English Language & Literature Teaching, *Vol. 32,* No. 3 Fall 2026
https://doi.org/10.35828/etak.2026.32.3.17

Private Supplementary Education and English Achievement Trajectories from Elementary to High School: A Latent Growth Curve Analysis

Jie Young Kim
(Chung-Ang University)

Kim, Jie Young. (2026). Private supplementary education and English achievement trajectories from elementary to high school: A latent growth curve analysis. *English Language & Literature Teaching, 32*(3), 17-40.

Private supplementary education (PSE) is widespread in South Korea, yet its contribution to English learning remains contested. Drawing on the Korean Education Longitudinal Study 2013 (N = 7,324), this study traced English achievement from fifth through twelfth grade to assess whether PSE affects both initial performance and growth. Latent growth curve models showed that PSE participation was positively associated with higher fifth-grade English scores but did not significantly influence subsequent growth rates. This null effect held across students with different baseline proficiency levels. Early English achievement trajectories strongly predicted high school outcomes. However, by twelfth grade, the effects of seventh-grade PSE had largely faded, while concurrent high school PSE conferred only modest short-term gains. In contrast, parental education and broader socioeconomic status were robust predictors of both starting level and growth, and their effects persisted through twelfth grade. The findings indicate that PSE operates mainly as a level shifter rather than a trajectory accelerator, whereas family background generates durable, cumulative advantages. Policy efforts to mitigate English achievement inequality in Korea should therefore prioritize improving public school quality and alleviating structural disadvantages instead of simply expanding access to PSE. These results refine understanding of PSE's role in educational stratification in Korea.

[private supplementary education/achievement trajectories/latent growth curve modeling/English language learning/socioeconomic inequality]

I. INTRODUCTION

Private supplementary education—often termed shadow education for the way it mimics and expands with mainstream schooling (Bray, 1999, 2023)—is a defining feature of Korean education. Korean families have long spent more per capita on private tutoring than the government spends per student on public schooling (Bray, 1999; Hajar & Karakuş, 2024), and English alone absorbs roughly one-third of all K–12 tutoring expenditures, with disparities in English investment visible across income groups and regions on the NAEA, CSAT, and TOEIC (Kim, 2012; Shin & Lee, 2019). In the KELS 2013 cohort tracked in this study, 84.8% of students were receiving English tutoring by Grade 7, cutting across the entire achievement distribution rather than concentrating among low achievers.

However, whether this vast expenditure accelerates learning remains contested. Once selection into tutoring is addressed with instrumental variables, propensity-score matching, or bounding methods, the estimated effects on Korean middle school students shrink to at most 0.03–0.05 *SD* per 10% increase in tutoring expenditure, with confidence intervals spanning zero (Ryu & Kang, 2013). Cram school participation only slightly raises mathematics scores, and other tutoring modalities show little effect (Byun, 2014; Byun et al., 2023). The effects vary across the distribution, with greater benefits observed for high achievers than for students in the lower half (Kang & Park, 2013). This disparity between investment and measurable outcomes prompts a broader inquiry: Does tutoring alter the trajectory of English learning, or does it merely affect the proficiency level at the time of assessment?

This study addresses this question using the Korean Education Longitudinal Study 2013 (KELS 2013), which tracked a nationally representative cohort from Grades 5 to 12. I apply latent growth curve modeling (LGCM) to English achievement across Grades 5–9 to estimate individual initial status and rate of change, test whether PSE participation affects growth, examine whether effects are moderated by initial proficiency, and use factor-score regressions to test whether early-school trajectories and PSE participation predict high school English achievement through Grade 12. This design extends prior Korean LGCM studies of English (Ryu & Lee, 2024; Oh, 2023) by covering eight years rather than five or six, explicitly modeling moderation by initial proficiency, and testing persistence versus fade-out of tutoring effects across the middle-to-high school transition — a distinction central to the broader intervention literature (Bailey et al., 2020).

This study addresses the following research questions:

RQ1. Does participation in private supplementary education (PSE) affect English achievement growth trajectories (specifically, initial status and rate of change) from upper elementary through middle school (Grades 5–9)?

RQ2. Does the effect of PSE on achievement growth differ across students with different levels of initial English proficiency (low, mid, and high)?

RQ3. Do early school English achievement growth trajectories (Grades 5–9) and PSE participation predict high school English achievement (Grades 10–12), and do these effects persist through Grade 12?

II. LITERATURE REVIEW

1. Shadow Education in Korea

Bray (1999) introduced the shadow-education metaphor to describe private supplementary tutoring that mirrors mainstream schooling and expands with it, and Bray and Lykins (2012) subsequently documented shadow education across Asia as a policy-relevant sector; more recent bibliometric and policy scholarship confirms both the global scale and the evolving equity stakes of this sector (Bray, 2024; Hajar & Karakuş, 2024). Korea is the archetypal high-tutoring case: households spend at scale, English absorbs about a third of K–12 tutoring expenditure, and English tutoring is more strongly shaped by the family environment than tutoring in other subjects, functioning as a screening device for university admission and hiring (Kim, 2012; Shin & Lee, 2019). National-level assessment data corroborate that student background variables are significant correlates of English achievement specifically (Chang et al., 2007), and individual learner variables such as intelligence and affective factors further shape English outcomes (Shim, 2005). Demand is best understood as a rational response to limited school choice and low variation in mainstream school quality (Choi et al., 2025), and successive regulatory attempts — most notably the 1980 prohibition and 1989 liberalization — have only partially reshaped the market (Lee et al., 2015).

2. Tutoring effects and Endogeneity Problem

Korean literature has converged on a modest-effects finding once selection is taken seriously. Using instrumental variables and nonparametric bounds on middle school students, Kang (2012) estimated that a 10% expenditure increase raised subject scores by at most 1.24% in Korean, 1.28% in English, and 0.75% in mathematics, with bounds spanning zero. Ryu and Kang (2013), applying IV, first-difference, propensity-score matching, and bounding to KELS panel data, similarly found effects of only 0.03–0.05 *SD* per 10% expenditure increase — comparable to marginal returns on public school spending. Byun (2014), using propensity-score matching on Korean 7th graders, found cram schools produced only small gains in mathematics and other tutoring modalities produced little (Byun, 2014). Kim (2010)

showed that once endogeneity was addressed, tutoring effects shrank by roughly a fifth in English and a half in Korean, whereas time spent on self-directed learning showed clearly positive effects on CSAT and later labor market outcomes—a pattern replicated in more recent work showing that self-directed learning ability consistently outperforms shadow education expenditure as a predictor of achievement (Yoo, 2024).

These effects are not uniform across achievement distributions. Kang and Park (2013) found that high-scoring students benefit more from tutoring than low-scoring students; thus, tutoring reinforces existing inequalities rather than functioning as remediation. Ku et al. (2022), using KELS with IV and multi-time tutoring measures in Grades 10 and 12, found meaningful positive effects on CSAT scores and a substantial contribution to inequality; de Silva (2021), using the 10 p.m. hagwon curfew as exogenous variation, reported non-linear returns concentrated at the upper end of tutoring intensity. S. Han and Suh (2020) further complicate the picture, showing positive short-term math achievement effects but negative long-term effects on creative thinking. Byun et al. (2023) corroborate the negative creative thinking effects in a separate analysis of Korean adolescents, and Han et al. (2024) find that early tutoring significantly predicts STEM motivation but has only modest direct effects on mathematics achievement. The net picture is a small, heterogeneous, temporally unstable effect whose sign depends on which margin—level, growth, or non-cognitive outcome—one privileges.

