



Exploring the alleged effect of lower academic achievement after the free semester in Korean Middle Schools

Hyekyung Jung¹ · Young-Eun Kim² · Insuk Kim²

Received: 9 February 2020 / Revised: 28 July 2020 / Accepted: 28 August 2020 / Published online: 14 September 2020
© Education Research Institute, Seoul National University, Seoul, Korea 2020

Abstract

The Enforcement Decree of the Elementary and Secondary School Act (2015) mandates that from 2016 all middle schools in Korea must administer one semester as a free semester (FS), called an exam-free semester or, alternatively, a free learning semester. Consequently, some Korean parents and students have expressed a feeling of anxiety about having no exams during the FS. This anxiety arises primarily from their concern about academic loss and neglect of study in the absence of exams, with ensuing lower achievement in future tests. There are many studies showing a positive effect of the FS on learning, but little is known about its practical effect on student achievement. Thus, this study explored if the FS affects students' achievement by estimating the average population treatment effect of the FS policy on academic achievement in Seoul schools. Based on the inverse probability of treatment weight (IPTW) approach with existing empirical data, the study found that, overall, no substantial differences were noticed in the academic achievement between the pilot FS schools and the non-FS schools. The findings provide empirical evidence for the effect of the FS on academic achievement, as well as critical information for policy makers.

Keywords Free semester · Achievement · Causal effects · Inverse probability of treatment weight

Introduction

The Enforcement Decree of the Elementary and Secondary School Act (2015.9.15) mandates that from 2016 all middle schools in Korea must administer a free semester (FS) for one semester, called an exam-free semester or, alternatively, a free learning semester. This act specifies that one semester among the first or second year of middle school should be implemented as a FS during which classroom assessment consists of a variety of performance tests instead of regular

paper and pencil tests such as the mid-term and final exams which are normally carried out (Ministry of Education 2015a). The purpose of the FS is to let students explore their aptitude and possible future career, experience the delight of learning, and develop the ability of self-directed learning (Ministry of Education 2015a). This policy is a brake on the time-honored school convention which depends on regular paper and pencil tests in evaluating student achievement.

The policy has, however, initiated a debate about the relationship between exams and student achievement. In Korea, the prevalent belief that academic achievement is identified with test scores underlies the view that there will be lower student achievement under the FS. Some believe that taking exams is the way to confirm what students are learning and students achieve more through having to take tests. Moreover, in the recent past, some Korean parents and students have expressed the view that they feel anxiety about having no exams during the FS (Yeo and Chae 2016). This concern arises primarily from their belief that the absence of exams results in academic loss and neglect of study and ensuing lower achievement when students are in upper grades.

Meanwhile, according to case studies of pilot schools which administered the FS in 2014, students' overall

✉ Young-Eun Kim
ey2015@kice.re.kr

Hyekyung Jung
hjungk@koreatech.ac.kr

Insuk Kim
insukkim@kice.re.kr

¹ Korea University of Technology and Education, 1600, Chungjeol-ro, Byeongcheon-myeon, Dongnam-gu, Cheonan-si, Chungcheongnam-do 31253, South Korea

² Korea Institute for Curriculum & Evaluation, 8, Gyohak-ro, Deoksan-eup, Jincheon-gun, Chungcheongbuk-do 27873, South Korea

satisfaction with their educational experience during the FS was significantly higher than with the usual test-based syllabus (Ministry of Education 2015b). This report shows that teachers make an effort to improve teaching strategies and to adapt the curriculums in their classes for the programs based on the FS activities. Furthermore, there are many studies about positive effects of the FS on self-directed learning, teacher development or career awareness (e.g., Choi et al. 2014; Ministry of Education 2015b; Kim 2017).

Nonetheless, there remain significant concerns about the FS's possible adverse effect on students' academic achievement (Shin and Park 2015). Also, Yeo and Chae (2016) revealed that students experiencing the FS showed contradictory emotions of satisfaction with various activities together with growing anxiety over upcoming tests, and became antipathetic towards the FS policy due to the absence of tests. As such, even if the aim of the FS is for students to explore their future career options and to improve their academic skills, some insist that the semester without exams negatively affects academic achievement in middle school and ends up impacting college admission. This argument is understandable in Korea where parents have overwhelming concerns about college entrance from when their children are at an early age.

There are several empirical studies that examine the impact of the FS on student achievement (Kim and Kang 2017; Cho et al. 2018; Kim 2018b; Kim et al. 2019). However, these studies imply limitations since their analyses were based on longitudinal secondary data (e.g., Kim and Kang 2017; Cho et al. 2018; Kim et al. 2019) or data from a single school (e.g., Kim 2018b).

Thus, this study aims to address if the FS policy has an unintended negative impact on student academic achievement. The uniqueness of the study is to use nationally representative data and attempt to estimate a causal effect with observational data under the causal inference framework (Rosenbaum and Rubin 1983). Given that student-level data were unavailable, this study used data that were publicly available: National Assessment of Educational Achievement

(NAEA). Since the NAEA data consist of a school-level data set, the unit of analysis is a school. Admittedly, the variability across schools regarding the implementation of the FS policy or the cross-level interaction influenced by student characteristics (Burstein 1980) was not able to be addressed in this study due to the data limitation. However, as Stuart (2007) pointed out, estimating a causal effect with the school-level data would be useful in a situation where student-level data are unavailable and the interventions of interest are implemented at a cluster level.

Also, this study explores the effects of the FS on students' academic achievement, which is primarily demonstrated by test scores. In the context of Korean education, a variety of tests strongly affect practices of teaching and learning and people traditionally think of tests as fair and objective. Thus, test scores have been widely used in college admissions and employment (Chang 2011; Kang 2007). As a result, students and parents are sensitive to test scores, particularly at the secondary level, so they try to maintain their relatively higher ranking through private academies (Lee 2011). In this circumstance in which students and parents are likely to become uncomfortable with the absence of tests during the Free Semester, the empirical evidence of this study is substantially meaningful not only to policy makers or educational practitioners but also to the students and parents.

Literature review

The policy of the free semester system

Along with the Enforcement Decree of the Elementary and Secondary Education Act (2015) and the Enforcement Decree of the Career Education Act (2015), the FS became a mandate in middle schools in Korea from the 2016 school year (See Table 1).

Starting with three-year demonstrations of an FS pilot program in schools nationwide, the FS was mandated to all middle schools in Korea from 2016. Starting at the outset

Table 1 The Legal Acts related to the FS Source: Enforcement Decree of the Career Education Act

Article 44 (Semesters)

(3) Among the semesters referred to in paragraph (1), the head of a middle school shall designate one semester as a free learning semester. In such cases, detailed matters concerning the designation of the free learning semester, such as the scope of semesters eligible for designation, shall be determined by the Ministry of Education

Article 48–2 (Methods, etc., of operating lessons during the Free Learning Semester)

(1) The head of a middle school shall operate types of classes encouraging student participation during the free learning semester and curricula for experiential activities for various types of experience, including the career search of each student.

