A Comparison of Two-Stage Least Squares (TSLS) and Ordinary Least Squares (OLS) in Estimating the Structural Relationship between After-School Exercise and Academic Performance

Kyulee Shin [1], Sukkyung You [2],* and Mihye Kim [1]

[1] Department of Sports Science, Seoul National University of Science & Technology, Seoul 01811, Korea; kyuleeshin@seoultech.ac.kr (K.S.); mihyekim0912@seoultech.ac.kr (M.K.)

[2] College of Education, Hankuk University of Foreign Studies, Seoul 130-791, Korea

* Correspondence: Skyou@hufs.ac.kr

Abstract: The current study examines the structural relationship between the academic performance exam scores of Korean middle school students and their after-school exercise hours. Although prior literature theoretically or experimentally predicts that these variables are positively associated, this association is difficult to empirically verify without controlling for mutual effects with other variables, or unless a full model is estimated by specifying the whole structure of all variables affecting the two variables in question. Unlike previous studies, this study estimates the structural relationship using two-stage least squares method, which does not require experimental observations collected for our particular purpose or estimating the full model. From this estimation, we empirically affirm that there is a positive structural relationship between students' after-school exercise hours and their academic performance exam scores, whereas the ordinary least squares method consistently estimates a negative relationship.

1. Introduction

Identifying the structural effect of regular exercise on mental health has been a popular research objective. For example, a study by Dishman et al. [1] found that physical exercise regulates the structure and function of the brain. In addition, Tomporowski et al. [2] reviewed theoretical works on this topic, identifying positive structural associations between physical exercise and academic achievement, and Cappelen et al. [3] reported experimental evidence that exercise has a spillover effect that improves the academic performance of college students. Nevertheless, prior empirical literature has not revealed a conclusive positive association between the two variables. For example, using randomly sampled data constructed from students aged between 7 and 15, Tremblay et al. [4] showed that the correlation between students' academic performance scores and physical fitness was not significantly different from zero. On the other hand, Castelli et al. [5] and Dwyer et al. [6] found empirical correlation coefficients strictly greater than zero between the two variables. In addition to these, we also find a number of empirical studies using longitudinal data with inconclusive evidence (e.g., Rees and Sabia [7]; Dills et al. [8]) or positive coefficient (e.g., You et al. [9]).

The lack of conclusive empirical investigations on the topic implies that the structural relationship between students' physical exercise and academic performance may not be adequately identified by estimating the correlation coefficient between the two variables. Likewise, the regression coefficient obtained by the ordinary least squares (OLS) estimation may not properly identify the structural relationship between the two variables, as the regression coefficient is an extended form of correlation coefficient.

The main goal of this study, therefore, is to empirically identify the structural relationship between students' physical activity and academic performance. To this end, we substitute the OLS-based approach with the two-stage least squares (TSLS) method. If proper instrumental variables are employed, the TSLS can reveal the structural relationship between physical exercise and academic performance. The TSLS estimation is a popular method for estimating the causal relationship between variables in statistics, economics, and ecology. In the current study, we use the TSLS method to estimate the structural relationship between the two variables.

In this study, we use data collected by the National Youth and Policy Institute (NYPI) in 2011 on Korean middle school students. The data set provides comprehensive information about Korean middle school students, including parental, academic, extra-curricular, physical, and environmental details. We extract information on the students' academic performance and physical activity from this data set, enabling us to estimate the structural relationship between the variables using the TSLS method. Specifically, we utilize data regarding the student's height, weight, and sleeping hours on weekends as our instrumental variables that are correlated with the after-school exercise hours but uncorrelated with the structural error. If we obtain the fitted after-school exercise hours by regressing the instrumental variables on the after-school exercise hours, the fitted after-school exercise hours become uncorrelated with the structural error, so that the effect of students' after-school hours on their academic performance scores can be identified by regressing the academic performance scores on the fitted after-school hours.

The TSLS model we employ in this study includes other explanatory variables that have been found to affect students' academic performance scores in the literature. In addition to parents' income levels and students' private lesson hours as our basic explanatory variables, we include students' gender, expenditure on after-school private lessons, number of siblings, and parents' education levels. According to Geary, Hoard, and Nugent [10] and Gneezy, Niederle, and Rustichini [11] among others, male and female students have relative advantages in different environments. Students' expenditure on after-school private lessons and number of siblings are included to explain the direct effects of parents' income on students' academic performance scores. Finally, parents' education levels are included to accommodate parental involvement with students' academic performance scores. Shumow, Lyutykh, and Schmidt [12] and Glick and Hohmann-Marriott [13] among others show that students' academic performance is related to parental involvement, and parents' education level can affect parental involvement.

To the best of our knowledge, the TSLS method has not been applied to the estimation of the positive structural relationship between students' academic performance scores and physical activity. We contribute to the literature by examining the causal relationship between Korean middle school students' physical activity and their academic performance.

The structure of this paper is as follows. In Section 2, we compare TSLS estimation with the OLS estimation method. In addition, we specify the models estimated by the OLS and TSLS methods using the variables for empirical analysis. In particular, we describe our specific strategy for estimating the structural relationship between students' academic performance score and after-school exercise hours by introducing instrumental variables. We also describe prior literature that is relevant to this study. In Section 3, we discuss the estimation and inference results and compare the results with hypotheses from the literature. Concluding remarks are provided in Section 4.