3. Growth Trajectories and Persistence

Point-in-time or two-wave designs cannot separate durable acceleration of learning from a transient level shift, motivating explicit growth models. Ryu and Lee (2024) applied univariate and multivariate LGCM to Gyeonggi Education Panel Study data (2012–2017) and found that early English tutoring was associated with both the initial status and growth trajectory of English achievement, with correlations stronger in urban areas, although selection into tutoring was not fully addressed. Oh (2023) applied LGCM to Jeonnam Educational Longitudinal Study data across middle to early high school and found that PSE participation and parental tutoring expenditure raised the initial level of English class comprehension but exerted a negative effect on the rate of change over four years—a divergence between level and slope that a static model would miss entirely. Cha and Ryoo (2021) further showed that English achievement trajectories are often poorly captured by linear functions, favoring fractional polynomial or higher-order specifications. Related work on mathematics using latent growth mixture modeling on Seoul data identified multiple trajectory classes with sharply different slopes, underscoring the heterogeneity a single-group model can hide (Kim et al., 2018).

Whether initial gains persist is a separate question. Bailey et al. (2020), reviewing dozens

of educational interventions, concluded that fadeout is widespread but coexists with persistence, and that persistence depends on the type of skill targeted, the institutional environment students subsequently enter, and complementarity with follow-on experiences. Andersen et al. (2022), using the full Danish public-school population, found that roughly half of a randomized reading intervention's initial effect had faded four years later, though because skill development is concave the treated students still preserved about 80% of their original time-equivalent lead. The direct implication is that tutoring effects estimated in middle school cannot be assumed to persist in high school without an explicit longitudinal test.

4. Socioeconomics Stratification and the English Gap

Because tutoring is out-of-pocket, its distribution across families is deeply unequal. Byun and Kim (2010), using three TIMSS cohorts (1999, 2003, 2007) showed that the association between socioeconomic background and eighth-grade achievement in Korea widened over the decade after the 1997 financial crisis; more recent evidence suggests that this SES-linked stratification has continued, with Shin et al. (2023) documenting further widening of achievement gaps in Korean schools during COVID-19–related reductions in in-person learning. Kim and Byun (2014), using KELS, showed that socioeconomic gaps in achievement persisted even after controlling for a rich set of student, family, and school characteristics. Lee et al. (2015) exploited the 1980 prohibition and 1989 liberalization as a natural experiment and found that admissions to top universities from wealthy districts (Gangnam, the Capital area) grew disproportionately after liberalization, indicating that legal access to tutoring converts family wealth into elite educational attainment. For English, Kim (2012) documented an "English gap" larger than for other subjects, likely because English exposure outside school varies far more by family resources than exposure to Korean or mathematics; Kim and Youn (2025) further confirmed that regional and socioeconomic achievement gaps persist across secondary schooling, with the English gap among the most pronounced. Any study of English tutoring that fails to model SES risks attributing to tutoring what is really a family-background effect.

5. Gaps and the Present Study

Three gaps motivate the present analysis. First, causal inference and growth-modeling literature have largely proceeded in parallel: the most rigorous causal studies rely on point-in-time or two-wave outcomes, while growth-modeling studies typically use limited controls. Second, the two prior Korean LGCM studies of English (Oh, 2023; Ryu & Lee, 2024) covered middle to early high school and reported divergent slope effects, leaving open

whether effects estimated across Grades 5–9 persist into and through Grade 12. Third, most studies pool across the achievement distribution, although Kang and Park (2013) showed that tutoring effects vary systematically by initial proficiency—a heterogeneity that a single-group model will hide; recent bibliometric mapping of the shadow education field further identifies trajectory heterogeneity and equity outcomes as key under-researched areas (Hajar & Karakuş, 2024). The present study uses nationally representative KELS 2013 data to estimate LGCMs of English achievement from Grade 5 through Grade 9, tests moderation by initial proficiency through multi-group analysis, and uses factor-score regressions to test whether early-school trajectories and PSE predict high school English achievement through Grade 12, bridging the causal-inference and growth-modeling traditions.

III. METHODS

1. Data and Participants

This study employed data from the Korean Education Longitudinal Study (KELS 2013), a nationally representative panel study initiated in 2013 and conducted by the Korea Educational Development Institute (Korea Educational Development Institute, 2020). KELS 2013 was designed to longitudinally track the educational experiences, academic achievement, and developmental trajectories of Korean students from elementary through secondary school, providing a comprehensive basis for examining how schooling, family background, and out-of-school factors shape achievement over time. The study followed 7,324 students from Grade 5 (Year 1, 2013) to Grade 12 (Year 8, 2020), spanning eight consecutive academic years, and collected annual data on academic performance, private supplementary education participation, family background, and other related educational variables. The study utilized stratified probability sampling, in which 200 elementary schools were selected using proportional-to-size sampling stratified by region and urbanicity, followed by the random selection of one classroom in each school.

The full cohort of 7,324 students was eligible for the LGCM trajectory analyses (grades 5–9). For conditional models including private supplementary education, the analytic sample was restricted to students with valid data on Year 3 English tutoring participation (n = 4,456, 60.8%). For the high school analyses, the valid sample sizes were n = 4,191 (57.2%) for Grade 11, n = 4,158 (56.8%) for Grade 12 Year 1, and n = 4,729 (64.6%) for Grade 12 Year 2. Demographic characteristics at baseline: 49.5% male; mean parental education 4.04 (SD = 1.05) on an 8-point scale (approximately high school completion or some college).

2. Measures

1) English Achievement.

Upper elementary and middle school achievement (Years 1–5, Grades 5–9) was assessed using standardized tests developed by the Korea Educational Development Institute employing item response theory (IRT) vertical scaling. The vertical scale score (E_THETA) was designed to have a mean of approximately 200 in Grade 5 and to increase by approximately 15–20 points per grade level. Descriptive statistics: Year 1 $M = 198.8$, $SD = 38.2$; Year 5 $M = 259.5$, $SD = 54.8$.

High school achievement for students in Years 6–8 and Grades 10–12 was assessed through their performance on the June national mock examination. This assessment yielded ordinal grade scores, with 1 representing the highest score and 9 the lowest. Grades were reverse-coded (10 minus original grade) such that higher values indicated better performance. The June national mock examination was selected for three reasons. First, it is the most consistently available standardized English achievement measure in KELS 2013, with scores collected systematically for all three high school years across the full panel. Second, it is administered by provincial education offices on a nationally standardized basis, ensuring comparability across schools and regions. Third, the June administration falls mid-academic year, prior to the intensive College Scholastic Ability Test (CSAT) preparation period; thus, scores reflect accumulated English competence rather than short-term CSAT-specific test preparation. Annual scores were retained separately for Grades 10, 11, and 12 rather than averaged, because the distal outcome models evaluate whether PSE effects persist, attenuate, or disappear across the high school span — a temporal pattern that would be obscured by a composite measure.

