Economics of Education Review 77 (2020) 101955

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

The impact of labor market work and educational tracking on student educational outcomes: Evidence from Taiwan☆

Fung-Mey Huang[a], Jen-Che Liao*,[b], Chin-Chun Yi[c]

[a] Department of Agricultural Economics, National Taiwan University, Taiwan
[b] Department of Economics, Fu Jen Catholic University, Taiwan
[c] Institute of Sociology, Academia Sinica, Taiwan

ARTICLE INFO

Keywords:
Early employment
Educational achievement
Educational tracking
Endogeneity
Selection

JEL classification:
C24
C26
I20
J13
J22

ABSTRACT

This study empirically investigates how working while enrolled in high school affects educational outcomes, while accounting for self-selected educational tracking. Using a longitudinal survey of Taiwanese youth and exploiting the inter-zip-code geographic variations in youth-preferred industries, we find a negative effect of school-year work on educational achievement, and the negative marginal impact is much stronger for academic-track than for vocational-track students. An exogenous increase in school-year hours worked of 10 hours per week lowers college entrance scores by a 0.117 (0.083) standard deviation for academic-track (vocational-track) students. The negative impact of school-year work tends to be overstated if the endogeneity arising from educational tracking is not acknowledged—with upward biases as large as 16% and 30% for academic- and vocational-track students, respectively. Among subjects, math scores suffer most from working during the school year.

1. Introduction

Working while enrolled full-time in senior secondary school is a common behavior for global youth—nearly 66% of 11th graders and 75% of 12th graders in the U.S. worked part-time during the period 1997–2003 (NLSY97), and 60% of U.K. full-time students at age 16–18 worked part-time in the early 1990s (Dustmann, Rajah, & Smith, 1997; Dustmann & van Soest, 2007). This raises questions about the impact of working while enrolled in school on educational attainment that have important implications for human capital investment and educational and training policy. The impact of early employment on educational achievement has been extensively discussed for Western youth.[1] This impact is, however, not clear for youth in Asia or Europe, where educational tracking plays an important role in allocating students to either academic-track or vocational-track education at the beginning of senior secondary school. Using Taiwanese data, the main goal of this paper is to examine the impact of school-year work in senior secondary school on educational achievement and the extent to which the impact of early employment interacts with educational tracking.

Estimating the effect of early employment on academic achievement depends crucially on whether the endogeneity of part-time work has been taken into account (Lee & Orazem, 2010; Oettinger, 1999;

☆ We are grateful to the Co-Editor Colm Harmon, two anonymous referees, and Tze-yi Yen for their constructive comments and suggestions on earlier versions of this article. We thank Shiushan Wang for helpful assistance with data processing. Fung-Mey Huang gratefully acknowledges financial support from Taiwan's Ministry of Science and Technology (under project 103-2410-H-002-015-). Any errors are our own.

* Corresponding author at: 510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City 24205, Taiwan.
E-mail address: jccepd@gmail.com (J.-C. Liao).

[1] Two opposing arguments regarding the relation between early employment and educational achievements are commonly expressed (Hotz, Xu, Tienda, & Ahituv, 2002; Light, 1999; Marsh, 1991). The experience of early employment during the high school years may help youth learn to allocate time more efficiently and motivate them to study harder and may thus produce positive educational outcomes. In contrast, youth employment may detract time and energy from study and homework and have an adverse effect on educational achievement. Time allocation between study and work is a joint decision that is confounded by many observed and unobserved factors (McCoy & Smyth, 2007).

https://doi.org/10.1016/j.econedurev.2020.101955
Received 21 December 2018; Received in revised form 13 December 2019; Accepted 7 January 2020
Available online 12 May 2020
0272-7757/

Rothstein, 2007; Ruhm, 1997). Early studies that ignored the endogeneity problem have yielded various findings on the effects of part-time work on educational outcomes.[2] Recent studies that attempt to address the problem of endogenous employment have found a harmful effect (Lee & Orazem, 2010; Stinebrickner & Stinebrickner, 2003; Tyler, 2003) or little/no effect on educational achievement (Buscha, Maurel, Page, & Speckesser, 2012).

Educational tracking has long been used to allocate students to different tracks at some point between primary and tertiary schools in the global educational system.[3] In most East Asian countries such as China, Hong Kong, Japan, Korea, Taiwan and Vietnam, compulsory education ends at the completion of grade 9 or 10.[4] In Taiwan, the High School Entrance Exam (HSEE) is a nationwide test held at the end of grade 9 that, along with self-selection, allocates 15-year-old students to either the academic track (senior high school) or the vocational track (senior vocational school). The assignment to different educational tracks is generally nonrandom, due to the potential correlation between educational tracking and educational achievement (particularly in a way not captured by observed covariates). For example, when students decide between an academic-track and a vocational-track senior secondary school, they may weigh observed and unobserved factors such as their own level of ability, their work preference, opportunities for future college entrance, and later labor market outcomes. These observed and unobserved factors likely, in turn, affect students' educational outcomes.

The relationship between early employment and academic performance becomes even more complicated when taking into account self-selected educational tracking. Both educational tracking and employment decisions are likely associated with some common observables and unobservables, thus making these two decisions endogenous and confounding the effect of part-time work. As a consequence, failing to account for the nature of the self-selection of educational tracking likely results in biased and inconsistent estimates of the effect of early employment. This type of endogeneity of early employment that arises from a nonrandomly *selected sample* of educational tracking is essentially different from the endogeneity of early employment in the *population*, as discussed earlier. It is then necessary to distinguish these two sources of endogeneity and address them separately in order to recover the causal effect of early employment. To our knowledge, there is no study that carefully examines the relationship of early employment and academic performance while simultaneously correcting for endogenous part-time work and self-selected educational tracking.

From a methodological perspective, several econometric issues arise in our analysis. First, school-year hours worked in senior secondary education is known to be (left-)*censored* and likely to be *endogenous*. The endogeneity problem of early employment has been well documented in the literature, as discussed above. The censoring problem of hours worked, however, is not directly applicable to correct for endogenous regressors. Second, selection for educational tracks is likely to be *endogenous* in the sense that the outcome of educational achievement for senior high school students, for example, is observed only for students who have chosen the academic track. Both employment and tracking decisions may be commonly affected by unobserved factors such as ability and motivation, making these two decisions endogenous. To address these econometric issues, our empirical framework relies on an endogenous switching regression model for educational tracking where school-year hours worked is treated as a censored endogenous regressor. Our estimation strategy is based on a simple three-stage parametric estimation procedure that is easy to implement and works well in a small Monte Carlo simulation experiment that is available upon request from the authors.

Throughout the paper, we use the terms "senior high school" ("senior vocational school") and "academic track" ("vocational track") interchangeably. The general term "high school" refers to senior secondary education/school and includes both senior high school and senior vocational school.

The paper is organized as follows. Section 2 describes our empirical strategy, consisting of a endogenous switching model with a censored endogenous regressor and a simple three-stage parametric estimation procedure. Section 3 discusses educational tracking in the Taiwanese schooling system. Section 4 describes the longitudinal Taiwanese youth data set and variables used. Section 5 presents the empirical results of the interaction effects of early employment on educational achievement between educational tracking and early employment. Section 6 concludes the paper.

2. Empirical framework

2.1. Model specification

To incorporate educational tracking behavior to examine the impact of school-year work on educational achievement between academic- and vocational-track students, we consider an endogenous switching regression with an endogenous regressor that is subject to the censoring (or corner solution) problem. Specifically, the proposed model is of the form:

$$\begin{cases} Y_{1i}^{*} = X_i'\beta_1 + H_{1i}\alpha_1 + \varepsilon_{1i} & \text{if } T_i = 1, \\ Y_{2i}^{*} = X_i'\beta_2 + H_{2i}\alpha_2 + \varepsilon_{2i} & \text{if } T_i = 0, \end{cases} \tag{1}$$

$$T_i = \mathbf{1}\{Z_{Ti}'\gamma < \varepsilon_{3i}\}, \tag{2}$$

$$(Y_{1i}, Y_{2i}) = (Y_{1i}^{*} \times T_i,\ Y_{2i}^{*} \times (1 - T_i)) \quad \text{observed}, \tag{3}$$

where (Y_{1i}^{*}, Y_{2i}^{*}) are latent continuous outcomes with (Y_{1i}, Y_{2i}) being the observed educational achievements of students in the academic and vocational tracks, respectively, and the subscript i indexes individuals. (H_{1i}, H_{2i}) respectively denote observed school-year hours worked per week for 11th graders in the academic and vocational tracks. X_i is a vector of exogenous explanatory variables that may include an intercept. T_i is a binary selection variable indicating the status of educational tracking: $T_i = 1$ (0) for academic- (vocational-) track students. $Z_{Ti} = (X_i', Z_{1i}')'$ is a vector of IVs that affect the probability of selection,

[2] D'Amico (1984) and Steinberg, Greenberger, & Garduque (1982) find no effect of working on grade point average or class ranking. Several studies found a nonlinear effect of working hours on academic outcomes, i.e., high school students who work moderate hours per week have better academic outcomes than students who do not work or who work a large number of hours. Lillydahl (1990) finds a positive effect up to 13.5 hours per week and a negative effect after that. Stern (1995) finds that the turning point is 15 hours per week, while Greenberg & Steinberg (1986) find a positive effect up to 20 hours per week. Oettinger (1999) reports a negative effect for Black and Hispanic youth who work more than 20 hours per week. Some studies find a negative effect. Marsh (1991) finds a negative and linear association between in-school working hours and a series of academic outcomes, such as test scores and senior secondary school completion. Eckstein & Wolpin (1999) find a negative but small effect of working on academic performance.

[3] In this line of literature, whether educational tracking reinforces or compensates for pre-existing differences in the educational outcomes of students with different abilities and family backgrounds has received a great deal of attention (e.g., Gamoran & Mare, 1989; Goldstein, Burgess, & McConnell, 2007; Hanushek & Wossmann, 2006; Pekkarinen, 2008; Vandenberge, 2006; VanderHart, 2006). There are two opposing arguments regarding educational tracking and educational achievement. Efficiency is the main positive argument for educational tracking. It is much easier for school teachers to teach students with lower variations in ability and preference. Peer effects, on the other hand, indicate a potential negative impact of early educational tracking (Hanushek & Wossmann, 2006; Pekkarinen, 2008). Through self-selected tracking, more capable students may benefit from interactions with similar peers, whereas lower-ability students may suffer from a lack of positive stimulus.

[4] The education system in European countries also allocates pupils to different tracks at some point between primary and tertiary schools (Meghir & Palme, 2005; Pekkarinen, 2008).

with Z_{1i} being the exclusion restrictions. The selection error ε_{3i} could be correlated with (ε_{1i}, ε_{2i}), resulting in an endogenous switching regression in which the outcomes (Y_{1i}^*, Y_{2i}^*) are observed only on a nonrandomly selected sample. As is standard, we assume here that the error terms (ε_{1i}, ε_{2i}, ε_{3i}) follow a trivariate normal distribution with mean vector zero and the variance of ε_3 is normalized to 1 for the identification of the parameter γ in the selection Eq. (2). We note that σ_{12} may not be defined, as Y_{1i} and Y_{2i} are not observed simultaneously.