2. Materials and Methods

2.1. Data

We use data collected on first-grade middle school students by NYPI in South Korea for our empirical analysis. The data set was not collected with the aim of identifying the structural relationship between students' academic performance scores and their after-school exercise hours; it was collected for a general purpose to understand the comprehensive actual conditions and environmental changes in youth and children in South Korea. Specif-

ically, the data set is called the Korean Children and Youth Panel Survey (KCYPS 2010), and KCYPS 2010 is a panel data set collected from elementary-school and middle-school students for 7 years from 2010 to 2016. For the middle school students, 2342 students participated in the survey in the first year, and 1881 students still remained in the sample pool by the final year. The data collect comprehensive aspects of students' daily lives in addition to their background information. Data are collected from the following categories: physical information, intellectual growth, social and emotional development, juvenile delinquency, daily time allocation, family background, friend group formation, educational environment, local community environment, media environment, and cultural environment. Due to this fact, we cannot control the correlation between after-school exercise hours and students' academic performance, necessitating the use of the TSLS method to estimate the structural relationship. The KCYPS 2010 panel survey is downloadable from the data archive described at the end of this paper. For the purpose of this study, we use the 2011-year wave that provides the observations of the variables in Models (1), (2), and (3) that are given below.

Before reporting the estimation results, we report the descriptive statistics of the variables in Table 1. There are 2280 students in the data set, but some variables have missing observations. Our estimations are made by removing missing observations from the data set. Although the data properties in Table 1 are straightforward, there are two factors that should be noted. First, about 6.4% of the 2280 students do not report their parents' yearly income. Other than the income variable, there are few missing observations for the other variables. Second, we conduct testing the normal distribution hypothesis using Jarque and Bera's [14] test statistic and report the test statistic values along with their p-values. They are reported in the final column of Table 1. Most variables are not distributed according to a normal distribution, although we cannot reject the normal distributional hypothesis for the students' height.

Table 1. Descriptive statistics.

Variables	Smpl. Size	Mean	Median	Max.	Min.	Std. Dev.	Skew.	Kurt.	J-B Test
Dependent Variable									
scor (in points)	2268	70.60	75.50	97.50	28.00	19.87	–0.61	2.15	208.27 (0.00)
Cause Variable									
phyh (in hours)	2280	7.20	0.00	480.00	0.00	42.39	7.04	57.08	296,782 (0.00)
Instrumental Variables									
heit (in meter)	2276	1.63	1.63	1.87	1.25	0.07	0.01	3.15	2.26 (0.32)
weit (in kilogram)	2265	54.09	52.00	120.00	30.00	10.66	1.10	4.99	841.84 (0.00)
slip (in minutes)	2.272	9.40	9.00	16.00	1.50	1.58	0.06	4.36	178.09 (0.00)
Other Explanatory Variables									
othh (in hours)	2280	103.69	84.00	786.00	0.00	102.62	1.52	7.43	2758 (0.00)
incm (in 10,000 KRW)	2133	4536	4000	30,000	300	2473	1.88	12.16	8727 (0.00)
gend (1 = male, 0 = female)	2280	0.50	1.00	1.00	0.00	0.50	–0.02	1.00	380.00 (0.00)
edct (in 10,000 KRW)	2265	27.07	28.00	300.00	0.00	23.94	1.72	13.39	11,314 (0.00)
nsib (no. of heads)	2279	1.11	1.00	6.00	0.00	0.65	1.06	7.69	2523 (0.00)

Table 1. *Cont.*

Variables	Smpl. Size	Mean	Median	Max.	Min.	Std. Dev.	Skew.	Kurt.	J-B Test
mote-m (1 = middle school, 0 = else)	2280	0.02	0.00	1.00	0.00	0.15	6.08	38.02	130,615 (0.00)
mote-h (1 = high school, 0 = else)	2280	0.48	0.00	1.00	0.00	0.49	0.05	1.00	380.00 (0.00)
mote-p (1 = polytechnic school, 0 = else)	2280	0.09	0.00	1.00	0.00	0.29	2.77	8.70	6025 (0.00)
mote-u (1 = university, 0 = else)	2280	0.27	0.00	1.00	0.00	0.44	1.00	2.00	476.10 (0.00)
mote-g (1 = graduate school, 0 = else)	2280	0.01	0.00	1.00	0.00	0.13	7.25	53.62	263,502 (0.00)
fate-m (1 = middle school, 0 = else)	2280	0.02	0.00	1.00	0.00	0.16	5.66	33.10	98,307 (0.00)
fate-h (1 = high school, 0 = else)	2280	0.36	0.00	1.00	0.00	0.48	0.57	1.33	390.52 (0.00)
fate-p (1 = polytechnic school, 0 = else)	2280	0.08	0.00	1.00	0.00	0.28	2.90	9.43	7147 (0.00)
fate-u (1 = university, 0 = else)	2280	0.37	0.00	1.00	0.00	0.48	0.53	1.12	387.71 (0.00)
fate-g (1 = graduate school, 0 = else)	2280	0.03	0.00	1.00	0.00	0.19	4.79	23.94	504,413 (0.00)

This table shows the descriptive statistics of the variables used for the empirical analysis. Figures in parentheses are the *p*-values of the Jarque–Bera test statistic.