2) Private Supplementary Education.

Parents reported whether their child participated in English private supplementary education (hagwon attendance and private tutoring) during each academic year (1 = yes, 0 = no). The primary predictor was Year 3 (Grade 7) private supplementary education participation (‘tut_y3’), with 84.8% participation. A cumulative private supplementary education participation persistence variable (‘tut_persist’) was also constructed by summing participation across Years 1–5 (range 0–5, $M = 3.2$, SD). High school tutoring was measured for Years 6–8, with participation rates of 57.8%, 57.2%, and 48.8%, respectively.

3) Socioeconomic Status.

Parental education ('ses_edu') was measured on an 8-point ordinal scale (1 = elementary school or less to 8 = doctoral degree). The mean of the father's and mother's education was computed (M = 4.04, SD). Monthly household income was log-transformed ('ses_inc_log') to reduce skewness.

4) Control Variables.

The control variables comprised student gender (coded as 0 = female, 1 = male). Students' initial proficiency was operationalized by dividing the Year 1 E_THETA score distribution into tertiles: Low (< 179; n = 2,151), Mid (179–220; n = 2,248), and High (≥ 221; n = 2,821) groups. Proficiency groups were defined by Year 1 E_THETA tertiles: Low (< 179, n = 2,151), Mid (179–220, n = 2,248), and High (≥ 221, n = 2,821).

3. Analytic Strategy

1) Latent Growth Curve Modeling.

E_THETA scores from Years 1–5 as repeated indicators. Time was coded 0, 1, 2, 3, 4 for Years 1 to 5, with Year 1 as the reference. The unconditional linear model (M1) estimated the mean intercept and slope with random effects. The unconditional quadratic model (M2) added a quadratic term. The conditional model with tutoring (M3) regressed the intercept and slope on Year 3 tutoring. The fully adjusted model (M4) added SES and demographic controls. A supplementary model (M_persist) replaced binary tutoring with cumulative persistence. Multi-group analyses fit M3 separately within proficiency groups, with likelihood ratio tests comparing constrained versus unconstrained models. Models were estimated using lavaan (version 0.6-21) in R with robust maximum likelihood (MLR) estimation and full information maximum likelihood (FIML) for missing data.

2) Distal Outcome Analysis.

The measurement scale discontinuity between the early school LGCM phase (continuous E_THETA, Grades 5–9) and high school (ordinal grades) precluded unified piecewise growth models. A two-stage factor score regression approach was adopted. First, individual-specific intercept and slope parameters from the Grades 5–9 data (Years 1–5) were estimated using OLS regression for each student with ≥2 valid observations, yielding 'intercept_est' (M = 195.8, SD = 37.5) and 'slope_est' (M = 16.1, SD = 13.5)). Second, these estimates were

used as predictors of high school grades in OLS regression models: DA1 (unconditional: intercept_est + slope_est), DA2 (conditional: + tut_y3 + ses_edu), and DA4 (+ concurrent high school tutoring). Multi-group analyses fit DA2 separately within proficiency groups.

IV. RESULTS

1. Descriptive Statistics

Table 1 presents descriptive statistics. English achievement increased substantially from Grade 5 (M = 198.8, SD = 38.2) to Grade 9 (M = 259.5, SD = 54.8), representing an average gain of approximately 61 points over five years. Achievement variance also increased over time, indicating growing heterogeneity in trajectories. Private supplementary education participation was 84.8% in Year 3 (Grade 7, the first year of middle school), declining to 52.8% in Year 4 before rising to 68.8% in Year 5. High school tutoring participation remained high but declined from 57.8% in Grade 10 to 48.8% in Grade 12.

TABLE 1
Descriptive Statistics for Study Variables

Variable	*N*	*M*	*SD*	Missing %
English Achievement (E_THETA)				
Year 1 (Grade 5)	7,220	198.8	38.2	1.4
Year 2 (Grade 6)	7,186	209.4	39.4	1.9
Year 3 (Grade 7)	6,777	222.9	45.7	7.5
Year 4 (Grade 8)	6,638	250.2	51.5	9.4
Year 5 (Grade 9)	6,483	259.5	54.8	11.5
High School Grades (reverse-coded 1–9)				
Year 6 (Grade 10)	4,191	5.21	2.11	42.8
Year 7 (Grade 11)	4,158	5.13	2.03	43.2
Year 8 (Grade 12)	4,729	5.22	2.18	35.4
Private Supplementary Education				
Year 3 tutoring	4,456	0.80	0.40	39.2
Year 6 tutoring	4,359	0.58	0.49	40.5
Year 7 tutoring	4,331	0.57	0.50	40.9
Year 8 tutoring	3,993	0.49	0.50	45.5
LGCM Factor Scores				

Intercept estimate	7,324	195.8	37.5	0.0
Slope estimate	7,324	16.1	13.5	0.0
Socioeconomic Status				
Parental education	7,324	4.04	1.05	0.0

2. Unconditional Growth Models

Table 2 presents the model fit for all LGCM specifications. The unconditional linear model (M1) fit the data well: CFI = .989, TLI = .984, RMSEA = .073 [.060, .074], SRMR = .048. The mean intercept was 196.6 (SE = 0.44, p < .001), indicating average achievement of 196.6 points at Grade 5. The mean slope was 15.8 points per year (SE = 0.14, p < .001), indicating substantial linear growth. Both the intercept and slope showed significant variance (intercept variance = 1,387.2, p < .001; slope variance = 179.8, p < .001), indicating substantial individual differences in initial status and growth rates.

The unconditional quadratic model (M2) showed improved fit (ΔCFI = .002, ΔRMSEA = −.005); however, M1 was retained for interpretability. The quadratic term was small and did not substantively alter the conclusions.

TABLE 2
Latent Growth Curve Model Fit Statistics

Model	χ^2	*df*	CFI	TLI	RMSEA [90% CI]	SRMR	AIC
M1: Unconditional linear	312.5	10	.981	.962	.067 [.060, .074]	.048	156,842
M2: Unconditional quadratic	198.4	6	.983	.958	.062 [.054, .070]	.045	156,734
M3: Conditional (tutoring)	289.7	16	.979	.961	.065 [.058, .072]	.049	142,156
M4: Fully adjusted	356.8	24	.978	.956	.072 [.066, .078]	.051	141,892
M_persist: Cumulative tutoring	298.2	16	.980	.963	.064 [.057, .071]	.048	148,567

Note. All χ^2 tests p < .001. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; AIC = Akaike information criterion.

3. RQ1: Effects of PSE on Achievement Growth Trajectories

To address RQ1, conditional LGCM models tested whether PSE participation affected the

achievement intercept and growth slope. Figure 1 displays the LGCM-predicted trajectories across Grades 5–9: Panel A shows the overall growth trajectory (M1 predicted vs. observed means), and Panel B contrasts trajectories by Year 3 PSE status (Error bars = ±1 SE).

