Three Essays on Labor Markets and Higher Education

ED185002

KASHIMA, Ryohei [1]

October 5, 2022

[1]E-Mail: rkashi1228@gmail.com

Abstract

This dissertation considers three topics under higher education and the labor market and extends three existing bodies of literature considering the endogeneity of higher education. Chapter 2 extends the existing literature on the theoretical consideration of vertical education mismatch. In the previous research, a Diamond-Mortensen-Pissarides-type (DMP-type) job search model analysis of a 2×2 high-low labor market was mainly conducted. Only the highly educated unemployed could apply twice, i.e., highly educated job seekers just need a half of marginal search costs. Previous studies considering theoretically analyzing educational mismatch mainly assumed that exogenous proportion of workers take education even though endogenous education decision mechanism is widely known. In chapter 2, I consider the conditions under which equilibrium solutions involving cross-skilled matching exists (over-education in empirical studies) and what is optimal educational policies when marginal search costs are identical. The results show that cross-skilled matching exists as an equilibrium only under limited conditions where education costs are sufficiently low, and two tightness are not too far apart. Cross-skilled matching is more likely to exist to avoid congestion as the number of people receiving education increases. Although the optimal educational policy was to discourage schooling in an ex-ante segregated equilibrium from previous research, I conclude that schooling should be discouraged if the schooling and vacancy costs are small enough.

Chapter 3 extends the existing literature on the hysteresis effect of labor market participation during recessions to test whether higher education can mitigate this negative effect. I include education and its interaction term in the standard estimation model of the hysteresis effect and conduct an IV estimation to handle the newly introduced ability bias. The results show that higher education mitigates part of the curse effect, but four years of education do not fully cover the curse effect of a 1% increase in the unemployment rate at entry. I also analyze whether negative shocks at entry lead to job polarization and find that the long-term rate of job polarization is not affected by a temporary recession, but the number of workers with intermediate skills declines in the short

run. This suggests a relationship between the effects of the curse and the recession-induced decline in the number of middle-skilled workers.

Chapter 4 examines whether the contents of higher education matter to get some positive impact from higher education. As an example, I checked the effect of participation in short-term study abroad (SSA) programs from randomly assigned data of applicants from a Japanese university in March and August 2014 (705 applicants and 300 participants). My results demonstrate that participation in SSA causally increased English test scores and long-term study abroad (LSA) participation rates. Regarding the firms' initial employment attributes, this chapter found that SSA participants tended to work for firms with significantly higher sales and foreign stock rates more than non-participants, although I did not find significant differences in initial monthly income.

acknowledgement

First, I would like to express my sincerest gratitude to my advisor, Dr. Ryo Kambayashi, for his continuous support, patience, motivation, and vast knowledge of my doctoral thesis and related studies. His constant guidance in the writing of this dissertation has been invaluable. I could not have imagined having a better advisor and mentor for my doctoral thesis.

In addition to my advisor, I would like to thank the other dissertation committee members. They provided insightful comments and encouragement and answered tough questions that encouraged me to expand my research from different perspectives.

I would also like to express special thanks to Dr. Maki Kato for engaging in the co-authoring of chapter 4 of this dissertation.

I would also like to thank my research colleagues in rooms 204 and 212, with whom I have shared stimulating discussions and good times over the past five years. Finally, I would like to express my gratitude to my family for their support over the years.

Contents

Chapter 1

Introduction

In recent years, much of the applied econometric research on the effects of higher education and years of education has focused on the endogeneity of education (ability bias). Endogeneity occurs in analytical procedures when there are items that depend on people's choices. In the case of education, students with higher ability levels tend to go on to higher education, especially in the higher education stage after the compulsory education period, and the ability levels of higher education students and high school graduates are different. If this endogeneity is not treated, especially if one is interested in the "effect of education (return)," it makes us impossible to distinguish whether what appears as a coefficient value is due to ability or the "effect of education," to estimate a consistent coefficients. For this reason, researchers use exogenous variation in education to estimate consistent estimators by using the law reforms of the increase in the number of years of compulsory schooling (Acemoglu and Angrist, 2000; Oreopoulos, 2005; Pischke and von Wachter, 2008; Grenet, 2013) or by using the difference in the month of birth (Angrist and Krueger, 1991), or by using the distance between the higher education institution and the individual's home(Kane and Rouse, 1993; Carneiro et al., 2011; Doyle and Skinner, 2016), and so on. Many researchers in the field of education and labor economics have addressed endogeneity by using exogenous variation in education to estimate the matching estimator (Angrist and Krueger, 1991; Ashenfelter and Krueger, 1994; Blackburn and Neumark, 1995; Harmon and Walker, 1995; Ashenfelter and Zimmerman, 1997; Cameron and

Heckman, 1998; Cunha et al., 2006; Carneiro and Lee, 2009; Carneiro et al., 2011; Kamhöfer and Schmitz, 2016; Doyle and Skinner, 2016; Kikuchi, 2017).

