「지방정부연구」 제27권 제3호(2023. 11): 289~312
The Korean Journal of Local Government Studies, Vol.27 No.3 (November 2023)
http://dx.doi.org/10.20484/klog.27.3.12

Delayed Graduation and Employment Decision

Joo, Yejin*

국문요약

This study analyzes the effect of delayed graduation on employment. To address the endogeneity issue caused by selection bias, which has not been addressed in previous studies, we employ Propensity Score Matching(PSM). Survival analysis is then used for truncated data. The results from Kaplan-Meier non-parametric survival analysis indicate that graduation postponement reduces the employment probability under same period of job searching. To examine the relationship between other factors and employment, Cox proportional hazards model is applied. The number of leave semesters, graduation postponement experience, possession of certifications, experience in English tests, and age are negatively associated with employment probability. Being males, being married, having a higher GPA, and choosing a double major result in higher employment probability compared to their counterparts. Graduates from universities in the Seoul metropolitan area have higher employment probabilities compared to those from universities in other regions.

The finding that delayed graduation is not advantageous for employment contradicts human capital theory, suggesting that it may not contribute to enhancing one's productivity. Therefore, local governments and universities should provide support to students in advance through initiatives such as employment training and job placement programs in order to reduce the opportunity costs associated with the graduation deferment period.

Keywords: Delayed Graduation, Employment Decision, Propensity Score Matching, Kaplan-Meier Survival Analysis, Cox Proportional Hazards Model

Ⅰ. Introduction

According to the Korean Statistical Office's Survey on Economic Activity, the youth unemployment rate in Korea(aged 15-29) was 7.1% during the 2008 financial crisis. It steadily increased and reached 9.8% in 2016 and 2017 but has since shown a declining trend, reaching

* Ph.D Candidate, Department of Economics, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea, (TEL: 010-5203-4642)

6.4% in 2022.[1)] The decline is attributed to a reduction in the youth population and an expansion of employment subsidies. However, as employment is still difficult to meet individual expectations, there has been an increase in the graduation postponement among prospective college graduates. Graduation postponement is motivated by the perception that graduates who are not in sync with the scheduled graduates might face disadvantages in the job market. The goal is to elevate individuals' human capital levels, aligning them with their desired employment outcomes. However, according to existing theories, it is not clear whether such action helps labor market performance.[2)]

The theory of human capital, developed by Mincer(1974), posits that the abilities and productivity of workers vary based on the amount of human capital investment. This, in turn, is connected to the outcomes in the labor market, specifically employment prospects and wage differentials. According to this theory, graduation delay can be viewed as an action involving more time invested in education, training, and practical experience, accumulating human capital.

The Screening Hypothesis, developed by Spence(1973), suggests that education and vocational training serve as signals to employers when information about the abilities of workers is lacking, rather than merely enhancing their abilities and productivity. According to this theory, among workers with similar educational levels, a longer time to graduation can act as a negative signal of their lack of abilities for employers.

Previous overseas studies primarily focused on analyzing the effect of delayed degree attainment on wages, shedding light on which theoretical models, such as human capital theory or signaling theory, are more applicable to reality. Witteveen & Attewell(2019), revealed that delayed degrees were unrelated to employment opportunities but were associated with lower income, providing evidence for signaling theory. Aina & Pastore(2012) also found that delayed graduation was strongly correlated with lower wages. Ma et al.(2016) showed that individuals with a longer time to graduation took more time to recover the cost of college compared to those with a shorter time to graduation in the United States.

Domestic studies primarily focused on analyzing the severity of youth unemployment and the unique situation in Korea, characterized by frequent graduation delays. These studies examined the relationships between taking a leave of absence and employment, delayed graduation and

1) Statistics Korea. (2022). *Youth Employment Trends: Employment Rate and Unemployment Rate.*. Retrieved from 'Survey on Economic Activity of the Population'.

2) Variations in the Employment Prospects of 'Postponing Graduation' Leading to Confusion and Dissatisfaction Among Job Seekers, as Perceived Differences in Survey Institutions. (2019, March 15). *U's Line*. Retrieved from http://www.usline.kr/news/articleView.html?idxno=12640

employment, and delayed graduation and wages in the labor market. Regarding the effect of delayed graduation on the probability of employment, the findings of various studies have been inconsistent, with some suggesting no significant influence, while others indicating negative or positive effects, leading to inconclusive conclusions.

Kim et al.(2016) found that the deferral of graduation for four-year college graduates had no significant effect on labor market outcomes. Byun(2017) revealed that voluntary delayed graduation increased the probability of employment by 16.8% compared to regular graduation. Kim et al.(2018) showed that deferred graduates had a 7% higher employment probability than non-deferred graduates. However, the starting wages for deferred graduates were 30% lower than those for non-deferred graduates. Lee(2019), using propensity score matching and a sample-selection model, uncovered that non-deferred graduates had a 1.15 times higher probability of employment, and deferred graduates had a 6.3% higher average real wage compared to non-deferred graduates. Kim(2020) found that postponement of graduation had a positive effect on wages.

Meanwhile, this paper explores various supply-side factors, in addition to graduation delay, and their influence on employment prospects. The various factors can be explained using human capital theory and signaling theory as well as the Statistical Discrimination Theory. This theory, introduced by Arrow(1973) and Phelps(1972), suggests that employers, confronted with information asymmetry, discriminate based on past information about the groups to which job seekers belong rather than predicting their actual abilities and productivity. Thus, in the real labor market, even among workers with similar human capital, disparities in employment prospects and wages exist.

Examining various factors influencing employment in domestic research, Kim & Seo(2013) found that major, certifications, language study abroad, employment-related courses, and career development programs had an effect on employment. Among the educational support factors, satisfaction with educational infrastructure affected employment. University conditions and financial factors did not have significant effect on employment. In a study by Lee et al.(2016), factors such as gender, certified English scores, major satisfaction, certifications, on-campus employment courses, and GPA were identified as key factors for the transition to the labor market, in the case of humanities and social science graduates. For natural science graduates, major satisfaction, gender, direct employment policies, and GPA were identified as key factors. Park & Moon(2018) revealed that young individuals with labor experience during enrollment had higher employment probability and higher probability of employment in large corporations compared to those without such experience.

Additionally, researchers such as Park & Ban(2006) and Chae & Kim(2009) have studied the impact of various factors on employment, including part-time work experience, GPA, English scores, family income, major, and alma mater. In particular, Chae & Kim(2009) revealed that the only factor that could change through a student's own efforts and had a positive effect on employment was the GPA.

Previous studies focusing on delayed graduation and labor market outcome have generally conducted regression analysis, logistic and probit analysis, and survival analysis without addressing the issue of selection bias, which is a limitation. If the treatment group and the control group are not randomly assigned but rather self-selected, the result of the analysis would be biased. Kim et al.(2018) attempted to address this issue for the first time by applying a selection bias model in estimating the wage function. Lee(2019) also aimed to resolve this problem by introducing propensity score matching(PSM) along with the selection bias model.