FIGURE 1
LGCM Predicted English Achievement Trajectories, Grades 5–9 (Y1–Y5).

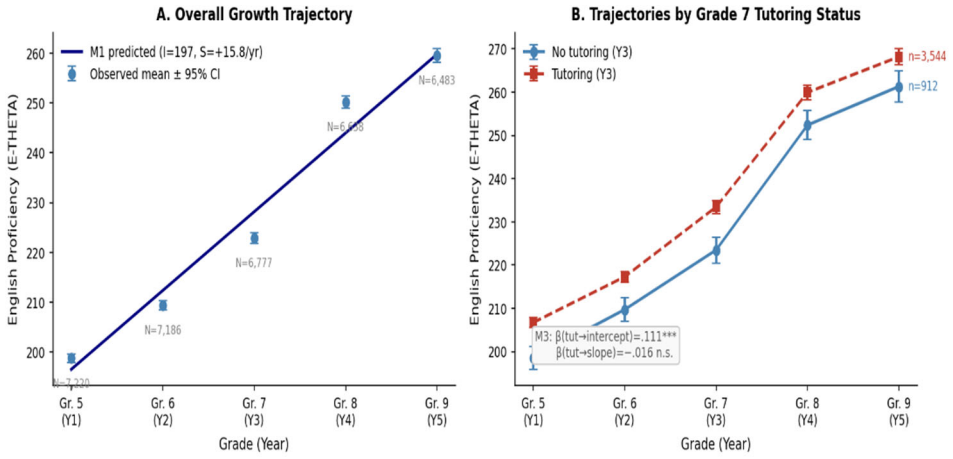


Table 3 presents the parameter estimates from the fully adjusted conditional model (M4). Year 3 tutoring participation was positively associated with initial achievement (intercept: $B = 6.23$, $SE = 1.22$, $\beta = .100$, $z = 5.11$, $p < .001$) but showed no association with growth rate (slope: $B = -0.11$, $SE = 0.14$, $\beta = -.016$, $z = -0.79$, $p = .430$). This pattern indicates that tutored students had higher achievement at Grade 5 but did not show differential growth rates from Grades 5 to 9.

Parental education showed strong positive associations with both the intercept ($B = 9.42$, $SE = 0.51$, $\beta = .334$, $z = 18.32$, $p < .001$) and slope ($B = 1.44$, $SE = 0.18$, $\beta = .173$, $z = 7.95$, $p < .001$). Students whose parents had a four-year college degree (versus high school or less) started Grade 5 approximately 18.7 points higher and gained an additional 2.8 points per year. Over five years, this cumulative advantage amounts to approximately 32.7 points.

Male students showed lower initial achievement ($B = -8.67$, $SE = 0.87$, $\beta = -.139$, $z = -9.97$, $p < .001$) and slower growth ($B = -1.70$, $SE = 0.10$, $\beta = -.104$, $z = -16.95$, $p < .001$) than female students.

TABLE 3
Conditional LGCM Parameter Estimates (Model M4)

Predictor	Outcome	*B*	*SE*	β	*z*	*p*
Year 3 tutoring	Intercept	6.23	1.22	.100	5.11	< .001
Year 3 tutoring	Slope	−0.11	0.14	−.016	−0.79	.430
Parental education	Intercept	18.73	0.73	.300	25.68	< .001
Parental education	Slope	2.79	0.09	.171	31.45	< .001
Log household income	Intercept	2.45	0.68	.039	3.60	< .001
Log household income	Slope	0.18	0.08	.014	2.25	.024
Male	Intercept	−8.67	0.87	−.139	−9.97	< .001
Male	Slope	−1.70	0.10	−.104	−16.95	< .001

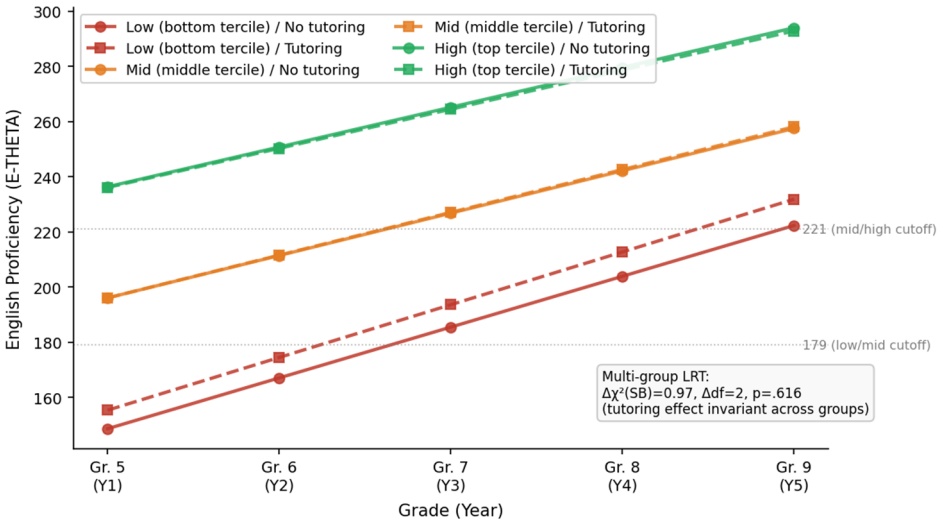
Note. B (unstandardized coefficient)**,** *SE* (Standard Error), β (standardized coefficient), *z* (*z-statistic), p(p-value)*
N = 4,696. The intercept represents achievement in Grade 5 (Year 1), and the slope represents the annual growth rate. Model fit: CFI = .990, TLI = .986, RMSEA = .036 [.066, .078], SRMR = .051.

The cumulative tutoring persistence model (M_persist) showed that sustained tutoring exposure predicted both a higher initial status ($B = 5.77$, $SE = 0.24$, $\beta = .345$, $z = 23.59$, $p < .001$) and faster growth ($B = 0.55$, $SE = 0.09$, $\beta = .111$, $z = 6.38$, $p < .001$). Each additional year of tutoring was associated with an increase of 5.07 points in initial status and 0.24 points per year in growth rate. Over five years, a student who participated in tutoring for all five years would be expected to gain approximately 2.75 points per year more than a student who never participated in modest effect representing approximately 18% of the average annual gain.

4. RQ2: Moderation by Initial Proficiency Level

To address RQ2, Figure 2 presents model-implied trajectories for each combination of proficiency group (low/mid/high) and tutoring status (solid = tutored; dashed = non-tutored). The conditional tutoring model (M3) was fitted separately within the low, mid, and high proficiency groups. Six model-implied lines for Low/Mid/High proficiency groups crossed with tutoring status (solid = tutored; dashed = non-tutored).