(2) Detailed matters concerning the types of class encouraging student participation and experiential activities under paragraph (1) shall be determined by the Minister of Education.

Article 6 (Operation of Intensive Academic Year of Semester for Career Education)

(3) A superintendent of education may manage a career semester system in connection with, or by integrating it with, the free learning semester prescribed in Article 44 (3) of the Enforcement Decree of the Elementary and Secondary Education Act

with 42 pilot schools in 2013, 811 schools (25%) in 2014 and 80% of middle schools in 2015 participated in applying the FS (Ministry of Education 2015a). Since the full implementation of the FS in 2016, many schools and districts have expanded the FS to the whole school year or introduced it to another grade (e.g., Gyeonggido Office of Education 2018; Jung et al. 2018), which implies public recognition of the positive influence of the FS. The Korean FS was enacted with a political catchphrase from the Park, Geun Hye administration that translates approximately to “Education for happiness with the development of students’ dream and talent”, an idea borrowed from the long established transition year in Ireland focusing on autonomy and choice and from the idea of a gap year in England (Kim 2013; Jung et al. 2018). The purpose of this policy is summarized in three points: raising students’ career awareness and opportunities for related career experience, improving teaching and learning methods, and strengthening curriculum autonomy for individual schools (Ministry of Education 2015a). Thus, students are encouraged to participate in a variety of activities in relation to their future career and learn without the burden of the so-called traditional tests and exams for one semester in middle school. This, nevertheless, does not signify the absence of classroom assessment during this period. Instead, classroom assessment depends entirely on performance assessment.

During this semester, 10 class hours a week are secured for the FS programs by reducing subject class hours and by consolidating creative experiential activity hours; namely, 2 periods daily, usually in the afternoon. The program mainly consists of four categories, i.e., career exploration activity, theme selecting activity, arts and physical activity, and club activity (Ministry of Education 2015a). The Free Learning Semester, a name used at the first introduction that is a literal translation of the policy, emphasizes free exploration and choice of activities.

For example, Table 2 shows that in 2016 ‘A’ middle school implemented a total of 170 h of FS activities during the second semester of the first grade in middle school with 34 h of career exploration, 51 h of theme selection, 68 h of arts & sports, and 17 h of club activity. In general, the school

surveyed students’ needs and preferences and then arranged programs for students to select.

The standards for operating the FS programs are specified in the 2015 Revised Curriculum, i.e., the national curriculum of Korea. It is important to note that the Korean education system, including curriculum and assessment, is controlled by the government (Ministry of Education 2017, p.7). Related to the national curriculum, the 2015 Revised Curriculum especially emphasizes student-centered strategies and key competencies needed for the future, which implies a paradigm shift in terms of curriculum policy (Ministry of Education 2017, p. 30).

Key features of the standards include: the purpose and the period of the FS implementation, the organization of the programs, and methods of student evaluation. In particular, the standards emphasize student participation, collaboration, discussion and project-based learning across the courses (see Table 3 for details).

Kim and Hong (2016) highlighted the main concepts of the FS in the national curriculum by describing its features as “self-directed learning, connection to the community, student participatory instruction, career education, and performance assessment” (p. 7) and pointed out that these do not significantly differ from concepts which have been emphasized through the guidelines up to now. Nonetheless, the FS operation is legally bound to the acts and the national curriculum. In addition, this policy provides a different paradigm of curriculum organization and implementation in middle schools of Korea, for the purpose of promoting students’ participation in learning and providing various opportunities for career exploration and self-directed learning.

Research on the free semester system

To date, research articles and publications on the FS have been steadily increasing. Previous research primarily deals with the definition of the Korean FS, its curriculum implementation, and program development, focusing on career education, student satisfaction, and teaching and learning methods.

Table 2 The FS model of ‘A’ middle school Source: Gwangsan Middle School (2016, p. 104)

	Monday	Tuesday	Wednesday	Thursday	Friday
1	Korean(3), Morals(1), Social Studies(2), Science(2), Tech(2), PE(2)				
2	Music(1), Arts(1), English(3), Career & Job(1), CEA(2) total 23 h				
3				Arts & Sports	
4				Activity	
5		Theme		Career	Arts & Physical
6		Selecting		Exploration	Activity
		Activity		Activity	
7	Creative				Club
	Experiential				Activity
	Activity				

Table 3 The FS guidelines in the 2015 revised curriculum Source: Ministry of Education (2015c, p. 11)

Schools should implement an Exam-Free semester to help students develop self-directed learning skills and attitudes by exploring their aptitudes and career plans as well as experiencing the joy of learning

- a. Middle schools should implement an Exam-Free semester for one semester
- b. During the Exam-Free semester, subjects and creative experiential activities are organized and implemented in accordance with the aims of the Exam-Free semester
- c. During the Exam-Free semester, schools should collaborate with local communities to offer experience-based activities such as career exploration activities, selective theme activities, club activities, arts/physical activities, and so forth
- d. The Exam-Free semester should promote student participatory learning such as collaborative learning, discussion and debate, project-based learning, and so forth
- e. The Exam-Free semester should make use of assessments focused on the process of learning to foster students' growth and learning, and avoid the use of standardized, paper-pencil tests such as midterms and finals
- f. The Exam-Free semester should use resources in and outside the school to support students' career designs and explorations
- g. Schools should make an effort to connect the aims of the Exam-Free semester with other semesters and grades

Cho (2017) analyzed 80 research articles about the FS published from 2013 to 2016 and noted that research focused on improving curriculum & instructional methods, evaluation methods, program development, overall management, and effects of the FS. Hwang et al. (2019) collected the data from 169 research articles published from 2013 to 2018 for the purpose of topic modeling analysis. As a result, the most frequent keywords for FS research were 'program', 'teacher', 'class', 'policy', 'activity', and 'career' in order and the main topics of research were policy implementation, school curriculums, development of teaching & learning methods, student evaluation, program development, and career experience activity (Hwang et al., pp. 308–309).

Regarding the school curriculum of the FS, Kim et al. (2014) searched the curriculum features of 42 pilot schools using document analysis. They found out that the arts & sports priority model was prevalent, and the portion of arts & sports activity was bigger among others. They discussed the implications of the FS in terms of school curriculum autonomy and students' options. Many studies explored what FS programs are being provided and how the FSs are operated. Thus, they analyzed its curriculum contents as well as teaching & learning styles.

Meanwhile, Park (2015) used critical discourse analysis targeting teachers, students, and parents of 10 pilot schools and revealed tensions between various identities surrounding the FS as well as their responses to policy implementation and the roles consequently given to them. This study, using a qualitative research method, has significance in that the main focus is given to the primary FS stakeholders, students and teachers. Shin et al. (2015) surveyed teachers, students, and parents from both pilot schools and non-pilot counterparts. This study specifically takes notice of the school members' understandings and attitudes in addition to the status quo of the FS adoption. Overall, the members showed positive understanding of the FS but a negative attitude toward policy continuity, as well as recognizing the potential of improving teaching & learning and burdens of assessment (p. 49).