2.2. Estimation Strategy and Prior Literature

The primary goal of this paper is to determine the structural relationship between Korean middle school students' academic performance scores and their physical activity. As there are many theoretical and experimental models in the literature, it would be inefficient to relate them to the empirical models relevant to the goal of this study. We instead specify the following empirical model and relate it to the other studies:

$$\log(scor_i) = \alpha + \beta\, phyh_i + \gamma\, othh_i + \delta \log(incm_i) + \pi\, gend_i + u_i \qquad (1)$$

where *scor* stands for the average score of Korean, English, mathematics, science, and social science exam scores as a measure of students' academic performance, and *phyh*, *othh*, and *incm* denote students' weekly after-school exercise hours; weekly average private lesson hours for Korean, English, mathematics, science, and society subjects; and parent's yearly income level measured in 10,000 Korean won, respectively. Here, students' exam scores for physical education are not included in *scor*, as our interest is in identifying the structural effect of the after-school exercise hours on the exam scores of the other subjects. We measure the students' physical activity by the variations in their after-school exercise hours. School curricula in different Korean middle schools in 2011 were more or less similar, making it difficult to capture different physical activities among students through regular school physical education hours. Instead, we noted that after-school exercise hours differed among students and captured this feature *phyh*. Here, the logarithm is taken to *scor* and *incm*, so that the coefficient of log(*incm*) can be understood as the income elasticity of the academic performance score, implying that it measures the percentage change of the score in response to the one-percentage increase of parents' yearly income. On the other hand, we cannot take the logarithm of *phyh* or *othh*, as many of them are zeros. Finally, the subscript *i* represents the *i*-th individual.

A positive structural relationship between academic performance and after-school exercise hours was predicted on a number of theoretic bases. For example, the impacts of exercise on mental health, affect, and cognition have been determined by Tomporowski et al. [2], de Greeff, Bosker, Oosterlaan, Visscher, and Hartman [15], Chang et al. [16], Álvarez-Bueno et al. [17], Xiang, et al. [18], and Tomporowski, McCullick, Pendleton, and

Pesce [19] among others. Furthermore, Dishman et al. [1] provides experimental evidence that regular physical activity results in a host of biological responses in both muscles and organs, modifying and regulating the structure and functions of the brain. Considering this theoretical and/or experimental evidence, we hypothesized that after-school exercise hours and academic performance scores would be positively associated.

Nevertheless, we cannot estimate the structural model (1) using the prediction model that is often estimated by the OLS method. It is mainly because the structural error term (u) can be correlated with after- school exercise hours. There are a number of reasons for this hypothesis. First, if a student spends time on after-school exercise, they may have to reduce the hours they spend studying other subjects. Korean middle school students typically spend their free time after school studying the school subjects in which they are relatively weak. Students may study their weak subjects themselves or take private lessons. Unless they reduce the amount of sleep they receive, additional time spent on after-school exercise will reduce the time students have for studying independently or taking private lessons on their weak subjects, implying that their academic performance will be indirectly affected by increasing their after-school exercise hours. In other words, the error term would be affected by the variation of students' after-school exercise hours. As another reason, if the student has a family or personal background in sports, this innate proficiency may act as a hidden variable behind the error term, so that after-school exercise hours become correlated with the error term. Note that students with the family or personal backgrounds in sports are more likely to take after-school exercise lessons. There may be many other reasons for the posited correlation.

The OLS estimation is incapable of estimating the structural relationship between the variables in question. This is mainly because the OLS method estimates a prediction model by presuming that the error term is uncorrelated with explanatory variables, letting the parameter coefficient be represented as a function of correlation coefficients between the variables in Model (1). Due to this fact, the OLS method is asymptotically biased in terms of the structural parameter estimation, leading to inconclusive empirical results. This has been demonstrated by the literature; for example, Tremblay et al. [4] empirically found that there was no correlation significantly different from zero between a child's academic achievement and their physical fitness by examining 6856 sixth-grade students in the US. On the other hand, Castelli et al. [5] and Dwyer et al. [6] found that the two variables were positively correlated using 259 and 7961 US students aged between 7 and 15.

We instead investigate the structural equation in (1) by directly estimating the structural parameter with the TSLS estimation method. The TSLS method is a two-step estimation that employs instrumental variables correlated with the cause variable but uncorrelated with the error term. The TSLS first regresses the cause variables on properly selected instrumental variables and the explanatory variables other than the cause variable to obtain the fitted cause variable. Second, the TSLS regresses the dependent variable on the fitted cause variable and the other explanatory variables in the model, estimating the unknown parameter in the structural model. Following this plan, after-school exercise hours and the other right-side variable in Model (1) are treated as the cause variable and the other explanatory variables. We first regress the after-school exercise hours against the instrumental and other explanatory variables to obtain the fitted value of the after-school exercise hours. In the second stage, we regress the academic performance score on the fitted value and the other explanatory variables so that the structural parameter can be estimated consistently. Note that the fitted value of the academic performance score keeps partial information of the after-school exercise hours while being uncorrelated with the error term from the fact that they are driven by the instrumental variables uncorrelated with the error term. Using the fitted after-school exercise hours, the second step now estimates the structural parameter by simply regressing the academic performance score on the fitted after-school hours and the other variables.