This study pays attention to how delayed graduation is associated with employment of college graduates in Korea. Additionally, it examines whether factors other than delayed graduation are related to the employment of college graduates. The statistical data used for estimation is the Graduates' Occupational Mobility Survey(GOMS). Particularly, this study employs PSM to resolve the endogeneity issue present in previous domestic studies analyzing the effect of graduation deferral on employment and takes advantage of survival analysis by utilizing the information of censored data simultaneously.[3]

The structure of the paper is as follows: Chapter 2 explains propensity score matching (PSM), Kaplan-Meier survival function, and Cox proportional hazards model. Chapter 3 describes the data used, presents definitions and basic statistics for each variable, and derives the analytical results. Finally, Chapter 4 provides the conclusion.

Ⅱ. Propensity Score Matching and Survival Analysis

1. Propensity Score Matching(PSM)

Propensity score matching(PSM), proposed by Rosenbaum and Rubin(1983), offers the advantage of efficiently finding matched individuals, even when there are many confounding variables or continuous variables. By using this method, it is possible to address the endogeneity issue related to self-selection bias and estimate the causal effect of a program implementation.

In general, the effect of a program can be determined by examining the difference in effects(τ)

3) The term 'censored data' refers to individuals in the dataset who have not experienced the event of being employed within the maximum specified job search period of 24 months.

for the same entity when it participates in the program and when is does not participate. τ refers to the average treatment effect on the treated(ATT).

$$\tau = E(Y_1 - Y_0 | Z = 1) = E(Y_1 | Z = 1) - E(Y_0 | Z = 1)$$

Z represents program participation status(1, 0), Y_1 represents the treated outcome after program participation, and Y_0 represents the untreated outcome after non-participation. However, it is impossible for the same entity to simultaneously participate and not participate in the program in reality. To solve this problem, the unobservable expectation of the counterfactual situation $E(Y_0|Z=1)$ is replaced with the observable non-participating group expectation $E(Y_0|Z=0)$. This causes sample selection bias because the treatment group and the control group have different characteristics. Statistical matching can resolve this problem, constructing a comparison group that closely resembles the counterfactual situation.

Rosenbaum and Rubin defined the propensity score of individual i(i=1,..,N) as the conditional probability of being in the treatment group($Z_i = 1$) given the observed vector of covariates x_i.

$$e(x_i) = pr(Z_i = 1 | X_i = x_i)$$

If X is given and assumed to be independent of Z_i, the following holds.

$$\Pr(z_1, \ldots, z_N | x_1, \ldots, x_N) = \Pi_{i=1}^{N} e(x_i)^{z_i} (1 - e(x_i))^{1 - z_i}$$

When there exist covariates, under the assumption of strong ignorability, the average difference between the treatment group and the control group for all individuals with propensity scores is the average treatment effect of the propensity scores. This assumption is satisfied when the following two conditions hold for individuals given covariates. First, when the criterion for selecting the program participation group is given, the program participation indicator(Z) and the outcome variable(Y_0, Y_1) are conditionally independent.[4] Second, the probability distributions of treatment assignment in the treatment group and the control group overlap in a common support region.[5]When these conditions are met, controlling for covariates can eliminate selection bias. For the case where various variables are included, estimating a one-dimensional

4) $(Y_0, Y_1) \perp Z | X$, X is the vector of covariates.

5) $0 < \Pr(Z=1|X) < 1$.

propensity score from a multidimensional set of covariates can solve the dimensionality problem.

Rosenbaum & Rubin(1983) demonstrated that if the treatment assignment is independent of covariates, and the two key assumptions mentioned above are satisfied, the same assumptions hold for propensity scores as well. In other words, they established that $(Y_0, Y_0) \perp Z \mid P(X)$, $0 < \Pr(Z=1 \mid P(X)) < 1$ holds along with $X \perp Z \mid P(X)$. This means that, if the strong ignorability assumption holds, individuals in the control group are matched to individuals in the treatment group based on the same propensity scores, and as a result, the observable characteristics of both groups have the same distributions. Therefore, it is possible to estimate the effect without selection bias, similar to a randomized experiment.

After estimating propensity scores, the probability that observed individuals in the experimental group and the paired individual in the control group will have exactly the same propensity scores is low. To address this issue, in this study, Caliper matching was used. Caliper Matching enhances the matching quality by utilizing only the individuals within a specified maximum distance, caliper, as control group. As the size of the caliper decreases, the matching quality improves, but the sample size decreases, leading to an increase in the variance of the estimator. To overcome this drawback of increased variance in the estimator, if there are multiple non-participants within the caliper and multiple individuals in the experimental group, these individuals are utilized as multiple neighbors in the matching. In this study, 1:1 nearest-neighbor matching using caliper was employed.

The evaluation of matching quality is based on assessing whether the distribution of propensity scores and control variables between the experimental and control groups is balanced. If there are significant differences between the two groups, it indicates the inadequacy of the propensity score estimation model in the first stage. Therefore, it is necessary to return to the first stage and re-estimate the model by including interaction terms or squared terms in the model.

2. Survival Analysis

Survival analysis can be divided into three methods based on the way of setting the survival function: non-parametric methods using the Kaplan-Meier survival function, semi-parametric methods using the Cox proportional hazards model, and parametric methods. This study employs both non-parametric and semi-parametric methods.

First, the survival of job seekers implies the continuation of being in the job-seeking state until the point when employment is obtained. The period from graduation to the point of employment can be defined as the job-seeking period. Representing the job-seeking period as a random variable τ and the observed job-seeking period values as t, the probability density function and cumulative distribution function for the employment probability at time t are as follows.

$$f(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \le \tau < t + \Delta t)}{\Delta t}, \; F(t) = \int_0^t f(s)ds = \Pr(\tau \le t)$$

Here, if we consider the probability of remaining in the job-seeking state, the survival function is defined as follows.

$$S(t) = 1 - F(t) = \Pr(\tau > t)$$

The Kaplan-Meier survival function at time t is as follows. The estimated value of this function represents the probability of finding employment at time i.

$$S(t) = \prod_{t(i) < t} \left(\frac{n_i - d_i}{n_i} \right)^{\delta(i)}$$

$t(i)$ represents the job search duration for sample i, which is arranged in ascending order based on the job-seeking period. n_i, d_i represent the number of samples and the number of samples employed at time i, respectively. $\delta(i)$ indicates whether the event is censored(unknown or survive at time i).

