

THE EFFECT OF EXTRA SCHOOL FUNDING ON STUDENTS' ACADEMIC ACHIEVEMENTS UNDER A CENTRALIZED SCHOOL FINANCING SYSTEM

Hosung Sohn

School of Public Service
Chung-Ang University
Seoul, South Korea
hsohn@cau.ac.kr

Heeran Park

Institute of Public Policy and
Administration
Chung-Ang University
Seoul, South Korea
walug71@cau.ac.kr

Haeil Jung

(corresponding author)
Department of Public
Administration
Korea University
Seoul, South Korea
hijk@korea.ac.kr

Abstract

This paper analyzes the effect of providing extra school funding on student achievement under the homogenous school funding system in South Korea. This study exploits an administrative cutoff rule that determines the provision of school funding and uses a regression discontinuity design to identify a causal impact of extra school funding. The analysis finds that a 20 percent increase in per pupil funding for underperforming schools reduced the number of below-average students in mathematics, English, social studies, and science by 19.7 percent, 17.0 percent, 16.1 percent, and 18.1 percent compared with the control-side means. The research findings suggest that additional funding for underperforming schools to promote vertical equity would improve students' academic outcomes if it is distributed directly to underperforming schools and used to provide new academic programs to students.

https://doi.org/10.1162/edfp_a_00375

© 2022 Association for Education Finance and Policy

1. INTRODUCTION

One of the main goals of school funding policy is to improve students' academic performance. Two distinct approaches in school funding policy are used to achieve this goal. One way is encouraging competition among schools or school districts in connection to funding levels, while the other way is providing more funding for underperforming schools or schools with economically disadvantaged students (McGuinn 2012). In the past decades, economic inequality seems to have exacerbated the disparity of educational performance and opportunity between students in poor and affluent families. Indeed, students with disadvantaged backgrounds face many barriers to performing as well in school compared to students from a more advantaged background, further amplifying socioeconomic inequality in society (Berliner 2005; Condron 2011; Dragoset et al. 2019). Therefore, attention has been focused on school funding to help underperforming schools and students from disadvantaged backgrounds (Jackson, Johnson, and Persico 2016).

School funding for underperforming schools is designed to provide quality education to disadvantaged and underperforming students. Socioeconomic status or individual characteristics such as family background or ethnicity should not be an obstacle to receiving a quality education. Further, school funding for underperforming schools creates a shorter ladder of social mobility for students in poor neighborhoods (Marks, Cresswell, and Ainley 2006; Condron 2011). A holistic strategy to organize resources in schools and share leadership among stakeholders, including schools, students, families, and community-based organizations, seems to be critical (Johnston et al. 2020). Nonetheless, providing grants without any competition may suppress motivation for schools to engage in innovative reform or creative teaching methods. For example, additional subsidies for teachers, which were not designed as an incentive scheme, failed to enhance teaching skills (Leuven et al. 2007; van der Klaauw 2008). Thus, it is essential to understand whether and why additional school funding for underperforming schools is effective in improving students' academic achievements.

The academic achievement gap among students has been a serious issue not only in Western countries, including the United States, but also outside of Western countries (Chmielewski 2019). Particularly, the achievement gap among elementary and middle school students in South Korea has worsened in recent decades (Byun and Kim 2010; Bae and Wickrama 2015; Choi and Park 2016). South Korea has been one of the Asian countries (including China, Singapore, Japan, and Hong Kong) that have sustained world-class academic performance by encouraging competition among students and schools (OECD 2018). However, the distribution of students' academic performance indicates widening inequality between high and low performers in recent years (You 2015; Lee and Ku 2019). Because elementary and middle schools in South Korea have been funded and tightly controlled by the central government, the school funding per pupil in poor neighborhoods is no less than in wealthy neighborhoods (Ryu 2013; Yang 2012). Additionally, teachers' ability is quite homogenous in South Korean public schools, on average, because public school teachers rotate among schools. Therefore, it seems that students' family resources and backgrounds are mostly responsible for underperforming students and schools. Thus, given the homogenous school funding structure in South Korea, the additional funding to underperforming schools is more about securing vertical equity.

Specifically, this study focuses on school funding policies for underperforming schools and their effectiveness in improving students' academic achievements in those schools. Therefore, this study examines indirectly whether additional school funding to underperforming schools makes up for students' academic shortcomings that are driven by family resources and backgrounds.

By taking advantage of a cutoff rule based on each school's average share of underperforming students in the provision of school funding in South Korea, this study uses regression discontinuity design (RDD).¹ In our analysis, we assume that schools falling right below or above the cutoff point are similar, on average, in observed and unobserved characteristics of students and schools. In other words, the cutoff rule in school funding is assumed to mimic random assignment by creating the treatment group of schools right above the cutoff point and the untreated schools right below the cutoff point.

Our RDD analysis shows that, except for reading results, most of the effect estimates in the other four subjects (mathematics, English, social studies, and science) are statistically significant at the 5 percent level and considerably large in magnitude. Also, those estimates are robust across different RDD specifications. Specifically, a 20 percent increase in per pupil funding for underperforming schools reduced the share of below-average students in mathematics, English, social studies, and science by 19.7 percent, 17.0 percent, 16.1 percent, and 18.1 percent compared with the control-side means. Moreover, by analyzing how the additional school funding is used post-treatment, we find that the funding was used for operating summer and after-school programs, as well as utilizing outside resources such as hiring college students as tutors. Hence, we argue that improvement in student achievement is driven mainly by such factors.

The rest of this paper proceeds as follows. The next section covers previous studies on school funding for underperforming schools. Section 3 explains the institutional background of our study. Section 4 describes the empirical strategy, the RDD using a natural experiment in South Korea, and the data. Section 5 provides the empirical results, including tests of identifying assumptions of the RDD analysis, tests of effect estimates, falsification tests, and robustness checks. The last section explains the results and discusses the study's implications.

2. PREVIOUS STUDIES

A large volume of research has been conducted to examine the effects of school finance reforms (SFRs), policies that provide additional funding to low-performing schools or schools with economically disadvantaged students. Many previous studies conducted evaluations of SFRs in the United States (e.g., Jackson, Johnson, and Persico 2016; Lafortune, Rothstein, and Schanzenbach 2018; Kreisman and Steinberg 2019). State supreme courts have overturned school finance systems in 28 states since 1971 due to increased demands for equality in school spending. The court-ordered SFRs led to legislative reforms aimed at reducing disparities in resources and funding across school

1. The Ministry of Education calculates the share of students whose achievement level is below basic (the share of underperforming students) for each of the five subjects: mathematics, reading, English, social studies, and science. According to a cutoff rule by the Ministry of Education, underperforming schools are those whose average share of underperforming students is equal to 5 percent or more.

districts between 1971 and 2010. Since 1990, funding has focused more on providing low-income districts with additional funding (Jackson, Johnson, and Persico 2016). It was a national policy effort to reduce a gap in school spending between wealthy and poor districts by increasing funding to disadvantaged schools or school districts (Jackson, Johnson, and Persico 2016; Lafortune, Rothstein, and Schanzenbach 2018).

Card and Payne (2002) found that SFRs increased funding to poorer districts and contributed to equalization in test score outcomes as well as in spending across districts. That study gauged the impact of policy using an instrumental variable to deal with endogeneity driven by unobserved characteristics that determine the decision to increase school spending, such as factors that led to a test scores gap. Lafortune, Rothstein, and Schanzenbach (2018) also concluded that SFRs led to enhancing the academic performance of students in low-income school districts. Jackson, Johnson, and Persico (2016) examined the effect of increased spending induced by SFRs on long-term outcomes, including adult income. They found that extended spending enhanced wages, family income, and educational attainment. Importantly, the magnitude of effects for students from low-income families was larger than for those from wealthier families.