For the purpose of the TSLS estimation, we employ three instrumental variables: students' height (*heit*), weight (*weit*), and hours spent sleeping on weekends (*slip*). Specifically,

we use the logarithms of *heit*, *weit*, and slip as our instrumental variables. Note that body characteristics such as height and weight are highly correlated with exercise hours (e.g., Beets, Beighle, Erwin, and Huberty, [20]), but there is little correlation between academic performance scores and these variables. Using this feature, we employ the height and weight as our instrumental variables. In addition, students' sleeping hours critically affect their academic performance, as Nihayah et al.'s [21] case investigation shows. Note that although students that spend regular hours on after-school exercise have a tendency to sleep more than other students to save energy, they must still attend school on time. Due to this fact, the student cannot sleep as long as they wish on weekdays, although they may be able to sleep as long as they wish on weekends. Therefore, we employ students' hours spent sleeping on weekends as our third instrumental variable. As we discuss below, these instrumental variables were found to be uncorrelated with the error term through a formal testing procedure, acting as proper instrumental variables.

We ensured our structural model estimation by conducting formal testing procedures that justified our use of the TSLS estimation. First, as described above, the key step of the TSLS estimation is in selecting proper instrumental variables that are uncorrelated with the error term. Nevertheless, the error term is unobserved, and it is difficult to estimate the correlation coefficient between the variables. Instead, we apply the *J*-test statistic to detect zero-correlation. Our instrumental variables (*heit, weit*, and *slip*) are selected with this statistical testing procedure. Next, we test whether the estimated structural coefficient is significantly different from zero. While testing this hypothesis with the *t*-test statistic as for the OLS estimation, we estimate the standard error of the coefficient with the robust standard error in White [22]. If the error term is conditionally heteroskedastic on the instrumental variables, the standard *t*-test statistic does not asymptotically follow a standard normal distribution under the null. We, therefore, modify the standard *t*-test statistic by modifying the standard error of the *t*-test statistic using the robust standard error so that the modified *t*-test statistic asymptotically follows a standard normal distribution under the null hypothesis. Using the modified *t*-test statistic, we test whether the structural coefficient is greater or less than zero to achieve the goal of this study. The weighting matrix of the *J*-test statistic is also modified by the robust covariance matrix consistent estimator so that it follows a standard chi-squared distribution under the hypothesis that the error term is uncorrelated with the instrumental variables.

Finally, the other variables on the right side of Model (1) are included as explanatory variables for students' academic performance scores. We first include the average private lesson hours of the other subjects as an explanatory variable on the right side by noting that private lessons are critically helpful in raising students' academic performance scores. Second, we note that family resources such as parents' income and wealth levels become useful resources for children's education. Parents' yearly income is included on the right side to capture this aspect.

2.3. Model Extensions

We next extend the model scope in Model (1) by including other explanatory variables on the right side. As our first extension, we specify the following model:

$$\log(scor_i) = \alpha + \beta\, phyh_i + \gamma\, othh_i + \delta \log(incm_i) + \pi\, gend_i + \xi\, edct_i + \rho\, nsib_i \\ + \eta\, mote\text{-}m_i + \theta\, mote\text{-}h_i + \kappa\, mote\text{-}p_i + \lambda\, mote\text{-}u_i + \mu\, mote\text{-}g_i + u_i \quad (2)$$

where *gend* is the gender of the student such that it is 1 if the student is male, and 0 otherwise; *edct* is the monthly expenditure on after-school private lessons measured in 10,000 KRW. Here, the private lessons include lessons for all subjects; *nsib* stands for the number of siblings; and *mote-x* is students' mothers' education level. Here, the attachment *x* to *mote* denotes the education level. That is, *m, h, p, u*, and *g* denote middle school, high school, polytechnic school, university, and graduate school, respectively. For example, *mote-m* is 1 if the student's mother receives up to middle school education, and 0 otherwise.

If a student's mother's education is less than the middle school level, we omit the associated dummy variable to avoid the dummy variable trap.

These additional explanatory variables are included in an effort to further detail the explanatory variables on the right side. First, according to Geary, Hoard, and Nugent [10] and Gneezy, Niederle, and Rustichini [11] among others, male and female students have relative advantages in different environments and may have different scores over different disciplines, yielding a statistically significant gender effect. Below, we empirically examine this aspect by including the gender dummy variable. Second, we include the monthly expenditure on private lessons (*edct*) and the number of siblings (*nsib*) to decompose the explanatory power of parent's income on academic performance scores in Model (1). If parent's income is spent on children's private lessons, we can partly explain the effect of income on academic performance, although it cannot replace the whole explanatory power of the parent's income. Likewise, if there are multiple children in a family, the family resources for children's education is restricted for each child, so that the effect of income on the academic performance score diminishes as the number of siblings increases. We, therefore, include the number of siblings as another explanatory variable. Finally, mother's education levels are included to detect the effect of parental involvement. According to Shumow, Lyutykh, and Schmidt [12] and Glick and Hohmann- Marriott [13] among others, students' academic performance is affected by parental involvement. There can be many forms of parental involvement, and we measure it through mothers' education levels. This is mainly because parental involvement is not easy to measure with an objective standard. Shumow, Lyutykh, and Schmidt [12] and Glick and Hohmann-Marriott [13] also measure parental involvement using the parents' education levels.