In models (1)-(5), we are interested in estimating the vector of finite-dimensional parameters $(\beta_1', \alpha_1, \beta_2', \alpha_2, \gamma')'$ based on n observations of a random sample $\{(Y_{ji}, X_i, H_{ji}, T_i, Z_{1i}, Z_{2i})\}_{i=1}^n$ for $j = 1, 2$. In particular, of primary interest is the estimation of (α_1, α_2), which measure the effects of hours worked on educational achievements in the academic and vocational tracks, respectively. The estimation of model (1) is potentially subject to two sources of bias: selection and endogeneity bias. First, the standard bias of track selection arises from the fact that the educational outcomes Y_1 (Y_2) are observed for the student only when they choose the academic (vocational) tracks. Second, in the outcome Eq. (1), the regressor H_{ji} is assumed to be endogenous in the sense that H_{ji} is possibly correlated with ε_{ji} for $j = 1, 2$. Moreover, the endogenous regressor H_{ji} is subject to corner solution problems, in the sense that in contrast to data observability problems, zero is the true value of H_{ji}, and it is the observed H_{ji}, instead of its latent version H_{ji}^*, that enters the outcome equation for Y_{ji}^* as the data generating process.

To account for the endogeneity and corner solution problems of H_{ji}, we construct an artificial outcome H_{ji}^* and slightly modify the conventional censored Tobit model as follows:

$$H_{ji}^* = Z_{Hi}'\theta_j + u_{ji}, \tag{4}$$

$$\begin{aligned} u_{ji} &= \rho_j \varepsilon_{ji} + \xi_i, \\ H_{ji} &= \max(H_{ji}^*, 0), \quad j = 1, 2, \end{aligned} \tag{5}$$

where H_{ji} is the censored version of the latent hours H_{ji}^*, $u_{ji}|Z_{Hi} \sim \text{Normal}\left(0, \sigma_{uj}^2\right)$, $\xi_i \sim \text{Normal}(0, 1)$ that is jointly independent of (ε_{1i}, ε_{2i}, ε_{3i}), and θ_j is the finite-dimensional track-specific parameter vector to be estimated. The error term u_{ji} is assumed to be correlated with ε_{ji}, as described in (5), making H_{ji}^* and hence H_{ji} endogenous to Y_{ji}^*. It can also be seen that u_{ji} is correlated with the selection error ε_{3i} through ε_{ji}, generating the endogenous selection problem when estimating (4) using only the sample of the individual tracks.

As usual, in addition to the exclusion restrictions Z_{1i}, we need excluded instruments Z_{2i} in Z_{Hi}, i.e., $Z_{Hi} = (X_i', Z_{2i}')'$, to enable us to use the IV approach to deal with the endogeneity problem of H_{ji}.

2.2. Estimation strategy

Based on the econometric model described above, our empirical strategy comprises three main stages. In the first stage, we conduct a probit estimation for the choice of educational track and compute the inverse Mills' ratios (IMRs), denoted by λ_1 and λ_2 for academic and vocational tracks, respectively. In the second stage, to address the problems of regime selection and corner solutions for the school-year hours worked, we plug the IMRs (obtained from the first stage) into a Tobit regression and calculate the fitted values of school-year hours worked (i.e., the estimates of $E(H_{1i}|Z_{Hi}, Z_{Ti}, T_i = 1)$ and $E(H_{2i}|Z_{Hi}, Z_{Ti}, T_i = 0)$) via a Tobit maximum likelihood estimation (MLE). Last, to simultaneously correct for endogenous educational tracking and endogenous employment, we conduct an IV estimation in the third stage for a switching regression for educational outcomes by plugging the IMRs (obtained from the first stage) and by replacing observed school-year hours worked with their predicted values (obtained from the second stage).

It is worth mentioning that one limitation of our second-stage Tobit-type estimation arises from a sign restriction that is inherent to the Tobit model, i.e., the sign of the regressor coefficients is required to be the same for the effects on the level of the outcome (i.e., part-time hours worked in our application) and on the probability that the outcome is positive (i.e., undertaking part-time work), and the violation of such restriction is commonly encountered in practice. We will examine this issue based on our data set in Section 5.2.

As an alternative to our first two stages of estimation, one solution that can relax the aforementioned sign restriction is the bivariate probit model for the two binary choices of educational track and part-time work as the first step, followed by a second step where a hours regression with two selection terms included for selection correction is estimated. We do not pursue this direction for two reasons. First, to explore the applicability of this alternative estimation approach in our context, we perform a bivariate probit estimation and a bivariate normality test based on our data set.[5] Our testing results clearly reveal that our data do not support the bivariate normality assumption. Second, the application of the bivariate probit model to our setting requires one more exclusion restriction for the binary choice of part-time work, which appears infeasible due to the absence of suitable instrumental variables given the data available to us, though this issue can be addressed by a one-step joint estimation.

3. Educational tracking in Taiwan

In 1968, the compulsory education reforms in Taiwan changed the junior secondary school system from a selective to a comprehensive structure and postponed educational tracking from the 7th grade (age of 12–13) to the 10th grade (age of 15–16). The school tracking system involves a nationwide HSEE and self-selection. The score from HSEE along with an individual's high school preference list allocate an junior high school graduate to either academic track or vocational track senior secondary school. As a consequence, the assignments to different tracks are nonrandom in nature. On the academic track, students spend three years in senior high school. On the vocational track, students enroll in either three years of senior vocational school or five years of junior college. The differences between academic and vocational education stem from both the curriculum offered and the average cognitive ability of the enrolled students. In addition to general courses (math, Chinese and English), vocational education includes courses in areas such as agriculture, industry, business, maritime studies, marine products, medicine, nursing, home economics, drama and art. In this study, we focus our comparisons on students at three-year senior high schools and senior vocational schools, and omit students at five-year junior colleges because they do not make their college entrance and employment choices at the same age as the three-year high school students.

Educational tracking is strict in Taiwan. Beyond senior secondary education, graduates of both senior high schools and senior vocational schools have various choices regarding the type of college they attend and their labor market participation. College entry options include attending four-year academic universities, four-year technical colleges/universities, and two-year junior colleges. Admissions to general four-year universities are based mainly on the General Scholastic Ability Test (GSAT),[6] while admissions to a two-year junior college or a four-year university of science and technology is based on the Technological & Vocational Education Test (TVET). Table 1 presents the distribution of college and labor market choices for senior secondary school graduates. Two salient phenomena can be observed. First, the majority of senior secondary school graduates pursue further tertiary education instead of entering the labor market. National data show that among graduates, 95% of academic-track and 82% of vocational-track students enter either a general university or a university of science and technology. Only

[5] The detailed model descriptions and the estimation and testing results are available upon request from the authors.

[6] Some students may also apply based on scores from the Advanced Subject Test.

Table 1
College and employment choices for senior secondary graduates from academic and vocational tracks.

	General University	University of S&T	Employment	Delayed
Longitudinal Survey of Taiwan Youth Project:				
Senior high school graduates	74.72%	9.80%	2.48%	13.01%
Senior vocational school graduates	4.73%	56.87%	21.53%	16.87%
National Data (2010):				
Senior high school graduates	95.50%		0.37%	4.11%
Senior vocational school graduates	81.91%		12.17%	5.92%

Notes: 1. S&T stands for science and technology. 2. The numbers in the longitudinal survey of the Taiwan Youth Project are computed from the J1 and J3 samples. 3. The national data are archived in Taiwan's Ministry of Education.

12%–21% of vocational-track graduates choose to participate in the labor market. Second, due to differences in the curricula offered, the probability of switching between the academic track and the vocational track is very low. The panel data show that 74% of academic-track graduates enter a general university, and only 10% enter a technical college. In contrast, only 4.7% of vocational-track graduates enter a general university, while 57% enter a technical college or a university of science and technology.

4. Data and variables

4.1. Data

The data used for this analysis are obtained from a longitudinal survey of the Taiwan Youth Project (TYP) Phase I, which was conducted yearly from 2000 to 2009 in northern Taiwan. The baseline sample used are students who were first interviewed in their last year of junior high school (i.e., 9th grade) in 2000 (referred to as the J3 sample hereafter). To examine the relationship between tracking selection at the end of 9th grade, part-time work during the high school years, and score on the college entrance examination, we use wave 1 - wave 6 of the J3 sample, which traced students from 9th grade to college entrance. To utilize the longitudinal data efficiently, we restricted it to students who had complete information on tracking selection and individual and family characteristics in wave 1 (9th grade). We further deleted individuals who did not enter senior secondary schools and those who entered five-year junior colleges.[7] The resulting sample has 2372 individuals, including 1076 academic-track students and 1296 vocational-track students.

For the sample size used in our three-stage estimation method, the first-stage estimation uses all 2372 students for the educational track selection. As longitudinal data, non-response behaviors occurred in the TYP survey when students entered high school and also when they entered college. As a consequence, the sample size in the second-stage estimation for part-time hours worked during high school years slightly drops to 2029 (with 973 academic-track and 1056 vocational-track students). In the third-stage estimation for the score of the college entrance examination, the sample size is either 526–541 (370–399) or 558–573 (409–440) for the academic (vocational) track, depending on the subjects under consideration and the observed or predicted hours used. Since the data on college entrance scores were collected in waves 5 and 6 during students' first year of college (some students delayed entering college), the sample size from the second- to third- step estimation sharply decreases due to nonresponses or missing values. To examine the potential attrition bias, both T test and Kolmogorov-Smirnov equality-of-distributions test are performed to compare the students' and their family characteristics between those who reported and those who did not report their scores of college entrance examination. We find no statistically differences of basic characteristics in terms of gender, parental education, family income, and intact family between these two groups.

The TYP provides rich information regarding students' and their parents' characteristics as well as the history of youths' education and employment. For educational history, we use academic records from both primary school and 9th grade as the proxy of academic ability. The TYP started to collect individual employment information in the summer right after graduation from junior secondary school at an average age of 15 due to the Labor Standard Act. The detailed information on hours worked per week, however, is only available from wave 3 for 11th graders.

4.2. Variables

For employment information during the senior secondary school years, the upper panel of Table 2 presents the part-time work experience of 11th graders by educational track. It shows that students on the vocational track engage in more part-time work than those in the academic track. During the entire senior secondary years (grades 10–12), 34% of senior high school students worked part-time, whereas 60% of senior vocational school students did so. The percentage of students working part-time drops between grade 10 and grade 11 for students in both educational tracks due to college entrance examinations. The GSAT and TVET examinations are held in grade 12. Since the examinations cover mainly the materials taught in the 10th and 11th grades, most senior secondary students decrease their labor supply when the college entrance examination is approaching. The 11th grade is critical for most senior secondary students. It is expected that students will tend to work more during the summer or winter vacation periods than during the school year. The part-time work variable in this study is measured by hours worked per week during the semesters of 11th grade. Conditional on working, the students on the vocational track also worked for a longer hours—the average is 31.25 hours per week for vocational-track students and 19.96 hours per week for academic-track students.