Several studies have investigated state school funding programs aligned with SFRs. Papke (2005) and Roy (2011) estimated the effect of Michigan's school finance equalization program, which aimed to increase funding for the least-funded districts. They found that the program not only allowed budgets to be equal within the state but also improved student performance in low-spending school districts. Papke (2005) found that increased funding had an impact on mathematics test pass rates, especially among schools with initially poor performance. On the other hand, Roy (2011) pointed out a negative effect on student performance in the highest-spending districts. Guryan (2001) examined the aid formula provided by the Massachusetts Education Reform Act of 1993 that intended to equalize funding for public schools across districts. It concluded that expanded funding to low-income districts increased their students' academic performance significantly. Van der Klaauw (2008) and Matsudaira, Hosek, and Walsh (2012) studied education policy under Title I, which is the federal law designed to provide more education funding to students from low-income households in the United States. They found that there was no impact on test scores at the individual level, as well as at the school level, when they used RDD.

Some studies investigated the impact of school funding policy by providing grants through competitive settings under eligibility criteria including poor academic performances, which usually comes with requesting strong reform plans and tight oversights. Goe (2006) evaluated the impact of California's Immediate Intervention/Underperforming Schools Program, which additionally offered financial supports to low-performing schools that were chosen through the application process. Goe concluded that schools receiving increased funding failed to enhance students' academic achievement. Carlson and Lavertu (2018) studied the effect of the federal school improvement grant (SIG), which was authorized under Title I of the Elementary and Secondary Education Act of 1965 in the United States. While the SIG targeted low-income groups, the grants were provided through competitive awards. Using RDD, Carlson and Lavertu found that Ohio's SIG program contributed significantly to students' academic achievement. In contrast to the results of Carlson and Lavertu's study of Ohio's

SIG program, however, Dragoset et al. (2019) found no impact of the SIG program on students' outcomes in about 460 schools from 50 school districts in 21 states. More recently, the Race to the Top in the United States, one of the federal education reforms for public schools, targeted a low-performance group but, at the same time, was executed in a competitive manner (McGuinn 2012).² Dragoset et al. (2016) found no clear effect of Race to the Top on students' outcomes.

Researchers in European countries have also conducted evaluations of education policies designed to increase funding for low-performing schools or schools with economically disadvantaged students. Ooghe (2011) studied education policy implemented in Flanders, Belgium, that provided extra personnel subsidies to schools based on the students' family background and found that this policy yielded positive effects in enhancing the academic scores of students from low-income families. By contrast, Bénabou, Kramarz, and Prost (2009) found no effect of the French Zones d'Education Prioritaire program, which provided extra funding to schools in disadvantaged areas based on student outcomes. Leuven et al. (2007) also showed that the additional funding designed to provide a grant to a disadvantaged group for primary schools in the Netherlands failed to enhance not only pupils' academic scores but also teachers' motivation.

In summary, it seems that previous studies provided inconsistent findings depending on the existing school funding system; how the new funding program was designed, distributed, and used under the local contexts; and the groups targeted by the new funding. As shown earlier, some programs provided funding under competitive settings. Also, there were almost no studies about school funding for underperforming schools or schools with economically disadvantaged students in Asian countries; studies were mostly conducted in Western countries and investigated education funding programs that were designed to narrow a funding gap across schools and school districts.

Moreover, the target of the additional funding was broadly set to be underfunded school districts or school districts with high concentrations of poverty. Because schools in such districts usually consisted of students from low-income families or families with limited educational resources, their research implicitly tested whether additional school funding to underperforming schools or school districts helps them overcome their deficits not only in school spending but also in family resources. As mentioned earlier, school funding and teachers' ability in South Korea are homogenous among elementary schools. Therefore, assuming the same existing resources per pupil in all schools, this study examines whether additional school funding to schools with underperforming students makes up for students' academic shortcomings, which are possibly driven by family resources and background.

3. INSTITUTIONAL BACKGROUND

Public education in South Korea is divided into three levels: six years of elementary school, three years of middle school, and three years of high school. Following the

2. Race to the Top is a competitive grant program intended to reward states that are "creating conditions for innovation and reform" instead of executing sanctions. It aims at changing to a federal education system from a decentralized education system. One of the goals is to turn around lowest-achieving schools by narrowing race- and income-based achievement gaps.

principle of autonomy of education adopted in Korea, local education is financed by a special account for educational expenses, which is established under the Local Education Government Act. It is separate from general local finance. The revenue source includes transfers from the central government (local education subsidies and subsidies from the treasury), transfers from local governments (local education tax, tobacco consumption tax, and city and province taxes), and independent revenue sources such as tuition and admission fees.

The school year relevant for this study began in March 2009, and the Ministry of Education in South Korea conducted a nationwide assessment of the educational achievement of students at every educational level in October 2009. The purpose of the assessment was to evaluate the performance of every student and examine whether variations in achievement exist across schools. Assessments were conducted in reading, mathematics, English, social studies, and science, though the assessment area varied by educational level and school years. Every sixth-grade student in elementary school—the focus of our study—was evaluated in all five subjects in October 2009 and July 2010. After the test, every student received a scale score based on their test performance, and each student was given one of four achievement levels for each subject: outstanding, average, basic, and below basic. For the outcome analysis, we define “below average” students as those whose achievement level is basic or below basic. For the assignment variable used for the RDD, we define “underperforming” students as those whose achievement level is below basic.

Based on the 2009 assessment, the Ministry of Education identified underperforming schools based on students’ test performance. To be more specific, the Ministry of Education calculated the share of students whose achievement level is below basic for each of the five subjects, and in March 2010 the Ministry identified schools whose average share was equal to 0.05 or above and determined that they were underperforming. The major purpose of the assessment was to help such schools promote student achievement and reduce educational inequality among schools in Korea. Accordingly, the Ministry of Education provided school funding to underperforming schools on the premise that these schools need assistance.

Figure 1 shows the simple correlation between the per pupil funding level and the share of students below average (binned scatterplot). As the figure suggests, there is a strong downward-sloping relationship between the two variables. Note that schools that received funding were required to use the money solely to promote student achievement. Unfortunately, there is no comprehensive report about how the extra funding for underperforming schools was used in each school. However, according to the Ministry of Education, the extra funding was given directly to underperforming schools (Ministry of Education, Science, and Technology 2010). Also, while there is no official government report on how schools spent their funding, anecdotal stories suggest that schools provided after-school and summer school programs and individual tutoring to students by hiring temporary teachers such as college students. In this study, we tested whether the funding was used for such factors using the survey data administered to principals. Using the post-treatment school-level data, we find that the shares of schools operating summer school and after-school programs were statistically and practically higher for the treated schools. Therefore, it appears that the extra school funding to

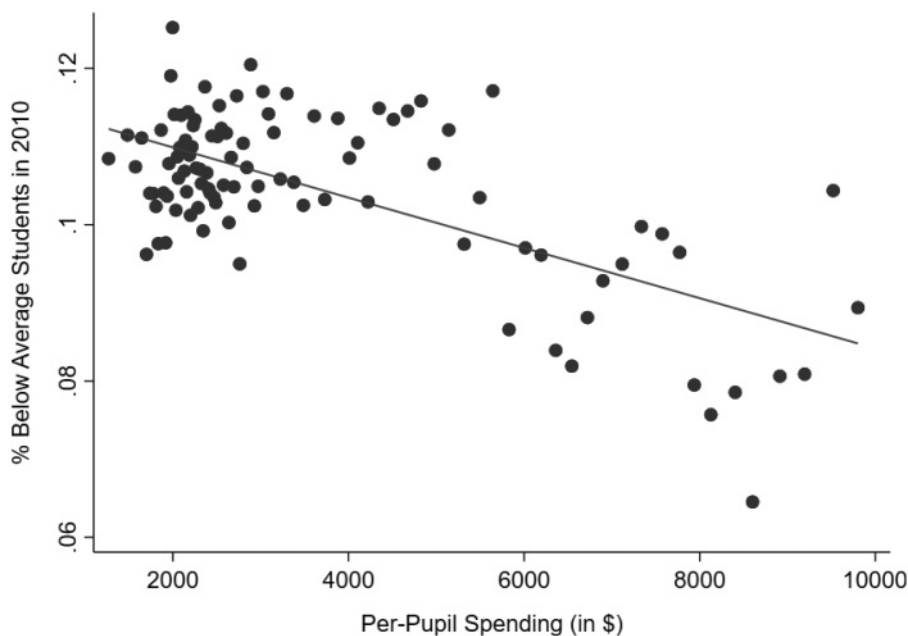


Figure 1. Per Pupil Spending vs. % Below Average (Binned Scatterplot)

underperforming schools was directly and properly used to improve students' academic outcomes.