We further extend Model (2) by including additional explanatory variables on the right side as follows:

$$
\begin{aligned}
\log(scor_i) = {}& \alpha + \beta\, phyh_i + \gamma\, othh_i + \delta \log(incm_i) + \pi\, gend_i + \xi\, edct_i + \rho\, nsib_i \\
& + \eta\, mote\text{-}m_i + \theta\, mote\text{-}h_i + \kappa\, mote\text{-}p_i + \lambda\, mote\text{-}u_i + \mu\, mote\text{-}g_i \\
& + \sigma\, fate\text{-}m_i + \tau\, fate\text{-}h_i + \varphi\, fate\text{-}p_i + \psi\, fate\text{-}u_i + \omega\, fate\text{-}g_i + u_i,
\end{aligned}
\tag{3}
$$

where *fate-x* indicates fathers' education levels. Here, the attachment *x* to fate represents the education level in the same way as the mother's education level.

As before, these additional indicator variables are expected to reduce the variation of the academic performance. In particular, fathers' education levels are included to detect the mutual effect of fathers' education on academic performance along with mothers' education in terms of parental involvement. As described above, Shumow, Lyutykh, and Schmidt [12] and Glick and Hohmann-Marriott [13] among others posit an effect of parental involvement with children's academic performance. We include fathers' education levels as another form of parental involvement to compensate for the effect of mothers' education levels on academic performance scores.

The main purpose of estimating these extended models using additional explanatory variables is in affirming the estimation results on the structural equation. If the structural Model (1) is consistently estimated with TSLS estimation, similar estimates should be obtained from Models (2) and (3). If the estimated signs of the structural parameter are substantially different among Models (1), (2), and (3), the TSLS estimation cannot provide a robust estimate. By comparing the estimated parameters, we can affirm the validity of our estimates obtained from the TSLS method.

3. Results and Discussion

3.1. Estimation and Inference Using Model (1)

We report the estimation results for Model (1) in Table 2. As described above, we obtain the TSLS parameter estimates by two-step regressions using 2099 observations. For comparison, we also report the estimation results obtained by the OLS method. The second and third columns report the OLS and TSLS estimation results, respectively. Figures in

parentheses are the *p*-values of the *t*-test statistics that test zero coefficients for the estimated parameters.

Table 2. Estimation results.

Variables	Model (1)	Model (1)	Model (2)	Model (2)	Model (3)	Model (3)
Estimation	OLS	TSLS	OLS	TSLS	OLS	TSLS
Dep. Var.	log(*scor*)	log(*scor*)	log(*scor*)	log(*scor*)	log(*scor*)	log(*scor*)
const	2.9007	2.8455	3.3814	3.2854	3.4531	3.4118
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
phyh	−0.0003	0.0045	−0.0003	0.0077	−0.0003	0.0078
	(0.0282)	(0.0206)	(0.0346)	(0.0172)	(0.0431)	(0.0174)
othh	0.0005	0.0005	0.0003	0.0003	0.0003	0.0003
	(0.0000)	(0.0000)	(0.0000)	(0.0015)	(0.0000)	(0.0025)
log(*incm*)	0.1516	0.1543	0.0830	0.0856	0.0706	0.0610
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0105)
gend			−0.0190	−0.0663	−0.0191	−0.0673
			(0.1555)	(0.0134)	(0.1525)	(0.0130)
edct			0.0018	0.0018	0.0018	0.0017
			(0.0000)	(0.0000)	(0.0000)	(0.0000)
nsib			−0.0180	−0.0217	−0.0199	−0.0264
			(0.1003)	(0.1645)	(0.0696)	(0.0955)
mote-m			−0.0606	0.0036	−0.0461	0.0465
			(0.2906)	(0.9630)	(0.4445)	(0.6247)
mote-h			0.0500	0.1105	0.0497	0.1023
			(0.1336)	(0.0506)	(0.1390)	(0.0687)
mote-p			0.1264	0.1121	0.1162	0.1038
			(0.0008)	(0.0904)	(0.0028)	(0.1140)
mote-u	(0.0000)	(0.0000)	0.1520	0.2095	0.1289	0.1794
			(0.0000)	(0.0003)	(0.0004)	(0.0023)
mote-g			0.1438	0.1598	0.1182	0.1169
					(0.0365)	(0.2275)
fate-m					0.0039	−0.0068
					(0.9457)	(0.9453)
fate-h					0.0285	0.0950
					(0.3963)	(0.0884)
fate-p			(0.0086)	(0.0882)	0.0384	0.0749
					(0.3355)	(0.2483)
fate-u					0.0650	0.1232
					(0.0689)	(0.0358)
fate-g					0.0757	0.1693
					(0.1064)	(0.0176)
Sample size	2099	2099	2083	2083	2083	2083
R^2	0.1175		0.1628		0.1651	
J-test stat.		7.2832		3.0864		3.3161
		(0.0262)		(0.2136)		(0.1905)

Figures in parentheses are the *p*-values of the test statistics computed by robust standard error. Each model is estimated by both OLS and TSLS estimations, where the TSLS estimation employs log(*heit*), log(*weit*), log(*slip*), and the other explanatory variables in the model except *phyh* as the instrumental variable.