The measurement of educational outcomes is based on the scores on college entrance examinations. For admission into a four-year general university, good scores on the GSAT are the main requirement. Similarly, TVET scores are critical when applying to a two- or four-year technical college/universities. These two college entrance examinations are on different subjects and have different maximum scores. The GSAT comprises five subjects: math, Chinese, English, natural science, and social science. For each subject, a student can earn a maximum of 15 points; thus, the total points that can be earned is 75. The TVET covers three general subjects (math, Chinese, and English) and special technical & science subjects. For each TVET subject, a student can earn a maximum of 100 points. To enable a comparison of these two college entrance examinations, we compare mainly the three general subjects, math, Chinese, and English, summing them to obtain a range of 0–45

[7] Less than 5% of the students in the J3 sample of the TYP enrolled in five-year junior colleges.

Table 2
Part-time work experience for secondary school students on different educational tracks.

Part-time work while enrolled in school	Academic track	Vocational track
Work experience:		
Ever worked during grades 10–12 (=1)	0.34	0.60
Ever worked in grade 10 (=1)	0.26	0.48
Ever worked in grade 11 (=1)	0.14	0.29
Hours worked per week in grade 11 when worked	20.76	34.37
Ever worked during school year (semester) in grade 11 (=1)	0.07	0.16
Hours worked per week during school year in grade 11 when worked	19.96	31.25
Distribution of part-time jobs by industry:		
Food industry (restaurants, fast food, beverage stores)	0.15	0.34
Cram school (servitors)	0.31	0.07
Convenience stores	0.04	0.06
Flyers/leaflets services	0.13	0.03
Promote product sales	0.09	0.02
Day care centers	0.00	0.004
Others (including gasoline stations)	0.28	0.47
Number of observations	1076	1296

Table 3
Scores for college entrance examinations for different tracks.

	Academic track		Vocational track	
	No part-time work	Part-time work	No part-time work	Part-time work
Three subjects	GSAT-3S		TVET-3S	
Chinese	11.60	11.50	65.25	65.08
	(502)	(62)	(325)	(115)
English	9.72	8.80**	48.54	48.48
	(511)	(62)	(326)	(114)
Mathematics	7.54	6.25**	52.97	49.63*
	(504)	(66)	(305)	(107)
Total scores	28.79	26.63**	166.98	164.61
	(496)	(62)	(303)	(106)
Five subjects	GSAT-5S			
Total scores	49.28	47.26*	-	-
	(785)	(100)	-	-

Notes: 1. Numbers of observations are in parentheses. 2. Part-time work refers to worked in school year of grade 11. 3. *, **, and *** denote the statistical mean difference of the two samples (no part-time work vs. part-time work) at the 10%, 5%, and 1% levels of significance, respectively.

points for the GSAT (referred to as the GSAT-3S hereafter) and 0–300 points for the TVET (denoted by TVET-3S).[8] In the later empirical analysis, we take the z-score of the GSAT-3S and TVET-3S to consider the impact of school-year hours worked on the changes in college entrance examination scores.

Table 3 presents the basic statistics for the college entrance examination scores from the different tracks by their part-time work experience. The *T* tests are performed to examine the statistical mean difference of the two groups (no part-time work versus part-time work) for students from different tracks. If we compare senior secondary students without part-time work to those with part-time work, we observe that the former tends to have higher college entrance examination scores than the latter. The differences tend to be more substantial for the academic-track students. Math scores, in particular, are significantly lower for senior secondary students with part-time work experience. Regarding the total five-subject GSAT scores (denoted by GSAT-5S), the bottom panel of Table 3 indicates that early employment also substantially lowers the GSAT-5S scores for academic-track students. Since very few vocational-track students take GSAT examination for their college entrance, we ignored their GSAT-5S scores.

The mean statistics of family background for working and nonworking students on different educational tracks are presented in Table 4. These statistics reveal that students on the academic track tend to have better-educated parents and higher family income than students on the vocational track. In addition, Table 4 indicates that students who have part-time work experience are more likely to come from families with lower levels of parental education, lower family income, and more likely to live in nonintact families. The detailed definition of all the variables used in this study and their basic statistics are presented in Table 12.

4.3. Instrumental variables

In this study, we use a rich set of geographic measures as instruments for endogenous tracking selection and endogenous part-time work.

4.3.1. IV for endogenous tracking selection

The excluded instrument for endogenous tracking selection is the density of senior high schools in a student's residential area measured by the number of senior high schools per square kilometer in a student's zip-code residential area, which is administrative data collected from the Ministry of Education in Taiwan. As an exclusion restriction, the density of senior high schools in a student's residential area should be correlated to the student's tracking selection but exogenous to his/her later college entrance scores.

On the one hand, the relevance of this excluded instrument for endogenous educational tracking relies on the role-model argument. In general, both students and parents in Taiwan prefer the academic track to the vocational track.[9] There may exist some potential role-model effects of senior high school students on junior high school students. When a junior high school student lives in an area with a higher density of senior high schools, he/she is exposed to the environment with more senior high school students and more activities associated with senior high schools, which may inspire the student to form a higher preference for the academic track versus the vocational track. This preference for the academic track may further inspire students to study harder to enter senior high school. Consistent with the role model effect discussed above, Fig. 1 shows that the density of senior high schools is significantly and positively correlated with students' preference for the academic track and also positively correlated with the subsequent likelihood of admitting into senior high schools.

On the other hand, the exogeneity of the density of senior high schools in our context can be justified by the following arguments. First, the establishment of both public and private senior high schools in Taiwan is regulated by educational authorities and is mainly determined by exogenous forces such as education policy, economic development, and demographic factors. To be more specific, the senior secondary educational policy has evolved, experiencing several major changes since 1968. Vocational education has been greatly expanded to meet the increased demand for medium-skilled workers stemming from Taiwan's high economic development between the 1960s and 1980s. Since the 1990s, however, the demand for tertiary education has substantially increased. The educational authorities accordingly reversed the ratio of senior high school students to senior vocational school students by expanding senior high schools and reducing senior

[8] The entrance examination for a general university is typically more difficult than that for a technical college. Comparing these two scores may overestimate the achievement of students taking the TVET, who are usually vocational-track students.

[9] In our data set, 59% of 9th graders prefer to attend senior high schools, and 41% prefer to attend senior vocational schools.

Table 4
Basic statistics of individual and family backgrounds.

Variables	Academic track		Vocational track	
	No part-time work	Part-time work	No part-time work	Part-time work
Male (=1)	0.50	0.33***	0.49	0.50
Academic scores in primary school	3.95	3.90	3.00	2.90*
Parents junior high graduates and below (=1)	0.19	0.25*	0.46	0.50*
Parents college graduates and above (=1)	0.26	0.14***	0.03	0.01**
Characteristics in grade 9				
Academic scores	3.68	3.42**	2.37	2.12***
Student's educational expectation	16.48	16.43	14.06	13.88*
Parents' educational expectation	18.55	17.86**	16.04	15.48***
Student's tracking preference (=1 if academic)	0.93	0.85***	0.34	0.30
Parents' attending School Day (=1)	0.62	0.50***	0.39	0.28***
Intact family (=1)	0.94	0.86**	0.90	0.80***
Family income (NT$ thousand)	71.78	58.63*	52.70	51.33
Characteristics in grade 11				
Intact family (=1)	0.89	0.79***	0.85	0.73***
Family income (NT$ thousand)	73.76	56.85***	51.64	48.70*
Number of observations	903	70	761	295

Notes: 1. Part-time work refers to work in school year of grade 11. 2. *, **, and *** denote the statistical mean difference of two samples (no part-time work vs. part-time work) at the 10%, 5%, and 1% level of significance, respectively.

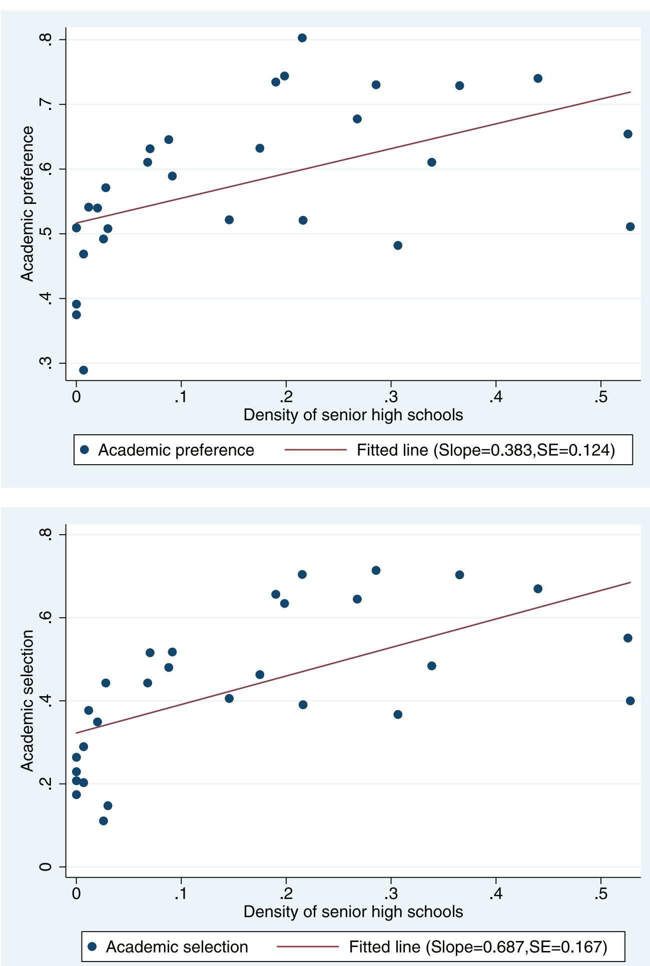

Fig. 1. Students tracking preference, senior high school selection, and density of senior high schools.

vocational schools (see Fig. 2). Meanwhile, a large expansion of four-year universities was undertaken to meet the demand for tertiary education for high school graduates.

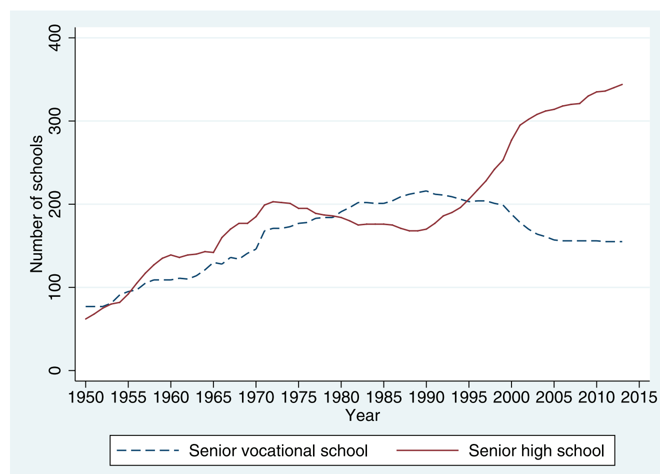

Fig. 2. The historical trend in the number of senior secondary schools in Taiwan.