Many studies have been undertaken using empirical data to determine the effects of the FS. A notable study is Kim (2017)'s analysis of FS-related outcomes and changes in relation to career maturity, core competencies, and school satisfaction. Students from two schools operating the FS in the second semester of 2016 participated in the study. The study showed that career maturity, social-core competencies, and satisfaction significantly increased after the FS was implemented (Kim 2017, p. 115). The results of this data from the two schools have implications as an empirical study which targets educational outcomes. Jeong et al. (2015b) argued that the students of the FS school were satisfied with their teachers, classes and overall school life following the accompanying curriculum transformation and various experiential activities. Kim (2018a) indicated that the multicultural education program as a theme selection activity in certain schools was effective in raising intercultural sensitivity. In addition, Lee (2016) focused on the changes at a pilot school in terms of both structural and psychological aspects. This study noticed that the school members developed comparatively positive attitudes towards the FS and the classroom changes featuring student-centered and cooperative learning. The data in these studies, however, are likely to be limited in the sense of its small size. While there are numerous studies based on one or a few schools, there are few studies into the educational impact of the FS that use large scale data.

Some studies tried to suggest gender differences from the effects of the FS. However, Jeong et al. (2015b) found no significant data representing gender differences regarding school satisfaction and Hyun (2017) could not highlight gender differences in creativity and career orientation. Meanwhile, gender differences were detected in certain subject studies. In the study of Hwang and Yoo (2016), boys in physical education classes were more satisfied with class time, methods of instruction, relationships with peers, and career exploration than girls. Kim and Lee (2019) discovered that there were gender differences in English class with respect to class satisfaction, participation, academic

improvement, self-directed learning, and understanding on private education.

The FS pilot schools reported growing satisfaction of students, teachers, and parents as well as an improvement in teaching & learning (Ministry of Education 2015b; Ministry of Education 2015c). It is probable that the students were able to choose programs according to their preference and experience a range of activities. At the outset, due to a lack of systematic support and quality programs, schools had to undergo trials. Consequently Jeong et al. (2015a) have raised questions about the FS in relation to the quality of student-centered teaching & learning, curriculum integration, student needs-based programs, and community supported activities.

Shin & Park (2015) list both positive and negative sides. One issue is the concern about academic ability, which is related to both sides. Whereas the positive argues that students can experience authentic learning through quality activities, the negative shows concerns about students' neglect of academic achievement (Shin and Park 2015, p. 312). Indeed a case study of three schools revealed that some students recognized the FS as "an idle semester" (Shin and Park 2015, p. 326). In addition, no exams and tests are likely to cause parents' anxiety over the loss of academic achievement. According to Shin et al. (2018), despite improvement in student autonomy and career awareness as well as student-centered teaching, parents and students expressed concern about being left behind academically and considered the free year to be a term without studying.

In fact, several studies analyzed the impact of the FS on academic achievement. Kim et al. (2019) highlighted academic achievement as the change after participating in the FS. This study, utilizing data from the 2013 Korean Education Longitudinal Study, explored the changes in achievement from participation in the FS and if students' SES affected academic achievement. They found that the students having participated in the FS showed higher growth in achievement in Korean, mathematics, and English compared to their counterparts, and with the exception of mathematics a still higher achievement in Korean and English in their second grade (p. 33). According to Kim and Kang (2017), who used the same data as above, the FS participation schools showed a bigger degree of academic improvement in Korean, mathematics, and English than non-participation schools. Kim (2018b) studied changes of academic records responding to satisfaction with the FS. He found that as the satisfaction increased, the achievement grew. The FS has a significantly positive effect on the low SES group. However, this study utilized data from the school records of different grade levels in one school. Cho et al. (2018) argued that unlike prevalent concerns about low achievement and high reliance on private academies during the FS, student achievement in Korean and mathematics grew significantly,

and the hours of private academy attendance did not change after participating in the FS based on the analysis from the Gyeonggi Longitudinal Study (p. 62).

In spite of active affirmation of the FS in areas such as reflective teaching, teacher-initiated curriculum adaptation, and student motivation, concerns about not having exams and consequent lower achievement have not been dispelled, even with the comprehensive implementation of the program all over the country. Also, as mentioned in the introduction, the previous studies that showed some positive effects of the FS on student achievement (Cho et al. 2018; Kim and Kang 2017; Kim 2018b; Kim et al. 2019) have limitations, particularly since the estimates capturing the effect of the FS based on the regression-based analysis with secondary data (i.e., multiple regression, two-level hierarchical analysis or t-test with sub-groups based on the SES levels) might be still confounded with selection bias. Thus, this study aims to add scientific evidence to the literature of student achievement after the implementation of the FS with an observational data by incorporating a propensity score weight strategy under the potential outcomes framework, which can help reduce selection bias with observational data, and is discussed in the following section.

Methods

In this study, the parameter of interest, i.e., estimand, is the average treatment effect (ATE) under the potential outcomes framework (Rosenbaum and Rubin 1983; Holland 1986). In general, the ATE targets the entire population and addresses the expected "mean value among all students in the population of these what-if differences in test scores" (Morgan and Winship 2007, p. 37). In this regard, the ATE estimates would provide valuable information to the public as well as educational policy makers and educators, by determining if, and to what extent, the students who experienced the FS, on average, underperform in academic achievement in comparison with those have not experienced the FS.

Given that the FS is not randomly assigned to schools, however, self-selection bias should be accounted for in order to draw a valid and reliable statistical inference about the impact of the FS policy on academic achievement. For this purpose, a propensity score-based approach was employed. The propensity score, $e(x)$, can be defined as the conditional probability of receiving the treatment based on pretreatment covariates (x) (Rosenbaum and Rubin 1983). This study took two analytical steps, based on the potential outcomes framework (Holland 1986), for the estimation of the impact of the FS. The first step is related to the treatment assignment and possibility of selection bias. Propensity scores were estimated based on a logit model for predicting the treatment indicator informing whether a school joined the FS or

not with observed key covariates as shown in Table 4. With respect to ways of improving the balance of pretreatment covariates with estimated propensity scores, this study used a weighting approach because the number of non-FS schools may not be sufficient for matching to construct comparable comparison pairs for individual FS schools. More importantly, as we used population data rather than sample data from a secondary data source (i.e., NAEA and ASA data), we wanted to keep all the FS schools in the analysis as much as possible, and so evaluate the impact of the FS for the target population of interest. To evaluate the group equivalence in pretreatment covariates, we examined the distributions of key covariates and the estimated propensity scores of the FS schools and non-FS schools with the original data and propensity score weighted data (see Table 5).