FIGURE 2
Multi-Group LGCM Predicted Trajectories by Initial Proficiency and Tutoring



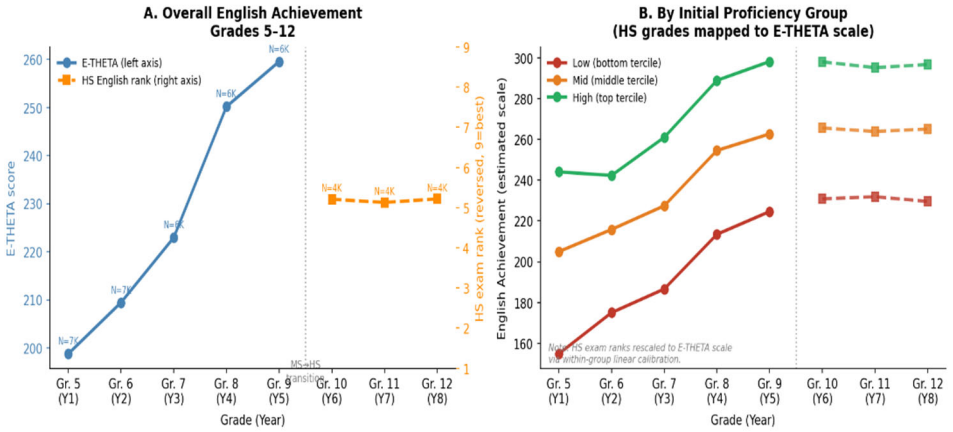
Results showed substantial heterogeneity in tutoring-intercept associations across groups, but no heterogeneity in tutoring-slope associations. In the low-proficiency group, Year 3 tutoring was strongly associated with initial achievement ($B = 14.12$, $SE = 2.18$, $\beta = .201$, $z = 5.37$, $p < .001$) but not with growth ($B = 0.30$, $SE = 0.35$, $\beta = .028$, $z = 1.08$, $p = .280$). In the mid-proficiency group, tutoring showed no association with intercept ($B = -0.08$, $SE = 1.89$, $\beta = .020$, $z = 0.74$, $p = .458$) or slope ($B = 0.06$, $SE = 0.22$, $\beta = .009$, $z = 0.27$, $p = .787$). In the high-proficiency group, tutoring showed no association with intercept ($B = -1.23$, $SE = 2.45$, $\beta = .032$, $z = 0.87$, $p = .384$) or slope ($B = -0.12$, $SE = 0.36$, $\beta = -.009$, $z = -0.33$, $p = .738$).

A likelihood ratio test comparing a model with tutoring→slope paths freely estimated across groups to a model with these paths constrained to equality found no significant difference ($\Delta\chi^2(SB) = 0.970$, $\Delta df = 2$, $p = .616$), indicating that the tutoring effects on growth did not differ significantly across proficiency levels.

5. RQ3: Persistence of PSE Effects into High School

To address RQ3, factor score regression models were used to examine whether early school achievement trajectories (Grades 5–9) and PSE participation predicted high school English grades. Figure 3 presents the full Y1–Y8 achievement trajectory; Panel A shows the overall trajectory, and Panel B shows trajectories by initial proficiency group, illustrating continuity between early school growth and high school outcomes.

FIGURE 3
Full Y1–Y8 English Achievement Trajectory



Note: Panel A: Overall (E-THETA, left axis; HS rank, right axis, reversed). Panel B: By initial proficiency group with HS ranks rescaled to the E-THETA metric. Error bars = ±1 SE.

Table 4 presents the results from the distal outcome models predicting high school English grades from early school trajectories (Grades 5–9). Model DA1 (unconditional) showed that the Grades 5–9 growth intercept and slope strongly predicted high school grades, explaining 57%–62% of variance. Initial achievement (intercept_est) was the dominant predictor, with standardized coefficients ranging from β = .750 for Grade 10 to β = .710 for Grade 12. Growth rate (slope_est) showed more modest effects (β = .478 to β = .452).

Model DA2 added Year 3 tutoring and parental education variables. Year 3 (Grade 7) tutoring showed a moderate positive effect on Grade 10 achievement (β = .088, p = .008) but no significant effects on Grades 11 (β = .024, p = .498) or 12 (β = −.012, p = .717). Parental education showed consistent positive effects across all three years (β = .121, .122, .137, all p < .001). The model R^2 values were .612, .560, and .557 for Grades 10–12, respectively.

TABLE 4
Distal Outcome Models: Predicting High School English Grades

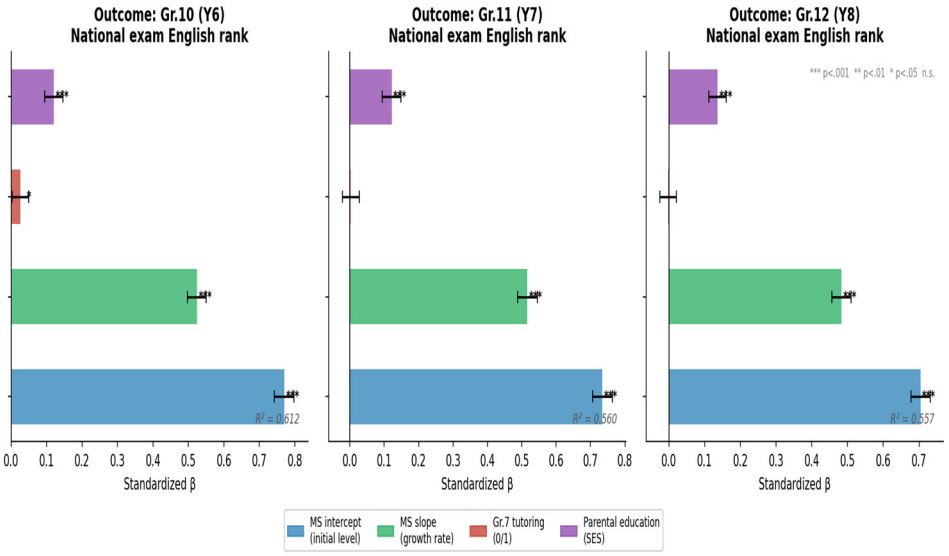
Model	Predictor	Grade 10 (hs6)	Grade 11 (hs7)	Grade 12 (hs8)
		β (*p*)	β (*p*)	β (*p*)
DA1: Unconditional				
	Intercept_est	.750 (< .001)	.713 (< .001)	.710 (< .001)
	Slope_est	.478 (< .001)	.469 (< .001)	.452 (< .001)
	R^2	.618	.567	.569
DA2: + MS tutoring				
	Intercept_est	.742 (< .001)	.706 (< .001)	.704 (< .001)
	Slope_est	.472 (< .001)	.464 (< .001)	.447 (< .001)

	Year 3 tutoring	.026 (.030)	.003 (.815)	−.002 (.872)
	Parental education	.121 (< .001)	.122 (< .001)	.137 (< .001)
	R^2	.612	.560	.557
DA4: + HS tutoring				
	Intercept_est	.731 (< .001)	.682 (< .001)	.672 (< .001)
	Slope_est	.463 (< .001)	.441 (< .001)	.416 (< .001)
	Year 3 tutoring	.019 (.135)	−.008 (.547)	−.013 (.297)
	Parental education	.115 (< .001)	.115 (< .001)	.129 (< .001)
	HS tutoring (concurrent)	.065 (< .001)	.089 (< .001)	.090 (< .001)
	R^2	.605	.546	.533

Note. MS = upper elementary and middle school (Grades 5–9); HS = high school. Standardized coefficients (β) and *p*-values are reported. Sample sizes: Grade 10 $n = 2{,}843$; Grade 11 $n = 2{,}825$; Grade 12 $n = 3{,}224$.