Second, the semi-parametric Cox proportional hazards model to be estimated is as follows.

$$\lambda(t) = \lim_{dt \to 0} \frac{\Pr(t < \tau < t + dt \mid \tau > t)}{dt} = \lim_{dt \to 0} \frac{\Pr(t < \tau < t + dt) S(t) dt}{} = \frac{f(t)}{S(t)}$$

Here, the hazard function(hazard rate), $\lambda(t)$, represents the probability of a college graduate who had not yet found employment until time t finding employment immediately after time t.

The Cox proportional hazards model, expressed by the hazard function, is represented as $\lambda_i(t) = \lambda_0(t) e^{\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k}$. Here, $\lambda_0(t)$ represents the baseline hazard when all explanatory variables x are zero, and β represents the regression coefficient. By transforming it as follows, we can estimate the coefficients of the explanatory variables and determine the effect of specific factors on employment.

$$\frac{\lambda_i(t)}{\lambda_0(t)} = e^{\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k}$$

$$\ln\left\{\frac{\lambda_i(t)}{\lambda_0(t)}\right\} = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

On the other hand, the Hazard Ratio, denoted as e^{β_k}, represents the change in the hazard rate when a predictor increases by one unit.

$$\frac{\lambda_i(tvertx_k = k+1)/\lambda_0(t)}{\lambda_i(tvertx_k = k)/\lambda_0(t)} = \frac{\lambda_i(tvertx_k = k+1)}{\lambda_i(tvertx_k = k)} = \frac{e^{\beta_1(k+1)+\beta_2 x_2+\cdots+\beta_k x_k}}{e^{\beta_1 k+\beta_2 x_2+\cdots+\beta_k x_k}} = e^{\beta_1}$$

The difference in survival between the treated and control groups can be tested using the log-rank test. The log-rank test statistically tests the null hypothesis of "no difference in survival time between groups at all time points." The test statistic for the log-rank test follows a chi-square distribution with (number of groups - 1) degrees of freedom, and it is calculated as follows.

$$\chi^2_{\log rank} = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}, \quad E_1 = \sum \frac{d_i}{n_i} n_{1i}, E_2 = \sum \frac{d_i}{n_i} n_{2i}$$

Here, O_j represents the observed frequency of employment in group j, and E_j represents the expected frequency of employment.

Ⅲ. Empirical Analysis Results and Interpretation

1. Data

The population of the GOMS(Graduates Occupational Mobility Survey) statistics released by the Korea Employment Information Service consists of graduates from 2-3 year programs, 4 year programs, and college of education. Every year, a sample of 18,000 graduates from the previous year is selected, and a survey is conducted once in September of the following year. Although the survey was started in 2006, originally designed as a panel survey conducted annually, only cross-sectional surveys have been conducted since 2012.

For this study, considering the unique circumstances caused by COVID-19, the year 2019, which is the year preceding the outbreak, was set as the reference year. Additionally, graduates of educational colleges were excluded from the analysis in that they have predetermined career path after graduation and only a small number of samples exists. Furthermore, if the information regarding the first job was missing in the survey, the current job(employment during the period from August 25th to August 31st, 2019) was used as a substitute for the first job. The variable names, content, and descriptive statistics for each variable are presented in 〈Table 1〉 and 〈Table 2〉.

〈Table 1〉 Variable Names and Descriptions

	Variable Definitions	Variable Description
grent1	Total Required Semesters for Graduation	(Number of months between graduation year-month and enrollment year-month) divided by 6.
grft	Total Enrollment Period in Semesters	Total Required Duration for Graduation - Total Leave of Absence Semesters
f103	Total Number of Leave of Absence Semesters	f1a=Language Leave, f1b=Advanced Studies Leave, f1c=Certificate Leave, f1d=Military Leave, f1e=Economic Leave, f1f=Health Leave
f124d	Graduation Delay	0=Unknown or None, 1=Exists
sexd	Gender	Male=1, Female=0
age	Age	Only include individuals below 35 years old
mard	Marital Status	0=Unmarried, 1=Married, Divorced, Separated.
provinced_#	Location of the School	1=Seoul, 2=Gyeonggi Province, 3=Chungcheong Province, 4=Gyeongsang Province, 5=Jeolla Province (Benchmark : Seoul)
majord_#	Major Field	1=Humanities, 2=Social Sciences, 3=Education, 4=Engineering, 5=Natural Sciences, 6=Medicine, 7=Arts and Physical Education (Benchmark : Humanities)
majorcat	Major	
gpa	Graduation GPA	Converted to a scale of 4.5
train	Vocational Training	0=None, 1=Exists
m001d	Possession of Certification	
i018d	English Exam	
i001d	Language Training	
f023d	Double Major Status	0=No Double Major, Unknown, 1=Double Major with Linked Major or Minor.
j001d	Setting Employment Goals	Experience of Setting Employment Goals 0=None, 1=Exists
edufd_	Father's Highest Educational Attainment	1=Unknown/High School or Below, 2=College Graduate, 3=Generate dummy variables for Graduate School Graduate.
edumd_	Mother's Highest Educational Attainment	
area	Location of University(by City)	1=Seoul, 2=Busan, 3=Daegu, 4=Daejeon, 5=Incheon, 6=Gwangju, 7=Ulsan, 8=Gyeonggi Province, 9=Gyeongsangnam Province (Benchmark : Seoul)
p045d	Military Service	0=None, 1=Exists.
b034d	Foreign Language Institute Attendance	
b043d	Field Practice Internship	
h001d	Part-Time Job while Enrolled	
k001d	Job Seeking Activities	

〈Table 2〉 Basic Descriptive Statistics of Variables

Variable	Obs	Mean	Std. Dev.	Min	Max
Total Required Semesters for Graduation	17,002	10.6255	3.5982	3.8	29.8[6)]
Total Enrollment Period in Semesters	17,002	7.6131	1.9695	3.8	29.8
Total Number of Leave of Absence Semesters	17,002	3.0124	2.7731	0	19
Experience of Graduation Delay	17,002	0.1356	0.3423	0	1
Gender	17,002	0.5598	0.4964	0	1
Age	17,002	26.1535	1.9609	21.3	34.9
Marital Status	17,002	0.0236	0.1518	0	1
School Location(Seoul)	17,002	0.2432	0.4290	0	1
School Location(Gyeonggi Province)	17,002	0.2237	0.4167	0	1
School Location(Chungcheong Province)	17,002	0.1710	0.3766	0	1
School Location(Gyeongsang Province)	17,002	0.2539	0.4353	0	1
School Location(Jeolla Province)	17,002	0.1082	0.3106	0	1
Major(Humanities)	17,002	0.1322	0.3387	0	1
Major(Social Sciences)	17,002	0.1997	0.3998	0	1
Major(Education)	17,002	0.0596	0.2367	0	1
Major(Engineering)	17,002	0.2974	0.4571	0	1
Major(Natural Sciences)	17,002	0.1421	0.3492	0	1
Major(Medicine)	17,002	0.0637	0.2442	0	1
Major(Arts and Physical Education)	17,002	0.1053	0.3070	0	1
Graduation GPA	17,002	3.6286	0.4333	1	4.5
Vocational Training Experience	17,002	0.2201	0.4143	0	1
Possession of Certification	17,002	0.5578	0.4967	0	1
Experience of English Exam	17,002	0.3743	0.4839	0	1
Experience of Language Training	17,002	0.1057	0.3075	0	1
Double Major Status	17,002	0.1614	0.3679	0	1
Experience of Setting Employment Goals	17,002	0.4568	0.4981	0	1
Father's Education Level (High School or Below)	17,002	0.4827	0.4997	0	1
Father's Education Level(College Graduate)	17,002	0.4293	0.4950	0	1
Father's Education Level (Graduate School Graduate)	17,002	0.0880	0.2833	0	1
Mother's Education Level (High School or Below)	17,002	0.6198	0.4854	0	1
Mother's Education Level(College Graduate)	17,002	0.3404	0.4738	0	1
Mother's Education Level (Graduate School Graduate)	17,002	0.0398	0.1955	0	1