However, despite the recognized importance, some fields treat education as an exogenous variable. This is mainly the case when one is interested in the differences in group returns between higher education students with relatively high ability levels and high school graduates with relatively low ability levels or for simplicity in analyses that explore developing fields or theoretical analyses. For example, Dolado et al. (2009) considered as an extension of the job search model the case where there are 2×2 high-skilled and low-skilled workers and jobs, but the skill level of the workers is exogenously determined by the structure. This is thought to be to eliminate the endogenously determined structure of education and focus the analysis on the presence or absence of the main on-the-job search function for the sake of simplicity of the model. However, while these studies are undeniably important, they are not suitable for considering the impact of educational decision-making and the effects of education.

In particular, higher education contains the most endogenous issue in education. In addition to the decision to take higher education, there is also the decision to choose a faculty. In such cases, it is difficult to ascertain the causal effect of which content of higher education has a positive effect on students' labor market outcomes. In addition, given that educational background is used for sorting in job offers, individual higher education decisions could affect the labor market equilibrium situation and dynamics of the labor market as a whole. When considering optimal higher education policy, it is necessary to consider how the labor market is affected by the existence of educational decisions.

With this background, this dissertation considers three topics related to higher education and the labor market and extends three existing bodies of literature to treat the endogenous nature of higher education. The analysis of each chapter and its results are as follows.

Chapter 2 extends the existing literature on the theoretical consideration of vertical education mismatch. In the previous research, a Diamond-Mortensen-Pissarides-type (DMP-type) job search model analysis of a 2×2 high-low labor market was mainly conducted. Only the highly educated unemployed could apply twice, i.e., highly educated job seekers just need a half of marginal search costs. Previous studies considering theoretically analyzing educational mismatch mainly assumed that exogenous proportion of workers take education even though endogenous education decision mechanism is widely known. In chapter 2, I consider the conditions under which equilibrium solutions involving cross-skilled matching exist (over-education in empirical studies) and what is optimal educational policies when marginal search costs are identical. The results show that cross-skilled matching exists as an equilibrium only under limited conditions where education costs are sufficiently low, and two tightness are not too far apart. Cross-skilled matching is more likely to exist to avoid congestion as the number of people receiving education increases. Although the optimal educational policy was to discourage schooling in an ex-ante segregated equilibrium from previous research, I conclude that schooling should be discouraged if the schooling and vacancy costs are small enough.

Chapter 3 extends the existing literature on the hysteresis effect of labor market participation during recessions to test whether higher education can mitigate this negative effect. I include education and its interaction term into the standard estimation model of the hysteresis effect and conduct an IV estimation to handle the newly introduced ability bias. The results show that higher education mitigates part of the curse effect, but four years of education do not fully cover the curse effect of a 1% increase in the unemployment rate at entry. Moreover, I analyze whether negative shocks at entry lead to job polarization and find that the long-term rate of job polarization is not affected by a temporary recession, but the number of workers with intermediate skills declines in the short run. This suggests a relationship between the effects of the curse and the recession-induced decline in the number of middle-skilled workers.

Chapter 4 examines whether the contents of higher education matter to get some positive impact from higher education. As an example, I evaluate the effect of participation in short-term study abroad (SSA) programs from randomly assigned data of applicants from a Japanese university in March and August 2014 (705 applicants and 300 participants). My results demonstrate that participation in SSA causally increased English test scores and long-term study abroad (LSA) participation rates. Regarding the firms' initial employment attributes, this chapter finds that SSA participants tended to work for firms with significantly higher sales and foreign stock rates more than non-participants, although I did not find significant differences in initial monthly income.

Finally, I conclude this chapter by summarizing and discussing the results of the entire results of the chapter. Chapter 2 examines the conditions under which workers voluntarily prefer vertical educational mismatch, a seemingly individually rational behavior, and the optimal educational policy under these conditions. Equilibrium is essentially determined by differences in worker productivity and labor market tightness. Chapter 3 examines whether workers' entry into higher education can reduce shocks to earnings when downward shocks to aggregate productivity are at work. Considering the downward shocks in chapter 3 as shocks to both sectors in the model of chapter 2, the level of education is expected to move little in the chapter 2 model. If the type of downward shock is different and the shock occurs in only one of the sectors, then Chapter 2 suggests that education levels would move. In other words, the estimates in Chapter 3 address the endogeneity of education only when there are downward shocks in all sectors. However, they suggest that the endogeneity problem recurs when there are different directions and types of shocks. In such cases, it is suggested that methods such as structural estimation, rather than reduced form estimation, are needed.

