U. S.

Among various forms of shadow education, I will focus on the effect of SAT preparation activities on SAT scores. Although SAT scores have long formed a significant component of college admission procedures, their importance has received increased emphasis in recent years. One of the key reasons is that selective postsecondary institutions increasingly rely on test scores to screen students. This trend has made it difficult for

institutions to achieve diversity (Alon 2009; Alon and Tienda 2007; Grodsky, Warren and Felts 2008).

In this study, I build on and augment prior research on shadow education and coaching effects on SAT scores. First, using propensity score matching technique, I will examine the causal relationship between shadow education and SAT achievement by carefully attending to issues of preexisting heterogeneity. Many prior studies emphasize that selection bias is a major concern when examining the causal effect of shadow education on SAT scores but only a few studies extensively utilized more advanced methods to deal with this issue in the U.S. contexts (Briggs 2001; Byun and Park 2012; Domingue and Briggs 2009; Hansen 2004; Powers and Rock 1999). Second, I will analyze variation in the effects on SAT scores based on the likelihood of receiving shadow education. That is, using the stratification-multilevel and smoothing-differencing method, I will summarize the systematic trend of heterogeneous effects across propensity strata. This result could reveal some masked aspects of shadow education in terms of educational stratification. For example, although average effects of shadow education are moderate, advantaged students might benefit significantly more than others from shadow education. In this case, shadow education is particularly beneficial for already advantaged students to enrich their educational achievement and thus widen the educational gap between social classes. To my knowledge, no prior research has examined this issue in shadow education studies. Third, along with these new methodological approaches, I will extensively examine the association between racial/gender background and the prevalence and effects of shadow education utilization. In sum, this study contributes to an understanding of how the prominence of shadow education, depending heavily on an individual's social and economic backgrounds, affects his or her academic achievement.

## **Background**

### *Shadow education in the U.S context*

In the U.S, even though shadow education has important implications for social stratification, it has received far less attention relative to other factors that contribute to educational inequality. This is due in part to the educational system in the U.S.. That is, shadow education is prevalent where formal examinations are used-- particularly centrally administered examinations--and there are 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, historically most East Asian countries strongly emphasize the value of college entrance exams for their admission process. Also, a college degree is one of the most important determining factors in an individual's success on the labor market over the course of his or her life (Baker and LeTendre 2005; Bray 1999). In contrast, in the United States, various non-cognitive factors such as extracurricular activities are also important, along with formal entrance exams (SAT or ACT), in college admission procedures (Bray 1999; Buchmann, Condron and Roscigno 2010; Byun and Park 2012). Such additional factors have somewhat relieved students of the pressures and competition to obtain high scores on standardized tests.

However, along with the expansion in secondary and post-secondary education, the importance of standardized tests has been growing in recent years in the U.S. (Alon 2009; Alon and Tienda 2007; Grodsky, Warren and Felts 2008). The growing importance of standardized tests in the admissions process may increase the pressure on students to score better on high-stakes tests, and this may lead to a growing demand for shadow education in

the U.S. (Buchmann, Condron and Roscigno 2010; Byun and Park 2012). The expansion of commercial coaching companies such as the Princeton review and Kaplan in recent years has confirmed this trend in the U.S. and Canada (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). Apparently, this private coaching business is one of the fastest growing and most lucrative multi-million dollar industries in the U.S.. Thus, understanding the implications and effects of shadow education has become an important task for stratification scholars and policymakers in the U.S..

Even though shadow education encompasses every educational activity occurring outside of school that enhances a student's educational achievement during every transition such as primary to secondary and secondary to post-secondary institutions, Buchmann et al. (2010) have called a variety of SAT/ACT test preparation behaviors the American style of shadow education. Compared to other nations, the most apparent and parallel forms of shadow education in the U.S are those activities that prepare students for SAT/ACT tests, using sources outside of formal schooling. In many countries, the preparation for college entrance exams is the most extensive and important form of shadow education (Bray 1999; Stevenson and Baker 1992). Therefore, examining the effects of various SAT test preparation activities from a comparative perspective could lead to a thorough discussion of the implications of shadow education in the United States.