2. Propensity Score Matching(PSM) Results

First, postponement of graduation experience was set as the dependent variable, and a logistic function was estimated. The estimated logistic function provides the probability (propensity score) for the dependent variable.

⟨Table 3⟩ Logistic Regression Estimation Results

Variables	Coef.	Std.Err	95% Conf.Interval	
Age	0.4625***	0.0143	0.4344	0.4906
Gender	−0.6525***	0.1239	−0.8953	−0.4097
Marital Status	−0.5046**	0.1480	−0.7947	−0.2145
School Location	−0.1380***	0.0185	−0.1743	−0.1019
Major Field	−0.1174***	0.0140	−0.1446	−0.0901
Graduation GPA	−0.4367***	0.0555	−0.5454	−0.3280
Language Training Experience	0.1991**	0.0696	0.0628	0.3356
Military Service	−0.3200*	0.1238	−0.5627	−0.0773
Possession of Certification	0.0110	0.0494	−0.0858	0.1078
Double Major	0.5086***	0.0585	0.3940	0.6233
Father's Education Level	0.0729	0.0435	−0.0122	0.1581
Mother's Education Level	0.0291	0.4919	−0.0673	0.1255
Vocational Training	0.3167***	0.5518	0.2086	0.4249
Constant Term	−11.6706***	0.4481	−12.5488	−10.7924

Note: * p⟨0.1, ** p⟨0.05, *** p⟨0.01.
Log likelihood = −5807.2047, Number of obs = 17,271, LR chi2(13) = 1961.33
Prob⟩chi2 = 0.0000, Psedo R2 = 0.1445

After estimating the logistic model, matching was performed. 269 unmatched samples, including 1 from the treated group and 268 from the untreated group which were deviated from the common support region, were excluded. As a result, the analysis was conducted using the matched dataset consisting of 17,002 observations.

⟨Table 4⟩ Common Support

treatment assignment	Common Support		total
	off support	on support	
untreated	268	14,697	14,965
treated	1	2,305	2,306
total	269	17,002	17,271

The quality of matching was evaluated by assessing the balance between the treated and control groups after propensity score matching. Rubin(2001) considered Rubin's B to be sufficiently balanced when it was less than or equal to 25(Rubin's R: [0.5 2]). Here, Matched sample shows Rubin’s B=5.6, indicating sufficiently balanced result.

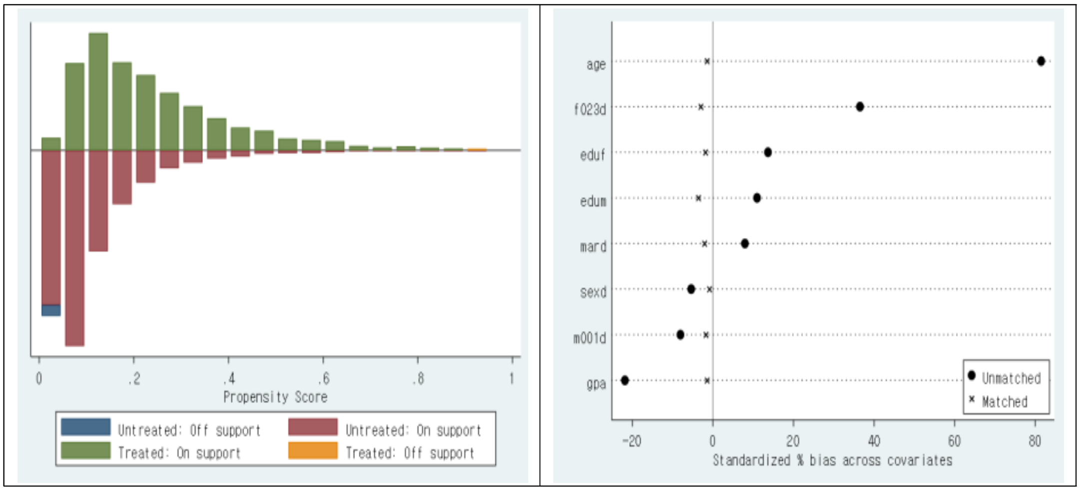
〈Table 5〉 Evaluation of Matching Results

Sample	Ps R2	LR chi2	p〉chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.137	1857.20	0.000	23.2	12.3	103.8[a)]	0.87	50
Matched	0.001	3.55	0.895	2.0	1.8	5.6	0.93	100

a) if B〉25%, R outside [0.5; 2]

The results of the matching can also be evaluated through graphical examination. 〈Figure 1〉 shows that the bias in the matched group is less pronounced compared to the unmatched group.

〈Figure 1〉 Propensity Score and Standarized % bias



〈Table 6〉 Estimated Average Treatment Effect Results

Variable	Sample	Treated	Controls	Difference	Std.Error	T-stat.
job	unmatched	0.8998	0.8634	0.0365	0.0076	4.82
	ATT	0.8998	0.8677	0.0321	0.0094	3.40

〈Table 6〉 shows the differences in the effect of the treated and control groups. The average treatment effect on the treated(ATT) shows a statistically significant result. Through bootstrapping with 100 iterations, the 95% confidence interval for the ATT was [0.0153, 0.0490] and the standard error was 0.0092.