Second, based on the estimated propensity scores, several weighted analyses (i.e., t-test, multiple regression, and the weighted analysis of covariance (ANCOVA) were conducted as a function of a treatment indicator and other covariates in predicting the school-average academic achievement scores in Korean, Mathematics and English. This approach anticipates that a propensity-based weighted sample would achieve similar distributions of baseline observed covariates between the weighted treatment and comparison groups (Austin and Stuart 2015). To estimate the population average treatment effect (PATE), the weights were created by using the inverse probability of a unit receiving the treatment condition that the subject actually received (Austin and Stuart 2015; Stuart, et al. 2001; Guo and Fraser 2010, p.197): $1/\hat{e}(x)$ for FS schools which participated in the FS, and $1/(1-\hat{e}(x))$ for the schools which did not adopt the FS. The propensity scored weights are then scaled so that the sums of the weights are equal to the original sample sizes of each treatment group and the average of the weights in each group becomes 1 (for more details, see Austin & Stuart

2015; Lunceford and Davidian 2004). Finally, the inverse probability of treatment weight (IPTW) approach can result in relatively extremely large values due to subjects in the treatment group with a very low probability of receiving the treatment (Austin and Stuart 2015). If that was the case, this study applied trimmed weights using the quantiles of the weight distribution (here, 95% percentile of the treatment group were used as the threshold for trimming). All the analyses were conducted with SAS 9.4, and for the ANOCVA, proc GLM was used.

Data and variables

This study used two existing data sources to conduct an empirical analysis of the impact of the FS on students' academic achievement in Korean middle schools. For the outcome measures, we used the data from the National Assessment of Educational Achievement (NAEA). The NAEA aims to monitor the quality of secondary education and assess school accountability at the national level by evaluating students' academic performance based on the national curriculum and relevant educational standards in Korea. The NAEA is annually administered for 3rd graders in middle school and 2nd graders in high school. The NAEA uses scaled scores with a mean of 200 and a standard deviation of 30 points. The NAEA data were publicly available at the school level until 2016. This study targets the 2014 FS cohort students who experienced the FS in grade 1 in 2014 and who took the NAEA in grade 3 in 2016.

Another data set this study used is the results of the Achievement Standards-based Assessment (ASA). The ASA is an assessment system that evaluates students' performance based on academic goals aligned with achievement standards instead of ranking students. The ASA results are generated at

Table 4 Baseline characteristics between the FS schools and non-FS schools

Variable	No FS (N = 230)		FS (N = 144)		Total	
	Mean	SD	Mean	SD	Mean	SD
Private school	0.40	0.49	0.11	0.32	0.29	0.45
High SES area (Gangnam)	0.14	0.34	0.22	0.41	0.17	0.37
School gender						
Boys only school	0.16	0.37	0.07	0.26	0.13	0.33
Girls only school	0.14	0.34	0.08	0.28	0.12	0.32
Coeducation	0.68	0.47	0.85	0.36	0.74	0.44
Proportion of students eligible for basic living security support	0.08	0.20	0.12	0.27	0.10	0.23
Number of students	679.7	241.5	692.4	244.3	684.5	242.3
2013 ASA school mean in grade 1						
Korean	72.09	5.26	73.03	4.07	72.45	4.86
Math	65.72	5.47	66.57	5.15	66.05	5.36
English	72.42	5.45	72.79	5.16	72.56	5.34

Table 5 Weighted descriptive statistics of key covariates between the FS schools and the non-FS schools

Variable	No FS (N = 230)		FS (N = 144)		Standardized Mean Diff.	Total	
	Mean	SD	Mean	SD		Mean	SD
Private school	0.29	0.46	0.25	0.44	0.06	0.28	0.45
High SES area (Gangnam)	0.16	0.37	0.16	0.37	0.00	0.16	0.37
School gender							
Boys only school	0.13	0.33	0.11	0.31	0.04	0.12	0.33
Girls only school	0.11	0.32	0.12	0.33	-0.02	0.12	0.32
Coeducation	0.76	0.43	0.77	0.43	-0.02	0.76	0.43
Proportion of students eligible for basic living security support	0.09	0.22	0.10	0.24	-0.03	0.09	0.22
Number of students	689.72	243.6	694.44	238.10	-0.01	691.54	241.21
2013 ASA school mean in grade 1							
Korean	72.35	5.00	72.47	4.15	-0.02	72.40	4.69
Math	65.97	5.14	66.08	5.02	-0.02	66.01	5.09
English	72.46	5.24	72.41	5.38	0.01	72.44	5.29

the school level and also publicly available. This study used the 2013 ASA data to examine the distributions between the FS schools and the non-FS schools in terms of student academic performance. This data were used because the 2013 ASA data provide outcomes from the closest school year prior to adopting the FS policy.

This study focused on 144 FS schools in Seoul (i.e., the largest metropolitan area in Korea). Given the FS policy was at the stage of dissemination in 2014, strategies and the impact of the FS could vary across provincial educational agencies, and so it was considered worthwhile to focus on the impact of the FS in Seoul. The 230 non-FS schools were included in the analysis for comparison. The total number of middle schools and their student population in Seoul in 2014 was 383 and 285,981, respectively (<https://www.sen.go.kr>). With the exception of specialized schools having no proper data, 374 schools have been analyzed in this study, which can be viewed as the target population of the middle schools in Seoul in 2014.

With little information known about how 144 pilot schools ended up with participation in the FS policy, the literature review along with the descriptive data analysis (as shown in Table 4) was employed in order to select the variables for propensity score estimation as a strategy to deal with selection bias. According to the literature, several studies (Kim and Kang 2017; Kim 2018b; Cho et al. 2018) informed that parents' social economic status (SES) is associated with students' achievement as well as implementation of the FS. Because the publicly available data do not contain a variable related to the SES index, we alternatively used an indicator informing whether a school is located in Gangnam province of Seoul, a high SES area. Another variable used as a proxy of the SES variable is the proportion of students eligible for basic living security support at the school level. Student gender is also an important variable in the field of education

in general. As mentioned in the literature section, student gender is also related to the implementation of FS policy or satisfaction of FS policy (Hwang and Yoo 2016; Kim and Lee 2019). This study used dummy variables informing school gender composition variable (i.e., boys only schools, girls only schools, or mixed-gender schools) as the unit of analysis is school. Finally, we examined 2013 ASA school mean variables in Korean, mathematics or English subjects for the first graders providing the information about academic performance prior to the FS policy participation in 2014. The reason we included the pretreatment academic information in the PS estimation is based the literature that the variables related to the outcome should be included regardless of whether or not they are associated to the receipt of treatment (Brookhart et al. 2006).

Table 4 presents the average distributions of key characteristics for the FS and non-FS schools. The results show that private schools, schools in the Gangnam area (which can be viewed as a high SES area) or coeducation schools were more likely to participate in the FS program. With respect to academic achievement, no substantial pretreatment differences were observed between the FS schools and non-FS schools based on the 2013 ASA school mean values by subjects for grade 1 in middle schools.

As a result, to estimate the propensity scores, we conducted a logistic regression as a function of a set of key covariates in predicting the FS participation at the school level. The final variables in the logistic model included indicators of private school and a high SES area, school gender composition indicators, the entire number of students, the percentage of students eligible for basic living security support, and the school means of first graders in 2013 in Korean, mathematics, and English based on the ASA data.