Figure 4 displays the standardized regression coefficients for all predictors across Grades 10–12, confirming the dominant contribution of the early school intercept and the rapid attenuation of PSE effects over time.

FIGURE 4
Standardized Regression Coefficients for Predictors of High School English Achievement



Note: DA2 Model, Grades 10–12. Bars show β for the five predictors across Grade 10/11/12 outcomes. Whiskers = ±1 SE.

Model DA4 included concurrent high school tutoring as well. Concurrent tutoring showed small but consistent positive associations with same-year achievement: Grade 10 tutoring predicted Grade 10 achievement (β = .123, $p < .001$), Grade 11 tutoring predicted Grade 11 achievement (β = .192, $p < .001$), and Grade 12 tutoring predicted Grade 12 achievement (β = .190, $p < .001$). When concurrent tutoring was added, the early school tutoring effects became non-significant. The model R^2 values were .605, .546, and .533 for grades 10–12.

Table 5 presents multi-group distal outcome analyses (DA2 specification) by initial proficiency. R^2 values varied systematically by proficiency group, with the middle group showing the highest variance explained (.449, .410, .385 for Grades 10–12) and the low group showing the lowest (.394, .340, .359). This pattern suggests that early school trajectories (Grades 5–9) were most predictive of high school achievement for students with moderate initial proficiency.

TABLE 5
Multi-Group Distal Outcome Analysis: Variance Explained by Initial Proficiency Group

Proficiency Group	*N* Grade 10	R^2 Grade 10	*N* Grade 11	R^2 Grade 11	*N* Grade 12	R^2 Grade 12
Low (< 179)	687	.394	669	.340	700	.359
Mid (179–220)	1,197	.449	1,190	.410	1,357	.385
High (≥ 221)	959	.426	949	.394	1,167	.379

Note. Each row represents Model DA2 estimated within that proficiency group. The models include intercept_est, slope_est, Year 3 tutoring, and parental education as predictors.

V. DISCUSSION

This study examined whether English private supplementary education (PSE) in Korea shifts the *trajectory* of achievement—its slope, not just its level—across the eight years from upper elementary school to the end of high school. Three findings are presented below. First, PSE in Grade 7 was positively associated with initial status but showed no association with the Grade 5–9 growth rate. Second, this null growth effect was consistent across low, mid, and high initial-proficiency strata. Third, early school PSE effects on high school achievement faded rapidly, while socioeconomic effects persisted through Grade 12. Each pattern connects directly to a specific claim in prior literature.

1. RQ1: PSE and Growth Trajectories

The signature result—a positive intercept association (β = .100, $p < .001$) coexisting with a null slope association (β = −.016, $p = .430$)—is the pattern that the causal-inference

literature on Korean tutoring would lead one to expect, but that a cross-sectional design cannot show. D. Ryu and Kang (2013), applying IV, first-difference, propensity-score matching, and bounding to KELS panel data, estimated tutoring effects on the order of 0.03–0.05 *SD* per 10% expenditure increase with confidence intervals spanning zero. Byun (2014) reported only small mathematics gains from cram schools and effectively no gains from other tutoring modalities. Kang (2012), using IV and nonparametric bounds, could not rule out zero effects in Korean, English, or mathematics. The present LGCM results add the missing longitudinal specification: the modest effects those studies identified appear to reside almost entirely in *level* rather than in *rate of change*.

The positive intercept coefficient itself is best read through the selection lens of Kim (2010), where controlling for endogeneity shrinks tutoring effects by approximately one-fifth in English and self-directed learning outperforms tutoring on longer-horizon outcomes such as CSAT scores and post-college wages. Because the Year 3 (Grade 7) PSE indicator postdates the Grade 5 intercept by two years, its positive association with the intercept cannot be causal; it more plausibly indexes prior tutoring exposure and family educational capital not captured elsewhere in the model. The persistence model reinforces this reading. Cumulative years of tutoring predicted both initial status ($\beta = .345$) and slope ($\beta = .111$, $p < .001$), but the practical slope magnitude was small — five continuous years of tutoring corresponded to roughly 18% of the mean annual gain. This is close to the dose-response signal Ku et al. (2022) recovered on KELS CSAT outcomes with multi-time IV specifications and consistent with de Silva's (2021) finding of diminishing marginal returns concentrated at the upper end of tutoring intensity.

The mechanism most compatible with these results is Han and Suh (2020) surface tutoring indirectly: tutoring may raise short-term test performance without building deeper competencies that would sustain accelerated learning over the years. Oh (2023) reported the same intercept-positive/slope-negative divergence for English class comprehension in the Jeonnam Educational Longitudinal Study over a five-year middle-to-high-school window. The two Korean LGCM datasets covering different regions and grade bands converging on the same qualitative pattern — a level shift without slope acceleration — increases confidence that this is a feature of Korean English tutoring rather than an artifact of any single design.

2. RQ2: Moderation by Initial Proficiency Level

Multi-group LGCMs showed no significant moderation of the PSE-slope association across low, mid, and high initial-proficiency groups (LRT $p = .616$). However, the tutoring-*intercept* association was substantially heterogeneous—large for low achievers ($\beta = .201$) and near zero for mid- and high-achievers—a pattern that runs opposite to Kang and Park's

(2013) distributional finding that high achievers benefit *more* from tutoring than low achievers. The apparent discrepancy is resolved by recognizing that Kang and Park estimated effects at different quantiles of the *outcome*, whereas the present multi-group analysis stratifies on initial proficiency. Within the low-achieving stratum, tutored students are a selected subgroup—likely from families with higher educational aspirations or greater willingness to invest in remediation—and the large intercept effect indexes selection rather than instructional impact. The absence of slope moderation in any group is a substantive result and shows that the null growth effect is not confined to a particular tier of achievers.

3. RQ3: Persistence into High School

Year 3 (Grade 7) PSE showed a small effect on Grade 10 English achievement ($\beta = .088$, $p = .008$) that disappeared by Grade 12 ($\beta = -.012$, $p = .717$), while concurrent high school PSE produced small proximal same-year effects ($\beta = .123$–$.192$). This pattern maps directly onto the fadeout framework in the broader intervention literature. Bailey et al. (2020) concluded that fadeout is widespread but coexists with persistence, and that whether an initial effect persists depends on the type of skill targeted and on whether subsequent environments complement the initial treatment. Andersen et al. (2022) showed on the full Danish public-school population that roughly half of a randomized reading intervention's effect had faded after four years, though because skill development is concave the treated students still preserved about 80% of their original time-equivalent lead.