4. METHODS AND DATA

Empirical Strategy

The provision of school funding (treatment) is a discontinuous function of each school's average share of students whose achievement level is below basic (henceforth, underperforming). Therefore, the setting allows us to use RDD that is favorable—under certain conditions—for securing the internal validity of research findings. Our treatment variable is the following indicator variable (D_s):

$$D_s = 1\{X_s \geq 0.05\}.$$

X_s indicates the average share of underperforming students in five subjects for school s . Hence, the treatment variable takes the value equal to 1 when the average share is equal to or greater than 0.05, and equal to 0 when the average share is less than 0.05.

The literature on RDD proposes estimating the functions nonparametrically using either global or local polynomial regression (Lee and Lemieux 2010); but local polynomial regression estimators are widely used in practice, as they are shown to provide a consistent estimator for treatment effects in the context of RDD (Hahn, Todd, and van der Klaauw 2001; Imbens and Lemieux 2008). To estimate the treatment effects, we follow the suggestions from the RDD literature by using local linear regression (i.e., degree of polynomial [p] equal to 1) with the triangle kernel function (Fan and Gijbels 1996; Gelman and Imbens 2019). Note that the results under uniform kernel specification and higher-order polynomials are qualitatively similar.

One important parameter that a researcher should determine is the bandwidth (h) choice. The bandwidth is arguably the most important parameter in the RDD study because it determines the analysis sample used in estimation, and discontinuity estimates are, oftentimes, sensitive to the bandwidth choice. As such, many methods are developed for deriving the optimal bandwidth choice in RDD (e.g., Imbens and Kalyanaraman 2012), but the RDD literature suggests providing discontinuity estimates under various bandwidth choices for the sake of transparency in research findings. We therefore discuss discontinuity estimates derived under the bandwidth choice provided by Imbens and Kalyanaraman (2012), and provide estimates derived under other bandwidth choices, available in a separate online appendix that can be accessed on Education Finance and Policy's Web site at https://doi.org/10.1162/edfp_a_00375.

In sum, using the student-level data, we estimate the following regression model using the observations within $c - h < X_s < c + h$:

$$Y_{is} = \alpha + \tau D_s + \beta(X_s - c) + \gamma(X_s - c)D_s + \varepsilon_{is},$$

where Y_{is} , D_s , $X_s - c$, $(X_s - c)D_s$, and ε_{is} denote a dependent variable, treatment, assignment variable, interaction term between the assignment and treatment variable, and the error term, respectively. The subscripts i and s indicate students and schools. The main outcome variable is a dummy variable indicating whether a student received the “below average” grade. The coefficient of interest is τ , the effect of school funding on Y_{is} .

Another issue is related to statistical inference. Errors are likely correlated within schools and not accounting for such serial correlation leads to underestimating the true standard errors (e.g., Moulton 1986; Lee and Card 2008). To obviate serial correlation issues, we conduct statistical inference using cluster-robust standard errors using the method proposed by Calonico et al. (2017). For the variables that vary at the school level, we conduct statistical inference based on conventional standard errors.

Data

This study uses restricted administrative data on the national assessment of educational achievement in 2009 and 2010. The data are administered by the Ministry of Education, and we obtained the population data from the Ministry by following the formal application process.³ The data consist of three datasets: student test scores (in achievement levels) for each of the five subjects, student survey data, and principal survey data. For estimating the treatment effects, we used student-level achievement levels, and for testing the identifying assumptions of an RDD, we used student- and school-level survey data administered to students and principals.

Table 1 shows student- and school-level descriptive statistics by treatment status and pre- and post-treatment periods (panels A and B). As can be seen from the two panels, the means of many of the student- and school-level baseline covariates are similar between untreated and treated groups when the sample is restricted within the bandwidth of 0.03 and 0.07. Panel C presents student achievements, and it is clear from the panel that the share of below-average students is higher for treated groups in the pretreatment

3. The population data are no longer disclosed. Only the sampled data (2 percent) are available for the purpose of research.

Table 1. Descriptive Statistics

Variables	Pretreatment (year 2009)				Post-treatment (year 2010)			
	Untreated		Treated		Untreated		Treated	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Panel A: Student Characteristics								
% female students	0.48	0.08	0.48	0.13	0.47	0.09	0.47	0.16
% preparing schoolwork	0.37	0.11	0.33	0.15	0.35	0.14	0.33	0.18
% reviewing schoolwork	0.43	0.11	0.41	0.15	0.47	0.14	0.49	0.21
% taking online lectures	0.53	0.11	0.51	0.17	0.64	0.16	0.50	0.22
Avg. no. of family members	3.18	0.22	3.26	0.45	3.17	0.25	3.24	0.37
Panel B: School Characteristics								
% master's degree or higher	0.22	0.15	0.23	0.19	0.23	0.16	0.24	0.19
% newly hired teachers	0.10	0.11	0.13	0.16	0.08	0.10	0.12	0.15
% operating after school	0.92	0.27	0.90	0.30	0.94	0.23	0.98	0.12
% operating summer school	0.81	0.39	0.88	0.33	0.82	0.38	0.94	0.23
% utilizing outside resources	0.63	0.48	0.53	0.50	0.63	0.48	0.72	0.45
% using customized materials	0.84	0.36	0.73	0.44	0.83	0.37	0.92	0.28
Panel C: Achievement (% below average students)								
Reading	0.30	0.09	0.36	0.13	0.28	0.11	0.25	0.17
Mathematics	0.23	0.08	0.29	0.12	0.33	0.13	0.27	0.19
English	0.28	0.10	0.36	0.14	0.27	0.12	0.26	0.18
Social studies	0.41	0.09	0.46	0.13	0.33	0.13	0.27	0.19
Science	0.18	0.07	0.22	0.10	0.21	0.10	0.16	0.14
Panel D: Other Information								
Average number of test takers	182	114	125	100	167	108	114	93
Average per pupil funding (2010)					\$3,122 (1,663)		\$3,761 (2,038)	
Total number of schools	594		196		594		196	

Notes: The numbers in parentheses are standard deviations.

period. The table also includes statistics on other information such as school funding level and the average number of test takers (Panel D). The average per pupil funding is about \$3,500. The total number of schools is 594 for untreated groups and 196 for treated groups. The average number of test takers for the untreated group is 182 and 167 in 2009 and 2010, respectively. For the treated group, the average number of test takers is 125 and 114 in 2009 and 2010, respectively. Note that the share of test takers is almost 100 percent due to the mandatory nature of the exam. Thus, our study is less likely to suffer from attrition bias.

5. EMPIRICAL RESULTS

Tests of Identifying Assumptions

Identification in RDD comes from the assumption that the relationship between the error term and assignment variable does not change discontinuously around the cutoff point on which the treatment turns. One way to ascertain such an assumption is to verify whether schools have imprecise control over the assignment variable (Lee and

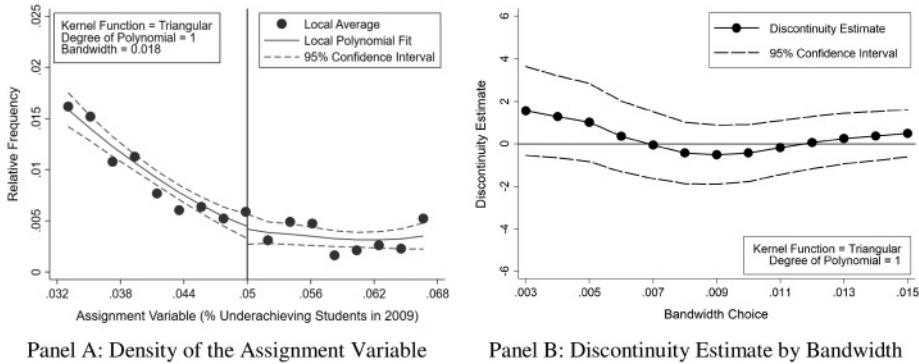


Figure 2. Results of Density Test

Lemieux 2010). If schools can control for the share of underperforming students, it is less likely that the provision of school funding is as good as random near the cutoff point. We present four facts to argue in favor of the assumption. First, the Ministry of Education did not announce the 0.05 cutoff point before the assessment date. Rather, the cutoff point was decided after the grading was done. Second, schools did not grade their students' exams. Each exam that a student took was sent to the Ministry right after the exams concluded. Third, the share of test takers is extremely high (about 99 percent) because of the mandatory nature of the exam, so it is less likely that there exists differential test-taking behavior. Fourth, the range of test scores that determine an achievement level (out of four levels) is decided after the grading. Therefore, a school can't engage in manipulating students' test scores to place them at a certain achievement level.