Results

Figure 1 shows the box and whisker plots of estimated propensity scores for the non-FS schools (coded as 0) and FS schools (coded as 1) based on a logit model. The left panel in Fig. 1 shows the distributions of the estimated probabilities with unweighted data. It clearly shows that FS schools have a higher probability of being in receipt of the policy compared to non-FS schools based on the observed covariates included in the logic model, and also indicates more homogenous groups than non-FS schools in 2014. On the other hand, as shown in the right panel of Fig. 1, when the IPTW was considered to yield the boxplot of the propensity scores between the two groups, the distributions were more similar in terms of median values and the overlap between the two boxes increased (the upper quartile minus the lower quartile). This presents the propensity score-based weight approach which helps improve the balance in pretreatment covariates between the FS schools and non-FS schools before the policy was implemented in 2014. In this matter, the propensity-based weight method allows researchers to approximate a cluster randomized trial with a non-experimental design.

In addition, we also examined the weighted descriptive statistics of each of the covariates used in Table 4 and the standardized mean differences between the two groups after the IPTW was involved in the analysis. Table 5 shows that the standardized mean differences were within the range of 0.05, showing the improved equivalence and reducing the selection bias related to the treatment assignment. Moreover, when we compared the distributions of the entire population in the total column of Table 4 and the weighted total column in Table 5, they are approximately identical in terms of both means and standard deviations for most variables. This implies that the estimate of the parameter based on the propensity score weighted data captures the average treatment effect targeting the entire population of interest, i.e., students in 1st grade resided in Seoul, Korea during the academic year of 2014.

Finally, we conducted a weighted outcome analysis to identify the impact of the FS policy on school-average achievement scores based on the 2016 NAEA for 3rd grade middle school students. The three NAEA school mean scaled scores in Korean, mathematics and English were used as outcome measures. First, we conducted a weighted *t*-test to assess the overall main effect of the FS policy. As shown in Table 6, the weighted average school means for the FS schools and non-FS schools essentially show identical results in all outcome measures, and no statistically significant differences were found based on *t*-statistics.

Next, we also conducted weighted multiple regressions. We included a variable indicating whether a school

participated in the FS in 2014 or not, after conditioning on some of the school variables included in Tables 4 and 5 (we only included variables (or a set of indicators) that were statistically significant in multiple regression analyses). Note that we also included the school mean pretest scores based on the 2013 ASA data in the regression models in order to further adjust the remaining imbalance between the FS and non-FS schools, even after the IPTW was applied in the outcome analyses.

As a result, the estimated regression coefficients of the FS participation in 2014 were statistically insignificant and almost close to 0 in the three outcome measures holding constant covariates listed in Table 7, in considering that the NAEA scaled scores ranged between 50 and 350 with a mean of 200 and an SD of 30. Finally, this study did not find statistically and substantially meaningful evidence that students who experienced the free semester during a semester of year 1 underperformed in NAEA scores in 3rd grade in comparison with students who attended regular middle schools located in Seoul that did not implement the FS policy in 2014. Moreover, the regression analysis results show that schools in a high SES area and schools that are girls only tend to show higher school means than other types of schools, and the school average in academic achievement scores appeared to indicate a positive association between the school size and prior ASA school means.

In addition, we also conducted a weighted ANCOVA analysis with an interaction term to explore if a differential effect of the FS policy exists depending on group characteristics. Thus attention was paid to categorical variables (high SES area and school gender), along with an interaction term between school gender and the indicator of FS participation in 2014, which would inform us about whether or to what extent the school-average achievement scores differ depending on the school location, student-gender composition or participation in the FS policy. Other interactions (i.e., high SES area FS) were not statistically significant, so in the final model we only included the interaction indicator to examine if the effect of FS participation is different depending on school gender characteristics.

Table 8 showed the results of ANCOVA in three academic subjects. With respect to the overall effectiveness of FS participation in 2014, the indicator of FS was still statistically insignificant in Korean and English, after the interaction was included. Furthermore, the interaction indicators (i.e., school gender FS), addressing the differential effects of FS participation depending on school characteristics related to student-gender composition, were also insignificant in Korean and English. The results are consistent with the results from the *t*-test and multiple regressions. However, the results are slightly different in mathematics. That is, after the interaction term was included in the analysis, the variables related to FS participation and the interaction between FS

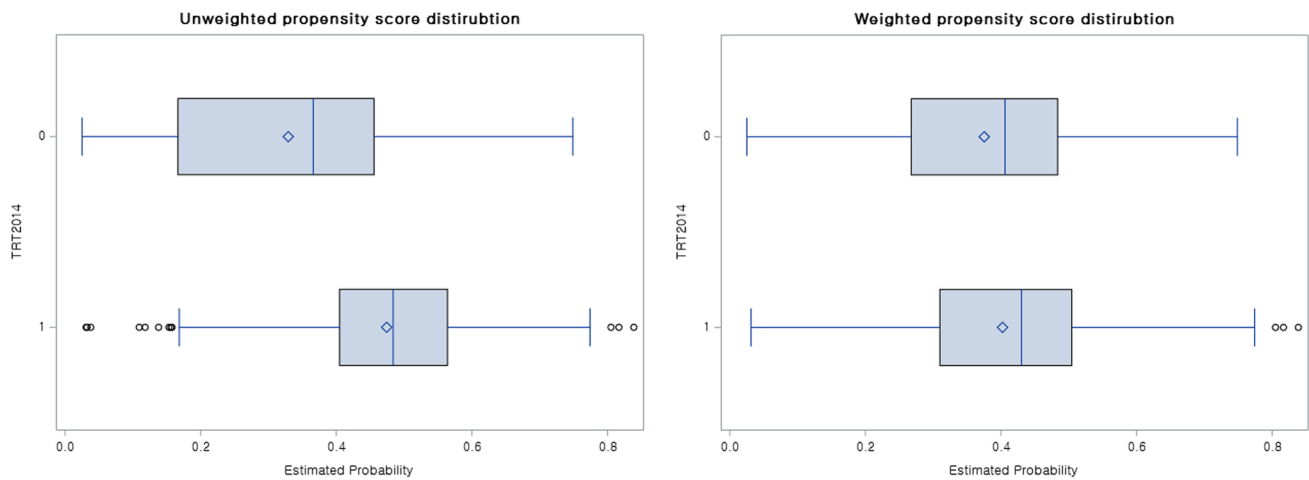


Fig. 1 The estimated probability of FS participation between the FS and non-FS schools with unweighted data (left) and weighted data (right)

Table 6 Weighted *t*-test results

NAEA outcomes	Treatment	Mean	SD	<i>t</i> Value	Pr. > <i>t</i>
Korean	Non-FS (0)	211.30	11.55	− 0.18	0.857
	FS (1)	211.50	10.79		
	Diff (0–1)	− 0.22	11.26		
Mathematics	Non-FS (0)	204.70	16.87	0.78	0.434
	FS (1)	203.30	15.52		
	Diff (0–1)	1.36	16.36		
English	Non-FS (0)	207.40	21.15	− 0.47	0.638
	FS (1)	208.40	19.43		
	Diff (0–1)	− 1.03	20.50		

and school gender became statistically significant in mathematics. To further investigate the results, Table 9 showed the least squares means (LS-means) in ANCOVA for certain sub-groups presented in categorical variables of Table 8. The *p*-values for the post hoc analysis were based on the Tukey–Kramer method, which yields the adjusted values for multiple pairwise comparisons. Note that LS-means produce the expected average values in a population if the levels of classification variables are balanced, i.e., the predicted population marginal means (Cai 2014).