3. Kaplan-Meier Analysis Results

〈Table 7〉 Survival Function Table for the Group without Graduation Delay

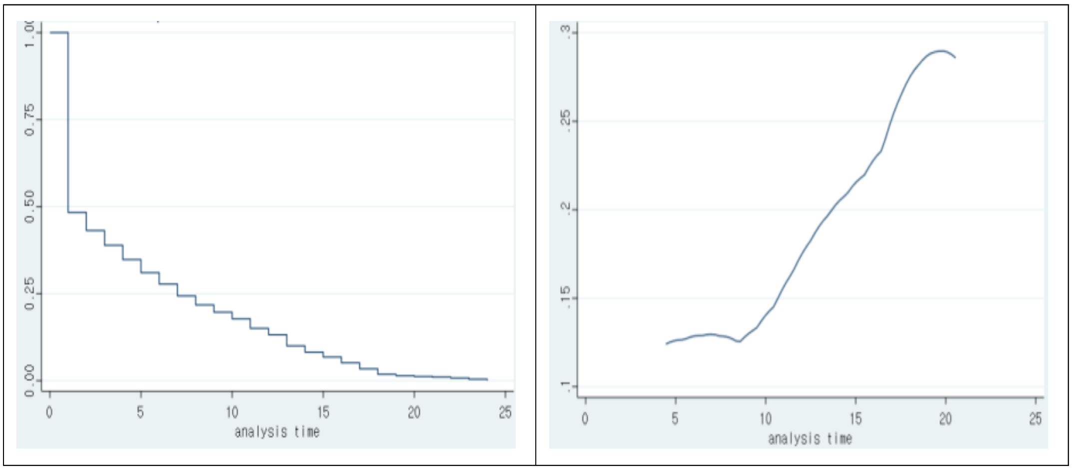
Time	Beg. Total	Fail	Net Lost	Suvivor Function	Std. Error	[95% Conf. Int.]	
1	12706	6557	0	0.4839	0.0044	0.4752	0.4926
5	4379	466	0	0.308	0.0041	0.3000	0.3160
10	2453	235	0	0.1746	0.0034	0.1680	0.1812
15	995	178	0	0.0643	0.0022	0.0601	0.0687
20	136	23	0	0.0089	0.0008	0.0074	0.0106
24	43	43	0	0	.	.	.

〈Table 8〉 Survival Function Table for the Group with Graduation Delay

Time	Beg. Total	Fail	Net Lost	Suvivor Function	Std. Error	[95% Conf. Int.]	
1	2074	1078	0	0.4802	0.011	0.4586	0.5015
5	762	91	0	0.3235	0.0103	0.3035	0.3437
10	456	46	0	0.1977	0.0087	0.1808	0.2151
15	217	32	0	0.0892	0.0063	0.0774	0.1020
20	73	7	0	0.0318	0.0039	0.0249	0.0400
24	22	22	0	0	.	.	.

The survival function tables for the samples with and without graduation deferment experience were truncated at 24 months. The summarized survival function tables provide the numbers of the unemployed individuals(survivors), which are represented in column Beg.total, and the numbers of the employed individuals(deaths), which are shown in column Fail, at monthly intervals. The number of employed individuals is calculated based on the endpoint, while the number of unemployed individuals is calculated based on the starting point of each interval. The estimated survival probability at time t, denoted as S(t), represents the cumulative survival rate and is the product of the survival rates for each interval.

〈Figure 2〉 Kaplan-Meier Survival Rate and Hazard Rate



The survival function curve illustrates the trend of cumulative survival rate over time. From this graph, it can be interpreted that the survival rate for individuals until the second period indicates a roughly 50% probability of remaining unemployed. In other words, individuals who are unemployed until the second period have a 50% chance of finding employment, and as the job search period extends, the probability of remaining unemployed decreases.

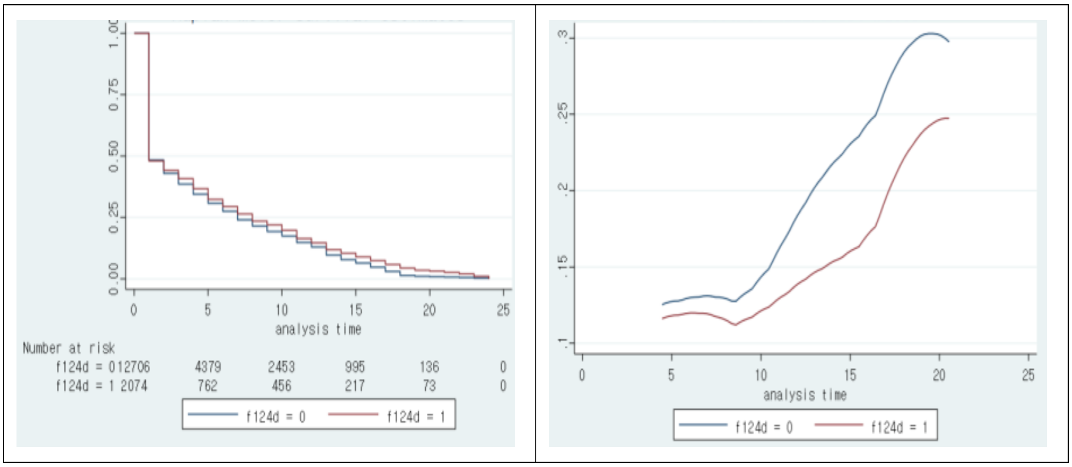
To investigate which variables influence employment, the data were divided into groups, and the survival probabilities for each group were estimated. In this case, the focus was on examining the association of graduation deferment experience(f124d) with employment. The analysis proceeded by initially conducting a Log-rank test to assess equality, followed by visually inspecting the graphs to identify differences.

〈Table 9〉 Log-rank test for equality of survival functions

failure_d : job analysis time_t : time		chi2(1) = 23.13 Pr>chi2 = 0.0000
Experience of Graduation Delay(f124d)	Events Observed	Events Expected
0	12706	12536.40
1	2074	2243.60
total	14780	14780.00

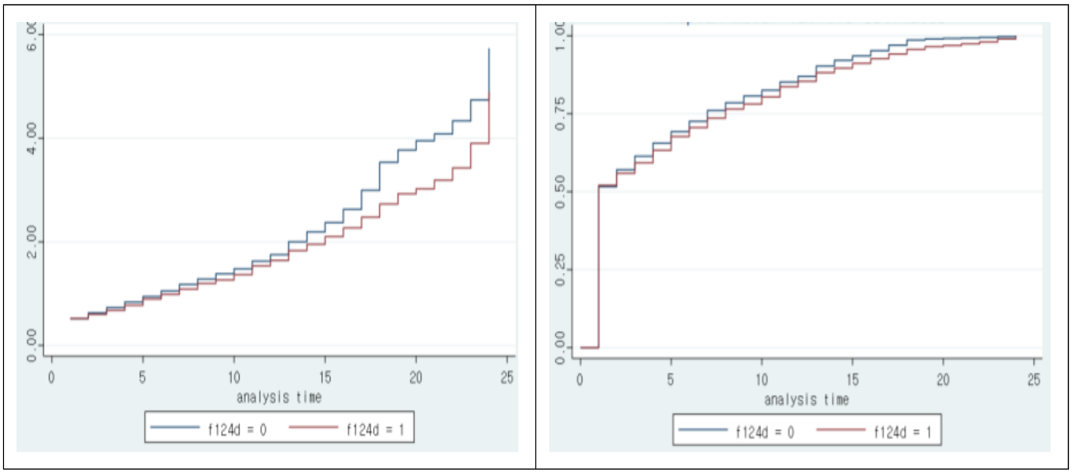
Since the p-value is 0.00, we reject the hypothesis that there is no difference in survival time between the groups at all time points. In other words, when dividing the groups based on graduation deferment experience(f124d), there exists a difference between the groups.