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

One of the most important questions about shadow education concerns its efficiency and equity. That is, in the context of students' investments, does shadow education have a

significant effect on their educational achievements and how does shadow education affect all students in terms of educational opportunity, equality, and achievement? 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, some studies have found strong positive effects (Buchmann 2002; Byun and Park 2012; Dang 2007; Tansel and Bircan Bodur 2005), other studies have found modest positive effects (Baker et al. 2001; Domingue and Briggs 2009), and even sometimes negative effects (Cheo and Quah 2005).

These mixed results may be attributed to various factors such as contextual differences of educational systems, broad divergences in the operational definitions of the shadow education variable, the choice of dependent variable, and applied statistical models (Byun [Forthcoming]). However, one of the key issues in estimating the causal relationship between shadow education and its outcome in prior research is the selection bias problem (Byun [Forthcoming]; Dang and Rogers 2008). That is, most of the research has limitations when it comes to addressing pretreatment heterogeneity. Many studies show that there are substantial preexisting differences between those students who use shadow education and those students who do not. These two groups differ by family background, prior educational achievement, educational aspirations and expectations. Among those students who use shadow education, the purposes of shadow education are different (remedial vs. enhancement). OLS regression with various control variables is the most common approach; however, this cannot fully address the issue of pretreatment heterogeneity (Brand and Xie 2010; Rosenbaum and Rubin 1983; Schneider and Association 2007). Among many studies on the effects of shadow education and coaching, only a few studies have dealt with this

problem and have used several advanced statistical models such as propensity matching estimation, fixed-effects, and difference in difference approach (Dang and Rogers 2008; Domingue and Briggs 2009; Lavy and Schlosser 2004; Mischo and Haag 2002). In this study, I use propensity score matching estimation with more complex counterfactual scenarios by comparing the use of private sources of shadow education to 1) Baseline counterfactual; 2) No use of shadow education; and 3) Use of only public sources, to estimate causal effects of shadow education.

### *Treatment effect heterogeneity 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 treatment effect heterogeneity in the effect of shadow education on academic achievements based on the likelihood that certain individuals will utilize shadow education. Treatment effect heterogeneity refers to estimation of the interaction between treatment and propensity of treatment (Brand 2010; Brand and Davis 2011; Brand and Xie 2010; Brand and Simon-Thomas [Forthcoming]; Xie, Brand, and Jann [Forthcoming]). While homogeneous shadow education effects are assumed in most prior research, it is reasonable to suppose that responses to shadow education differ. That is, individuals differ not only in preexisting background attributes but also in their behavioral response to shadow education.

In terms of shadow education, treatment effect heterogeneity can answer such questions as who is most likely to utilize shadow education, and who benefits most from shadow education: is shadow education particularly beneficial for already advantaged students to get ahead of others (i.e., the enhancement function), or do students from less privileged backgrounds, considering their poor counterfactuals, benefit from a diminished the

educational gap (i.e., the remedial function)? Also, how do different kinds of shadow education (private sources vs. public sources) show different patterns of treatment effect heterogeneity? 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 policy making to deal with specific sub-populations.

Despite many prior studies on shadow education, only one study shows some evidence of treatment effect heterogeneity. Domingue and Briggs (2009) estimated the effects of coaching on SAT scores using propensity score matching method with the Educational Longitudinal Survey of 2002. They utilized one independent variable (commercial courses) among various variables related to shadow education. In this analysis, they estimated propensity score stratum-specific coaching effects on SAT score using Linear regression models. While they did not explicitly summarize the trend in the variation of effects by propensity score strata using the stratification-multilevel method, they found some evidence of treatment effect heterogeneity and argued that “the effects vary quite dramatically with the highest effect estimates found in the higher subclasses.” (Domingue and Briggs 2009:19). In my analysis, I build on their research and expand this by addressing more complex counterfactuals and testing for systematic trends in the treatment effect heterogeneity using SM and SD model.

There are two competing theoretical models for interpreting patterns of treatment effect heterogeneity of shadow education. First, individuals who are most likely to use shadow education also benefit most from shadow education, which is called *positive selection*. Human capital theory supports this pattern. That is, individuals who think that they will have the highest returns from shadow education are most likely to use shadow education.



So, their utilization of shadow education is indeed a rational behavior based on a utility-maximizing strategy. Also, we can consider the quantity and quality aspects of shadow education. Since the quantity and quality of shadow education varies by economic investments, there is a possibility that those who are most likely to use shadow education use shadow education longer and of higher quality than others.