We can summarize the estimation results as follows. First, the most significant result is that the OLS method estimates the coefficient of the after-school exercise hours differently from that obtained by the TSLS method. Note that the TSLS method estimates a positive coefficient with *p*-value almost equal to zero, whereas the OLS method estimates a negative coefficient with almost zero *p*-value. Although the estimated coefficients are close to zero, they are statistically very significant. This implies that the OLS estimation cannot reveal the structural relationship between students' after-school exercise hours and their academic performance scores due to the correlation between the after-school exercise hours and the error term. On the other hand, the TSLS method estimates a positive coefficient, implying

that the direct causal effect of the after-school exercise hours on the academic performance score is positive, as predicted by Tomporowski et al. [2], de Greeff, Bosker, Oosterlaan, Visscher, and Hartman [15], Chang et al. [16], Álvarez-Bueno et al. [17], Xiang, et al. [18], and Tomporowski, McCullick, Pendleton, and Pesce [19] among others. These different estimates imply that other indirect factors must affect the students' academic performance that is not captured by the variables in the model. That is, the other factors are associated with the error term. The negative result obtained from the OLS method implies that the indirect effect associated with the error term dominates over the direct effect obtained by the TSLS estimation, reversing the positive sign obtained from the TSLS method. Given the reality of standard Korean middle school students, this negative value is quite plausible. Most Korean middle students spend their after-school time studying independently or taking private lessons to compensate for weak school subjects. Given this fact, if they spend their after-school time on exercise, it will reduce the likelihood that the scores of other subjects would increase, leading to an overall negative sign in the model. In addition, many other factors could have caused the negative sign in the OLS estimates.

Second, the parameter estimates of the other explanatory variables are not very different between the OLS and TSLS estimations. Note that the coefficients of *const*, *othh*, log(*incm*) are similar in the OLS and TSLS estimates. Furthermore, all of the coefficients are statistically significant, implying that they have the power to explain the variation of academic performance scores. The coefficient of *othh* is strictly positive and significant for the model estimated by both OLS and TSLS estimations. This implies that taking private lessons is helpful in raising students' academic performances. In addition, when *phyh* and *othh* are given, the model estimates by both OLS and TSLS methods imply that a 1 percentage increase of parents' yearly income increases students' academic performance scores by about 15%. Although the academic performance score is inelastic to parent's yearly income, this does not imply that parents' income is insignificant to students' academic performance scores. The wealthier parents are, the more opportunities are available to raise the student's academic performance score. For example, students with wealthier parents can take more expensive but more efficient private lessons to raise their academic performance scores. This example also conversely implies that income elasticity may change if other explanatory variables are included that affect the effect of income on academic performance scores. For example, if the cost of private lessons is included in the right side, it can partially explain the effect of parents' income on academic performance scores. As another example, if the number of siblings is included in the right side, it will reduce the effect of parents' income on the academic performance score. We below examine how the estimated income elasticity responds to the existence of other variables on the right side.

Third, we examine the goodness-of-fit aspect of our model. When the OLS estimation is applied, R^2 is approximately 11.75%, implying that the variation of the academic performance score can be explained by further including other appropriate explanatory variables. We below examine how R^2 is modified by including the right-hand-side variables in Models (2) and (3). When the TSLS estimation applies, we examine the *J*-test statistic as described above and obtain 0.0262 as its *p*-value that appears to be in a grey area. Using this *p*-value, it is difficult to accept or reject the hypothesis that the error term is uncorrelated with the instrumental variables. We postpone this examination for the moment until we discuss the other model extensions.

3.2. Estimation and Inference Using Model (2)

We next discuss the estimation results for Model (2) reported in Table 2. The third and fourth columns in Table 2 contain the estimation results made by the OLS and TSLS methods, respectively. The estimation results can be summarized as follows. First, the same signs are obtained for the structural parameter as for the estimates using Model (1) by both OLS and TSLS methods. Even the coefficient values are more or less similar between the OLS and TSLS estimations. This implies that the model estimations are robust despite the model extension. In particular, the same coefficient value is obtained for the

after- school exercise hours from the OLS estimation. On the other hand, the positive relationship obtained by the TSLS estimation is further intensified and becomes more statistically significant. This implies that the positive relationship between after-school exercise hours and students' academic performance scores is better affirmed by this model estimation. Furthermore, the negative coefficient estimated by the OLS method also affirms the correlation between the structural error and after-school exercise hours.

Second, the estimated income elasticity of the academic performance score is significantly different from those in Model (1). Note that the estimated elasticity drops about half from about 0.15 to about 0.08, implying that the newly included explanatory variables in Model (2) partly explain the income effect in Model (1). As described above, the expenditure on after-school private lessons (*edct*) is correlated with parents' income level, and the number of siblings (*nsib*) also partly explains the effect of parents' income. If the number of siblings increases, we should expect that the effect of parents' income on academic performance scores diminishes. With these additional explanatory variables, parents' income effect is believed to decrease.

Third, the effects of the expenditure on private lessons and the number of siblings on academic performance scores are obtained as expected. When parent's income is spent on private lessons for children, its effect on children's academic performance scores increases, although the income effect decreases. On the other hand, given the level of parent's income, if the number of siblings increases, the available income effect is reduced, implying that the academic performance score decreases. Although the TSLS estimator reports that the effect of the number of siblings is statistically insignificant, the OLS estimator reports that the effect is relatively more significant.

Fourth, mothers' education on the students' academic performance scores also provides meaningful information. If a student's mother receives education up to the middle school level, it does not have a statistically significant effect on the student's academic performance score. For both models estimated by the OLS and TSLS methods, the estimated coefficients of *moth-m* and *moth-h* are statistically insignificant. On the other hand, if the student's mother has undertaken higher education, it positively and significantly affects the student's academic performance score. When the OLS method is applied, the coefficients of *mote-p*, *mote-u*, and *mote-g* are strictly positive and significant. If the student's mother receives education at the university level, the effect is the most influential as the coefficient value of *mote-u* is greater than the others. If the TSLS method is applied, the estimation results are slightly different from those found with the OLS method. Although estimation results are more or less similar to the OLS estimation case, the estimated coefficient of *mote-p* is in a statistically grey area. We postpone this examination until we discuss the estimations using Model (3).