One way to statistically test for the argument that schools have imprecise control over the assignment variable is to conduct a density test proposed by McCrary (2008) and Cattaneo, Jansson, and Ma (2020). The idea behind the density test is that if schools have less control over the assignment variable, the densities of the assignment variable are smooth across the cutoff point. Panel A of figure 2 shows the densities of the assignment variable. As expected, we do not see any discernible discontinuity in the densities at the cutoff point. To examine the statistical significance of the discontinuity in the densities at the cutoff, we formally derive discontinuity estimates under various bandwidth choices. The results are shown in panel B of figure 2. The horizontal axis indicates the bandwidth choice, and the corresponding discontinuity estimates are displayed on the vertical axis. We also juxtaposed a 95 percent confidence interval to see whether the estimated discontinuities are statistically significant at the 5 percent level. As can be seen from panel B, all of the estimated discontinuities are statistically insignificant at the 5 percent level as the confidence interval encloses the zero horizontal line. We argue, therefore, that manipulation of the assignment variable is less likely given the four facts and density test results.

Another potential threat to identification in our study is that parents may choose to move their children to schools that did or did not receive additional school funding, which may lead to a sample selection issue. While we cannot test for such an issue, we argue that such behavior is less likely. Student mobility rates across schools are low in

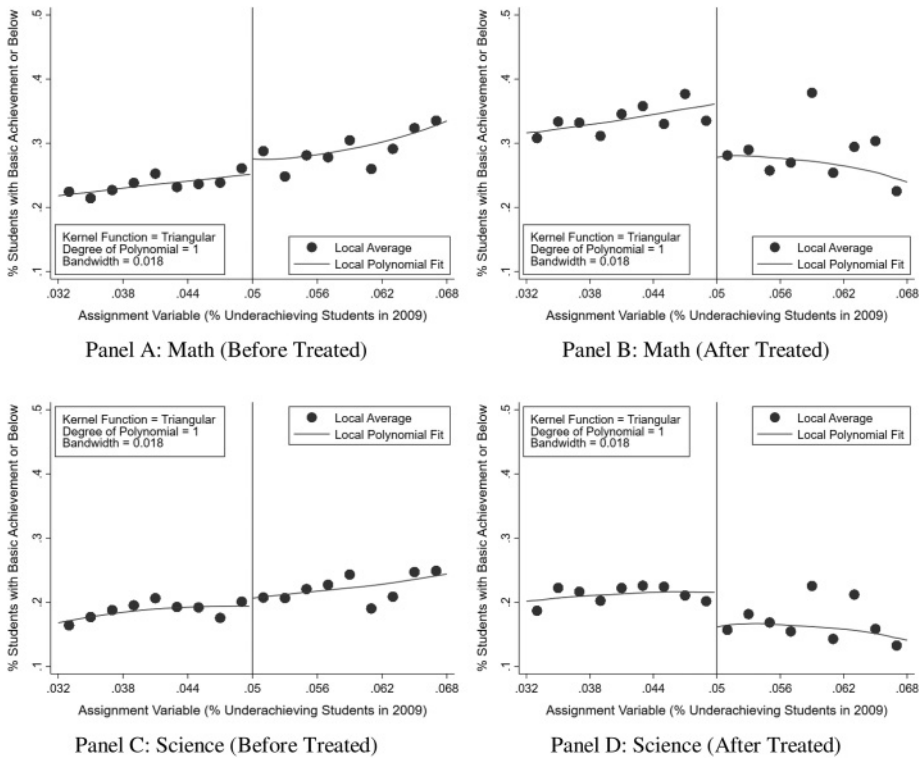


Figure 3. Density of Math and Science Performance Before and After Treatment

Korea, and it is hard to believe that parents would send their children to other schools in the middle of the school period just to benefit from such funding. Moreover, the list of schools that received school funding was not publicized by the Ministry of Education, so the mobility issue is less likely to be salient in our setting.

As a final way to verify our identifying assumptions, we test whether baseline characteristics such as gender, family composition, teacher characteristics, and school characteristics are systematically correlated with the assignment variable, especially near the cutoff point. In the context of an RDD, we should not observe any statistically significant discontinuities in the densities of these variables at the cutoff point. While we cannot test for the differences in unobservable characteristics between the two groups, the prevalence of statistically significant discontinuities in observable covariates may confound our treatment effects. Figure A1, available in the online appendix, shows the densities of the baseline student covariates by the assignment variable. The densities of all the variables are very smooth across the assignment variable, and we do not see any salient discontinuities in these variables at the cutoff point. In figure 3 (panels A and C) and figure 4 (panels A, C, E), we present densities of baseline test performance by the assignment variable. The share of underperforming students is, in general, increasing in the assignment variable without any discernible discontinuities at the cutoff point. Densities of school characteristics (panels A, C, and E in figure A2 and panels A and C in figure A3, available in the online appendix) are very smooth across the assignment variable, and no discernible discontinuities are observed at the 0.05 cutoff point.

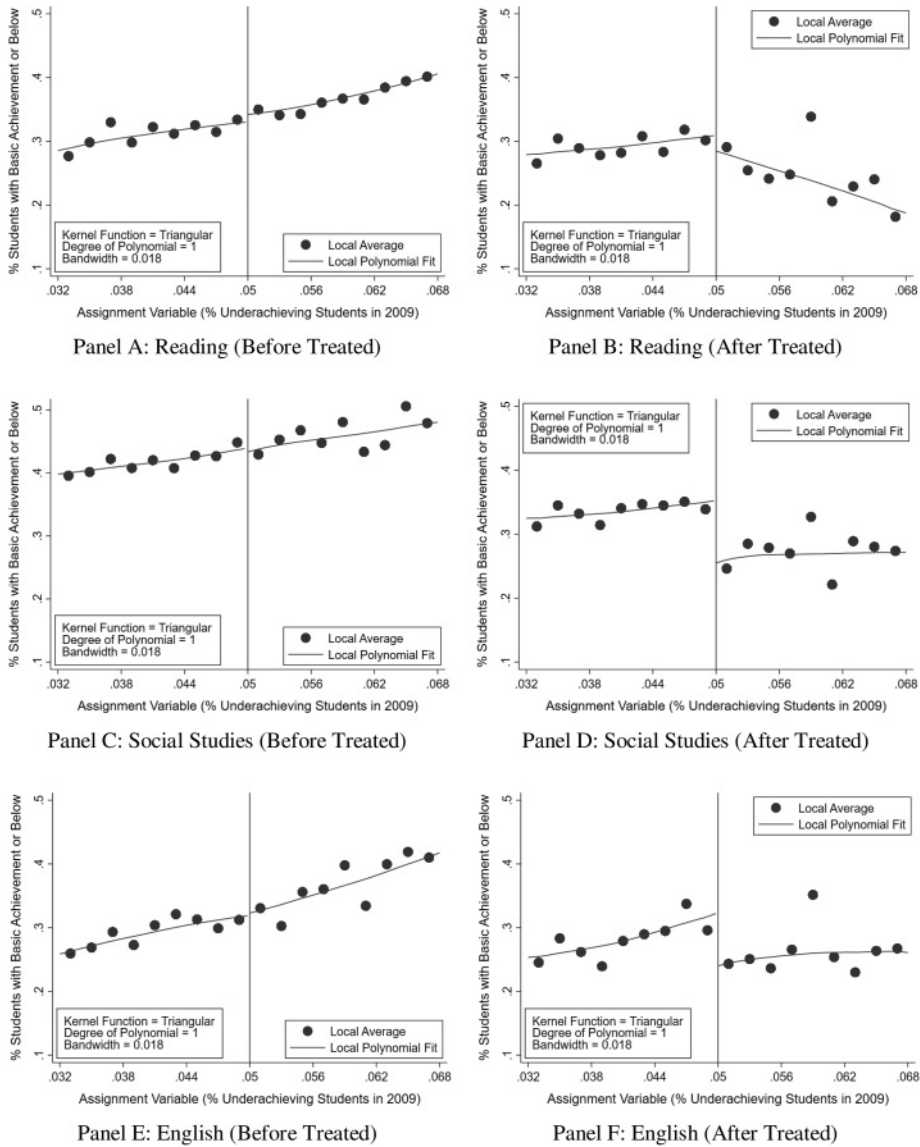


Figure 4. Density of Reading, Social Studies, and English Performance Before and After

Note, however, that it is not the case for the share of schools providing customized instructional materials (panel E in figure A3, available in the online appendix). While we cannot test the statistical significance of the observed discontinuity in the share at the cutoff point from the figure, the difference in the ratio seems to be approximately 0.1. Finally, we tested whether there is a discontinuity in the number of test takers at the cutoff point. The densities are presented in online appendix figure A4. As can be seen from the figure, while the densities are downward trending across the assignment variable, there is no discernible discontinuity at the cutoff point.