〈Figure 3〉 Survival Probability Graph and Hazard Rate between Two Groups



The Kaplan-Meier survival probability graph shows that the group without graduation deferment experience(represented by the blue line) has a lower probability of unemployment. This indicates that graduation deferment experience lowers the employment probability under same job searching period. The table below the graph represents the hazard table, indicating the number of unemployed individuals(survivors) at each time point.

〈Figure 4〉 Cumulative Hazard Rate and Failure Function between Two Groups



By considering the group-specific hazard rate graph, cumulative hazard rate graph, and failure function graph, it can be concluded that the group without graduation deferment experience(f124d=0) has a higher hazard and failure rate, while the group with graduation deferment experience(f124d=1) has lower hazard and failure rates. Since both hazard and failure represent employment, it can be inferred that graduation deferment experience has a negative effect on employment.

4. Cox Proportional Hazards Model Analysis Results

The Cox Proportional Hazards Model is a multivariate analysis technique that estimates the relationship between the event probability and one or more predictor variables which can be categorical or continuous. Under the assumption that other predictor variables included in the model are constant, it allows us to estimate the coefficients representing the influence of each predictor variable on the event probability.

The table below presents the results of the overall Cox proportional hazards model analysis. Each model shows statistical significance. Model (Ⅰ) represents the results of regression analysis using only five variables. In models (Ⅳ), (Ⅴ), and (Ⅵ), the coefficients for the total number of academic semesters, regional location in Chungcheong province, and major field in the categories of social sciences, natural sciences, and arts are not statistically significant at the 10% significance level.

Model(Ⅵ) shows that the total number of academic leave semesters, graduation deferment experience, possession of qualifications, experience of English tests, and age have a negative effect on the employment probability. It can be observed that being male, being married, having a high GPA, and having a double major lead to a higher employment probability. Looking at the influence of the location of the university in model (Ⅵ), the Gyeonggi and Chungcheong regions have 4.97% [(exp(-0.051)-1)*100] and 43.48% lower employment probabilities than Seoul area, respectively. It is also revealed that the Gyeongsang and Jeolla regions have 47.84% and 8.97% lower employment probabilities than Seoul area, respectively. In other words, it can be inferred that graduates from universities located in non-metropolitan areas have relatively lower employment probabilities compared to those from universities in the Seoul metropolitan area.

As for the major fields, the employment probabilities for the social sciences, engineering, natural sciences, medical and health sciences, and arts fields are higher than that of humanities by 0.89%, 6.3%, 2.51%, 27.97%, and 21.08%, respectively. In the case of education majors, the employment probability is lower by 10.77% compared to the humanities field. This demonstrates the reality that employment is more challenging in the humanities field compared to most other fields.

〈Table 10〉 Cox Proportional Hazards Model Analysis Results (Coef.)

Variable		(Ⅰ)	(Ⅱ)	(Ⅲ)	(Ⅳ)	(Ⅴ)	(Ⅵ)
Total Enrollment Period (Number of Semesters)					−0.0054 (0.0033)	0.0047 (0.0053)	0.0044 (0.0053)
Total Required Semesters for Graduation(Number of Semesters)		−0.0085*** (0.0033)	−0.0684** (0.0033)	−0.0056 (0.0033)			
Total Number of Leave of Absence Semesters							−0.1203* (0.0043)
Age		−0.0134** (0.0062)	−0.0156** (0.0062)	−0.0147** (0.0063)	−0.0146** (0.0063)	−0.0219*** (0.0058)	−0.015** (0.0063)
Experience of Graduation Delay		−0.0557** (0.0251)	−0.0553** (0.0253)	−0.0463* (0.0254)	−0.0462* (0.0254)	−0.0596** (0.0267)	−0.0673** (0.0269)
Gender		0.1108*** (0.0198)	0.1081*** (0.0211)	0.1049*** (0.0214)	0.1042*** (0.0214)	0.0988*** (0.0211)	0.1260*** (0.0232)
Marital Status		0.2910*** (0.0532)	0.2935*** (0.0532)	0.2841*** (0.0533)	0.2835*** (0.0533)	0.2913*** (0.0533)	0.2850*** (0.0533)
Seoul Area (Bench-mark)	Gyeonggi Province		−0.0475* (0.0247)	−0.0531** (0.0251)	−0.0544** (0.0251)	−0.0514** (0.0252)	−0.0510** (0.0252)
	Chungcheong Province		−0.0279 (0.0264)	−0.3300 (0.0270)	−0.0336 (0.0270)	−0.0358 (0.0270)	−0.3610 (0.0270)
	Gyeongsang Province		−0.5593** (0.0240)	−0.0636*** (0.0244)	−0.0652*** (0.0246)	−0.6633*** (.02456)	−0.6510*** (0.0246)
	Jeolla Province		−0.0794** (0.0307)	−0.0902*** (0.0314)	−0.0918*** (0.0315)	−0.9259*** (.03150)	−0.0937*** (0.0315)
Humanities (Bench-mark)	Social Sciences		−0.0055 (0.0295)	0.0070 (0.0300)	0.0066 (0.0300)	0.0067 (.03003)	0.0089 (0.0300)
	Education		−0.1090*** (0.0419)	−0.1077** (0.0428)	−0.1063** (0.0429)	−0.1081** (0.0429)	−0.1139*** (0.0430)
	Engineering		0.0343 (0.0286)	0.0600** (0.0298)	0.0600** (0.0298)	0.0606** (0.0298)	0.0611** (0.0298)
	Natural Sciences		0.0063 (0.0323)	0.0273 (0.0328)	0.2737 (0.0329)	0.0272 (0.0329)	0.0248 (0.0329)
	Medicine		0.2243*** (0.0395)	0.2542*** (0.0409)	0.2543*** (0.0409)	0.2582*** (0.0409)	0.2466*** (0.0411)
	Arts & Physical Education		0.0164 (0.0341)	0.0181 (0.0352)	0.0190 (0.0352)	0.0203 (0.0352)	0.1913 (0.0352)
Graduation GPA				0.0820*** (0.0195)	0.0820*** (0.0195)	0.0861*** (0.0194)	0.0818*** (0.0195)
Vocational Training Experience				0.0062 (0.0198)	0.0059 (0.0198)	0.0042 (0.0198)	0.0057 (0.0198)
Possession of Certification				−0.0251 (0.0176)	−0.2560 (0.0176)	−0.0250 (0.0176)	−0.0256 (0.0176)
Experience of English Exam				−0.0723*** (0.0186)	−0.7088*** (0.0187)	−0.7439*** (0.0187)	−0.7502*** (0.0187)
Experience of Language Training				0.0065 (0.0274)	0.0070 (0.0274)	0.0041 (.0273379)	0.0096 (0.0274)
Double Major Status				0.0549** (0.0241)	0.0553** (0.0241)	0.4982** (0.0241)	0.0510** (0.0241)
Experience of Setting Employment Goals				0.0028 (0.0171)	0.0029 (0.0171)	0.0021 (0.0171)	0.0025 (0.0171)
Father's Edu. Level (High School or Below)	College Graduate				−0.1306 (0.0198)	−0.0149 (0.0198)	−0.1510 (0.0198)
	Graduate School Graduate				−0.0045 (0.0347)	−0.0067 (0.0347)	−0.0078 (0.0347)
Mother's Edu. Level (High School or Below)	College Graduate				0.0027 (0.0207)	0.0022 (0.0207)	0.0017 (0.0207)
	Graduate School Graduate				−0.0276 (0.0472)	−0.0268 (0.0472)	−0.0283 (0.0472)