One could argue that the exogeneity of the density of senior high schools might be weakened by the potential competition effects that could arise from a high density of senior high schools in an area; because such competition effects may raise teaching and educational standards and ultimately students' educational attainment. To examine this possibility, we use the additional J1 sample in the TYP survey, which collected complete HSEE scores and admitted high schools for junior high school graduates.[10] Using the students' HSEE scores and their corresponding admitted high schools, we construct the school ranks of high schools in Northern Taiwan. In general, senior high schools would compete with others with similar ranks based on the HSEE scores of their admitted students. We expect that, if the competition effect exists within a zip-code area, the ranks of senior high schools in the same area would be very similar (i.e., small variance of school ranks). That said, the density of senior high schools must be negatively associated with the dispersion of senior high school ranks in

[10] The J1 sample is another baseline sample in the TYP survey that consists of students who were in the first year of junior high school (i.e., 7th graders with an average age of 13 years old) in 2000.

the same zip-code area for the competition effect to exist. To investigate this condition based on our data set, we regress the standard deviation of senior high school ranks on the density of senior high schools by zip-code areas and find a significant and positive association between the senior high school density and the rank dispersion. This evidence indicates that the existence of competition among senior high schools within the same zip-code area does not seem to be supported by our data.

Second, even if competition effects exist, this should not affect our results. In Taiwan, the score a student achieves on the HSEE is the main criterion used to assign him/her to the academic or vocational tracks and to elite or non-elite senior secondary schools. According to our data, there is a high likelihood (79.9% in our data set) of a junior high school graduate being allocated to a senior secondary school outside of his/her residential zip-code area. This indicates that a student's place of residence, and hence the density of senior high schools in his/her residential area, should not exert much influence on his/her eventual score on the college entrance examination.

4.3.2. IV for part-time work

To deal with the endogeneity problem of part-time work, previous studies frequently used instrumental variable (IV) estimation and/or fixed effects estimation to take into account the correlation between individuals' unobserved heterogeneity and part-time work, and to obtain consistent estimates of school-year employment effects. It has been argued that the following IVs are uncorrelated with the unobserved heterogeneity in educational achievement but sufficiently correlated with the high school labor supply: the child labor law (Lee & Orazem, 2010; Tyler, 2003), the job types (Stinebrickner & Stinebrickner, 2003), and the local labor market conditions (Lee & Orazem, 2010; Oettinger, 1999; Rothstein, 2007; Ruhm, 1997).

In this study, we use the density of various local store types to measure local labor market opportunities in various businesses relevant to high-school students in students' residential zip-code areas. Specifically, the four main IVs of this type are the densities of convenience stores, food stores (including restaurants, beverage, and fast-food stores), gasoline stations, and flyer/leaflet distribution services located in a student's residential zip-code area, which are administrative data obtained from the Ministry of Economic Affairs in Taiwan. To reflect the heterogeneous size of the zip-code areas, the densities of local youth employment opportunities are taken either per squared kilometer or per 1000 population aged 15–19.

The rationale for using the densities of various store types in students' residential zip-code areas as appropriate IVs for part-time work is as following. Most of the high school students are marginal workers who do not have sufficient skills in the labor market and cannot commit to full-time work. As a result, they can only match with certain businesses and industries (such as convenience stores, restaurants, beverage stores, fast-food stores, gasoline stations, cram school tutors, and flyer/leaflet services) that require limited skills, provide a flexible working schedule, and pay the hourly minimum wage. The bottom panel of Table 2 presents the relative frequency of these workplaces for high school students by educational track. It shows that an academic-track student is more likely to work as a servitor at a cram school if he/she participates in part-time work, whereas a vocational-track student tends to work in the food industry (such as restaurants, beverage stores, and fast-food outlets) or at a gasoline station.

We note here that although the cram school accounts for the largest share of part-time jobs for academic-track students, we do not incorporate the density of cram schools as an instrument for part-time work. This is because not only are the cram schools the most common workplace for academic-track students but they could also serve as an important source of assistance for students' academic outcome, thus invalidating the instrumental exogeneity.

With regard to the exogeneity of the IVs for part-time work, we note that the establishment of food stores, gasoline stations, flyer/leaflet services, and convenience stores typically depends on the local population size and business conditions, and thus should be independent of underlying factors that also affect an individual's score on a college entrance examination. The variation in the density of various store types, which are taken either per squared kilometer or per 1000 population aged 15–19, is largely out of the individualâs control and should not affect academic performance, except through work behavior.

The county-level unemployment rates for youth aged 15–19 are included as an additional IV for endogenous part-time work. The youth unemployment rates of the student's resident county are commonly used to satisfy exclusion restrictions (e.g., Darolia, 2014; Lee & Orazem, 2010; Oettinger, 1999; Rothstein, 2007; Ruhm, 1997; Triventi, 2014). The justification of its relevance, as argued by Triventi (2014), is that the local youth unemployment rates can serve as a proxy of local labor demand for young people.

One could argue that to pursue a better occupational outcome upon graduation from high school, high local unemployment rates may provide an "incentive" for high school students to work hard on their academic studies in order to achieve higher educational attainment outcomes, thus invalidating our identification strategy. This argument seems plausible, but it should not be the case in our context. First, the goal of high school students in Taiwan is to enroll in university instead of to turn to the labor market. In fact, based on the national data in Taiwan, the net enrollment rate of youth aged 15–18 is 96%, and upon graduation only 0.37% (12.17%) of senior high school (senior vocational school) graduates enter the labor market (Table 1). Further, as noted by Triventi (2014), there are no strong theoretical reasons that can justify the influence of unemployment rates on academic achievement.

Before closing this section, we would like to stress that the exogeneity of our IVs can be further strengthened by a robustness check, which will be discussed in Section 5.5, where we find no evidence of significant effects of part-time work during summer/winter vocations, lending credibility to the main results based on our estimation strategy and the IVs used.

5. Empirical results

5.1. Specification and normality tests

To check the appropriateness of the specifications of our first-stage probit and second-stage Tobit models, we perform a number of specification and normality tests. For the first-stage probit model for tracking selection behavior, we employ the probit specification test proposed by Hosmer and Lemesbow (1980) (hereafter, the HL test) for both the linear index $Z_{Ti}'\gamma$ and the normality of the error term, and the Lagrange-multiplier (LM) type normality tests (denoted by LM_n^P) proposed by Bera, Jarque, and Lee (1984) for probit models. The upper panel of Table 5 shows that, based on the standard probit model, both the HL test on the linearity and normality specifications and the LM_n^P test on the normality specification are rejected at the 1% significance level. However, using the series probit that includes the squared and cross-product terms of the covariates, we do not reject the series probit specification (with the HL test statistic being 5.196) and the normality assumption (with $\mathrm{LM}_n^P = 0.936$).[11]

In the second-stage Tobit model for part-time hours worked, we perform the LM type normality tests for Tobit models (denoted by LM_n^T), as proposed by Bera et al. (1984). For comparison purposes, we also compute the LM normality test (LM_n^{T*}), as suggested by Jarque and Bera (1980), based on the OLS linear regression that ignores the non-negative nature of hours worked. In addition, we consider several series

[11] We use polynomials as the series functions to allow for more flexible functional forms while maintaining the index as linear in parameters.

Table 5
Specification and normality test results in the first-stage (series) probit and second-stage Tobit models.

Models	LM normality tests LM_n^P (BJL1984)	HL specification tests HL_n (HL(1980))
First-stage probit	41.900***	37.158***
First-stage series probit	0.936	5.196
Number of observations	2372	2372

Models	LM normality tests LM_n^T (BJL1984)	LM normality tests LM_n^{T*} (JB1980)	LOOCV values
Second-stage (series) Tobit			
1. Academic track			
Specification A1	0.466	87194.8***	525.83
Specification A2	0.021	87198.9***	525.91
Specification A3	0.631	86791.2***	527.13
Number of observations	973	973	973
2. Vocational track			
Specification V1	14.816***	1036.3***	3754.36
Specification V2	2.840	1034.2***	3747.05
Specification V3	2.821	1035.2***	3749.08
Number of observations	1056	1056	1056

Notes: 1. We employ the probit specification test proposed by Hosmer and Lemesbow (1980) (HL(1980)) (the HL test for short), which is based on a diagnostic measure assessing the match between observed binary outcomes and predicted probabilities. 2. LM_n^P and LM_n^T represent the LM normality test statistics for the probit and Tobit models, respectively, which are suggested by Bera et al. (1984) (BJL(1984)). LM_n^{T*} represents the LM normality test statistic suggested by Jarque and Bera (1980) (JB(1980)) based on OLS regression models. The LM_n^{T*} normality test is included for comparison purposes to point out the appropriateness of the second-stage Tobit model for part-time hours worked. 3. The critical values for the probit specification test are $\chi^2_{10}(0.90) = 15.99$, $\chi^2_{10}(0.95) = 18.31$, and $\chi^2_{10}(0.99) = 23.21$. The critical values for both the probit and the Tobit LM normality tests are $\chi^2_2(0.90) = 4.61$, $\chi^2_2(0.95) = 5.99$, and $\chi^2_2(0.99) = 9.21$. 4. Specification A1 includes all demographic and instrumental variables in a linear manner, corresponding to standard Tobit models. Specifications A2 and A3 additionally enter squared terms of demographic variables and squared terms of demographic and instrumental variables, respectively. Specifications V1, V2, and V3 are defined analogously. 5. LOOCV stands for the leave-one-out cross validation. 6. *, **, and *** denote a significance level of 10%, 5%, and 1%, respectively.

Tobit specifications that allow for nonlinear effects of demographic factors and/or instrumental variables on hours worked as follows. Specification A1 includes all demographic, instrumental, and IMR variables in a linear manner, corresponding to the standard Tobit model. Specifications A2 and A3 additionally add squared terms of the demographic variables and of the demographic and instrumental variables, respectively. Specifications V1, V2, and V3 are defined analogously. The lower panel of Table 5 reveals that for the LM normality tests (LM_n^{T*}), normality based on an OLS linear regression for hours worked is clearly rejected in all cases when ignoring the non-negative nature of working hours. This suggests that OLS is an inadequate estimation method for hours worked in the second stage and thus justifies the use of the Tobit-type models. When it comes to the standard Tobit models (i.e., with a linear index) in the second stage, we respectively do not reject and do reject the Tobit normality assumptions for the academic ($LM_n^T = 0.466$) and vocational ($LM_n^T = 14.816 > \chi^2_2(0.99) = 9.21$) tracks. By further relaxing the linearity restriction imposed on the Tobit models, similar to the findings from the first-stage standard probit versus series probit estimation, we do not reject Tobit normality when series estimation for Tobit models is used for both educational tracks.

Table 6
The marginal effects in the first- and second-stage estimation: educational tracking and school-year hours worked.