In Chapter 4, we analyze the returns to short-term study abroad, where the use of subsidies is prominent. However, causality is unclear in many aspects of the content of the returns to higher education. The results suggest that the returns are linked to the accumulation of human capital in the form of language. Regarding the returns of individuals who participated in the program, the effect of changing employment attributes was observed only for some students. Individual returns may have arisen from comparisons with students who did not participate in the program. Policymakers should consider the rise in language aspects as an accumulation of human capital. Whether the current amount of subsidies is truly appropriate is highly debatable. In addition, in recent years, Japan has been growing momentum to encourage all students to study abroad. However, the idea that if it is good, everyone should do it, as seen in these results, is the same as the argument that everyone should have higher education. Chapter 2 suggests that there may be an optimal level of education, which is the level of education most appropriate for all students and the level of education most appropriate for all students. Chapter 2 suggests that there may be an optimal level of education. In other words, making higher education available to everyone is not necessarily socially desirable. Similarly, it is important to consider that there is an optimal amount of permission to study abroad.

While considering the endogenous nature of education did not significantly change the story concerning Chapter 3, the story has changed in Chapters 2 and 4. For this reason, I believe that I have demonstrated the importance of considering and analyzing the endogenous nature of education in this dissertation.

Chapter 2

Voluntary Vertical Educational Mismatch in a Single-Application DMP Job Search Model

Abstract

This chapter extends the existing literature on the theoretical consideration of vertical education mismatch. In the previous theoretical analysis, a DMP-type job search model analysis of a 2×2 high-low labor market was mainly conducted. Only the highly educated unemployed could apply twice. It is identical as marginal search costs were half. For educational mismatch, education is considered exogenous even though an endogenous education decision mechanism exists. In this chapter, I consider the conditions under which equilibrium solutions involving cross-skilled matching exist (over-education in empirical studies) and what is optimal educational policies when marginal search costs are identical. The results show that cross-skilled matching exists as an equilibrium only under limited conditions where education costs are sufficiently low, and two tightness are not too far apart. Cross-skilled matching is more likely to exist to avoid congestion as the number of people receiving education increases. Although the optimal educational policy was to discourage schooling in an ex-ante segregated equilibrium from previous research, I conclude that schooling should be discouraged if the schooling and vacancy costs are small enough.

2.1 Introduction

While education creates avenues to acquire jobs with high educational requirements, often with high average incomes, some people experience a disconnect between their educational level and the job requirements. Several empirical analyses have shown that so-called vertical educational mismatches (also called overeducation), where workers are excessively educated for the level of education required for the job, result in lower job satisfaction and wages and higher separation rates than adequate workers (Verdugo and Verdugo, 1989; Hartog, 2000; McGuinness, 2006). These analyses emphasize the high mismatch rate of highly educated workers in developed countries. Combined with individual adverse effects, the complete elimination of vertical educational mismatch would improve the individual utility, suggesting that reducing the proportion of mismatched workers to zero may be better for social welfare as a whole. Most empirical research is dominated by the idea that over-education occurs as a result of inefficient misallocation due to the presence of some stigma.

It is also possible that vertical educational mismatches are occurring due to voluntary decision-making by workers. Vertical cross-skilled matching is considered in theoretical considerations of when vertical skill mismatch occurs, mainly using a DMP-type job search model(Albrecht and Vroman, 2002; Gautier, 2002; Dolado et al., 2009; Flinn and Mullins, 2015). Gautier (2002) assumed that in a 2×2 high-low type labor market, the highly educated unemployed could apply to both sectors simultaneously. At the same time, the less educated could only be matched with non-skilled jobs. Albrecht and Vroman (2002) and Dolado et al. (2009) also considered a single matching market. However, they assumed that the highly educated were matched with two types of jobs and that the low educated had zero productivity when matched with skilled jobs, and they were only matched with non-skilled jobs. Dolado et al. (2009) also considered an on-the-job search (OJS) mechanism. Under these assumptions, workers accepted a mismatch when the benefits of obtaining a job through a vertical mismatch were greater than unemployment as insurance. The existence and characteristics of such an equilibrium were discussed.

While searching for possible individual equilibria is important, these studies neglected to consider two points. First, they made strong assumptions about marginal search costs, and second, they ignored the endogenous decision structure of education. Regarding the first point, as described in the previous paragraph, highly educated workers can be matched with either job, whereas less educated workers are only non-skilled. In this case, it would be equivalent to doubling the time available for searching for highly educated. The oddness is especially apparent when OJS is included, as in Dolado et al. (2009). The marginal search cost is half that of the low-skilled unemployed only when the high-skilled are unemployed. However, the marginal search cost is identical when the low-skilled are OJS (strictly speaking, the marginal search cost is lower by the OJS efficiency). Second, they considered education to be determined by the exogenous fraction for the tractability of their model. When considering vertical educational mismatch, the endogenous structure of education has not been seriously considered. This is important in understanding how vertical educational mismatch occurs under voluntary educational decisions. In DMP-type search models, models that consider the endogenous decision structure of education have already existed(Charlot and Decreuse, 2005, 2010; Flinn and Mullins, 2015). Charlot and Decreuse (2005) considered the self-selection mechanism and concluded that a negative education subsidy is the optimal education policy.