Finally, the goodness-of-fit properties of the model have been substantially enhanced from the estimation of Model (1). Note that R^2 of the model estimated by the OLS method is about 0.16, which is greater than that obtained from Model (1), implying that the newly added explanatory variables have additional explanatory powers for the variation of the academic performance score. Furthermore, the *J*-test statistic reports 0.2136 as its *p*-value, implying that the employed instrumental variables are not correlated with the error term. From this aspect, the structural coefficient of after-school exercise hours must be properly estimated.

3.3. Estimation and Inference Using Model (3)

We finally discuss the estimation results for Model (3) reported in Table 2. The fifth and sixth columns in Table 2 report the estimation results obtained by the OLS and TSLS methods, respectively.

The estimation results can be summarized as follows. First, we can maintain the same model interpretation as for Model (2) estimation for *const, phyh, othh*, log(*incm*), *gend, edct, nsib*, and most dummy variables for mothers' education. We obtain exactly the same signs as before for these variables, and the magnitudes of the estimated coefficients are only

slightly different from the earlier case. Their statistical significance measured by *p*-values remains the same; the only difference is the coefficient of *mote-g*. For the model estimated with the OLS method, the obtained coefficient is statistically significant, but the TSLS method reports it as an insignificant coefficient. With this observation, the only statistically significant coefficient becomes that of *mote-u*.

Second, fathers' education levels have various effects on students' academic performance scores. When the OLS method is applied, all levels of fathers' education are found to be statistically insignificant. On the other hand, if the TSLS method is applied, the estimated coefficients of *fate-u* and *fate-g* are statistically significant. In particular, the coefficient of *fate-g* is more significant, and its magnitude is similar to that of *mote-g* in Model (2). Given the fact that the coefficient of *mote-g* in Model (3) is insignificant, the explanatory power of *mote-g* in Model (2) appears to have shifted to that of *fate-g* in Model (3).

Finally, the goodness-of-fit aspects of Model (3) are the same as those of Model (2). Note that the *J*-test statistic is about 3.3162 with *p*-value equal to 0.1905, implying that the model is estimated with TSLS estimation using proper instrumental variables. Therefore, we can be confident of the positive structural relationship between students' after-school exercise hours and academic performance scores. When the model is estimated with the OLS method, R^2 is approximately the same as that in Model (2). This aspect is already implied by the fact that the newly included dummy variables on fathers' educations are statistically insignificant. They are statistically meaningful only when the model is estimated by the TSLS estimation.

Before closing this section, we provide the final remark on the data. Our data set is collected from the 1st grade middle school students in the year of 2011, and the structural relationship might have changed since then as Lee [23] points out the causes of the recent rank decline of Korean students in PISA scores. Thus, it could be an interesting research topic to compare the structural relationship of this study with that using the most recent data to examine how the structural relationship evolves over time. Nevertheless, the structural model may not be identical to those of this study as there might have been external changes to the Korean educational environments since 2011. For such a case, new structural equations need to be specified and/or new instrumental variables may have to be employed to use the TSLS estimation. We leave this as a topic for future research.

4. Conclusions

This study estimates the structural relationship between the academic performance scores of Korean middle school students and their after-school exercise hours. Prior literature theoretically or experimentally predicts that these variables are positively associated, but it is difficult to empirically verify this association unless experimental data are collected by controlling for the mutual effects with other variables or estimating a full model by specifying the whole structure of all variables affecting the two variables in question. Unlike previous studies, this study estimates the structural relationship with a two-stage least squares estimation that does not require experimental observations collected for this particular purpose or estimating the full model. From this estimation, we empirically affirm that there is a positive structural relationship between after-school exercise hours and students' academic performance scores, whereas the ordinary least squares method consistently estimates a negative relationship. In addition, we examine the roles of other variables that affect students' academic performance scores, such as private lesson hours for school subjects other than exercise, parents' income levels, students' gender, private lesson cost, the number of siblings, and parents' education level. The estimated coefficients are interpreted from the viewpoints of prior literature and the current environment for Korean middle school students.

From the structural coefficient estimated as positive, we can have an insight on the physical education policy for students along with other subjects taught in the middle school. In Korea, the physical education class is often substituted to others such as reading, mathematics, and science that are believed to be more crucial for students' performance.

This conventional practice might have been justified by the negative relationship between the physical education and students' performance as revealed by OLS. Nevertheless, the current study using the TSLS estimation shows that their structural relationship is indeed positive, implying that the physical education class is a complement to other classes in increasing students' academic performance as it can be inferred from the fact that even the after-school exercise is helpful to students' academic performance. Furthermore, it also implies that education policy needs to be implemented to increase the regular physical education class hours in the middle school education in a way to comply with various demands from students so that the positive effect of the after-school exercise can be shared by the other students not attending the after-school exercise.

Author Contributions: Conceptualization, formal analysis, writing—original draft preparation, K.S.; methodology, supervision, S.Y.; writing—review and editing, M.K. All authors have read and agreed to the published version of the manuscript.