Second, individuals who are less likely to use shadow education benefit more from shadow education, which is called *negative selection*. If an individual actually uses shadow education, even though he/she may have a weak propensity to use shadow education, he/she might have strong educational motivation and thus benefit more from shadow education. In other words, the remedial function may be more important than the enhancement function for educational success. Grodsky argued that “shadow education in preparation for college entrance exams is a relatively ineffective means of improving test scores but nonetheless quite important in terms of its psychic benefits .... [which] serve to at least partially assuage the needs of the affluent (parents) to feel they have done everything they can to pave the way for their children” (Grodsky 2010:475). That is, students from affluent families might use shadow education without strong motivation to increase their test scores. They just use shadow education since their parents force them to do it. In this case, shadow education is likely to be an ineffective tool for increasing test scores. Also, we should consider the different counterfactuals between high propensity and low propensity goers. If shadow education is an effective tool to increase SAT scores, extremely poor achievements of low propensity non-shadow education goers will mean a larger difference than high-propensity goers. Since high propensity non-shadow education goers have diverse social and cultural capital, and thus potentially alternative means for securing high test scores, they can diminish the gap to their counterfactuals. Therefore, the main goal of this study is to examine these two

competing scenarios (positive vs. negative selection) and to reveal the systematic pattern of the treatment effect heterogeneity of shadow education.

*What is shadow education?*

Along with methodological issues, I consider two key analytical issues that relate to variation in effects: 1) definition of shadow education, and 2) racial and gender differences.

First, the operational definitions of shadow education are diverse and controversial. Buchmann et al. (2010) constructed an indicator of *highest-level test preparation* in their research which includes four kinds of test preparation activities: “(0) used no preparation of any kind; (1) used test preparation books and/or computer software and/or videos, but no other type of test preparation; (2) took a high school course either alone or in combination with test prep in category 1, but no other type of test prep; (3) took a private course either alone or in combination with test prep in categories 1 and 2, but had no private tutoring; (4) used a private tutor either alone or with any type of test prep” (Buchmann, Condron and Roscigno 2010:444). They included all four activities as a form of shadow education and did not make a public vs. private distinction. However, Grodsky criticized these operational definitions of shadow education and argued that “Shadow education... exists in the private as opposed to the public sphere... has economic costs that bar most disadvantaged families from participation” (Grodsky 2010:476). Thus, Grodsky assumed that the private sources such as private courses and private tutoring should only be considered as shadow education in the United States.

Alon (2010) similarly emphasized the importance of the distinction between private and public sources in her comments and supplementary analysis on Buchman et al.’s work. In contrast to Grodsky, she considered both private and public sources to be aspects of shadow

education and argued that distinguishing private vs. public is analytically important for its economic aspects and effectiveness (Alon 2010).

For these critics, Buchman et al. responded that their definition of shadow education is context specific and their main goal is to extend to the United States the utility of a general theoretical concept developed in other national contexts (Buchmann, Condron, and Roscigno 2010). They argued that the distinction between public and private is somewhat ambiguous in the United States compared to other nations and many students use multiple strategies simultaneously.

Careful consideration should be given to the operationalization of shadow education. In this study, following Alon's perspective (2010), I propose to use the dichotomous definition of shadow education by means of distinction between public and private. This decision is motivated by my effort to understand the degree to which shadow education contributes to educational inequality. More detailed information of measurement of shadow education will be introduced in the measurement section below.

### *Racial differences in shadow education*

The last key issue in the shadow education studies is racial difference. Buchmann et al. found that there are substantial racial and ethnic differences in the use of test preparation: blacks and Hispanics are more likely to utilize test preparation than whites from comparable backgrounds (Buchmann, Condron & Roscigno 2010). This pattern is surprising given black's lower levels of social and financial resources and lower test achievements than whites. Alon (2010) specifically examined the racial disparity in test preparation and concluded "racial disparity in the structure of educational opportunity, as a result of affirmative action policies at selective colleges and universities, is a key to understanding

strategic decisions about test preparation” (Alon 2010:6). That is, different educational opportunities of post-secondary education between races gives different motivations and expectations for test preparation and this induces racial disparities in shadow education.