The Korean tutoring case appears to be closer to full fadeout than to the Andersen benchmark. Two features of the mechanism plausibly explain this. First, if PSE primarily teaches assessment-specific strategies rather than transferable language competence — the reading of Han and Suh's (2020) short-term positive, long-term negative pattern — then the initial effect has nothing to attach itself to once curriculum demands and assessment formats shift at the high school transition. Second, the environment students subsequently enter is a much more intense tutoring environment in its own right: the small proximal effects of concurrent Grade 10–12 PSE observed here are consistent with de Silva's (2021) evidence that hagwon intensity has non-linear returns concentrated at the upper end. In effect, high school tutoring resets the meter, absorbing whatever advantage Grade 7 tutoring confers.

By contrast, the Grades 5–9 growth intercept and slope together explained 57%–62% of the variance in high school English grades, with the intercept dominating ($\beta = .710$–$.750$) and slope contributing modestly ($\beta = .452$–$.478$). Early trajectories, in other words, carry forward strongly, but PSE does not form them.

4. Socioeconomic Stratification: A Persistent, Structural Gap

The contrast between fading PSE effects and persistent SES effects is the most policy-relevant finding of the study. Parental education predicted both initial status (β = .334) and growth rate (β = .173) across Grades 5–9 and continued to predict high school achievement (β = .121–.137, all $p < .001$) even after conditioning on the early school trajectory. This is precisely the widening-gap pattern documented by Byun and Kim (2010) across three TIMSS cohorts and confirmed by Kim and Byun (2014) on KELS after extensive controls, extended here into an eight-year within-cohort growth model. Lee et al. (2015), using the 1980 prohibition and 1989 liberalization of tutoring as a natural experiment, showed that legal access to tutoring converts family wealth into elite university admission; the present results identify the growth-rate channel through which that conversion likely operates.

The "English gap" Kim (2012) documented, larger than for other subjects, driven by variation in out-of-school English exposure across income groups and regions, takes concrete form in these estimates. Students whose parents held a four-year degree began Grade 5 approximately 18.7 points ahead and gained an additional 2.8 points per year, accumulating a roughly 32.7-point advantage over five years, or more than half a standard deviation. Because PSE participation is itself stratified by family resources, and because it produces at best small proximal effects rather than growth-rate effects, the SES–PSE–achievement chain functions less as a mobility channel than as a mechanism through which existing inequality reproduces itself.

5. Implication and Limits

Read against the literature, these findings support three claims. Substantively, PSE in Korea's English domain functions primarily as a level shifter rather than as a growth accelerator, with concurrent effects that do not accumulate and Grade 7 effects that fade by Grade 12. Methodologically, the divergence between intercept and slope estimates confirms Cha and Ryoo's (2021) argument that longitudinal designs recover information that point-in-time analyses cannot and shows that at least some of the "no effect" verdicts in the causal-inference literature are more precisely "no *slope* effect" verdicts. For policy, if roughly 1.4% of GDP in PSE spending produces measurable proximal gains but no durable trajectory shift, while SES-linked factors produce a widening cumulative advantage across the same period. Bray's (1999) original framing of shadow education as a system that mirrors and inherits the stratification of the mainstream remains the most parsimonious account of the trajectories in this cohort.

This study has several limitations. PSE was measured as participation rather than intensity, quality, or curriculum content, which likely attenuates the estimated effects. The Grade 5

intercept is a *starting* observation, not a baseline free of prior tutoring, so what is labeled "initial status" already embeds several years of earlier PSE. Although the eight-year window is longer than any Korean LGCM English study to date, it does not extend into post-secondary or labor-market outcomes, where H. Kim (2010) found the most durable returns to accrue to self-directed learning rather than tutoring. Extending the trajectory to these outcomes and modeling PSE intensity and modality rather than binary participation are the natural next steps.

VI. CONCLUSION

This study used the Korean Education Longitudinal Study 2013 to estimate eight years of English achievement trajectories in a nationally representative Korean cohort and to determine whether private supplementary education (PSE) accelerates learning or merely shifts its measured level. Three results converge on a consistent answer. First, PSE was positively associated with initial achievement status but showed no meaningful association with the Grade 5–9 growth rate; even sustained multi-year exposure produced only a modest dose-response effect on slope—a small fraction of the mean annual gain. Second, multi-group LGCMs revealed no moderation of this null slope effect by initial proficiency, meaning the absence of trajectory acceleration holds uniformly across the achievement distribution. Third, Grade 7 PSE effects on high school English faded substantially within three years, while concurrent high school PSE produced only small proximal same-year effects that did not accumulate over time.

Read against the literature reviewed above, these patterns cohere. They extend, in a longitudinal specification, the modest-effects verdicts of Ryu and Kang (2013), Byun (2014), Kang (2012), and Kim (2010) by showing that whatever effect tutoring has resides almost entirely in level rather than in rate of change. They replicate the intercept-positive/slope-flat divergence Oh (2023) reported for the Jeonnam cohort and align with the fadeout framework Bailey et al. (2020) and Andersen et al. (2022) established for the broader intervention literature. Where the present study advances the field is in combining eight-year LGCM estimation with explicit multi-group testing and distal-outcome regressions, which together isolate PSE's contribution from the far larger and persistent contribution of family background.

That contribution is the study's second main finding. Parental education was a strong and persistent predictor of both initial achievement status and growth rate across Grades 5–9, and continued to exert a significant influence on high school English performance through Grade 12, even after conditioning on early school trajectories. This is the widening-gap dynamic documented by Byun and Kim (2010) and Kim and Byun (2014) in earlier Korean

cohorts, expressed here as a within-cohort growth process. Because PSE participation is itself stratified by family resources and produces at best a level shift, it functions less as a mobility channel than as a mechanism through which existing inequality reproduces itself — a finding consistent with Park (2023) and Lee et al. (2015) counterfactual simulations on KELS. The "English gap" Kim (2012) documented across income and region takes concrete cumulative form in these trajectories, with the Grade-5 achievement gap widening substantially through Grade 9.

For policy, these findings challenge the tacit assumption that expanding access to PSE is an effective route to closing achievement gaps. This does not imply that PSE serves no legitimate function; concurrent high school tutoring demonstrates a real, if modest, proximal role in helping students navigate high-stakes assessments. However, it does imply that the case for PSE as a driver of learning trajectories is weak, and that the case for public investment in mainstream English instruction and in addressing SES-linked determinants of learning is correspondingly strong.

Three limits define the natural next steps. PSE was measured as participation rather than intensity, quality, or curriculum content, likely attenuating estimated effects; modality-specific and dose-response designs would sharpen these estimates. The Grade 5 intercept is a starting observation, not a pre-tutoring baseline; thus, the level-vs-slope decomposition depends on assuming that prior tutoring exposure is captured by the intercept term. The trajectory ends at Grade 12; extending it into university and labor-market outcomes, where Kim (2010) found returns to accrue to self-directed learning rather than tutoring, would complete the eight-year story with a fifteen- or twenty-year one. In the framing Bray (1999) proposed at the beginning of this literature, the shadow follows the shape of the mainstream. The evidence here suggests that it does not reshape it.