	Educational tracking	School-year hours worked/10 Academic track	School-year hours worked/10 Vocational track
Characteristics in grade 9:			
Academic scores in primary school	0.031***	0.210	−0.071
	(0.009)	(0.353)	(0.277)
Academic scores	0.056***	−0.410	−0.426**
	(0.007)	(0.333)	(0.201)
Parents' junior high graduates and below	−0.078	−0.157	0.349
	(0.136)	(0.708)	(0.396)
Parents' college graduates and above	0.068*	−0.269	−2.641
	(0.041)	(0.982)	(3.179)
Student's educational expectation	0.018***	0.172	0.010
	(0.004)	(0.134)	(0.103)
Parents' educational expectation	0.010***	−0.064	−0.219**
	(0.003)	(0.097)	(0.087)
Student's tracking preference	0.361***		
	(0.028)		
Parents attending School Day	0.025*		
	(0.014)		
Intact family	0.040		
	(0.026)		
Family income	0.001*		
	(0.000)		
Characteristics in grade 11:			
Intact family		−1.209	−1.488***
		(0.822)	(0.427)
Family income		−0.017**	−0.010
		(0.008)	(0.008)
Instrumental variables (IVs): geographic variables in residential areas			
Density of senior high schools	0.325***		
	(0.056)		
Density of convenience stores		0.477***	
		(0.180)	
Density of gasoline stations (per km^2)		2.006**	
		(1.002)	
Density of flyers/leaflets services		−0.191***	
		(0.064)	
Density of food stores/10			0.014***
			(0.005)
Density of gasoline stations (per 1000 population)			0.575**
			(0.255)
Youth unemployment rate		−0.496	−0.133
		(0.483)	(0.279)
IMRs (λ_1 or λ_2)		−0.396	0.283
		(0.601)	(0.292)
Variance of ε_1 (ε_2)		17.664***	19.208***
		(3.543)	(1.401)
Demographic factors	yes	yes	yes
Squared all regressors (series)	yes	no	yes
Interacted all regressors (series)	yes	no	no
LR χ^2 tests			
1. H_0: Joint nonsignificance of IVs	49.720***	11.557**	12.106***
[*p*-values]	[0.000]	[0.021]	[0.002]
2. H_0: Joint nonsignificance of covariates and IVs	1702.31***	45.43***	53.43***
[*p*-values]	[0.000]	[0.000]	[0.000]
Pseudo *R*-squared	0.521	0.061	0.022
Correctly classified	86.05%		
Number of observations	2372	973	1056

Notes: 1. Bootstrap standard errors are in parentheses. 2. Constants are included. 3. Demographic factors include gender and race. 4. *, **, and *** denote a significance level of 10%, 5%, and 1%, respectively.

Last, the series Tobit specifications are then chosen based on the leave-one-out cross validation (LOOCV) criterion. As reported in Table 5, the preferred specifications that minimize the corresponding LOOCV criteria are A1 (with a CV value of 525.83) for the academic track and V2 (with the CV value of 3747.05) for the vocational track. These two CV-determined specifications pass the Tobit normality tests (with $\mathrm{LM}_n^T = 0.466$ and $\mathrm{LM}_n^T = 2.840$ for the academic and vocational tracks, respectively) and fail the OLS normality checks (with respective $\mathrm{LM}_n^{T*} = 87194.8$ and $\mathrm{LM}_n^{T*} = 1034.2$ for the academic and vocational tracks).

5.2. Educational tracking and early employment behaviors

Using the chosen specifications, Table 6 presents the marginal effects on educational tracking and part-time work from the first- and second-stage estimation. The first column in Table 6 reports the first-stage estimation results from a series probit model for tracking selection. The characteristics of both individuals and families during the junior high school years are controlled for. As mentioned in the previous subsection, the series probit model for tracking selection includes the linear form, squared form, and the interaction of all the regressors. The marginal effects shown in Table 6 depict that academic records in the primary school years and junior high school years that serve as the proxy variables for academic ability play an important role in determining educational tracking. The significant and positive effects of academic records reveal that a student with higher academic ability is more likely to be selected into the academic-track system. Parental education and parents' educational expectations for their junior high school child and a student's own educational expectations promote a higher probability of entering the academic track. A student's preference for the academic track significantly and strongly raises his/her probability of entering that track. Demographic factors (such as gender and race), however, play no role in tracking selection. Last, the goodness of fit of the tracking selection model reveals that the prediction achieves correct classification of 86% of the cases. The LR χ^2 test indicates strong instrumental relevance of the density of senior high schools at the 1% level of significance.

The second-step estimation results are presented in the second and third columns in Table 6, where the standard Tobit (Specification A1) and series Tobit (Specification V2) models for the school-year hours worked are estimated for academic- and vocational-track students, separately. Given that the track is selected, academic record and parental education play a nonsignificant role in determining early employment. As expected, family characteristics during the senior high school years affect students' part-time work. In particular, living in an intact family significantly decreases school-year hours worked, and higher family income discourages a student's part-time work, especially for the academic-track students.

We next turn to the estimation results for the five IVs that are used to address the endogeneity of school-year hours worked. First, the results of the LR tests (Table 6) show that these IVs jointly and significantly affect the part-time hours worked for academic- and vocational-track students at the 5% and 1% levels of significance, respectively. In particular, since convenience stores and gasoline stations provide flexible working hours and are particularly attractive for youth, the local densities of these employment opportunities are significantly and positively associated with part-time work for academic-track students. However, the density of flyer stores in a student's residential neighborhood plays a negative role in influencing his/her part-time work. The unexpected sign of this IV may be due to its high correlation with the density of convenience stores and gasoline stations.

Moreover, vocational school students are more likely to undertake a part-time job in restaurants, fast food, or beverage stores. The density of these food stores in a student's residential zip-code neighborhood is positively and significantly associated with early employment at the 1% level of significance. The density of gasoline stations significantly and positively influences the early employment of vocational school students, reflecting the fact that part-time work in gasoline stations attracts vocational-track students. As expected, local economic conditions (measured by the youth unemployment rates of the population aged 15–19) are negatively and nonsignificantly associated with early employment.

As mentioned earlier in Section 2.2, in addition to imposing distributional assumptions, the Tobit models also restrict the sign of the regressor coefficients to be the same for the effects on part-time hours worked and on the probability of undertaking part-time work (i.e., the probability of hours being positive). To investigate this sign restriction inherent in the Tobit model in our application, we employ a two-step Heckman type estimation procedure to allow the level and probability effects to be distinct. We find that among significant control variables, 8 out of 10 (5 out of 7) coefficients for the academic- (vocational-) track share the same signs for the level and probability effects. This evidence indicates that the violation of the sign restrictions does not seem to be a serious problem in our application.[12]

5.3. Main results: the effects of early employment on scores on college entrance examinations

To investigate the role played by endogeneity from educational tracking, from part-time work, or from both in estimating the relationship between school-year hours worked and the scores on college entrance examinations, we consider four scenarios for endogeneity correction: (1) no endogeneity correction; (2) endogenous switching correction; (3) employment endogeneity correction; and (4) correction for both endogenous employment and endogenous educational tracking. These four scenarios correspond to the OLS, Switching, Censored IV, and Censored IV Switching estimators.

Since math, Chinese, and English are three general subjects that are required in both GSAT and TVET examinations, we mainly compare the scores of each of these three general subjects and the total score of these three (denoted by GSAT-3S and TVET-3S). We standardize the scores to allow us to compare the impacts of school-year hours worked on the z-score change in scores. Previous academic records, educational expectations, demographic factors, and family characteristics during the senior high school years are commonly controlled in each of the four scenarios for endogeneity correction.

Table 7 presents the main results for the effects of early employment on scores on college entrance examinations by Chinese, English, math, and the 3 subjects combined under four scenarios for endogeneity correction. Several important findings emerge in the third-stage estimation results. First of all, the estimation results for three subjects under the four specifications consistently reveal that school-year hours worked have negative impacts on the college entrance examination scores for students in both the academic and the vocational tracks after controlling for observed academic ability, demographic factors, and family characteristics. To be more specific, when no endogeneity (the OLS estimator) or only endogenous tracking (the Switching estimator) is corrected, the estimated effects of hours worked are not statistically significant, except for the Switching estimate for academic-track students. Yet, after correcting for the endogeneity of part-time work, the negative impacts of part-time work on the scores for college entrance examinations are largely strengthened. We find a significant and negative impact for students in both the academic-track and the vocational-track regimes. Furthermore, when compared to the Censored IV estimate that controls for only censored endogenous employment, the OLS estimates tend to understate the negative effect of school-year work on academic performance. A quantitatively similar pattern can also be found when one moves from Switching to Censored IV Switching estimates of the impact of part-time hours worked. These results clearly suggest that instrumenting early employment appears to be important in our specifications.

There are two possible explanations that are consistent with the aforementioned under-biasedness of the OLS and Switching estimates. First, the decision of how many hours to work appears to be positively related to unobserved characteristics (e.g., academic-related shocks or

[12] These results are available upon request from the authors.

Table 7
The third-stage estimation results: linear specifications for the z-scores of college entrance examinations by subject.

	Academic track				Vocational track				Difference between tracks
	OLS	Switching	Censored IV	Censored IV Switching	OLS	Switching	Censored IV	Censored IV Switching	Censored IV Switching
Total 3 subjects									
Observed or predicted	−0.084	−0.063***	−0.139**	−0.117**	−0.020	−0.016	−0.118**	−0.083*	−0.039***
hours worked/10	(0.054)	(0.025)	(0.055)	(0.046)	(0.033)	(0.018)	(0.049)	(0.049)	(0.013)
IMRs	–	0.197***	–	0.181***	–	0.089*	–	0.092*	0.065***
	–	(0.028)	–	(0.127)	–	(0.050)	–	(0.050)	(0.011)
Number of observations	526	526	558	558	370	370	409	409	967
Chinese									
Observed or predicted	−0.020	−0.012	−0.027	−0.012	−0.006	−0.008	−0.057	−0.030	−0.023
hours worked/10	(0.057)	(0.024)	(0.046)	(0.040)	(0.034)	(0.021)	(0.044)	(0.045)	(0.038)
IMRs	–	0.113***	–	0.112***	–	0.081	–	0.078*	0.025**
	–	(0.022)	–	(0.023)	–	(0.050)	–	(0.046)	(0.011)
Number of observations	532	532	564	564	399	399	440	440	1004
English									
Observed or predicted	−0.068	−0.048	−0.113**	−0.095**	−0.001	−0.004	−0.076	−0.059	−0.050
hours worked/10	(0.058)	(0.030)	(0.044)	(0.045)	(0.034)	(0.017)	(0.050)	(0.048)	(0.035)
IMRs	–	0.173***	–	0.157***	–	0.035	–	0.037	0.061***
	–	(0.027)	–	(0.026)	–	(0.042)	–	(0.043)	(0.011)
Number of observations	541	541	573	573	399	399	440	440	1013
Mathematics									
Observed or predicted	−0.093	−0.075**	−0.151***	−0.136***	−0.034	−0.024	−0.138***	−0.106**	−0.097**
hours worked/10	(0.060)	(0.036)	(0.052)	(0.049)	(0.035)	(0.024)	(0.050)	(0.053)	(0.040)
IMRs	–	0.144***	–	0.127***	–	0.090**	–	0.095*	0.039***
	–	(0.026)	–	(0.026)	–	(0.045)	–	(0.050)	(0.012)
Number of observations	536	536	570	570	373	373	412	412	982

Notes: 1. Bootstrap standard errors are in parentheses. 2. Control variables include prior academic scores, gender, races, parental education, family income and intact family in grade 11, educational expectations, and a constant. 3. The last column reports the coefficient estimates of the interacted predicted hours (IMRs) with the track dummy. We note that these numbers are not exactly the same as the coefficient differences between the academic and vocational tracks because the ridge regression is used in the third stage to address the multicollinearity problem. 4. *, **, and *** denote a significance level of 10%, 5%, and 1%, respectively.

private information such as responsibility and maturity that also affect academic performance), making the OLS estimated effect of early employment biased toward zero. The implication is that students experiencing stronger academic-related shocks or private information likely work more while in school and perform better academically. Second, the understatement of the negative impact of school-year hours worked is consistent with what is known as attenuation bias, which arises from the problem of measurement error and is likely to appear for self-reported hours of work. We note that this finding is also consistent with previous studies (e.g., Stinebrickner & Stinebrickner, 2003; Tyler, 2003) focusing on Western youth.