By considering a DMP-type model in which two of the problems identified in previous studies are included in the model, this chapter derives the conditions for an economy with highly educated workers who voluntarily choose the non-skilled sector. Charlot and Decreuse (2005) also concluded that the optimal education policy is to discourage education in a labor market completely segregated by education. However, in a society that allows access to both sectors and allows educated workers to choose the sector where they get their jobs, the optimal education policy is determined by education cost and vacancy cost. The results suggest that the optimal policy will change depending on schooling and vacancy cost.

This study stands on DMP-type job search model (Diamond, 1982; Mortensen and Pissarides, 1994; Mortensen, 1986; Mortensen and Pissarides, 1999). The standard DMP-type model derives equilibrium conditions for the homogenous workers seeking homogenous jobs and conducts comparative statics. Various extensions have been conducted starting from a DMP-type model, such as changing how wages are determined, introducing a job search mechanism, considering heterogeneity between jobs and labor, having an endogenous decision structure for education, etc. This study combines two models into one and uses them in the analysis: Dolado et al. (2009)'s model that introduces an OJS mechanism while considering the heterogeneity of jobs and workers, and Charlot and Decreuse (2005)'s model that considers the endogenous decision structure of education. This chapter also related to previous empirical studies summarized by Hartog (2000), (McGuinness, 2006), and Leuven and Oosterbeek (2011) that examined the impact of vertical educational mismatch (also called overeducation).

The remainder of this chapter is structured as follows: Section 2 explains the model structures. Section 3 defines steady-state individual equilibrium, derives existence conditions, and conducts comparative static. Subsequently, Section 4 considers social optimal and optimal education policy. Finally, Section 5 concludes the chapter.

2.2 Models

2.2.1 Basic assumptions and settings and notations

I consider a steady-state economy in the model with continuous-time two-sided search and matching, where there are two types of jobs, skilled (index:$_s$) and non-skilled (index:$_n$), and two education types of workers, high educated (index:^{h}) and low educated (index:^{l}). I basically combine Dolado et al. (2009)'s model with Charlot and Decreuse (2005)'s model. Dolado et al. (2009) present equilibrium conditions for ex-post segmentation and cross-skill matching. However, they primarily show social welfare, comparative statics analysis, and numerical analysis in the cross-skill matching equilibrium. Charlot and Decreuse (2005) proposes a structure of educational choice in a job search model and suggests the existence of an optimal education cost policy. The inborn ability of workers (a) is continuously distributed $a \in [0, 1]$ (pdf: $\phi(a)$ cdf: $\Phi(a)$). Individual workers can observe their own abilities. Firms cannot observe workers' abilities before employment, and firms can observe workers' abilities after employment. Before entering the labor market, workers have to decide whether or not to receive an education. Additionally, I assume that highly educated workers must declare in which sector they search for jobs before entering the labor market to be fare marginal search cost between low and highly educated unemployed job searchers. About education, I assume that schooling takes a fixed exogenous cost (C_0) and schooling subsidy (C_1), which does not depend on individual ability. I also assume that education does not affect workers' ability itself, but education can change access to job sectors, and skilled sectors have higher productivity.[1]. There is an education restriction for production. Even if low-educated workers are matched with a skilled job, firms' production becomes 0[2]. Therefore, firms fill vacancies in skilled jobs with only high-educated workers, and vacancies in non-skilled jobs are filled by both low-educated and highly-educated workers. Workers know such a strategy of firms' employment. Thus, workers only apply for jobs they are likely to be hired. In other words, low-educated workers'do not apply for skilled jobs. At each instant, the proportion λ of all workers are born as unemployed workers. Each individual faces the same mortality risk λ and the total number of workers is constant (normalized to 1). I assume skilled sector firms can identify whether applicants come from unemployment or OJS.

2.2.2 Matching functions

Unemployed workers and firms' vacancies meet through a random match process. Let u_j^i denote the numbers of unemployed workers with i education who search in market j, $e_j k^i$ denotes the numbers of employed workers with i education working in market j who search k sector under unemployment status, and v_j denotes the numbers of the vacancy on market j. Even though Dolado

[1]If I assume the human capital feature, worker productivity increases with education. In other words, the gains from workers' schooling would be greater than in a world of fixed ability. In addition, firms' profits from employing highly educated workers would also increase. Both sides' profits would increase. Therefore, workers would take more education in the human capital world than in the Signaling world. There are some differences between the Signaling World and the Human Capital World, but the implications of the job search model are not significantly different. In this chapter, I consider the case of no human capital accumulation.