REFERENCES

Andersen, S. C., Bodilsen, S. T., Houmark, M. A., & Nielsen, H. S. (2022). *Fade-out of educational interventions: Statistical and substantive sources*. CESifo Working Paper Series 10094. CESifo. https://doi.org/10.2139/ssrn.4286967

Bailey, D. H., Duncan, G. J., Cunha, F., Foorman, B. R., & Yeager, D. S. (2020). Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions. *Psychological Science in the Public Interest, 21*(2), 55-97. https://escholarship.org/content/qt2t78t0kh/qt2t78t0kh.pdf

Bray, M. (1999). *The shadow education system: Private tutoring and its implications for planners*. UNESCO International Institute for Educational Planning.

Bray, M. (2023). Understanding private supplementary tutoring: Metaphors, diversities and agendas. *Journal for the Study of Education and Development: Infancia y Aprendizaje, 46*(4), 728-773. https://doi.org/10.1080/02103702.2023.2194792

Bray, M. (2024). Still in need of confronting: Shadow education and its implications in the sustainable development agenda. *International Journal of Educational Development, 104*, 102967. https://doi.org/10.1016/j.ijedudev.2023.102967

Bray, M., & Lykins, C. (2012). *Shadow education: Private supplementary tutoring and its implications for policy makers in Asia*. Asian Development Bank & Comparative Education Research Centre, University of Hong Kong.

Byun, S.-y. (2014). Shadow education and academic success in Republic of Korea. *Education in the Asia-Pacific Region, 23*, 39-58.

Byun, S.-y., Jung, J., & Shin, T.-S. (2023). Does shadow education discourage or encourage creative thinking? Evidence from South Korea. *Journal for the Study of Education and Development: Infancia y Aprendizaje, 46*(4), 774-808. https://doi.org/10.1080/02103702.2023.2230042

Byun, S.-y., & Kim, K.-k. (2010). Educational inequality in South Korea: The widening socioeconomic gap in student achievement. *Research in Sociology of Education, 17,* 155-182.

Cha, I., & Ryoo, J. H. (2021). Application of fractional polynomials to model non-linear trajectories of English achievement. *Korean Journal of Educational Evaluation, 34*(2), 385-408.

Chang, K., Lee, E., & Kim, M. (2007). Students' variables and educational achievement in English. *English Language & Literature Teaching, 13*(3), 253-273.

Choi, S., Jo, J., & Kim, Y. C. (2025). Towards the SKY (Korean Ivy League): Educational strategies of Korean university-aspirant families. *Anthropology & Education Quarterly, 56*(3), 1-11. https://doi.org/10.1111/aeq.70007

de Silva, T. (2021). The impact of private tutoring on higher education outcomes: Evidence from South Korea. *Research in Labor Economics, 49,* 239-276.

Hajar, A., & Karakuş, M. (2024). A bibliometric analysis of shadow education in Asia: Private supplementary tutoring research trends (2000−2022). *International Journal of Educational Development, 108*, 103075. https://doi.org/10.1016/j.ijedudev.2024.103075

Han, J., Cho, H., Lee, S., & Chang, A. (2024). Early learning, tutoring, and STEM motivation: Impact on Korean students' mathematics achievement. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(6), em2460. https://doi.org/10.29333/ejmste/14659

Han, S., & Suh, H. (2020). The effects of shadow education on high school students' creative thinking and academic achievement in mathematics: The case of the Republic of Korea. *Educational Studies, 49*(2), 314-333.

Kang, C. (2012). The effect of private tutoring expenditures on academic performance: Evidence from middle school students in South Korea. *KDI Journal of Economic Policy, 34*(2), 139–171.

Kang, C., & Park, Y. (2013). To whom does money matter? Distributional effects of private tutoring expenditures on student performance. *KDI Journal of Economic Policy, 35*(1), 3-23.

Kim, H. (2010). *Analysis of the effects of private education on academic achievement and labor market performance*. KDI Research Monograph. Korea Development Institute. https://www.semanticscholar.org/paper/2ba92268cced4135d0287592247cfa50e51fd3dd

Kim, H. (2012). *Equity and efficiency of Koreans' English education investments* (KDI Policy Forum 245). Korea Development Institute. https://doi.org/10.22740/kdi.forum.e.2012.245

Kim, H. M., Han, S., & Kim, Y. (2018). A longitudinal analysis on the relationships among mathematics academic achievement, affective factors, and shadow education participation. *School Mathematics, 20*(2), 287-306.

Kim, K.-k., & Byun, S.-y. (2014). Determinants of academic achievement in Republic of Korea. In H. Park & K.-k. Kim (Eds.), *Korean education in changing economic and demographic contexts* (pp. 13-37). Springer.

Kim, S., & Youn, M. (2025). Unequal trajectories: Regional and socioeconomic gaps in student achievement in Korea. *Korean Journal of Educational Research, 63*(5), 131-157. https://doi.org/10.30916/kera.63.5.131

Korea Educational Development Institute (KEDI). (2020). *Korean Education Longitudinal Study 2013 (KELS 2013)* [Data set]. Korea Educational Development Institute. https://kels.kedi.re.kr

Ku, I., Lee, H., & Kim, J.-E. (2022). Does shadow education contribute to inequality? *Asia Pacific Journal of Education, 44,* 773-791.

Lee, K.-w., Hahn, S., & Sung, T. (2015). Nature or nurture: Private tutoring and education inequality. *The Korean Journal of Economics, 22*(1), 26-37.

Oh, Y.-k. (2023). Analysis of parental factors influencing growth trajectories in English class comprehension among middle and high school students. *The Journal of Learner-Centered Curriculum and Instruction, 23*(21), 617-632.

Park, H. (2023). Parental involvement and growing education inequality in South Korea. *Journal of Student Research, 12*(4), 1-5.

Ryu, D., & Kang, C. (2013). Do private tutoring expenditures raise academic performance? Evidence from middle school students in South Korea. *Asian Economic Journal, 27*(1), 59-83.

Ryu, J., & Lee, B. (2024). A longitudinal study of English achievement of Korean EFL young adolescent students: Focusing on initial English proficiency, English private tutoring, and urbanicity. *English Teaching, 79*(1), 69-92.

Shim, S.-H. (2005). A study on the relationship between student variables and English accomplishment of high school students. *English Language & Literature Teaching, 11*(4), 261-286.

Shin, C.-K., An, Y., & Oh, S.-Y. (2023). Reduced in-person learning in COVID-19 widens student achievement gaps in schools in Korea. *Asia Pacific Education Review, 25*(1), 45-55. https://doi.org/10.1007/s12564-023-09862-0

Shin, H., & Lee, B. (2019). "English divide" and ELT in Korea: Towards critical ELT policy and practices. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 1-19). Springer. https://doi.org/10.1007/978-3-319-58542-0_5-1

Yoo, J. (2024). Self-directed learning ability and shadow education expenditures: A comparison of impact on student achievement in Korea during COVID-19. *Sage Open, 14*(2). https://doi.org/10.1177/21582440241245206

Examples in: English

Applicable Languages: English

Applicable Levels: Secondary

Jie Young Kim (professor)
Dept. of English Education, College of Education
Chung-Ang University
(06974) 84 Heukseok-ro, Dongjak-gu, Seoul, Korea
Email: jk221@cau.ac.kr

Received: July 13, 2026
Reviewed: August 7, 2026
Revised version received: September 21, 2026