Note : * p〈0.1, ** p〈0.05, *** p〈0.01. (.)s are Std. Err. The degrees of freedom and LR values for each model were (5)76.41, (15)140.03, (22)176.37, (26)177.34, (26)175.44, (27)183.27, respectively.

The experience of English exams has a negative association with employment, while possession of certifications appears to have a negative effect although it is not statistically significant at the 10% significance level. These results reflect the recent hiring trend of companies prioritizing competencies, suitability for the job, and passion over qualifications or English proficiency. Regarding the parents' education level, a negative relationship with the employment probability is observed. This can be attributed to the expectation that increases as parents' education level rises, leading their children to make greater efforts to find good job opportunities even if it prolongs the job search period.

Ⅳ. Conclusion

The purpose of this study is to analyze the effect of graduation deferment on employment. There exist a number of previous studies that identify the effect of delayed graduation on employment. Nevertheless, most of these papers either fail to resolve the endogeneity issue or, even in cases where they have addressed selection bias, they simply conduct the difference-in-differences(DID) analysis without taking into account truncated data. This study addresses both of these issues using Propensity Score Matching-Survival Analysis.

The data used in this study is derived from the Graduates Occupational Mobility Survey(GOMS) published by the Korean Employment Information Institute. First, by using propensity score matching(PSM) which consists of caliper matching and logistic regression analysis, 332 unmatched samples were excluded. Then, survival analysis was conducted with a dataset of 17,002 matched data points to examine the influence of graduation deferment and other factors on post-graduation employment. Both non-parametric and semi-parametric survival analysis methods were employed. The result of comparing the groups with non-parametric anlaysis reveals that the probability of unemployment(survival probability) of the group without graduation deferment is lower than its counterpart. This suggests that graduation deferment reduces the employment probability in each job-seeking period.

Furthermore, to investigate the influence of factors other than delayed graduation on the probability of employment, Cox proportional hazards model was employed. Model (Ⅵ) shows the total number of leave of absence semesters, graduation deferment, possession of certifications, experience in English exams, and age have a negative effect on the probability of employment. It is also observed that the probability of employment is higher for males, married individuals, those with high GPAs, and those with double majors compared to their counterparts.

The probability of employment in the Chungcheong region is found to be lower than in the

Seoul region, although this difference is not statistically significant at the 10% significance level. On the other hand, the Jeolla, Gyeongsang, and Gyeonggi regions show statistically significant lower employment probabilities compared to the Seoul region. This phenomenon can be attributed to the concentration of employment opportunities, occupational training activities that university students can participate in, voluntary work, internships, and other opportunities in the Seoul region. It is also due to the preference for graduates from the Seoul region in the job market. These are the reasons why the majority of high school students from non-capital regions choose to attend universities in Seoul.

The probability of employment for all majors, except for the education field, is higher than that for humanities field. A higher level of parental education is associated with the lower employment probability, indicating that as parental education level increases, individuals are willing to endure longer job search periods in order to secure better job opportunities.

Contrary to the human capital theory perspective which suggests that graduation deferment can enhance productivity by improving credentials, the finding that graduation deferment is not advantageous for employment aligns better with the signaling theory, indicating a lack of preparation for employment. It can be interpreted as graduation deferment increases opportunity costs at the societal level as well as the personal level. Therefore, it is necessary to provide pre-employment training for students at the school level and support them with career roadmaps to reduce opportunity costs associated with delayed graduation. Additionally, the local government, the central government, and the universities should cooperate to offer support for various factors that can increase employment probabilities.

Particularly, the lower employment probabilities in non-capital regions compared to the capital region, found in the result of Cox regression analysis, is a serious problem. It is because the concentration of talent in the capital region can lead to inefficiencies in resource allocation and can potentially harm the country's long-term competitiveness according to national balanced development perspective. Therefore, in order to promote national balanced development, it is necessary to increase job opportunities for college graduates from non-capital regions in their respective areas. Currently, the government has announced a comprehensive plan for the era of local autonomy, including four special zones which are Opportunity Development Zones, Education Development Zones, Urban Convergence Zones, and Cultural Zones, to attract and nurture talent in the regions and create workplaces for them. Through the Opportunity Development Zones, companies will be attracted to the region, and the talent for these companies will be attracted and nurtured through the Education Development Zones. Based on these central government plan, local governments should expedite the enhancement of interconnections among regions, industries, and universities to focus on increasing employment prospects for regional universities as quickly as possible.

Due to the instability of statistical data related to the rapid changes in the environment surrounding the youth employment caused by the COVID-19 pandemic, this study utilized data from the year before the pandemic. Therefore, further comparative analysis using data from the COVID-19 period is necessary to examine the employment effects caused by non-face-to-face work.

References

Kim, B., Jo, D., & Choi, H. (2018). Analysis of the Employment and Wage Effects of Postponing Graduation. Korean Journal of Vocational Skills Development, 21(2):, 119-156.

Kim, B., & Seo, H. (2013). Analysis of Individual and University Factors Affecting the Employment of New College Graduates. Korean Journal of Education Finance and Economics, 22(4).

Kim, J., Lee, S., & Lee, S. (2016). A Study on the Current Status of Graduation Postponement among University Students and Support Measures. Research Report, Korea Youth Policy Institute.

Kim, J., & Jeong, I. (2002). Analysis of Awareness on the Economic Effects of Human Capital Investment. Journal of Education Finance and Economics, 11(1): 133-162.

Kim, J. (2019). A Study on the Causes of Youth Unemployment: Focusing on the Analysis of Labor Market Transitions. KDI Policy Study, 2019-14.

Kim, H. (2013). Experience of Taking a Leave of Absence during University Education and Employment. Korea Labor Institute Labor Policy Study, 13(3).