In table 2, we present regression discontinuity estimates for student-level baseline covariates. As we mentioned in the empirical strategy subsection, we provide

Table 2. Regression Discontinuity Estimates for Baseline Student-level Covariates

Baseline outcomes	Bandwidth (h)			
	$h = 0.009$	$h = 0.012$	$h = 0.015$	$h = 0.018$
Share of female students	0.007 (0.017) [25,297]	0.004 (0.015) [36,164]	0.007 (0.013) [50,954]	0.008 (0.012) [69,739]
Share of students preparing schoolwork	-0.027 (0.024) [25,186]	-0.025 (0.019) [35,996]	-0.026 (0.016) [50,712]	-0.022 (0.014) [69,408]
Share of students reviewing schoolwork	-0.027 (0.027) [25,190]	-0.026 (0.021) [35,985]	-0.023 (0.018) [50,726]	-0.017 (0.016) [69,423]
Share of students taking online lectures	-0.004 (0.024) [25,131]	-0.001 (0.021) [35,918]	-0.003 (0.018) [50,621]	0.001 (0.016) [69,299]
Number of family members	0.074 (0.051) [25,311]	0.066 (0.042) [36,181]	0.046 (0.036) [50,973]	0.034 (0.032) [69,765]

Notes: Standard errors clustered at the school level are in parentheses. The number of observations is in brackets. Regression discontinuity estimates are derived using local linear specification (i.e., $p = 1$) and the triangular kernel function.

effect estimates under various bandwidth choices. As can be expected from the densities of these covariates, discontinuity estimates at the cutoff point are statistically and practically insignificant, regardless of the bandwidth choice. We also tested whether baseline test performance is significantly different, and table 3 presents the results (“Pretreatment” column). All the discontinuity estimates are statistically and practically insignificant. Table 4 presents discontinuity estimates for baseline school characteristics (“Pretreatment” column). As can be inferred from online appendix figure A3 (panel E), estimated discontinuities are significant only for the share of schools providing customized instructional materials. Though the discontinuities are statistically significant for this variable, we argue that this does not invalidate our identifying assumption because the estimated discontinuities are very small (approximately 0.14). While we cannot test for the differences in other school-level covariates due to data limitations, we further argue that school characteristics are less likely to be significantly different between the two groups given that elementary education is mandatory in Korea and that the Ministry of Education makes a significant effort to homogenize school-level characteristics across the schools. All in all, we proceed with the assumption that school funding is as good as random at the cutoff point, given the results presented in figures 3 and 4, online appendix figures A1 to A3, and tables 2 to 4.

Effect Estimates

To examine whether school funding led to a meaningful increase in student achievement, we first conduct a graphical analysis using regression discontinuity-type graphs presented in figures 3 and 4. The fitted line is derived from local linear specification using the triangular kernel function. The bandwidth used for fitting the line is 0.018. We use this specification across the figures for the sake of consistency. On the left side of the figures, we show densities of test performance using pretreatment test results. We should not observe any significant discontinuities for pretreatment test results, as

Table 3. Regression Discontinuity Estimates for Pre- and Post-treatment Achievement Level

Outcome variables	Effect Estimates	
	Pretreatment	Post-treatment
Reading	-0.021 (0.014) [69,678]	-0.023 (0.022) [63,794]
Mathematics	0.007 (0.013) [69,683]	-0.065** (0.026) [63,779]
English	-0.013 (0.017) [69,684]	-0.046* (0.024) [63,794]
Social studies	-0.003 (0.015) [69,664]	-0.053** (0.025) [63,808]
Science	-0.001 (0.010) [69,667]	-0.038** (0.019) [63,806]

Notes: An outcome variable is a dummy variable that indicates whether a student received the "below average" grade. The effect estimate is derived from a regression of this indicator variable on the treatment variable that varies at the school level. The effect estimate indicates the average difference in the proportion of students scoring "below average" between treated and untreated schools. Regression discontinuity estimates are derived under the bandwidth choice of 0.018. Standard errors clustered at the school level are in parentheses. The number of observations is in brackets. Regression discontinuity estimates are derived using local linear specification (i.e., $p = 1$) and the triangular kernel function. ** and * indicate statistical significance at the 5% and 10% levels, respectively.

there was no treatment. On the right side of the figures, we present densities of test performance using post-treatment test results. Presentation of the pre- and post-treatment densities side by side, we argue, would facilitate the examination of treatment effects.

Panels A and B of figure 3 show densities of mathematics results before and after treatment, respectively. Before the treatment, the densities are increasing smoothly by the assignment variable value, and there does not appear to be a significant discontinuity at the 0.05 cutoff point. After the treatment, however, the densities are smooth up to the cutoff point, followed by a huge drop in the density at the cutoff point, and the densities are quite noisy thereafter. Turning to the science results (panels C and D of figure 3), we see a similar pattern for the pretreatment test results and similar discontinuity for the post-treatment test results. Panels C to F in figure 4 show the same information for social studies and English exams. That is, densities are smooth across the assignment variable for the pretreatment test results. Discontinuities are, however, salient for the post-treatment outcomes.

The pattern of the densities of reading test results is, however, different from the other four subjects. As can be seen from panel A of figure 4, the pattern of the densities for the pretreatment period is similar with other subjects, but that is not true for the post-treatment period (panel B). The densities observed for the treated group are located, in general, at the lower portion of the graph, implying that the average share

Table 4. Regression Discontinuity Estimates for Baseline and Post-treatment School Characteristics

	Effect Estimates	
	Pretreatment	Post-treatment
Share of teachers with master's degree or higher	−0.028 (0.030) [679]	0.043 (0.031) [674]
Share of newly hired teachers	−0.006 (0.026) [676]	0.014 (0.025) [672]
Share of schools operating after school	−0.040 (0.060) [680]	0.047 (0.040) [675]
Share of schools operating summer school	−0.040 (0.054) [680]	0.114** (0.054) [675]
Share of schools utilizing outside resources	−0.104 (0.086) [680]	0.085 (0.082) [675]
Share of schools providing customized materials	−0.143* (0.081) [680]	0.048 (0.066) [675]

Notes: Regression discontinuity estimates under the bandwidth choice of 0.018. Standard errors are in parentheses. The number of observations is in brackets. Regression discontinuity estimates are derived using local linear specification (i.e., $p = 1$) and the triangular kernel function. ** and * indicate statistical significance at the 5% and 10% levels, respectively.

of underperforming students in the treated group is lower than that of the untreated group. Note, however, that we do not see any significant discontinuity near the cutoff point where the treatment effect is identified. In other words, the results imply that, while the average share of underperforming students is lower for the treated group, no difference in the share is observed when the comparison is based on schools around the cutoff point. Hence, it seems that the effect of school funding on reading achievement is insignificant.