As shown in Table 9, the differences in the LS-mean values in Korean and English were less than 2 points (1.85 and 0.78, respectively). In considering that the NAEA scaled scores ranged between 50 and 350 with a mean of 200 and an SD of 30, the average differences appeared trivial and insubstantial. In terms of mathematics, the LS-mean value was slightly higher in the non-FS group than in the FS group. When we look closely into the interaction results, however, the difference between the non-FS and FS schools in mathematics was only significant among schools with boys only (218.63 vs. 207.12) with a *p*-value of 0.019, whereas there

appeared to be no substantial differences between the two groups in other school gender types.

In summary, with respect to Korean and English, this study did not find any statistically and substantially meaningful evidence that students who experienced the free semester during a semester of year 1 underperformed in NAEA in the 3rd grade compared to students who attended regular middle schools located in Seoul that did not implement the FS policy in 2014. Also, there was no overall difference between students in FS schools and students in non-FS schools in mathematics based on the results of the weighted *t*-test and multiple regression analysis. However, when we further classified the school types based on the student-gender composition and compared the achievement scores between the FS group and non-FS group, holding student-gender composition constant, the predicted average score for the male population who attended boys only schools and participated in the Free Semester in 2014 was expected to be slightly lower than the predicted value for the male population who attended boys only schools but did not receive the FS policy.

Discussion and further research

This study was motivated by academic curiosity about bridging the gap between the education policy and its practical effects on student learning and achievement. The results add academic implications to the FS policy with contributions to the construction of timely evidence in relation to the FS. Based on empirical census data analysis and by incorporating an advanced statistical approach using propensity score-based weighting to deal with selection bias in an observational study, this study attempts to evaluate the

Table 7 Weighted multiple regression results

NAEA outcomes	Variable	Estimate	SE	t Value	Pr > t
Korean	Intercept	156.13	6.73	23.19	< .0001
	FS participation 2014	0.01	0.88	0.02	0.988
	High SES area	11.82	1.18	10.01	< .0001
	Girls only school	10.81	1.40	7.72	< .0001
	Boys only school	- 0.57	1.35	- 0.42	0.675
	Number of students	0.01	0.00	6.94	< .0001
	2013 ASA school mean	0.59	0.10	6.23	< .0001
Mathematics	Intercept	125.18	7.95	15.74	< .0001
	FS participation 2014	- 1.42	1.24	- 1.15	0.252
	High SES area	19.26	1.67	11.52	< .0001
	Girls only school	3.87	1.95	1.99	0.047
	Boys only school	5.29	1.90	2.78	0.006
	Number of students	0.02	0.00	8.77	< .0001
	2013 ASA school mean	0.90	0.12	7.33	< .0001
English	Intercept	75.03	9.65	7.78	< .0001
	FS participation 2014	1.04	1.42	0.73	0.466
	High SES area	25.37	1.90	13.33	< .0001
	Girls only school	11.30	2.24	5.04	< .0001
	Boys only school	2.47	2.18	1.13	0.259
	Number of students	0.02	0.00	7.49	< .0001
	2013 ASA school mean	1.53	0.14	11.18	< .0001

average effect of the FS policy on academic achievement in Seoul, the capital of Korea. The study aimed to produce significant and fruitful information in terms of responding to prevalent concerns about having no exams during a semester of middle school.

For analysis, we used the entire number of schools in Seoul, so that the analysis results in this study can be interpreted as the average population treatment effect of the FS policy. Overall, this study found that for the entire population for the middle schools in Seoul of 2014, no substantial differences were found in the academic achievement between the pilot FS schools and the non-FS schools because the average mean differences between the FS schools and non-FS schools appeared close to zero across subjects considering the measurement scores. This result is consistent with the findings from some previous studies (e.g., Kim 2016; Cho et al. 2018) and will relieve concerns about the FS's adverse effect on students' academic achievement.

First, contrary to mounting concerns as shown above, the FS does not appear likely to degrade academic achievement. More examination with cross-validation analysis, however, is needed, since the results showing no difference from implementing the FS may be a function of consistent dependence on private tutoring or more invigorated private academies seeking to focus on academic results to compensate for the schools' perceived lack in this area. Therefore, utilizing data such as private tutoring expenses, achievement trends, or changes during or after the FS is needed for comparison. One study (Kim 2016) empirically showed that when other covariates were controlled for, there were no statistically significant differences between students who participated in the FS and those who did not in private tutoring expenses. However, the study had a limitation because the analytical data included FS students who had experienced the free semester for a few months in 2015. Additionally, Park (2017) raised an important issue by addressing the wider gap in expenditure on private tutoring between high-income and middle-income households based on survey data from 2009–2016 from Statistics Korea. The study pointed out that while the FS policy had no impact on the 'average' expenditure of private tutoring between the free semester group and their non-FS counterpart, high-income families were more likely to participate in private tutoring as well as spend more money on private education, which may signal an unintended side-effect of the FS policy such as increasing inequality in educational opportunities.

It is also important to note that when the schools were classified by student-gender composition, unlike other types of schools, the school-average scores for FS schools with boys only were slightly lower than the average for non-FS schools with boys only. Given the fact that the data were not balanced in terms of school gender, most middle schools were identified as mixed-gender type (about 75%), and only 10 boys only schools participated in the FS pilot program; the results about the interaction effects should be interpreted with caution and cannot be generalized. However, the results may signal that the free semester can have a different influence according to school or student characteristics. Thus, further in-depth quantitative and qualitative research should be conducted addressing such questions as what kinds of activities male students did during the free semester and whether there were different experiences related to academic performance among boys in comparison with the experiences among girls in the period during and after the free semester. Along with this point of view, one drawback of the current study is that the analysis does not capture the process of FS policy under the variously different school contexts such as student and teacher compositions, school resources, FS curriculum development and operations, and teacher and parent cooperation/understanding, etc.