Kim, H. (2020). The Impact of Characteristics of University Graduates with Postponed Graduation on the Gap between Expected and Actual Wages. Korea Labor Institute Labor Policy Study, 20(2): 105-141.

Byun, J. (2017). Types of Graduation Delay among University Students and Employment Outcomes: Focusing on the Effects of Leave of Absence for Employment Purposes. Journal of the Korean Society for Industrial and Applied Mathematics, 18(6).

Chae, C., & Kim, T. (2009). Analysis of Factors Determining the Employment Outcomes of College Graduates. Vocational Education Research, 28(2): 89-107.

Choi, H., & Min, B. (2022). A Study on the Determinants of Survival of Social Cooperatives Using Survival Analysis, The Korean Journal of Local Government Studies, 26(3): 155-195.

Yang, J. (2015). Labor Market Conditions and Decision to Delay Graduation among University

Students. Applied Economics: 17(4).

Lee, G. et al. (2001). Empirical Study on Factors of Human Capital Formation in Companies and Knowledge Management Approaches for Index Development: Focusing on Domestic Listed Companies. Human Resource Management Research, 25(2): 29-63.

Lee, J., Kim, D., & Kim, J. (2016). Labor Market Outcomes of Leave of Absence: Focusing on Women Under 39 Years Old among University Graduates. Journal of Women's Economic Research: 13(1).

Lee, P. (2019). The Impact of University Students' Graduation Delay on Labor Market Outcomes. Journal of Education Finance and Economics, 28(3): 1-27.

Yun, M., & Kim, H. (2022). A Study on the Reemployment Experience and Policy Recognition of the Middle-Aged, The Korean Journal of Local Government Studies, 26(3): 281-305.

Jung, S., & Park, K. (2013). Employment and Wage Effects of Leave of Absence Periods. Proceedings of the 2013 Korea Labor and Employment Panel Academic Conference: 1-26.

Park, J., & Moon, Y. (2018). Determinants of Labor Market Outcomes for College Graduates. Korean Journal of Regional Sociology, Regional Studies, 26(4): 151-170.

Park, M. (2023). Reemployment Effect of the Female Internship System:Focusing on the Heckman Model and Sequential Mediating Effect, The Korean Journal of Local Government Studies, 27(1): 111-139.

Park, S., & Ban, J. (2006). The Status and Outcomes of Employment Preparation Efforts among College Graduates. Korean Journal of Population Studies, 29(3): 29-50.

Aina, C., & Pastore, F. (2012). Delayed Graduation and Overeducation: A Test of the Human Capital Model versus the Screening Hypothesis, IZA discussion paper, n..6413 (http://ftp. iza. org/dp6413. pdf).

Aina, C., & Pastore, F. (2020). Delayed graduation and overeducation in Italy: A test of the human capital model versus the screening hypothesis. Social Indicators Research, 152(2): 533-553.

Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Behr, A., & Theune, K. (2016). The causal effect of off-campus work on time to degree. Education Economics, 24(2): 189-209.

Bhaskaran, S. S., Lu, K., & Aali, M. A. (2017). Student performance and time-to-degree analysis by the study of course-taking patterns using J48 decision tree algorithm. International Journal of Modelling in Operations Management, 6(3): 194-213.

Böckerman, P., Haapanen, M., & Jepsen, C. (2018). More skilled, better paid: labour-market returns to postsecondary vocational education. Oxford economic papers, 70(2): 485-508.

Cunningham, S. (2021). Causal inference: The mixtape. Yale university press.
Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. Review of Economics and statistics, 84(1): 151-161.
Douglas, D., & Attewell, P. (2019). The relationship between work during college and post college earnings. Frontiers in sociology, 4: 78.
Ferrer, A. M., & Menendez, A. (2014). The puzzling effects of delaying schooling on Canadian wages. Canadian Public Policy, 40(3): 197-208.
Giani, M. S., Attewell, P., & Walling, D. (2020). The value of an incomplete degree: Heterogeneity in the labor market benefits of college non-completion. The Journal of Higher Education, 91(4): 514-539.
Heckman, J. J. (1979). Sample selection bias as a specification error. Econometrica: Journal of the econometric society: 153-161.
Ma, J., Pender, M., & Welch, M. (2016). Education Pays 2016: The Benefits of Higher Education for Individuals and Society. Trends in Higher Education Series. College Board.
Mincer, J. (1974). Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1): 41-55.
Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology, 66(5): 688.
Spence, M. (1973). Job market signaling. The Quarterly Journal of Economics, 87(3): 355-374. MIT Press, August.
Yu, X. (2021). Later timing but informed decision? Delayed postgraduate attainment and US college graduates' earnings. Social Science Research, 98: 102583.

주예진(周예진): 서울대 경제학과 대학원 박사과정을 수료하였다. 주요 관심분야는 정책 분석 및 평가, 인구 문제, 응용미시계량 등이다. 주요 논문은 “Do External Impacts Cause Differences between Regional and National Business Cycles?: Focusing on the Dongnam Metropolitan Area(2023)”, “Have Risk Management and Digitalization Enhanced the Efficiency of the Banking Industry in Korea?(2023)”이 있다. yjjoo@snu.ac.kr

〈논문접수일: 2023. 9. 23 / 심사개시일: 2023. 9. 27 / 심사완료일: 2023. 11. 7〉

국문 요약

대학 졸업 유예기간과 취업 결정

주예진

본 연구는 졸업 유예 경험이 취업에 미치는 영향을 분석한다. 선행 연구는 선택-편의와 절단된 정보를 동시에 고려하지 못해 이를 개선하기 위해 성향점수매칭-생존분석을 사용하였다. 분석 대상은 매칭된 17,002개의 GOMS 데이터이다. Kaplan-Meier 생존함수 분석 결과, 졸업 유예는 구직기간별 취업확률을 낮췄다. Cox 비례모형 분석 결과, 휴학 학기 수, 졸업유예 경험, 자격증 소지 여부, 영어시험 경험, 연령이 취업률에 음의 영향을 미쳤다. 남성이거나 혼인을 한 경우, 학점이 높거나 복수전공을 한 경우는 반대의 경우보다 취업률이 높았으며 서울권 대학 출신이 타지역 대학 출신자보다 취업확률이 높았다. 부모 학력이 높을수록 취업확률은 낮았다. 졸업 유예가 취업에 불리하다는 결과는 유예를 통해 생산성을 높일 수 있다는 인적자본이론의 관점과 대치된다. 대학과 지방정부는 사전에 직업 훈련 등으로 학생들의 취업을 지원하여 학생 개인과 사회 차원의 기회비용을 감소시켜야 한다.

키워드 : 졸업 유예, 취업 결정, 성향 매칭 점수, Kaplan-Meier 생존분석, Cox 비례위험 모형