Byun and Park (2012) examined the effects of private course and private tutoring on SAT scores particularly focused on comparing East Asian and other racial and ethnic groups. They found that East Asian students were most likely to use a private course for enrichment purposes and benefited most from this. They concluded that “growing demand for shadow education may also lead to an increase in racial/ethnic inequalities in educational achievement” (Byun & Park 2012:56).

In sum, the racial/ethnic difference in the prevalence and effects of shadow education is one of the crucial factors in understanding the implications of shadow education in the U.S. In this study, I will examine the effect of shadow education separately for three racial groups: white, black and Asian. I will extend prior research by examining how the pattern of heterogeneous effects of shadow education on SAT scores differs by racial background. If there are notable racial differences in shadow education utilization as has been seen in prior studies, I could expect the different pattern of treatment effect heterogeneity of shadow education between races. By examining these trends, I could suggest a more profound understanding of racial disparity that exists in utilization and effects of shadow education. Lastly, along with racial differences, I will also look at the gender differences in shadow education.

## **Research Questions**

In this study, I will examine the following research questions:

1. What is the causal relationship between shadow education and SAT scores when we consider pre-existing heterogeneity and complex counterfactual states (described below)?
2. How do the effects of shadow education on SAT scores differ by the individual's likelihood of using shadow education? In other words, who benefits most from shadow education?
3. How do the effects of shadow education on SAT scores differ by race and gender?

I will examine these questions by comparing the use of private sources of shadow education to three alternative counterfactuals: 1) Baseline counterfactual (anything else except use of private sources); 2) No use of shadow education; and 3) Use of only public sources

## **Analytical Strategy: Methods, Data, and Measurement**

### *Methods*

Considering two methodological issues, pretreatment heterogeneity and treatment effect heterogeneity, I will use propensity score matching, stratification-multilevel (SM) and, smoothing-differencing (SD) methods in this study to shed light on: 1) the average effect of shadow education on SAT achievement and 2) the heterogeneous effect of shadow education by the likelihood that individuals will use shadow education on SAT achievement. My analysis proceeds in four steps. First, I will estimate individuals' propensity scores based on probability of receiving shadow education. Second, using estimated propensity scores, I will examine the treatment effects for the treated (TT) who utilize shadow education under three different counterfactual scenarios (I describe the counterfactual scenarios in greater detail in the measurement section below). Third, using SM methods, I will summarize a systematic

pattern of heterogeneous treatment effects in response to shadow education across propensity strata. The SM method is similar to propensity score matching estimation, but it differs in that it examines the variation in treatment effects as a function of the propensity strata and summarizes the systematic trend using a hierarchical linear model (Brand and Davis 2011; Brand and Xie 2010). Finally, I will conduct a sensitivity test using the Smoothing-differencing method. Using the Smoothing-differencing method, I can test for sensitivity to the parametric and strata-specific homogeneity assumptions imposed in SM methods (Brand, Pfeffer and Goldrick-Rab 2012; Xie, Brand, and Jann [Forthcoming]). The SD method consists of the three steps: “1) estimate propensity scores for each unit; 2) for each group (the treatment and control group), fit separate nonparametric regressions of the dependent variable on the propensity score 3) take the difference in the nonparametric curves between treated and untreated to obtain the pattern of treatment effect heterogeneity as a function of the propensity score” (Brand, Pfeffer and Goldrick-Rab 2012:14)

These analytical strategies could demonstrate not only a more robust causal relationship between shadow education and SAT achievements than the conventional OLS approaches, but could also provide evidence of treatment effect heterogeneity and its systematic pattern over the propensity strata.

### *Data*

In this analysis, I will use a panel dataset containing extensive information about respondents’ social and economic backgrounds and educational information such as grades, educational activities and aspirations: the Educational Longitudinal Survey of 2002 (ELS).

ELS followed a nationally representative sample of U.S high school sophomores in 2002 through their senior years in 2004 and beyond in 2006. All of my measures come from

the 2006 ELS data (base year through second follow-up). I will restrict the sample to those who, in their senior year of high school, reported that they already took or were planning to take the SAT or ACT in 2004. Also, to see the effect of shadow education on SAT score, I further restrict the sample to those students who had valid SAT or ACT scores.