Secondly, our main contribution to the existing literature lies in the finding that endogeneity from educational tracking plays an important role in estimating the causal effects of early employment on educational achievement. We demonstrate the importance of educational tracking from two aspects. First, after correcting for the endogeneity from educational tracking, the negative impact of school-year work on college entrance examination scores becomes weaker but remains statistically significant. The results from the third and fourth columns of Table 7 imply that after controlling for the endogeneity from educational tracking, on average, exogenously increasing school-year hours worked by 10 hours per week lowers GSAT-3S scores for academic-track graduates from a −0.139 standard deviation to a −0.117 standard deviation. In other words, the negative impact of part-time work on academic achievement tends to be overestimated, with the bias growing as large as 16% for academic-track students if the endogeneity arising from educational tracking is not acknowledged. The implication is that the selection into educational tracking induces a positive correlation between part-time work and academic-related unobservables, thus offsetting the negative impact of early employment. These findings also apply to vocational-track students—the analogous effects of part-time work for vocational-track students are −0.118 versus −0.083 standard deviations, with the upward bias being 30% when only endogenous part-time work is controlled for. Second, the tracking correction terms (IMRs) are statistically significantly associated with the scores on college entrance examinations. It indicates that the selection into educational tracks in Taiwan is likely nonrandom, and thus the correlation between tracking behavior and academic performance on college entrance examinations cannot be ignored when estimating the effect of early employment on academic achievement. Furthermore, the coefficient estimates of selection correction terms are positive for both academic- and vocational-track students. Along with the fact that the selection correction term λ_{2i} for vocational-track students is negative, (i.e., $\lambda_{2i} = E\left(\varepsilon_{3i} | \varepsilon_{3i} \leq Z'_{Ti}\gamma\right) = -\phi\left(Z'_{Ti}\gamma\right)/\Phi\left(Z'_{Ti}\gamma\right) < 0$) this shows that students who choose the academic track are positively selected relative to those who choose the vocational track.

Thirdly, working while in academic school has a larger negative marginal impact on academic achievement than working while in vocational school, except for the nonsignificant effects on the Chinese subject. For example, the estimated effects on the math subject for the academic and vocational tracks are −0.136 versus −0.106 standard deviation, respectively. To examine whether part-time hours worked plays a significantly different role for academic- versus vocational-track students, the third-stage Censored IV Switching regression is further performed by pooling samples of both tracks and additionally interacting the tracking dummy variable with all regressors and the selection term. The last column of Table 7 presents the results on the differences in the coefficient estimates of hours worked and those of IMRs between the academic and the vocational tracks. The results show that part-time hours worked has significantly larger negative marginal effects on the scores of college entrance examinations for academic-track students than for vocational-track students, especially for the subjects of English, math, and the 3 subjects combined. In addition, the results for the IMR coefficients also reveal significantly distinctive effects between academic- and vocational-track students. These results again highlight the importance of the interaction between students' labor market work and educational tracking and the need for a separate treatment of academic and vocational tracks when examining the effects of part-time work.

Table 8
The third-stage estimation results: linear specifications for the z-scores of three subjects of college entrance examinations.

	Academic track: z(GSAT-3S)				Vocational track: z(TVET-3S)			
	OLS	Switching	Censored IV	Censored IV Switching	OLS	Switching	Censored IV	Censored IV Switching
Observed or predicted hours worked/10	−0.084	−0.063***	−0.139**	−0.117**	−0.020	−0.016	−0.118**	−0.083*
	(0.054)	(0.025)	(0.055)	(0.046)	(0.033)	(0.018)	(0.049)	(0.049)
Academic scores in primary school	0.146***	0.104***	0.127***	0.098***	0.120**	0.100***	0.103***	0.089**
	(0.044)	(0.029)	(0.025)	(0.025)	(0.059)	(0.035)	(0.032)	(0.033)
Academic scores in grade 9	0.305***	0.157***	0.195***	0.151***	0.286***	0.177***	0.184***	0.176***
	(0.037)	(0.020)	(0.020)	(0.020)	(0.046)	(0.028)	(0.028)	(0.029)
Male	−0.051	−0.040	−0.091*	−0.089*	−0.055	−0.044	−0.047	−0.051
	(0.072)	(0.049)	(0.047)	(0.053)	(0.090)	(0.058)	(0.057)	(0.053)
Parents' junior high graduates and below	0.064	0.091*	0.029	0.090	0.030	0.020	−0.004	0.007
	(0.098)	(0.052)	(0.055)	(0.058)	(0.094)	(0.064)	(0.057)	(0.060)
Parents' college graduates and above	0.235**	0.077	0.170***	0.073	0.589**	0.354**	0.264*	0.282**
	(0.092)	(0.060)	(0.056)	(0.060)	(0.275)	(0.158)	(0.139)	(0.137)
Intact family in grade 11	0.200	0.143**	0.105	0.105	−0.123	−0.040	−0.110	−0.096
	(0.124)	(0.070)	(0.070)	(0.072)	(0.131)	(0.078)	(0.073)	(0.075)
Family income in grade 11	0.003***	0.001**	0.001**	0.001	0.003*	0.002*	0.001*	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Student's educational expectation in grade 9	0.034**	0.015	0.032***	0.017	0.058**	0.043***	0.051***	0.046***
	(0.016)	(0.011)	(0.010)	(0.011)	(0.024)	(0.015)	(0.014)	(0.014)
Parents' educational expectation in grade 9	0.027**	0.012	0.020**	0.012	0.038**	0.029**	0.027**	0.025**
	(0.013)	(0.009)	(0.009)	(0.009)	(0.018)	(0.012)	(0.011)	(0.011)
IMRs (λ_1 or λ_2)		0.197***		0.181***		0.089*		0.092*
		(0.028)		(0.027)		(0.050)		(0.050)
Adjusted R-squared	0.275	0.296	0.265	0.294	0.209	0.207	0.200	0.208
Number of observations	526	526	558	558	370	370	409	409

Notes: 1. z(·) denotes z-scores. 2. GSAT-3S (TVET-3S) refers to the sum of the scores of the Chinese, English, and math subjects on the GSAT (TVET). 3. Bootstrap standard errors are in parentheses. 4. Constants and race variables are included. 5. *, **, and *** denote a significance level of 10%, 5%, and 1%, respectively.

Fourthly, labor supply during high school years has substantially heterogeneous effects on the scores of Chinese, English, and math subjects. To uncover important heterogeneities in student performance across the three different subjects, we further investigate the impact of part-time work on the scores of each subject under the same four scenarios for endogeneity correction. The results presented from Chinese to math in Table 7 show that the effects of part-time work on student performance are consistently negative across the four scenarios for endogeneity correction and across the Chinese, English, and math subjects. The magnitudes and significance, however, provide substantial heterogeneities in the effects of part-time work on student performance across these three subjects. For the Chinese subject, small and nonsignificant labor supply effects were found for students in both the academic and the vocational tracks. The panel for the English subject in Table 7 shows significant negative impacts of part-time work for academic-track students, but nonsignificant effects for vocational-track students. Quantitatively, for academic-track students, increasing work by 10 hours per week during school year lowers English scores by 0.095 standard deviation in the college entrance examination. Traditionally, vocational-track education has paid less attention to the English subject, which may explain why there is a nonsignificant difference in the English scores between employed and nonemployed students. Among subjects, math experiences the largest negative impact from working during high school—increasing the hours worked per week by 10 hours in school year lowers math scores in the college entrance examination by 0.136 (0.106) standard deviation for academic-track (vocational-track) students, respectively. Although school-year hours worked has a larger marginal effect on the math score for academic-track students than for vocational-track students, the total effects, however, may be reversed. Considering that, on average, academic- and vocational-track students spent 19.96 and 31.25 hours working per week, respectively, we find that conditional on being employed, undertaking part-time jobs leads to a 0.271 (0.303) standard deviation decrease in GSAT (TVET) math scores for academic-track (vocational-track) students.

Lastly, the impacts of individual and family characteristics on the scores on college entrance examinations are consistent with various endogeneity controls. Table 8 further presents more detailed estimation results along with other control variables. Previous academic scores substantially and positively affect scores on college entrance examinations. In particular, academic scores in the primary school years have a very long-lasting impact on college entrance examinations. Males perform worse in college entrance examinations than females. Parents' education and family income play a positive role for students in both the academic and the vocational track. Educational expectations from both students and parents significantly increase students' educational achievements, especially for students in the vocational track.

5.4. Linear or nonlinear effects of early employment?

This subsection further investigates if early employment affects students' educational achievement in a linear or a nonlinear way. Previous findings regarding the way in which early employment linearly or nonlinearly affects educational outcomes are rather mixed. Some studies find linear and negative effects of youth working while enrolled in school on academic outcomes (e.g., Eckstein & Wolpin, 1999; Marsh, 1991; Tyler, 2003). Evidence is also found that high school students who work a moderate number of hours per week tend to perform better academically than students who do not work or who maintain heavy work schedules (e.g., D'Amico, 1984; Greenberger & Steinberg, 1986; Lillydahl, 1990; Oettinger, 1999; Stern, 1995).

To capture the potential nonlinear effect of early employment, we employ the cubic B-spline approximating function,[13] which are one popular choice for splines in practice, of hours worked. We set 5 internal knots to the minimum, 0.25th quantile, 0.50th quantile, 0.75th quantile, and maximum of predicted hours. For the purpose of comparison with the linear effects, we normalize by equalizing the

[13] We thank an anonymous referee for suggesting this to us.