Table 3 displays the regression discontinuity estimates. Each student received one of the achievement categories. Our outcome variable is an indicator variable that indicates whether a student received the “below average” grade. We conducted a regression of this indicator variable on the treatment variable that varies at the school level, and so our effect estimate indicates the average difference in the proportion of students scoring “below average” between treated and untreated schools. In the table, we present both the pre- and post-treatment effect estimates. All the discontinuity estimates are derived from student-level data and local linear specification ($p = 1$) using the triangular kernel function. Note that discontinuity estimates under local quadratic specification and other kernel functions are qualitatively similar (available upon request). In the table, we present discontinuity estimates under our preferred bandwidth choice provided by Imbens and Kalyanaraman (2012), which is approximately 0.018. The choice of $h = 0.018$ implies that the discontinuity estimate is derived from comparing untreated schools, whose share of underperforming students is between 0.032 and 0.049, with treated schools, whose share is between 0.050 and 0.067. The discontinuity estimates under other bandwidth choices are available in online appendix table A1.

As can be expected from the graphical analysis in figures 3 and 4, all the estimated discontinuities for pretreatment outcomes are not statistically significant. The magnitude of the effect estimates is also very small, ranging from -0.021 to 0.007 , which indicates that the share of underperforming students is similar near the cutoff point before the treatment. To give an example, the discontinuity observed under the bandwidth choice of 0.018 for mathematics is 0.007 , with a standard error of 0.013 . The effect estimate indicates that the average share of underperforming students in the 2009 exam for treated schools (schools whose share in the 2009 exam is between 0.050 and 0.067) is 0.7 percentage points higher than that of untreated schools (schools whose share in the 2009 exam is between 0.032 and 0.049). The number implies that the share of underperforming students is similar between untreated and treated groups. As mentioned previously, results presented in the “Pretreatment” column of Table 3 are in favor of the identifying assumption of our RDD setting, given that all the discontinuity estimates are statistically and practically insignificant.

The last column of table 3 shows the results of the post-treatment outcome analysis. For reading, the estimated discontinuities are around -0.023 , and the effect estimate is statistically insignificant. The results imply that school funding was not effective in reducing below-average students in reading. Contrary to the reading results, the effect estimates for the other four subjects are statistically significant. Moreover, the magnitude of the estimated discontinuities is practically significant. To be more specific, we find that the share of below-average students in mathematics decreased by 6.5 percentage points. For English, the estimated discontinuity is approximately 4.6 percentage points, though the statistical significance is slightly weaker compared with the other three subjects. The estimated effects on social studies and science are 5.3 and 3.8 percentage points, respectively.

To benchmark our effect estimates in terms of the effect of change in school funding, we estimated the percent change in per pupil funding driven by the policy. Specifically, we conducted an RD analysis using per pupil funding as an outcome variable and estimated the discontinuity in the funding at the cutoff point. The result shows that the change in average per pupil funding at the cutoff point is about 20 percent under the optimal bandwidth choice of about 0.019 . Note that the control-side means for each subject are 0.33 , 0.27 , 0.33 , and 0.21 . Accordingly, the effect of a 20 percent increase in per pupil funding is equivalent to a decrease of 19.7 percent ($= 0.065/0.33$, math), 17.0 percent ($= 0.046/0.27$, English), 16.1 percent ($= 0.053/0.33$, social studies), and 18.1 percent ($= 0.038/0.21$, science) in the proportion of students who received the below-average achievement level.

Relating our effect estimates to previous studies, note that Kreisman and Steinberg (2019) find that a 10 percent increase in expenditures yields about 0.1 standard deviation increase in reading scores and 0.08 increase in math. Papke (2005, 2008) finds that a 10 percent spending increase led to a nearly 4 percentage point increase in the fraction of students scoring as proficient on a fourth-grade math test. Our effect estimates are relatively larger than those found in previous studies, and we argue that such estimates are driven mainly by the three mechanisms. First, schools had a strong reason to focus on reducing the share of students “below basic” in the subsequent year, because focusing resources on low-achieving students was strongly encouraged by the Ministry of Education. Schools that received funding were required by the Ministry to use the

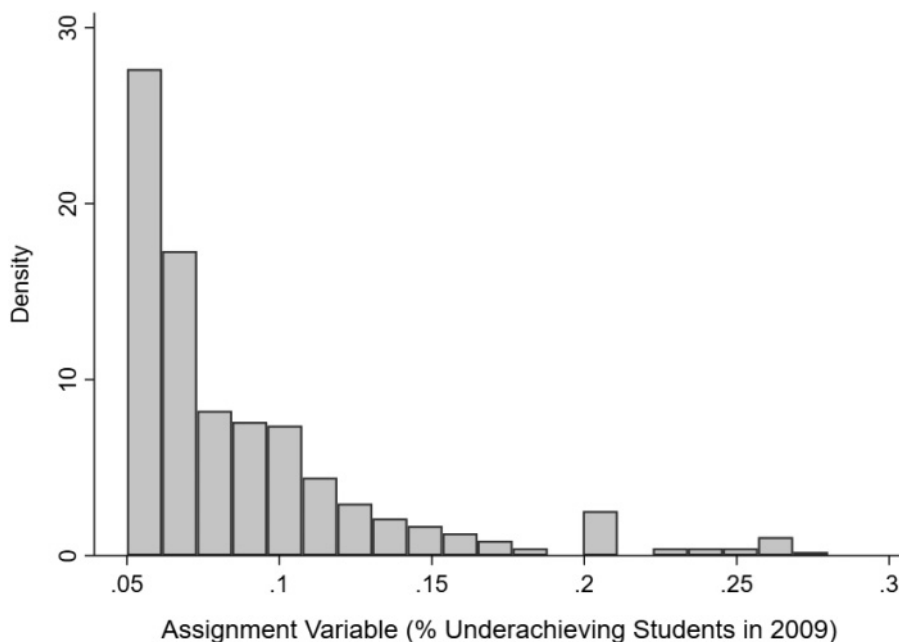


Figure 5. Density of Treated Schools

money solely to promote student achievement. Schools were encouraged to hire teachers, mentors, and operate after-school programs and were refrained from using the funding on things such as buildings, and so forth, which are relatively less effective in promoting student achievement. Thus, it is very likely that schools assigned additional teachers to low-performing students and focused most of their attention on improving the achievement level of these students.

Second, Korea is famous for “education fever,” as evidenced in various sources. Given the competitive educational environment that surrounds Korea’s education system, parents put significant pressure on teachers. Hence, it is very likely that schools put considerable effort into freeing themselves from being labeled as underperforming schools. While the list of underperforming schools was not publicized, the list may have been shared anecdotally by teachers, students, and parents. Given the first and second reasons mentioned above, many of the treated schools might have focused efforts on the lowest performers such as by doing a so-called teaching to the test practice (Lazear 2006). Third, we plotted the density of the distribution of school performance and examined whether there is a lot of density in the performance distribution very close to the cutoff for “below basic.” If there is such density, the small improvement in learning might be enough to lift a large share of students from below basic to above. As can be seen from figure 5, we find that most of the treated schools were clustered around 0.05 and 0.1. Thus, we believe that the relatively large effect estimates observed in our setting were driven by the combined effect of additional funding and the three factors mentioned above.

To examine whether schools focused resources on low-achieving students and whether some of the mechanisms mentioned above are pervasive, we analyzed the effect on other performance margins. Specifically, we estimated the regression

discontinuity effects on the other two achievement categories (i.e., whether a student received an outstanding achievement or average achievement). We find that the reduction in the share of below-average students in mathematics is accompanied mostly by increases in the share scoring the average grade. Also, we find that the reduction in the proportions of below-average students in social studies and science are accompanied mostly by increases in the share scoring the outstanding grade. We argue that the decrease in the share scoring below average being mapping to movements along at least one of the other score thresholds implies plausibly that many of the mechanisms mentioned above are pervasive.

One question that arises from the discussion above is how much of the estimated impact could be due to the increase in school spending per se. Isolating the causal impact of each mechanism mentioned above is difficult because measuring each mechanism reliably is a challenging task. We argue, however, that the estimated impact is driven mostly by the spending for two reasons. First, the national assessment of educational achievement in 2009 was the first nationwide assessment that relaxed informational constraints about students' achievement. Accordingly, such assessment was a wake-up call for many schools. Second, school funding based on this assessment was one of the first that increased spending constraints. That is, schools were able to spend such funding solely on promoting student achievements. Hence, we argue that these two factors interacted to produce the impact of school spending.