[2]Some jobs require a specific education level. For example, if you want to be a high school teacher in Japan, you must graduate from university (and need to take some specific classes in university). This assumption is usual in 2×2 matching models(Albrecht and Vroman, 2002; Gautier, 2002; Dolado et al., 2009).

et al. (2009) and Albrecht and Vroman (2002) consider non-directed type matching, I assume directed type matching function as Charlot and Decreuse (2005) and Gautier (2002) as following:

$$M_n(v_n, u_n^l + u_n^h) \tag{2.1}$$

$$M_s(v_s, u_s^h + e_{nn}^h) \tag{2.2}$$

I assume the matching functions M_j ($j = \{n, s\}$) follow ordinal search theory's matching function assumptions[3]. Also, these matching functions satisfy (2.3), (2.4) Inada conditions:

$$\lim_{v_j \to \infty} \frac{\partial M_j}{\partial v_j} = \lim_{(u^l + u^h) \to \infty} \frac{\partial M_n}{\partial (u^l + u_n^h)} = \lim_{(u_s^h + e_{nn}^h) \to \infty} \frac{\partial M_s}{\partial (u_s^h + e_{nn}^h)} = 0 \tag{2.3}$$

$$\lim_{v_j \to 0} \frac{\partial M_j}{\partial v_j} = \lim_{(u^l + u^h) \to 0} \frac{\partial M_s}{\partial (u^l + u_n^h)} = \lim_{(u_s^h + e_{nn}^h) \to 0} \frac{\partial M_s}{\partial (u_s^h + e_{nn}^h)} = \infty \tag{2.4}$$

From assumption of homogeneity, matching functions (the number of matching : M_j) can be transformed as matching rate functions depending on one argument of labor market tightness as $\theta_n = \frac{v_n}{u^l + u_n^h}$ or $\theta_s = \frac{v_s}{u_s^h + e_{nn}^h)}$, then $\frac{M_n(v_n, u^l + u_n^h)}{u^l + u_n^h} = M_n(\theta_n, 1) = m_n(\theta_n)$, and $\frac{M_s(v_s, u_s^h + e_{nn}^h)}{u_s^h + e_{nn}^h} = M(\theta_s, 1) = m_s(\theta_s)$. These $m_j(\theta_j)$ is job finding rate for workers' view ($m_j(\theta_j) \in [0, 1]$). $m_j(\theta_j)$ is increasing function for θ_j. From firms' point of view, matching rate is $\frac{M_n(v_n, u^l + u_n^h)}{v_n} = \frac{m_n(\theta_n)}{\theta_n} = p_n(\theta_n)$ and $\frac{M_s(v_s, u_s^h + e_{nn}^h)}{v_s} = \frac{m_s(\theta_s)}{\theta_s} = p_s(\theta_s)$. p_j ($\in [0, 1]$) is decreasing function for θ_j.

2.2.3 Value functions

This section presents a full set of value functions if a cross-skilled matching exists under a single application setting. In the case of ex-post segmentation, the model can be written using some equations in those systems.

[3]1: Homogeneous of degree one (constant return to scale(CRS)), 2: $M_n(0, u^l + u^h) = M_s(0, u^h + e_{nn}^h) = M_j(v_j, 0) = 0$ and must $Min(v_n, u^l + u^h) \geq M_n(v_n, u^l + u^h)$ and $Min(v_s, u^h + e_{nn}^h) \geq M_s(v_s, u^h + e_{nn}^h)$, 3: Non decreasing for each factors and concave (Petrongolo and Pissarides (2000)), 4: Twice continuously differentiable for easy mathematical reasoning (Charlot and Decreuse (2005)). I assume these 4 assumptions in this chapter.

Let $U_j^i(a)$ be the asset value of unemployment for i education individuals who search for jobs in the sector j. I also define $W_{jk}^i(a)$ as the asset value of getting a job in sector j for i education individuals who search for jobs in sector k when they are unemployed. The unemployment value function becomes as followings:

$$rU_b^l(a) = b \tag{2.5}$$

$$rU_n^l(a) = b + m_n(\theta_n)[W_{nn}^l(a) - U_n^l(a)] \tag{2.6}$$

$$rU_n^h(a) = b + m_n(\theta_n)[W_{nn}^h(a) - U_n^h(a)] \tag{2.7}$$

$$rU_s^h(a) = b + m_s(\theta_s)[W_{ss}^h(a) - U_s^h(a)] \tag{2.8}$$

Here, b denotes constant unemployment benefit. There exist some low-educated unemployed workers who always prefer unemployment benefits to work, and $j = b$ implies such a case. Highly-educated job seekers decide which sector they search for depending on their ability before entering the labor market by choosing equations (2.7) or (2.8). Let $w_{jk}^h(a)$ denote wages working in sector j for an individual with ability a who searches for a job in the sector k when they are unemployed. Working values are:

$$rW_{nn}^l(a) = w_{nn}^l(a) - \lambda[W_{nn}^l(a) - U^l(a)] \tag{2.9}$$

$$rW_{nn}^h(a) = w_{nn}^h(a) - \lambda[W_{nn}^h(a) - U_n^h(a)] + \eta m_s(\theta_s)[W_{sn}^h(a) - W_{nn}^h(a)] \tag{2.10}$$

$$rW_{sn}^h(a) = w_{sn}^h(a) - \lambda[W_{sn}^h(a) - U_n^h(a)] \tag{2.11}$$

$$rW_{ss}^h(a) = w_{ss}^h(a) - \lambda[W_{ss}^h(a) - U_s^h(a)] \tag{2.12}$$

Highly educated workers in the non-skilled sector take benefit from OJS. That is the third term of equation (2.9). I assume that OJS workers have less time searching for jobs than unemployed individuals. Hence there is exogenous penalty on OJS matching rate ($\eta \in [0, 1]$). The instantaneous benefit is wage ($w_j^i(a)$) depending on individual ability.

Next, let's consider firm-side values. Firms face vacancy cost (O_j) and firm side matching rate ($p_j(\theta_j) = \frac{M_j}{v_j} = \frac{m_j}{\theta_j}$). Before hiring, firms cannot observe applicants' abilities but can observe applicants' educational backgrounds and actions under unemployment status among highly educated. Based on this information, firms use three types of expected ability for hiring. $E[a^l]$ indicates the expected ability of low educated, and $E[a^h_j]$ indicates the expected ability of highly educated who search for jobs in the j sector under unemployment status. Let $J^i_{jk}(a)$ denote the asset value of a firm in the j sector hiring i education workers who seek work in the k sector when they are unemployed. Also, let V_j be the asset value of firms that post vacancies in the j-sector. Then, vacancy values are:

$$rV_n = -O_n + p_n(\theta_n)\xi_n[J^l_{nn}(E[a^l]) - V_n] + p_n(\theta_n)(1-\xi_n)[J^h_{nn}(E[a^h_n]) - V_n] \tag{2.13}$$

$$rV_s = -O_s + p_s(\theta_s)\xi_s[J^h_{ss}(E[a^h_s]) - V_s] + \eta p_s(\theta_s)(1-\xi_s)[J^h_{sn}(E[a^h_n]) - V_s] \tag{2.14}$$

Here $\xi_n = \frac{u^l}{u^l + u^h_n}$ and $\xi_s = \frac{u^h_s}{u^h_s + e^h_{nn}}$. Firms only accept if employment surplus is positive for mismatched workers. After employing, firms can observe the ability of workers from their production. Employment values are:

$$rJ^l_{nn}(a) = y_n(a) - w^l_{nn}(a) - \lambda[J^l_{nn}(a) - V_n] \tag{2.15}$$

$$rJ^h_{nn}(a) = y_n(a) - w^h_{nn}(a) - \lambda[J^h_{nn}(a) - V_n] - \eta m_s(\theta_s)[J^h_{nn}(a) - V_n] \tag{2.16}$$

$$rJ^h_{sn}(a) = y_s(a) - w^h_{sn}(a) - \lambda[J^h_{sn}(a) - V_s] \tag{2.17}$$

$$rJ^h_{ss}(a) = y_s(a) - w^h_{ss}(a) - \lambda[J^h_{ss}(a) - V_s] \tag{2.18}$$

where $y^i_{jk}(a)$ and $w^i_{jk}(a)$ indicates the productions and wages of i education individual working in sector j who search jobs in k sector when they are unemployed. Here, production does not depend on outside options of workers. Therefore, $y^l_{nn}(a) = y^h_{nn}(a) = y_n(a)$ and $y^l_{sn}(a) = y^h_{ss}(a) = y_s(a)$. Production functions become $y_s(a) = \tau(a)y_n(a), \tau(a) \geq 1$, and $\tau(a)$ is productivity gap rate depending on individual ability. Here, I assume production functions for both sectors are twice differentiable and monotone increases depending on individual ability. The instantaneous benefit is $y_j^i(a) - w_j^i(a)$, which is profit for firms, and other terms are expected values of transition. The highly educated employees in the non-skilled sector have higher production and job destruction rate than those lower educated because of OJS. Firms in the non-skilled sector enjoy the benefits of higher productivity when they hire highly educated workers, but the negative side is higher turnover because of the OJS mechanism.