In this study, I will conduct analysis for each racial/ethnic group of students separately to fully examine racial variation. Most prior research of shadow education suggested substantial differences between races in terms of their social backgrounds, purpose of shadow education and its consequences. Also, we can anticipate a different pattern of treatment effect heterogeneity of shadow education by race from prior studies. Thus, I will use three racial categories: white, black and Asian and separately estimate the causal effects of shadow education and treatment effect heterogeneity to understand racial variation.

### *Measurements*

In this study, I will distinguish public test preparation sources that include books, computer software, videos and high school courses from private sources such as private courses and private tutors to measure variants of shadow education. Since both datasets have similar questions about test preparation, I can utilize the same strategy for the shadow education variable from both. This operationalization of the shadow education variable is supported by Grodsky (2010) and Alon (2010) regarding the differences between those two in terms of its effectiveness and economic aspects. However, as Buchman et al. pointed out, this definition has some limitations. First, many students use multiple sources from both private and public sections. Second, they pointed out that without more detailed data, this distinction between public and private may well be erroneous since many of public sources may actually come from private sections. Lastly, if we use the dichotomous approach, we

can't measure the specific effect for each activity (Buchmann, Condrón & Roscigno 2010). I generally agree with these arguments. However, since the private resources (private courses and private tutoring) have an apparent economic bar for those with limited economic capital (Buchmann, Condrón & Roscigno 2010; Byun & Park 2012), its implications for stratification are substantially different from those of other resources. Therefore, in the analysis, I will adopt a dichotomous approach and generate three categories for the shadow education variables: 1) those who do not use any shadow education, 2) those who use *only* public resources, and 3) those who use only private resources or those who use private resources with other public resources.

Since I use these three categories for the shadow education variable, I will generate multiple contrasts. The effect of private test preparation is my key interest and thus I will examine 1) the effect of private test preparation vs. anything else (the all-inclusive, or baseline, comparison), 2) private test preparation vs. no test preparation, 3) private test preparation vs. public test preparation. Domingue and Briggs (2009) estimated the effect of private courses on SAT score by using the propensity score matching method. However, since they used only one counterfactual, "anything else," they might underestimate the effect of private sources. The "anything else" category includes those who do not use any test preparation, and those who use public resources and/or private tutoring. Thus, if we consider all test preparation activities to be shadow education, we should think about the presence of complex counterfactuals.

Using various background variables, I will estimate a student's propensity to use private test preparation compared to three counterfactuals. I am going to use a rich array of background factors, including family background, prior educational achievement, educational



expectation/aspiration, relationship with parents, urban/rural, private/public school, AP/IB classes, and so on, to estimate an individual's propensity for shadow education.

Lastly, SAT score is my outcome variable. If an individual only has the ACT scores, I will convert the ACT score to the SAT scale, ranging from 400-1600.

## **Implications**

In this study, I want to analyze the causal effect of shadow education and its heterogeneous treatment effects using the stratification-multilevel and smoothing-differencing models with a large panel dataset: ELS02. As many prior studies have pointed out, shadow education has important implications for educational opportunity and the process of social stratification. Since access to shadow education varies according to social class, it could be a key lever in the social transmission of privilege (Alon 2010). Therefore understanding causal relationships and the patterns of treatment effect heterogeneity could provide significant suggestions for policymakers and stratification scholars. The causal relationship tells us about the efficacy and equity aspects of shadow education such as whether shadow education is *really* helpful to increase SAT scores, while patterns of treatment effect heterogeneity (positive selection vs. negative selection) implicate different responses on the part of policymakers to shadow education. Overall, I hope that this study could reveal the emerging mechanism of social stratification in the United States and also provide some insights that could help to alleviate inequalities in educational opportunity.