Robustness Checks

To further assess the robustness of our estimated results, we conducted falsification tests. The idea behind our falsification tests is that if the discontinuity estimates observed at the 0.05 cutoff point are driven mainly by school funding, we should not observe any discontinuities around cutoffs where there is no variation in the treatment. For example, we should not observe statistically and practically significant discontinuity at the cutoff point, such as 0.02, because school funding is not provided for schools whose share of underperforming students is around 0.02. If the statistically significant discontinuity observed for the 0.02 cutoff point is of a similar magnitude as that of the 0.05 cutoff point, it indicates a serious threat to the internal validity of the results.

Online appendix figure A5 shows the densities of post-treatment outcomes for all the subjects around the false cutoffs. Specifically, we examined the densities around 0.02 and 0.03 cutoff points. As can be seen from all the panels in figure A5, all the densities are smooth across the values of the assignment variable. The estimated discontinuities at the false cutoffs are presented in online appendix table A3. All the estimates are derived under the same specification that we used for the true cutoff point. Panel A displays discontinuity estimates at the 0.02 cutoff point, and panel B shows discontinuity estimates at the 0.03 cutoff point. Notably, all the discontinuity estimates are statistically insignificant. The size of the discontinuities is also very small. While not shown in the paper, discontinuity estimates under the local quadratic specification for English turned out to be statistically significant at the 5 percent level. Note, however, that the magnitude of the estimates is very small (i.e., 0.014). We, therefore, argue that such statistical significance is driven by a small variance in the densities. In conclusion, we infer that the results of the falsification tests support the internal validity of our findings.

Another potential threat in this study is mean reversion. According to Chay, McEwan, and Urquiola (2005), if observable and unobservable factors that drive mean-reversion are continuous at the cutoff point, then the regression discontinuity design will allow for isolating the effect estimates that are free of mean reversion because the design will effectively cancel out the effect of mean-reversion. As a matter of course, we cannot test whether unobservable factors that affect mean-reversion are similar across the cutoff point. We argue, however, that our effect estimates are less likely to be biased by mean reversion because we are using the same schools between the two periods and because many observable school characteristics are continuous at the cutoff point. Moreover, Kane and Staiger (2002) note that the most significant factor that induces mean reversion is the number of students. Because there is no discernible discontinuity in the number of test takers at the cutoff point, we argue that the RDD setting in our context is less likely to be biased by the mean reversion issue. Furthermore, the number of schools is much larger. Thus, we argue that the mean reversion at the school level is less likely to be salient in our context.

6. DISCUSSION AND CONCLUSION

Many studies in the United States and Europe have examined education policies that provided additional funding for underperforming schools or schools with economically disadvantaged students to improve students' academic outcomes. These studies investigated the effect of school funding to narrow a gap in school funding across schools and school districts under a localized school financing system. The findings were inconsistent due to how the new funding was distributed and used under the local contexts, and what groups were targeted by the new funding program. Compared with those previous studies, our study is unique and informative because it examines a new targeted funding program that directly financed underperforming schools for academic programs on specific subjects in a homogenous school funding context. Specifically, taking advantage of the cutoff rule of the provision of school funding to underperforming schools in South Korea, this study conducts an RDD analysis and consistently estimates the impact of school funding on sixth-grade elementary students in five subjects: reading, mathematics, English, social studies, and science.

The results show that school funding for underperforming schools was effective in improving students' test outcomes in mathematics, English, social studies, and science, but no significant effect was found in reading. Specifically, a 20 percent increase in school funding decreased the share of below-average students in mathematics, English, social studies, and science by 19.7 percent, 17.0 percent, 16.1 percent, and 18.1 percent, respectively, compared with the control-side means. These results are not only statistically significant but also considerable in magnitude.

From a policy perspective, it would be informative if a researcher could investigate how schools used their funding. Although there is no official information on how each school used the extra funding, anecdotal evidence shows that many schools hired temporary teachers to provide after-school and summer school programs and individual tutoring to students. To test statistically whether the funding was used for such programs, we present the densities of school covariates in online appendix figures A2 and A3. In the figures, panels A, C, and E correspond to the pretreatment densities, while panels B, D, and F correspond to the post-treatment densities. For example, we see

that many of the densities of the share of schools operating after-school programs increased after the treatment (panel F in online appendix figure A2). Also, the shares of schools utilizing outside resources (panel D in figure A3) and using customized instructional materials (panel F in figure A3) increased significantly when compared with the pretreated period. The discontinuity estimates for these school covariates are presented in table 4. As can be seen from the “Pretreatment” column, while most of the discontinuity estimates are statistically and practically insignificant, all the estimates are negative, indicating that the shares of schools operating after-school and summer school programs, utilizing outside resources, and providing customized instructional materials, are higher in the untreated schools. Note, however, that discontinuity estimates all become positive after the treatment, implying that the shares in these factors are higher for treated schools, though the discontinuity estimates are imprecise due to the small sample size and possibly due to large variance in the densities. Nevertheless, it seems that the extra school funding was used properly to improve students’ academic outcomes.⁴

Some of the previous studies of additional funding for underperforming schools argued several reasons the intervention was not effective in improving students’ academic achievements. First, the money might not directly go to the underperforming schools in some of the funding interventions, so that the impact of the interventions was negligible or null (Goe 2006). Second, the additional funding was not used to initiate new academic programs and resources for students (Leuven et al. 2007; van der Klaauw 2008). Third, existing local funding of schools was reduced after the extra school funding, implying crowd-out across school districts (Gordon 2004; Matsudaira, Hosek, and Walsh 2012). Fourth, the additional school funding is relatively small and often not enough to make up the economically disadvantaged students’ limitations from their lower-income family backgrounds and resources (van der Klaauw 2008; Bénabou, Kramarz, and Prost 2009).

The first and second explanations imply that school funding for underperforming schools in our study is found to improve students’ academic outcomes because it is given directly to those targeted schools and used solely to provide new academic programs and resources for student academic improvement. The third and fourth explanations are closely related to how public schools are funded and how much underperforming schools were underfunded. Most public schools in the United States are locally funded by property taxes, and their school spending varies by whether school districts are located in poor or wealthy neighborhoods. Thus, programs focusing on horizontal equity, school finance reform, or school funding for underperforming schools in the United States were designed to equalize uneven funding across school districts but might not be enough to provide extra help for underperforming schools beyond equalization in funding.

On the other hand, public elementary and middle schools in South Korea are funded homogeneously, and funding is controlled closely by the central government. Under

4. Using student-level data, we also tested whether the share of students who participated in after-school programs are higher for the treated schools during the post-treatment period (i.e., 2010). We find that the share of after-school program participants is about 9 percentage points (24 percent) higher in the treated schools and the estimates are statistically significant.

such a policy environment, the additional school funding in our analysis works as extra resources for students' academic programs in underperforming schools. Therefore, our findings suggest that a school funding program to promote vertical equity would improve students' academic outcomes under the policy condition of homogenous school funding. Also, our findings imply that additional funding for underperforming schools would work if it were given directly to those targeted schools and used solely to provide new academic programs and resources for student academic improvement.

Some limitations and caveats exist in our study. First, it is not possible to know whether and how underperforming schools are sustaining their new programs and resources after their extra funding ended. Second, this study is not able to evaluate the long-term impact on students' test scores and their learning behavior, as the administrative data are available only for two years. Third, we must be cautious about generalizing our findings to other contexts such as secondary school students because our analysis is based on sixth-grade elementary students in South Korea. Also, RDD analysis may constrain our results to local interpretation related to the cutoff rule of the school funding provision in our study. Fourth, students' academic outcomes in our study are relatively simple compared with other studies. As explained earlier, after the test in each subject, students were given one of four achievement levels: outstanding, average, basic, and below basic.

Although there are limitations and caveats, this study is meaningful because it is one of the first studies on such policies in Asia, and its policy circumstances, as well as implementations, were different from other previous studies in Western countries. Thus, the findings of this study encourage new studies to examine the impact of school funding on underperforming schools across diverse countries and under different school funding schemes.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A3A2924956). The authors declare that they have no conflict of interest.