Table 8 ANCOVA analysis results

Subject	Source	DF	Type III SS	Mean Square	F Value	Pr > F
Korean	FS	1	147.4	147.4	2.17	0.142
	High SES area	1	6385.3	6385.3	94.02	<.0001
	School gender	2	3985.1	1992.5	29.34	<.0001
	School gender FS	2	296.0	148.0	2.18	0.115
	Number of students	1	3314.4	3314.4	48.80	<.0001
	2013 ASA school mean	1	2702.0	2702.0	39.79	<.0001
Math	FS	1	1021.7	1021.7	7.71	0.006
	High SES area	1	16,563.3	16,563.3	124.95	<.0001
	School_gender	2	740.9	370.4	2.79	0.063
	School_gender FS	2	1313.5	656.7	4.95	0.008
	Number of students	1	10,560.5	10,560.5	79.67	<.0001
	2013 ASA school mean	1	7397.4	7397.4	55.80	<.0001
English	FS	1	26.0	26.0	0.15	0.703
	High SES area	1	30,617.2	30,617.2	171.01	<.0001
	School gender	2	4219.8	2109.9	11.78	<.0001
	School gender FS	2	304.5	152.3	0.85	0.428
	Number of students	1	10,166.8	10,166.8	56.79	<.0001
	2013 ASA school mean	1	21,523.0	21,523.0	120.21	<.0001

Table 9 Post hoc results: adjusted means by sub-groups

Variable\subjects	Korean			Math			English			
	LS-mean	SE	P-value	LS-Mean	SE	P-value	LS-mean	SE	P-value	
FS	0	218.05	0.85	0.142	214.07	1.19	0.006	219.44	1.38	0.703
	1	216.20	1.09		209.21	1.53		218.66	1.76	
High SES area	0	211.37	0.66	<.0001	202.33	0.93	<.0001	206.51	1.08	<.0001
	1	222.88	1.18		220.96	1.66		231.59	1.90	
School Gender	1	214.06	0.65	0.655*	209.28	0.91	0.160	214.80	1.06	0.738
	2	212.83	1.35	<.0001**	212.88	1.90	0.999	216.49	2.20	0.005
	3	224.49	1.37	<.0001***	212.77	1.88	0.176	225.87	2.19	<.0001
<i>Interaction (School gender × FS)</i>										
1	0	213.55	0.76	0.910	209.06	1.07	1.000	213.79	1.24	0.816
	1	214.57	0.88		209.50	1.23		215.81	1.42	
2	0	214.71	1.56	0.691	218.63	2.18	0.019	218.21	2.53	0.965
	1	210.95	2.14		207.12	2.99		214.78	3.49	
3	0	225.90	1.68	0.879	214.52	2.34	0.922	226.33	2.72	1.000
	1	223.09	2.03		211.02	2.82		225.40	3.28	

*p-value for 1 (coeducation) vs 2 (boys only schools) comparison of school gender; **p-value for 2 vs 3 (girls only schools) comparison; ***p-value for 1 vs 3 comparison

Second, more studies with various research questions should be conducted rather than only researching student satisfaction. Consistent studies on self-directed learning skills, career maturity, and teachers' professional development with both qualitative and quantitative analyses are necessary for the FS policy implementation and its expansion. A recent study (Jung et al. 2018) shows how the Free Semester policy can be formatted and expanded to a Free

School-Year Program. Also, the Gyeonggido Office of Education announced a Free School-Year Semester policy for 2019, which implies the implementation of the Free School-Year program in grade 1 and aligned programs for grades 2 & 3 in middle school. Their focus is in having a competency-based curriculum, learning-centered classes, growth-based assessment, and student choice (Gyeonggido Office of Education 2018, p.1). Therefore, the further analysis needs to

be focused on competency, student growth, and a variety of learning in both quantitative and qualitative terms.

Third, various indicators or assessment tools need to be developed to understand student growth in competencies such as collaboration, social skills, communication, self-directed learning, problem solving, etc. Some high school teachers mention that students experiencing the FS in middle school are more likely to show better presentation or debate skills than those who do not. Nonetheless, there are no reliable tools to evaluate these skills. As the purpose of the FS is closely connected to competencies, more inquires on competencies in instruction and assessment should be conducted to understand the process and product of the FS.

Fourth, as the paradigm of classroom assessment has been changing from product-oriented to process-focused, it is necessary to look for ways to evaluate core competencies through classroom-level curriculum competency, which is the most appropriate form of performance evaluation. The 2015 revised curriculum specifies the core competencies to be addressed in school education and emphasizes process-focused assessment aligned with learning activities in the classroom. In addition, some research has been conducted on how to evaluate core competencies at the national and classroom levels. To evaluate core competency in teaching and learning situations, a systematic evaluation system should be established to check core competencies proposed in the curriculum. That is, it is necessary to support the evaluation of core competencies at the classroom level by developing items or problems from a real life context that can measure such competencies and use them in teaching and learning situations.

Fifth, admittedly the concept of ‘student achievement’ is used in a limited way with academic achievement or test scores. However, recognizing abilities from multi-dimensions and reconceptualizing ‘achievement’ is a way to transform the status quo understandings of achievement and practices in education. Moreover, studies reflecting on the concept of achievement should be conducted.

Nonetheless, this study had certain limitations. First, the generalization of the results should be limited, because this study focused only on the impact of the FS policy in the Seoul area, and used the pilot FS schools in 2014, which can be viewed as an initial period of the policy implementation. Therefore, more comprehensive analyses should be conducted using national data that cover the entire population of Korea. Second, the impact of the FS should be further examined for different periods of time (i.e., from 2015 through 2019) to make sure no negative effect of the free semester policy is sustained across the different school years. Third, the unit of analysis in this study is school-level data, and thus the study was not able to address how and to

what extent student characteristics were interacted with the free semester policy.

Lastly, attention was paid only to academic performance in order to investigate the effectiveness of the FS. Thus, it is necessary to further examine the impact of the FS on other outcome measures such as students’ affective domains or well-being. To do so, it would be valuable to have assessment using student-level data from the NAEA, or international data such as the PISA for Korean students. By combining different data sets at the student-level, longitudinal data can be created to assess if the FS policy has an impact on student growth. Also, the usage of such existing data allow various outcome measures such as collaborative problem solving, student well-being or subject-specific affective domains to be obtained. Finally, if possible, it would be worthwhile to examine whether student experiences of the Free Semester in middle school have a noticeable relationship with college entrance exam scores in the long run (e.g., Korean SAT).

References

- Austin, P. C., & Stuart, E. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate treatment effects in observational studies. *Statistics in Medicine*, *34*, 3661–3679. <https://doi.org/10.1002/sim.6607>.
- Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology*, *163*(12), 1149–1156. <https://doi.org/10.1093/aje/kwj149>.
- Burstein, L. P. (1980). The analysis of multilevel data in educational research and evaluation. *Review of Research in Education*, *8*, 158.
- Cai, W. (2014). *Making comparisons fair: How LS-means unify the analysis of linear models* (Paper SAS060-2014). Cary, NC: SAS Institute Inc.
- Chang, E. J. (2011). The rise of the meritocracy in Korean society and its education problem: In searching for the democratic justice. *Society and Philosophy*, *21*, 71–106.
- Cho, H. Y. (2017). A study on analysis and suggestions on research relations with Free Semester System. *The Journal of Humanities and Social Science*, *8*(3), 352–376.
- Cho, S. K., Kim, M. J., Park, K. A., Seo, J. E., Lee, B. M., & Jeong, D. W. (2018). Policy effect analysis of Free Semester program: Focusing on school factors. *Korean Journal of Educational Administration*, *36*(4), 41–68.
- Choi, S. D., Lee, S. E., Kim, So, K. H., B. C., Park, S. Y., & Hong, C. N. (2014). Comprehensive Report on Case Studies of Pilot Schools Implementing the Free Semester Program in 2013. Korea Educational Development Institute. 2014-19.
- Enforcement Decree of the Career Education Act. <https://www.law.go.kr/LSW/eng/engLsSc.do?menuId=2§ion=lawNm&query=%EC%A7%84%EB%A1%9C%EA%B5%90%EC%9C%A1%EB%B2%95%EC%8B%9C%ED%96%89%EB%A0%B9&x=0&y=0#liBgcolor0>. (retrieved from National Information Center, 2019. 3. 10).