2.2.4 Search sector decision for highly educated unemployed job seekers

To be the fair marginal cost of search, I assume a single application for highly educated unemployed workers. Educated workers can apply to both sectors, but they need to choose in which sector they search with unemployment status. Highly-educated unemployed job seekers choose the non-skilled sector if the payoff stream of starting a non-skilled sector is greater than starting a skilled sector. That is:

$$U_n^h \geq U_s^h \tag{2.19}$$

$$\Leftrightarrow m_n(\theta_n)[W_{nn}^h(a) - U_n^h(a)] \geq m_s(\theta_s)[W_{ss}^h(a) - U_s^h(a)] \tag{2.20}$$

is satisfied. Let $\sigma_1 \in (0,1)$ denote the endogenous cut-off point below which educated unemployed individuals search for jobs in the non-skilled sector.

2.2.5 Schooling decision for workers

This subsection follows Charlot and Decreuse (2005)'s setting. Workers take education if the payoff stream for high educated minus the direct cost of education is greater than the payoff stream for low educated. That is:

$$\max_a \{U_n^h(a), U_s^h(a)\} \geq U_n^l(a) + C_0 - C_1 \tag{2.21}$$

where C_0 denotes the exogenous cost of education and C_1 denotes the government's lump-sum subsidy for education. Let $\sigma_2 \in (0, 1)$ denote the endogenous cut-off point below which individuals do not take education. Comparing to Charlot and Decreuse (2005), this chapter needs to maximize form to account for search sector decisions. Suppose there exists a cross-skilled matching, $\max_a\{U_n^h(a), U_s^h(a)\} = U_n^h(a)$ for $\sigma_2 \leq a \leq \sigma_1$ and $\max_a\{U_n^h(a), U_s^h(a)\} = U_s^h(a)$ for $\sigma_1 < a \leq 1$.

2.2.6 Steady state condition based on worker flows and stocks

I focus on the steady-state situation, which implies raw of motion is equal to 0. The size of the labor force normalized to 1. Let u_j^i be the mass of the number of unemployed i education job seekers who search for a job in the sector j. Let e_{jk}^i be the number of i education workers in j sector who search for a job in the sector k when they are unemployed. u_j^i and e_{nn}^h obey raw of motion as follows:

$$\mathrm{d}u_n^l/\mathrm{d}t = \lambda(\Phi(\sigma_2) - \Phi(\sigma_3)) - (m_n(\theta_n) + \lambda)u_n^l = 0 \tag{2.22}$$

$$\mathrm{d}u_n^h/\mathrm{d}t = \lambda(\Phi(\sigma_1) - \Phi(\sigma_2)) - (m_n(\theta_n) + \lambda)u_n^h = 0 \tag{2.23}$$

$$\mathrm{d}u_s^h/\mathrm{d}t = \lambda(1 - \Phi(\sigma_1)) - (m_s(\theta_s) + \lambda)u_s^h = 0 \tag{2.24}$$

$$\mathrm{d}e_{nn}^h/\mathrm{d}t = m_n(\theta_n)u_n^h - (\eta m_s(\theta_s) + \lambda)e_{nn}^h = 0 \tag{2.25}$$

2.2.7 Expected ability

Let a_j^i denote the ability of i education individuals who search for a job in the sector j when they are unemployed. The mean ability of each type of worker becomes as follows:

$$E[a_n^l](\sigma_2, \sigma_3) = \int_0^{\sigma_2} \frac{\phi(a)}{\Phi(\sigma_2) - \Phi(\sigma_3)} a da \tag{2.26}$$

$$E[a_n^h](\sigma_1, \sigma_2) = \int_{\sigma_2}^{\sigma_1} \frac{\phi(a)}{\Phi(\sigma_1) - \Phi(\sigma_2)} a da \tag{2.27}$$

$$E[a_s^h](\sigma_1) = \int_{\sigma_1}^{1} \frac{\phi(a)}{1 - \Phi(\sigma_1)} a da \tag{2.28}$$

$E[a_j^i]$ is an increasing function of the dependent cut-off. This is the composition effect described by Charlot and Decreuse (2005).

2.2.8 Model dynamics

Based on above settings, I can graphically show the model dynamics and settings as Figure 2.1. Above boxes indicate firm side vacancy values, bottom boxes indicate worker side unemployment values and middle boxes indicate matching values for both. Below line indicates workers' ability distribution in 0 to 1. Each arrows indicates workers' flow.