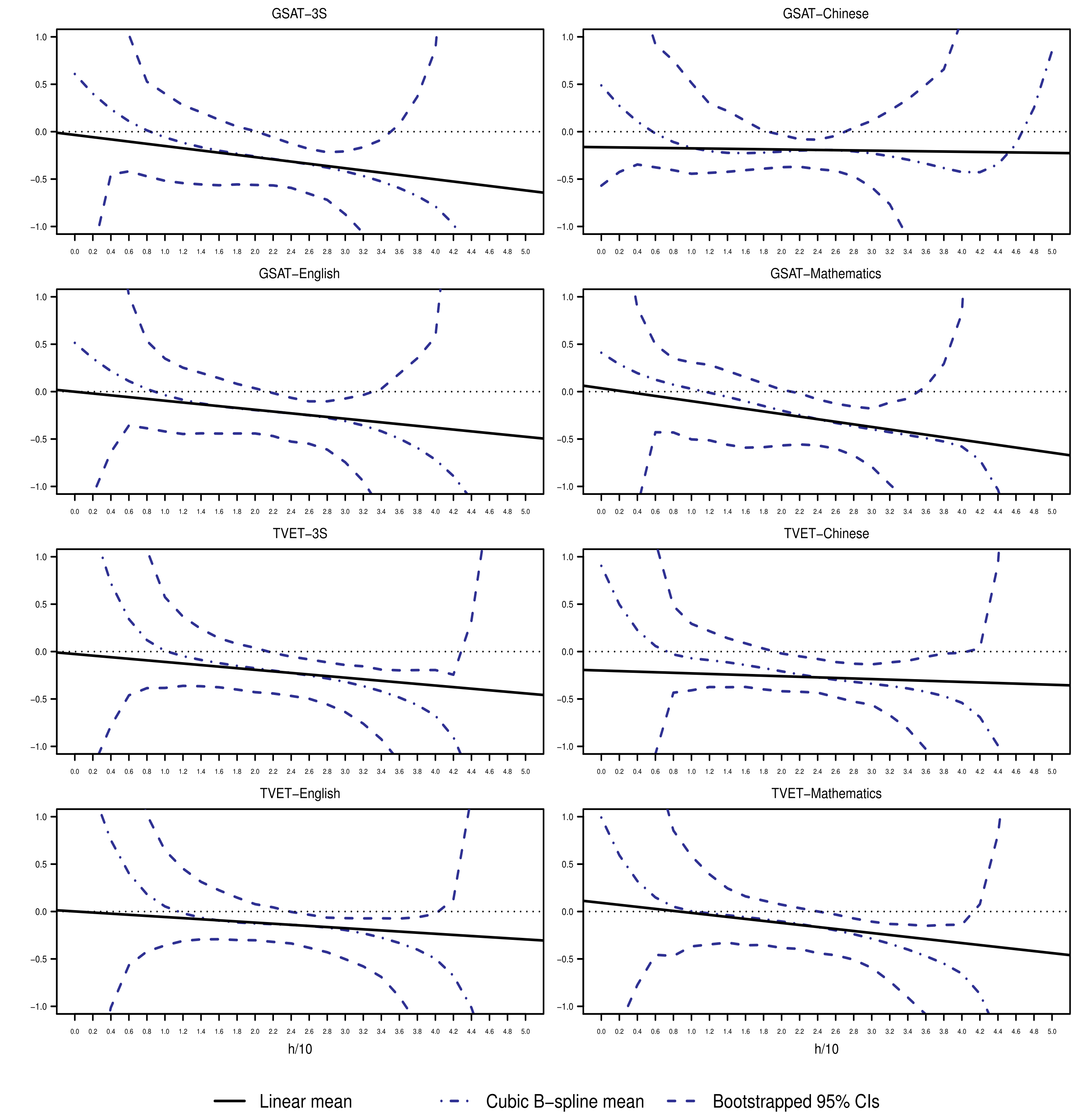

Fig. 3. Linear versus nonlinear effects of hours worked while in school on GSAT and TVET scores by subject.

estimated cubic B-spline function and the estimated linear function when evaluated at $h = 24$ hours worked. The results of the estimated cubic spline and linear mean functions, along with bootstrapped 95% confidence intervals for the former, are presented in Fig. 3. Fig. 3 reveals that while the cubic B-spline is nonlinear in nature, the estimated cubic B-spline mean functions of the educational attainment outcomes with respect to hours worked are essentially linear. Moreover, Fig. 3 also reveals that the estimated linear functions are entirely covered by the bootstrapped 95% confidence intervals based on the cubic B-spline estimation. These findings are robust across the two educational tracks and the three subjects.

5.5. Further results and robustness checks

The impacts of part-time work on performance in college entrance examinations may vary with academic ability. Table 9 presents the estimation results for the specification where part-time work is additionally interacted with academic record in grade 9. The results reveal that the linear negative impacts of early employment on GSAT-3S scores become slightly larger for academic-track students, while the interaction term has similar coefficient estimates across the three general subjects and is significant at the 5% level only for the English subject. In contrast, different patterns emerge for students in vocational

Table 9
The third-stage estimation results: interacting predicted hours worked with academic scores in grade 9 by subject.

	Academic track				Vocational track			
Dependent variables	z(GSAT-3S)	z(GSAT-C)	z(GSAT-E)	z(GSAT-M)	z(TVET-3S)	z(TVET-C)	z(TVET-E)	z(TVET-M)
Predicted hours worked/10	−0.135***	−0.024	−0.111***	−0.147***	−0.079*	−0.027	−0.056	−0.103**
	(0.043)	(0.039)	(0.040)	(0.047)	(0.045)	(0.046)	(0.047)	(0.048)
Predicted hours worked/10	0.013**	0.009*	0.011**	0.008	0.019***	0.014***	0.014***	0.015***
× academic scores in grade 9	(0.006)	(0.005)	(0.005)	(0.006)	(0.004)	(0.003)	(0.004)	(0.004)
IMRs (λ_1 or λ_2)	0.175***	0.108***	0.152***	0.124***	0.083*	0.071*	0.030	0.088*
	(0.026)	(0.024)	(0.025)	(0.024)	(0.050)	(0.043)	(0.041)	(0.045)
Control variables	yes	yes	yes	yes	yes	yes	yes	yes
Adjusted *R*-squared	0.292	0.154	0.225	0.204	0.213	0.174	0.120	0.142
Number of observations	558	564	573	570	409	440	440	412

Notes: 1. z(·) denotes z-scores. 2. GSAT-C, GSAT-E, GSAT-M, and GSAT-3S refer to the score of the subjects of Chinese, English, math, and the sum of the scores of these three subjects on the GSAT. The notations for TVET are defined analogously. 3. Bootstrap standard errors are in parentheses. 4. Control variables include prior academic scores, gender, races, parental education, family income and intact family in grade 11, and educational expectations. 5. *, **, and *** denote a significance level of 10%, 5%, and 1%, respectively.

Table 10
The third-stage estimation results: z-scores of total five-subject GSAT scores for academic-track students.

	Academic track: z(GSAT-5S)			
	OLS	Switching	Censored IV	Censored IV Switching
Observed or predicted hours worked	−0.076* (0.045)	−0.057*** (0.020)	−0.125*** (0.043)	−0.105*** (0.039)
IMRs (λ_1 or λ_2)		0.177*** (0.024)		0.162*** (0.022)
Control variables	yes	yes	yes	yes
Adjusted *R*-squared	0.358	0.391	0.357	0.385
Number of observations	830	830	885	885

Notes: 1. z(·) denotes z-scores. 2. Bootstrap standard errors are in parentheses. 3. GSAT-5S refers to the sum of the scores of the Chinese, English, math, natural science, and social science subjects on GSAT. 4. Control variables include prior academic scores, gender, races, parental education, family income and intact family in grade 11, and educational expectations. 5. *, **, and *** denote a significance level of 10%, 5%, and 1%, respectively.

schools. The higher a vocational school student's academic ability, the smaller is the negative impact of early employment on him/her. Our explanation is that it is well known that, in general, the GSAT examination is more difficult and more competitive than the TVET examination. As a consequence, to prepare for the GSAT, students on the academic track are more likely to attend a cram school after their regular school schedules than students on the vocational track. The time constraints likely outweigh the benefit of participating in the labor market for students with a higher academic record. The interaction of part-time work with academic ability plays a nonsignificant role. However, the vocational-track students with higher academic ability may benefit more from participating in the labor market when their time constraints are relatively light.

To check the robustness of our results, we use a total five-subject GSAT score (GSAT-5S) to investigate the role of endogeneity in educational tracking and part-time work for academic-track students. Table 10 presents the estimated effects of part-time work on the *z*-scores of the GSAT-5S under four specifications as noted earlier. The results are quantitatively similar to those for the three-subject GSAT-3S scores reported in Table 7—the marginal effect of part-time work on college entrance examination scores is −0.105, which is slightly lower than the −0.117 reported in Table 7 using GSAT-3S as an educational outcome.

More essentially, to examine for "placebo" effects, we further investigate the impacts of part-time work for those students who worked only during summer or winter vacations during high school year, which is expected to have smaller effects on academic performance. The results presented in Table 11 clearly reveal that in contrast to the estimated effects presented in Table 7, working part time only during summer and/or winter vacations has a nonsignificant and positive impact on college entrance examinations for both academic-track and vocational-track students. The estimated "placebo" effects are 0.050 (with a standard error of 0.106) and 0.059 (with a standard error of 0.057) standard deviations for academic- and vocational-track students,

Table 11
The third-stage estimation results: the impact of part-time work during summer and/or winter vacations while in high school.

	Academic track: z(GSAT-3S)				Vocational track: z(TVET-3S)			
	OLS	Switching	Censored IV	Censored IV Switching	OLS	Switching	Censored IV	Censored IV Switching
Observed hours worked/10	−0.036	−0.032			0.010	0.016		
	(0.033)	(0.022)			(0.030)	(0.019)		
Predicted hours worked/10			−0.061	0.050			0.037	0.059
			(0.121)	(0.106)			(0.054)	(0.057)
IMRs (λ_1 or λ_2)		0.202***		0.192***		0.093**		0.202***
		(0.029)		(0.029)		(0.047)		(0.029)
Control variables	yes	yes	yes	yes	yes	yes	yes	yes
Adjusted *R*-squared	0.277	0.300	0.257	0.290	0.204	0.200	0.198	0.207
Number of observations	516	516	558	558	334	334	409	409

Notes: 1. z(·) denotes z-scores. 2. GSAT-3S (TVET-3S) refers to the sum of the scores of the Chinese, English, and math subjects on GSAT (TVET). 3. Bootstrap standard errors are in parentheses. 4. Control variables include prior academic scores, gender, races, parental education, family income and intact family in grade 11, and educational expectations. 5. *, **, and *** denote a significance level of 10%, 5%, and 1%, respectively.

Table 12
Variable definitions and descriptive statistics.