- Enforcement Decree of the Elementary and Secondary Education Act. <https://www.law.go.kr/LSW/eng/engLsSc.do?menuId=2§ion=lawNm&query=%EC%B4%88%EC%A4%91%EB%93%B1%EA%B5%90%EC%9C%A1%EB%B2%95&x=0&y=0#liBgcolor13>. (retrieved from National Information Center, 2019. 3. 10).
- Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. Thousand Oaks, CA: Sage Publications Inc.
- Gwangsan Middle School. (2016). *School Curriculum*.
- Gyeonggi Office of Education. (2018). 2019 Plan for Gyeonggi Free School-Year System (Press release, 2018. 12).
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81, 945–960.
- Hwang, E. H., Jang, J. H., & Yang, H. G. (2019). An analysis of the research trends in Free-Semester using text-mining techniques. *Journal of Education & Culture*, 25(3), 299–318.
- Hwang, J. Y., & Yoo, J. A. (2016). An analysis on student's perception of physical education activity in Free Semester System at Korean middle school. *Asian Journal of Physical Education of Sport Sciences*, 4(2), 39–52.
- Hyun, E. R. (2017). The development of a creative career education program for a Free Semester Operation media-convergence design. *Research of Digital Design*, 15(2), 509–518.
- Jeong, Y. G., Min, Y. S., Byun, H. H., Kim, Y. M., Seol, H. S., Lim, J. Y., & Jang, H. J. (2015). *The study on the implementation of the Exam-Free Semester curriculum*. KICE, Research Report, RRC 2015-7-1.
- Jeong, Y. D., Yoon, Y. I., Lee, K., T. (2015). A study of school life satisfaction through free semester curriculum. *The Journal of Education Assignment Institute*, 21(1).
- Jung, Y. S., Lee, Y., & Park, D. (2018). Analysis on Policy Formation of Free Semester Program and Policy Change to Free School-Year Program Based on Kingdon's Policy Streams Model. *The Journal of Politics of Education*, 25(1), 143–172.
- Kang, C. D. (2007). The study on the social historical change process and feature in the university entrance system of Korea. *The Journal of Research in Education*, 28, 83–113.
- Kim, D. S. (2017). Difference of career maturity, core competencies, and school satisfaction according Free Learning Semester. *The Journal of Curriculum and Evaluation*, 20(3), 101–121.
- Kim, H. G. (2018). The effect of the multicultural education program on intercultural sensitivity in middle school FLS. *Multicultural Education Studies*, 11(3), 147–168.
- Kim, J. S. (2013). A study of transition year curriculum in Ireland. *The Journal of Curriculum Studies*, 31(4), 165–193.
- Kim, J. Y. (2018). A study on changes to students' academic records according to their satisfaction with the Free Semester System before and after its implementation (2015–2017). *Journal of Learner Centered Curriculum and Instruction*, 18(18), 293–316.
- Kim, M. J., & Hong, H. J. (2016). Research on the sustainability of the Free Learning Semester and its implication for the identity of the middle school curriculum. *The Journal of Curriculum Studies*, 34(1), 1–28.
- Kim, M. J., Hong, H. J., Lee, J. C., & Lim, J. I. (2014). An analysis of school reports on Free Learning Semester: Focusing on development and operation of school curriculum. *The Journal of Curriculum Studies*, 32(2), 215–241.
- Kim, N. K., & Lee, S. M. (2019). Middle school students' perceptions towards the effectiveness of the free-year program. *Studies in Foreign Language Education*, 33(2), 161–187.
- Kim, Y. B. (2016). *Comparisons between students experienced the Free Semester and students who don't on academic achievements and the expenses of private education*. Korean Educational Development Institute, Issue Paper 2016-01.
- Kim, Y. B., & Kang, H. S. (2017). Student achievement growth among middle school students. *The Journal of Korean Education*, 44(1), 33–61.
- Kim, Y. B., Namgung, J. Y., & Park, K. H. (2019). An analysis of the changes in academic achievement for middle school students based on their participation in the Free-Semester System and the socio-economic status of their families. *The Journal of Korean Education*, 46(1), 5–39.
- Lee, G. S. (2011). A thought on ineducability in current schools. In G. S. Lee (Ed.), *The Era of Ineducability* (pp. 15–30). Seoul: Educational Community Friend.
- Lee, I. H. (2016). An analysis of school change of a middle school according to the Free Semester policy. *Journal of Digital Convergence*, 14(1), 33–42.
- Lunceford, J. K., & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine*, 23(19), 2937–2960.
- Ministry of Education. (2015a). *The plan for the Exam Free Semester in middle school (press release)*.
- Ministry of Education. (2015b). *Cases of the pilot and exemplar schools for the Exam Free Semester*.
- Ministry of Education. (2015c). *The 2015 revised curriculum*.
- Ministry of Education. (2017). *Education in Korea*.
- Morgan, S., & Winship, C. (2007). *Counterfactuals and causal inference: Methods and principles for social research*. New York: Cambridge University Press.
- Park, H. Y. (2015). A critical discourse analysis regarding the tensions between the identities on the application of the Free Semester Program. *The Journal of Curriculum Studies*, 33(3), 1–27.
- Park, Y. (2017). *Study on the effect of the Free Semester policy* (pp. 2017–2113). Policy Report Series: Korea Development Institute.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55.
- SAS Institute Inc. (2013). *SAS @9.4 Statements: Reference*. Cary, NC: SAS Institute Inc.
- Shin, C. K., Hwang, E. H., & Kim, E. Y. (2015). An analytical study on the operation and actual state of the Free-Semester Program. *Asian Journal of Education*, 16(3), 27–55.
- Shin, C. K., & Park, M. J. (2015). The case study of three schools with Free-Semester Policy. *Korean Journal of Educational Administration*, 33(1), 309–338.
- Shin, H. S., Lee, Y. S., & Jung, Y. S. (2018). Analysis on students and parent involvement experience in Free School Year program: Based on the grounded theory paradigm model. *The Journal of Research in Education*, 31(3), 53–98.
- Stuart, E. A. (2007). Estimating causal effects using school-level data sets. *Educational Researcher*, 36(2), 187–198.
- Stuart, E. A., Cole, R. S., Bradshaw, C. P., & Leaf, P. J. (2001). The use of propensity scores to assess the generalizability of results from randomized trials. *Journal of the Royal Statistical Society: Series A*, 174(2), 369–386.
- Yeo, J. Y., & Chae, S. M. (2016). The School experiences of middle-school students in the Free Learning Semester. *Journal of the Korea Contents Association*, 16(5), 698–709.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.