References

1. Dishman, R.K.; Berthoud, H.-R.; Booth, F.W.; Cotman, C.W.; Edgerton, R.; Fleshner, M.R.; Gandevia, S.C.; Gomez-Pinilla, F.; Greenwood, B.N.; Hillman, C.; et al. Neurobiology of exercise. *Obesity* **2006**, *14*, 345–356. [CrossRef] [PubMed]
2. Tomporowski, P.D.; Davis, C.L.; Miller, P.H.; Naglieri, J.A. Exercise and children's intelligence, cognition, and academic achievement. *Educ. Psychol. Rev.* **2008**, *20*, 111–131. [CrossRef] [PubMed]
3. Cappelen, A.W.; Charness, G.; Ekström, M.; Gneezy, U.; Tungodden, B. Exercise improves academic performance. Working paper series. *Inst. Ind. Econ.* **2017**. [CrossRef]
4. Tremblay, M.S.; Inman, J.W.; Willms, J.D. The Relationship between physical activity, self-esteem, and academic achievement. *Pediatr. Exerc. Sci.* **2000**, *12*, 312–323. [CrossRef]
5. Castelli, D.M.; Hillman, C.H.; Buck, S.M.; Erwin, H.E. Physical fitness and academic achievement in third- and fifth-grade students. *J. Sport Exerc. Psychol.* **2007**, *29*, 239–252. [CrossRef] [PubMed]
6. Dwyer, T.; Sallis, J.F.; Blizzard, L.; Lazarus, R.; Dean, K. Relation of academic performance to physical activity and fitness in children. *Pediatr. Exerc. Sci.* **2001**, *13*, 225–237. [CrossRef]
7. Rees, D.I.; Sabia, J.J. Sports participation and academic performance: Evidence from the national longitudinal study of adolescent health. *Econ. Educ. Rev.* **2010**, *29*, 751–759. [CrossRef]
8. Dills, A.K.; Morgan, H.N.; Rotthoff, K.W. Recess, physical education, and elementary school student outcomes. *Econ. Educ. Rev.* **2011**, *30*, 889–900. [CrossRef]
9. You, S.; Shin, K.; Kim, H. *Longitudinal Effects of Physical Education on Attention and Academic Achievement*; Mimeo: Memphis, TN, USA, 2013.
10. Geary, D.C.; Hoard, M.K.; Nugent, L. Boy's visuospatial abilities compensate for their relatively poor in-class attentive behavior in learning mathematics. *J. Exp. Child. Psychol.* **2021**, *211*, 105222. [CrossRef]
11. Gneezy, U.; Niederle, M.; Rustichini, A. Performance in competitive environments: Gender differences. *Q. J. Econ.* **2003**, *118*, 1049–1074. [CrossRef]

12. Shumow, L.; Lyutykh, E.; Schmidt, J.A. Predictors and outcomes of parental involvement with high school students in science. *Sch. Community J.* **2011**, *21*, 81–98.

13. Glick, J.; Hohmann-Marriott, B. Academic performance of young children in immigrant families: The significance of race, ethnicity, and national origins. *Int. Migr. Rev.* **2007**, *41*, 371–401. [CrossRef]

14. Jarque, C.; Beta, A.A. Test for normality of observations and regression residuals. *Int. Stat. Rev.* **1987**, *55*, 163–172. [CrossRef]

15. de Greeff, J.W.; Bosker, R.J.; Oosterlaan, J.; Visscher, C.; and Hartman, E. Effects of physical activity on executive functions, attention and academic performance in preadolescent children: A meta-analysis. *J. Sci. Med. Sport* **2018**, *21*, 501–507. [CrossRef]

16. Chang, Y.K.; Chi, L.; Etnier, J.L.; Wang, C.C.; Chu, C.H.; Zhou, C. Effect of acute aerobic exercise on cognitive performance: Role of cardiovascular fitness. *Psychol. Sport Exerc.* **2014**, *15*, 464–470. [CrossRef]

17. Álvarez-Bueno, C.; Pesce, C.; Cavero-Redondo, I.; Sánchez-López, M.; Garrido-Miguel, M.; Martínez-Vizcaíno, V. Academic achievement and physical activity: A meta-analysis. *Pediatrics* **2017**, *140*, e201714988. [CrossRef]

18. Xiang, M.; Gu, X.; Jackson, A.; Zhang, T.; Wang, X.; Guo, Q. Understanding adolescents' mental health and academic achievement: Does physical fitness matter? *Sch. Psychol. Int.* **2017**, *38*, 647–663. [CrossRef]

19. Tomporowski, P.D.; Mccullick, B.; Pendleton, D.M.; Pesce, C. Exercise and children's cognition: The role of exercise characteristics and a place for metacognition. *J. Sport Health Sci.* **2015**, *4*, 47–55. [CrossRef]

20. Beets, M.W.; Beighle, A.; Erwin, H.; Huberty, J. After-school program impact on physical activity and fitness: A meta-analysis. *Am. J. Prev. Med.* **2009**, *36*, 527–537. [CrossRef]

21. Nihayah, M.; Ismarulyusda, I.; Syarif, H.L.; Nurzakiah, M.S.; Baharudin, O.; Fadzil, M.H. Sleeping hours and academic achievements: A study among biomedical science students. *Procedia Soc. Behav. Sci.* **2011**, *18*, 617–621. [CrossRef]

22. White, H. A Heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* **1980**, *48*, 817–838. [CrossRef]

23. Lee, K.H. Analysis on the hypotheses on the causes of decline of basic leaning capabilities and discussion on the solutions. *Politics Educ.* **2021**, *28*, 37–61.