Variable	Definition	Mean	SD	N
Dependent variables				
Educational Tracking	1 = senior high school, 0 = senior vocational school	0.45	0.50	2372
Hours worked	Hours worked per week in school year of grade 11	5.21	13.57	2030
GSAT-3S	Total scores in three subjects (Chinese, English, and math) of General Scholastic Ability Test (GSAT), ranging from 0 to 45, with each subject having a maximum of 15 scale points.	28.14	6.93	589
GSAT-C	Score in Chinese subject on GSAT, ranging from 0 to 15	11.37	2.20	597
GSAT-E	Score in English subject on GSAT, ranging from 0 to 15	9.50	3.33	605
GSAT-M	Score in math subject on GSAT, ranging from 0 to 15	7.32	3.54	602
GSAT-5S	Total scores in five subjects (Chinese, English, math, natural science, and social science) on GSAT, ranging from 0 to 75	47.45	10.74	1002
TVET-3S	Total scores in three subjects (Chinese, English, and math) of Technological & Vocational Education Test (TVET), ranging from 0 to 300, with each subject having a maximum of 100 points	170.44	44.37	447
TVET-C	Score in Chinese subject on TVET, ranging from 0 to 100	66.30	14.03	481
TVET-E	Score in English subject on TVET, ranging from 0 to 100	50.06	19.78	482
TVET-M	Score in math subject on TVET, ranging from 0 to 100	53.81	21.76	450
Independent variables				
Demographic factors and parental education				
Male	1 = male, 0 = female	0.49	0.50	2372
Minnan	1 = Minnan, 0 = other	0.77	0.42	2372
Hakka	1 = Hakka, 0 = other	0.07	0.26	2372
Mainlander	1 = Mainlander, 0 = other	0.13	0.34	2372
Native	1 = Native, 0 = other	0.01	0.09	2372
Parents junior high graduates and below	1 = Both parents have education levels equal to or below junior high school, 0 = other	0.35	0.48	2372
Parents college graduates and above	1 = Both parents have education levels equal to or above college, 0 = other	0.12	0.33	2372
Characteristics during junior high school years				
Academic score in primary school	The average academic score in primary school years	3.39	1.03	2372
Academic score in grade 9	Academic score in grade 9	2.88	1.31	2372
Student's tracking preference	1 = academic track, 0 = vocational track	0.59	0.49	2372
Parents attending School Day	1 = parents attending School Day in grade 9, 0 = otherwise	0.47	0.50	2372
Student's educational expectation	Years of schooling, student's own educational expectation	15.04	2.64	2372
Parents' educational expectation	Years of schooling, parents' expectation for their child	16.98	3.04	2372
Intact family in grade 9	1 = living in an intact family in grade 9, 0 = otherwise	0.88	0.32	2372
Family income in grade 9	Family income in grade 9 in thousand NT$	60.65	38.62	2372
Characteristics during senior high school years				
Intact family in grade 11	1 = living in an intact family in grade 11, 0 = otherwise	0.84	0.39	2372
Family income in grade 11	Family income in grade 11, in thousand NT$	60.31	38.96	2372
Instrumental Variables (IVs)				
Density of senior high schools in residential area	Number of senior high schools per square kilometer in students' zip-code residing neighborhood	0.18	0.15	2372
Density of food stores in residential area	Number of restaurants, fast-food, and beverage stores per square kilometers in students' zip-code residing neighborhood	31.67	42.09	2372
Density of gasoline stations (per km^2) in residential area	Number of gasoline stations per square kilometer in students' zip-code residing neighborhood	0.40	0.28	2372
Density of gasoline stations (per 1000 population) in residential area	Number of gasoline stations per thousand population age 15–19 in students' zip-code residing neighborhood	0.42	0.61	2372
Density of convenience stores in residential area	Number of convenience stores per thousand population aged 15–19 in students' zip-code residing neighborhood	7.88	6.68	2372
Density of flyer/leaflet services in residential area	Number of flyer/leaflet services per thousand population aged 15–19 in students' zip-code residing neighborhood	19.70	20.87	2372
Youth unemployment rate in residential area	Unemployment rate of young population aged 15–19 in students' zip-code residing neighborhood	14.07	0.70	2372

respectively, whereas their counterparts reported in Table 7 are −0.117 (with a standard error of 0.046) and −0.083 (with a standard error of 0.049) standard deviations. These statistically nonsignificant effects are consistently present across the four specifications we consider. These results provide further evidence that our estimation strategy has effectively addressed the selection and endogeneity problems.

6. Conclusions

This study has examined the association between educational tracking and school-year hours worked for 11th graders and their subsequent scores on college entrance examinations. Our results reveal that school-year work has significantly and substantially adverse effects on later educational achievement for both academic- and vocational-track students, and endogeneity from both part-time work and educational tracking can not be ignored. After accounting for the endogeneity of part-time work, the negative impacts of school-year work are largely strengthened. Such negative impacts of part-time work, however, tend to be overestimated, with the bias reaching as high as 16% and 30% for academic- and vocational-track students, respectively, if the endogeneity arising from educational tracking is not acknowledged. Among subjects, math scores suffer most from working during high school.

The policy implications that emerge from our empirical analysis are worth highlighting. First, early employment during the senior secondary school years will adversely impact college entrance. Reducing hours worked or prohibiting employment during the high school stage may preserve students' long-term educational achievement, especially for those who are interested in math-related fields. Second, working may be an undesirable necessity for low-income students, in the sense that their part-time work serves as the main financial source for their tuition and other expenses. For such low-income students, our results suggest that working during summer/winter vacation may be a better choice, and government tuition subsidies can provide additional support.

Our econometric strategy can also be applied to other interesting empirical settings. For example, an important topic in labor economics that is similar in concept to this study is to estimate the impacts of labor supply during college years when the choice between two- and four-year college is taken into account (e.g., Kane & Rouse, 1995; Light, 1999). Another potential application is in health economics and the investigation of the effects of the breastfeeding duration on later child health outcomes while accounting for the self-selection of delivery in baby-friendly hospitals (e.g., Howe-Heyman & Lutenbacher, 2016; Pérez-Escamilla, Martinez, & Segura-Pérez, 2016 for reviews of the literature).

CRediT authorship contribution statement

Fung-Mey Huang: Conceptualization, Methodology, Software, Data curation, Writing - original draft, Writing - review & editing. **Jen-Che Liao:** Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. **Chin-Chun Yi:** Conceptualization, Writing - review & editing.

References

Bera, A. K., Jarque, C. M., & Lee, L.-F. (1984). Testing the normality assumption in limited dependent variable models. *International Economic Review, 25*(3), pp.563–578.

Buscha, F., Maurel, A., Page, L., & Speckesser, S. (2012). The effect of employment while in high school on educational attainment: A conditional difference-in-differences approach. *Oxford Bulletin of Economics and Statistics, 3*, pp.380–396.

D'Amico, R. (1984). Does employment during high school impair academic progress? *Sociology of Education, 57*(3), pp.152–164.

Darolia, R. (2014). Working (and studying) day and night: Heterogeneous effects of working on the academic performance of full-time and part-time students. *Economics of Education Review, 38*, pp.38–50.

Dustmann, C., Rajah, N., & Smith, S. (1997). Teenage truancy, part-time working and wages. *Journal of Population Economics, 10*(4), pp.425–442.

Dustmann, C., & van Soest, A. (2007). Part-time work, school success and school leaving. *Empirical Economics, 32*(2–3), pp.277–299. https://doi.org/10.1007/s00181-006-0086-1.

Eckstein, Z., & Wolpin, K. I. (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica, 67*(6), pp.1295–1339.

Gamoran, A., & Mare, R. D. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality? *American Journal of Sociology, 94*(5), pp.1146–1183.

Goldstein, H., Burgess, S., & McConnell, B. (2007). Modelling the effect of pupil mobility on school differences in educational achievement. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 170*(4), pp.941–954.

Greenberger, E., & Steinberg, L. D. (1986). *When teenagers work : The psychological and social costs of adolescent employment.* New York: Basic Books.

Hanushek, E. A., & Wossmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal, 116*(510), pp.C63–C76.

Hosmer, D. W., & Lemesbow, S. (1980). Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics - Theory and Methods, 9*(10), pp.1043–1069. https://doi.org/10.1080/03610928008827941.

Hotz, V. J., Xu, L. C., Tienda, M., & Ahituv, A. (2002). Are there returns to the wages of young men from working while in school? *The Review of Economics and Statistics, 84*(2), pp.221–236.

Howe-Heyman, A., & Lutenbacher, M. (2016). The baby-friendly hospital initiative as an intervention to improve breastfeeding rates: A review of the literature. *Journal of Midwifery & Women's Health, 61*(1), pp.77–102.

Jarque, C. M., & Bera, A. K. (1980). *An efficient large-sample test for normality of observations and regression residualsTechnical Report*. Australian National University.

Kane, T. J., & Rouse, C. E. (1995). Labor-market returns to two- and four-year college. *The American Economic Review, 85*(3), pp.600–614.

Lee, C., & Orazem, P. F. (2010). High school employment, school performance, and college entry. *Economics of Education Review, 29*(1), pp.29–39. https://doi.org/10.1016/j.econedurev.2009.03.004.

Light, A. (1999). High school employment, high school curriculum, and post-school wages. *Economics of Education Review, 18*(3), pp.291–309. https://doi.org/10.1016/S0272-7757(99)00007-2.

Lillydahl, J. H. (1990). Academic achievement and part-time employment of high school students. *The Journal of Economic Education, 21*(3), pp.307–316.

Marsh, H. W. (1991). Employment during high school: Character building or a subversion of academic goals? *Sociology of Education, 64*(3), pp.172–189.

McCoy, S., & Smyth, E. (2007). So much to do, so little time: Part-time employment among secondary students in ireland. *Work, Employment & Society, 21*(2), pp.227–246. https://doi.org/10.1177/0950017007076630.

Meghir, C., & Palme, M. (2005). Educational reform, ability, and family background. *American Economic Review, 95*(1), pp.414–424. https://doi.org/10.1257/0002828053828671.

Oettinger, G. S. (1999). Does high school employment affect high school academic performance? *Industrial & Labor Relations Review, 53*(1), pp.136–151. https://doi.org/10.1177/001979399905300108.

Pekkarinen, T. (2008). Gender differences in educational attainment: Evidence on the role of tracking from a finnish quasi-experiment. *The Scandinavian Journal of Economics, 110*(4), pp.807–825.

Pérez-Escamilla, R., Martinez, J. L., & Segura-Pérez, S. (2016). Impact of the baby-friendly hospital initiative on breastfeeding and child health outcomes: A systematic review. *Maternal & Child Nutrition, 12*(3), pp.402–417.

Rothstein, D. S. (2007). High school employment and youths' academic achievement. *Journal of Human Resources, XLII*(1), pp.194–213. https://doi.org/10.3368/jhr.XLII.1.194.

Ruhm, C. J. (1997). Is high school employment consumption or investment? *Journal of Labor Economics, 15*(4), pp.735–776.

Steinberg, L., Greenberger, E., & Garduque, L. (1982). Effects of working on adolescent development. *Development Psychology, 18*, pp.383–395.

Stern, D. (1995). *School to work: Research on programs in the United States.* Washington and London: Taylor and Francis, Falmer Press.

Stinebrickner, R., & Stinebrickner, T. R. (2003). Working during school and academic performance. *Journal of Labor Economics, 21*(2), pp.473–491.

Triventi, M. (2014). Does working during higher education affect studentsacademic progression? *Economics of Education Review, 41*, pp.1–13.

Tyler, J. H. (2003). Using state child labor laws to identify the effect of school year work on high school achievement. *Journal of Labor Economics, 21*(2), pp.381–408.

Vandenberge, V. (2006). Achievement effectiveness and equity: The role of tracking, grade repetition and inter-school segregation. *Applied Economics Letters, 13*(11), pp.685–693. https://doi.org/10.1080/13504850500404944.

VanderHart, P. G. (2006). Why do some schools group by ability? Some evidence from the NAEP. *American Journal of Economics and Sociology, 65*(2), pp.435–462.