REFERENCES

- Bae, Dayoung, and Kandauda A. S. Wickrama. 2015. Family socioeconomic status and academic achievement among Korean adolescents: Linking mechanisms of family processes and adolescents' time use. *Journal of Early Adolescence* 35(7): 1014–1038. 10.1177/0272431614549627
- Bénabou, Roland, Francis Kramarz, and Corinne Prost. 2009. The French zones d'éducation prioritaire: Much ado about nothing? *Economics of Education Review* 28(3): 345–356.
- Berliner, David C. 2006. Our impoverished view of educational research. *Teachers College Record* 108:949–995. 10.1111/j.1467-9620.2006.00682.x
- Byun, Soo-yong, and Kyung-keun Kim. 2010. Educational inequality in South Korea: The widening socioeconomic gap in student achievement. *Research in the Sociology of Education* 17:155–182. 10.1108/S1479-3539(2010)0000017008
- Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. 2017. rdrobust: Software for regression discontinuity designs. *Stata Journal* 17(2): 372–404. 10.1177/1536867X1701700208

- Card, David, and Abigail A. Payne. 2002. School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics* 83:49–82. 10.1016/S0047-2727(00)00177-8
- Carlson, Deven, and Stéphane Lavertu. 2018. School improvement grants in Ohio: Effects on student achievement and school administration. *Educational Evaluation and Policy Analysis* 40(3): 287–315. 10.3102/0162373718760218
- Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2020. Simple local polynomial density estimators. *Journal of the American Statistical Association* 115(531): 1449–1455. 10.1080/01621459.2019.1635480
- Chay, Kenneth Y., Patrick J. McEwan, and Miguel Urquiola. 2005. The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review* 95(4): 1237–1258. 10.1257/0002828054825529
- Chmielewski, Anna K. 2019. The global increase in the socioeconomic achievement gap, 1964 to 2015. *American Sociological Review* 84(3): 517–544. 10.1177/0003122419847165
- 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. 10.1016/j.rssm.2016.01.002
- Condron, Dennis J. 2011. Egalitarianism and educational excellence: Compatible goals for affluent societies? *Educational Researcher* 40(2): 47–55. 10.3102/0013189X11401021
- Dragoset, Lisa, Jaime Thomas, Mariesa Hermann, John Deke, Susanne James-Burdumy, Cheryl Graczewski, Andrea Boyle, Courtney Tanenbaum, Jessica Giffin, Rachel Upton, and Thomas E. Wei. 2016. *Race to the Top: Implementation and relationship to student outcomes*. Washington, DC: National Center for Education Evaluation and Regional Assistance.
- Dragoset, Lisa, Jaime Thomas, Mariesa Hermann, John Deke, Susanne James-Burdumy, and Dara Lee Luca. 2019. The impact of school improvement grants on student outcomes: Findings from a national evaluation using a regression discontinuity design. *Journal of Research on Educational Effectiveness* 12(2): 215–250. 10.1080/19345747.2019.1571654
- Fan, Jianqing, and Irene Gijbels. 1996. *Local polynomial modelling and its applications: Monographs on statistics and applied probability* 66. Vol. 66. Boca Raton, FL: CRC Press.
- Gelman, Andrew, and Guido Imbens. 2019. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics* 37(3): 447–456.
- Goe, Laura. 2006. Evaluating a state-sponsored school improvement program through an improved school finance lens. *Journal of Education Finance* 31(4): 395–419.
- Gordon, Nora. 2004. Do federal grants boost school spending? Evidence from Title I. *Journal of Public Economics* 88(9–10): 1771–1792. 10.1016/j.jpubeco.2003.09.002
- Guryan, Jonathan. 2001. Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts. NBER Working Paper No. 8269.
- Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69(1): 201–209.
- Imbens, Guido, and Karthik Kalyanaraman. 2012. Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies* 79(3): 933–959. 10.1093/restud/rdro43

- Imbens, Guido W., and Thomas Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142(2): 615–635. 10.1016/j.jeconom.2007.05.001
- Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. 2016. The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *Quarterly Journal of Economics* 131(1): 157–218. 10.1093/qje/qjv036
- Johnston, William R., John Engberg, Isaac M. Opper, Lisa Sontag-Padilla, and Lea Xenakis. 2020. *Illustrating the promise of community schools: An assessment of the impact of the New York City Community Schools Initiative*. Research Report No. RR-3245-NYCCEO. RAND Corporation.
- Kane, Thomas J., and Douglas O. Staiger. 2002. The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives* 16(4): 91–114. 10.1257/089533002320950993
- Kreisman, Daniel, and Matthew P. Steinberg. 2019. The effect of increased funding on student achievement: Evidence from Texas's small district adjustment. *Journal of Public Economics* 176:118–141. 10.1016/j.jpubeco.2019.04.003
- Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. 2018. School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics* 10(2): 1–26. 10.1257/app.20160567
- Lazear, Edward P. 2006. Speeding, terrorism, and teaching to the test. *Quarterly Journal of Economics* 121(3): 1029–1061. 10.1162/qjec.121.3.1029
- Lee, David S., and David Card. 2008. Regression discontinuity inference with specification error. *Journal of Econometrics* 142(2): 655–674. 10.1016/j.jeconom.2007.05.003
- Lee, David S., and Thomas Lemieux. 2010. Regression discontinuity designs in economics. *Journal of Economic Literature* 48(2): 281–355. 10.1257/jel.48.2.281
- Lee, Inwha, and Namwook Ku. 2019. Analysis of PISA 2015 reading achievement characteristics of Korean students and influence of educational context variables. *Journal of Reading Research* 50:113–144. 10.17095/JRR.2019.50.4
- Leuven, Edwin, Mikael Lindahl, Hessel Oosterbeek, and Dinand Webbink. 2007. The effect of extra funding for disadvantaged pupils on achievement. *Review of Economics and Statistics* 89(4): 721–736. 10.1162/rest.89.4.721
- Marks, Gary N., John Cresswell, and John Ainley. 2006. Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation* 12(2): 105–128. 10.1080/13803610600587040
- Matsudaira, Jordan D., Adrienne Hosek, and Elias Walsh. 2012. An integrated assessment of the effects of Title I on school behavior, resources, and student achievement. *Economics of Education Review* 31(3): 1–14. 10.1016/j.econedurev.2012.01.002
- McCrary, Justin. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142(2): 698–714. 10.1016/j.jeconom.2007.05.005
- McGuinn, Patrick. 2012. Stimulating reform: Race to the Top, competitive grants and the Obama education agenda. *Educational Policy* 26(1): 136–159. 10.1177/0895904811425911
- Ministry of Education, Science, and Technology. 2010. *Schools in Need of Achievement Improvement Support Program: Master Plan*. Seoul: Republic of Korea.

- Moulton, Brent R. 1986. Random group effects and the precision of regression estimates. *Journal of Econometrics* 32(3): 385–397. 10.1016/0304-4076(86)90021-7
- Ooghe, Erwin. 2011. The impact of “Equal Educational Opportunity” funds: A regression discontinuity design. IZA Discussion Paper No. 5667.
- Organization for Economic Co-Operation and Development (OECD). 2018. *PISA 2018 results: Executive summary*. Paris: OECD.
- Papke, Leslie E. 2005. The effects of spending on test pass rates: Evidence from Michigan. *Journal of Public Economics* 89(5–6): 821–839. 10.1016/j.jpubeco.2004.05.008
- Papke, Leslie E. 2008. The effects of changes in Michigan’s school finance system. *Public Finance Review* 36(4): 456–474. 10.1177/1091142107306287
- Roy, Joydeep. 2011. Impact of school finance reform on resource equalization and academic performance: Evidence from Michigan. *Education Finance and Policy* 6(2): 137–167. 10.1162/EDFP_a_00030
- Ryu, Min Jung. 2013. A study on the reform of financial grants for local education. *Korean Journal of Local Government Studies* 17(3): 315–334.
- Van der Klaauw, Wilbert. 2008. Breaking the link between poverty and low student achievement: An evaluation of Title I. *Journal of Econometrics* 142(2): 731–756. 10.1016/j.jeconom.2007.05.007
- Yang, Won Young. 2012. An analysis of financing in rural schools focused on public elementary and middle schools in Jeollabuk-Do. *KNU Journal of Educational Research* 27:1–22.
- You, Hyesun. 2015. Do schools make a difference?: Exploring school effects mathematics achievement in PISA 2012 using hierarchical linear modeling. *Journal of Educational Evaluation* 28(5): 1301–1327.