## REFERENCE

- Alon, S. 2009. "The evolution of class inequality in higher education." *American Sociological Review* 74(5):731-55.
- . 2010. "Racial differences in test preparation strategies." *Social Forces* 89(2):463-74.
- Alon, S., and M. Tienda. 2007. "Diversity, opportunity, and the shifting meritocracy in higher education." *American Sociological Review* 72(4):487-511.
- Baker, D., and G.K. LeTendre. 2005. *National differences, global similarities: World culture and the future of schooling*: Stanford University Press.
- Baker, D.P., M. Akiba, G.K. LeTendre, and A.W. Wiseman. 2001. "Worldwide shadow education: Outside-school learning, institutional quality of schooling, and cross-national mathematics achievement." *Educational Evaluation and Policy Analysis* 23(1):1-17.
- Brand, Pfeffer, and Goldrick-Rab. 2012. "Interpreting community college effects in the presence of heterogeneity and complex counterfactuals." *CCPR Working paper*.
- Brand, J.E., and D. Davis. 2011. "The impact of college education on fertility: Evidence for heterogeneous effects." *Demography*:1-25.
- Brand, J.E., and Y. Xie. 2010. "Who Benefits Most from College?" *American Sociological Review* 75(2):273-302.
- Brand, Jennie E., and Juli Simon-Thomas. [Forthcoming]. "Causal Effect Heterogeneity." in *Handbook of Causal Analysis for Social Research*, edited by Stephen L. Morgan: Springer Series.
- Bray, M. 2006. "Private supplementary tutoring: Comparative perspectives on patterns and implications." *Compare* 36(4):515-30.
- Bray, M., Institut international de planification de l'éducation, and Unesco. 2009. *Confronting the shadow education system: What government policies for what private tutoring?*: United Nations Educational, Scientific and Cultural Organization.
- Bray, TM. 1999. "The shadow education system: private tutoring and its implications for planners."
- Buchmann, C. 2002. "Getting ahead in Kenya: Social capital, shadow education, and achievement." *Research in Sociology of Education* 13:135-61.
- Buchmann, C., D.J. Condrón, and V.J. Roscigno. 2010. "Shadow education, American style: Test preparation, the SAT and college enrollment." *Social Forces* 89(2):435-61.
- Byun, S. [Forthcoming]. "'Shadow Education and School Success in South Korea'." in *Korean Education in Economic and Demographic Contexts*, edited by H. Park and K. Kim: Springer Publishing.
- Byun, S., and H. Park. 2012. "The Academic Success of East Asian American Youth." *Sociology of Education* 85(1):40-60.
- Cheo, R., and E. Quah. 2005. "Mothers, maids and tutors: an empirical evaluation of their effect on children's academic grades in Singapore." *Education Economics* 13(3):269-85.
- Dang, H.A. 2007. "The determinants and impact of private tutoring classes in Vietnam." *Economics of Education Review* 26(6):683-98.
- Dang, H.A., and F.H. Rogers. 2008. *How to interpret the growing phenomenon of private tutoring: human capital deepening, inequality increasing, or waste of resources?*: The World Bank.

- Davies, S., and J. Aurini. 2006. "The franchising of private tutoring: a view from Canada." *Phi Delta Kappan* 88(2):123.
- Domingue, B., and D.C. Briggs. 2009. "Using linear regression and propensity score matching to estimate the effect of coaching on the SAT." *Multiple Linear Regression Viewpoints* 35(1):12-29.
- Grodsky, E. 2010. "Commentaries: Learning in the Shadows and in the Light of Day." *Social Forces* 89(2):475-81.
- Grodsky, E., J.R. Warren, and E. Felts. 2008. "Testing and social stratification in American education." *Annu. Rev. Sociol* 34:385-404.
- Lavy, V., and A. Schlosser. 2004. "Targeted remedial education for under-performing teenagers: costs and benefits." National Bureau of Economic Research.
- Lawrence, I., G. Rigol, T. Van Essen, and C. Jackson. 2004. "A historical perspective on the content of the SAT." *Rethinking the SAT: the future of standardized testing in university admissions*:57-74.
- Mischo, C., and L. Haag. 2002. "Expansion and effectiveness of private tutoring." *European Journal of Psychology of Education* 17(3):263-73.
- Mori, I., and D. 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(1):36-48.
- Rosenbaum, P.R., and D.B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70(1):41-55.
- Schneider, B., and American Educational Research Association. 2007. *Estimating causal effects: Using experimental and observational designs*: American Educational Research Association Washington, DC.
- Stevenson, D.L., and D.P. Baker. 1992. "Shadow education and allocation in formal schooling: Transition to university in Japan." *American Journal of Sociology*:1639-57.
- Tansel, A., and F. Bircan Bodur. 2005. "Effect of private tutoring on university entrance examination performance in Turkey." *IZA Discussion Paper No. 1609*.