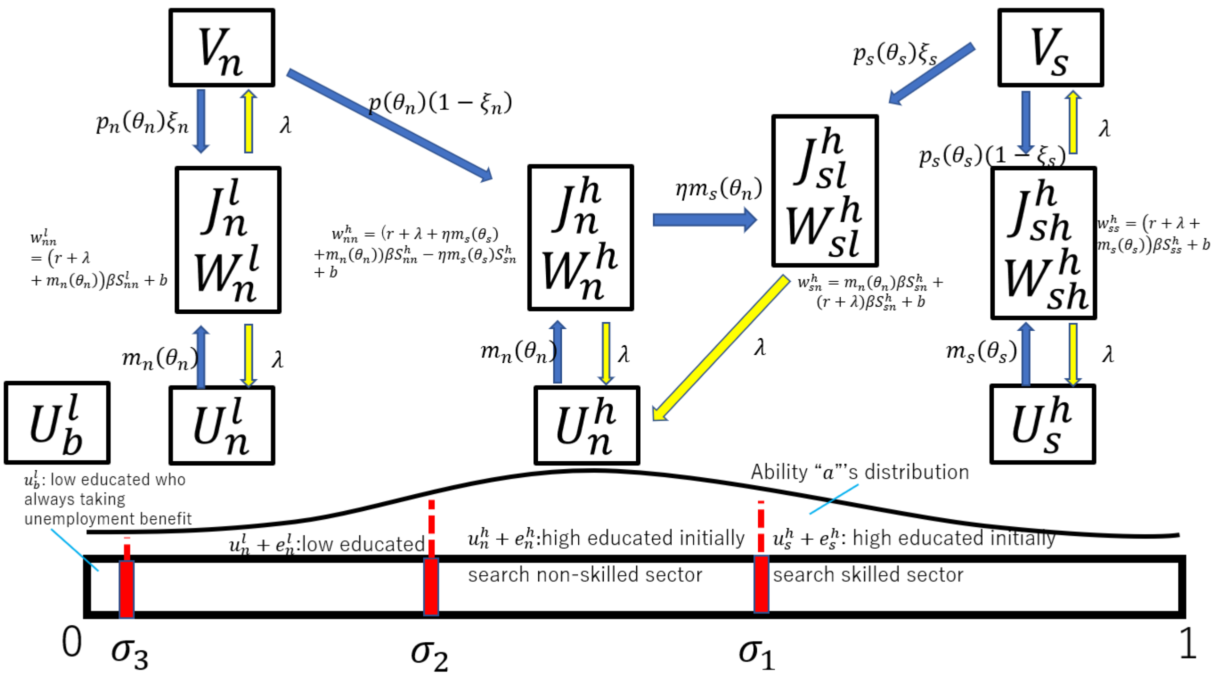

Figure 2.1: Model dynamics and situation existing cross-skilled matching and 4 types workers

2.3 Individual Equilibrium

2.3.1 Closing the model

Wage decision through bargaining

I assume wages are determined through Nash bargaining over the matching surplus as:

$$\beta(J^i_{jk}(a) - V_j) = (1-\beta)(W^i_{jk}(a) - U^i_j) \tag{2.29}$$

Here, $\beta \in (0,1)$ indicates the exogenous bargaining power of the workers. Based on equations (2.6)-(2.18) and (2.29), workers' wage schedules are determined. Here, I show wage equations using matching surplus $S^i_{jk}(a) = W^i_{jk}(a) - U^i_j(a) + J^i_{jk}(a) - V_j$ as follows:

$$w^l_{nn}(a) = (r + \lambda + m_n(\theta_n))\beta S^l_{nn}(a) + b \tag{2.30}$$

$$w^h_{nn}(a) = (r + \lambda + m_n(\theta_n) + \eta m_s(\theta_s))\beta S^h_{nn}(a) - \eta m_s(\theta_s)\beta S^h_{sn}(a) + b \tag{2.31}$$

$$w^h_{sn}(a) = (r + \lambda)\beta S^h_{sn}(a) + \beta m_n(\theta_n) S^h_{nn}(a) + b \tag{2.32}$$

$$w^h_{ss}(a) = (r + \lambda + m_s(\theta_s))\beta S^h_{ss}(a) + b \tag{2.33}$$

Free entry and job creation conditions

I can rearrange employment values (2.15)-(2.18) based on free entry conditions $V_j = 0$,

$$J^l_{nn}(a) = \frac{y_n(a) - w^l_{nn}(a)}{r + \lambda} \tag{2.34}$$

$$J^h_{nn}(a) = \frac{y_n(a) - w^h_{nn}(a)}{r + \lambda + \eta m_s(\theta_s)} \tag{2.35}$$

$$J^h_{sn}(a) = \frac{y_s(a) - w^h_{sn}(a)}{r + \lambda} \tag{2.36}$$

$$J^h_{ss}(a) = \frac{y_s(a) - w^h_{ss}(a)}{r + \lambda} \tag{2.37}$$

Steady state conditions (22)-(25) imply $\xi_n = \frac{u^l}{u^l+u^h_n} = \frac{\Phi(\sigma_2)-\Phi(\sigma_3)}{\Phi(\sigma_1)-\Phi(\sigma_3)} = \xi_n(\sigma_1,\sigma_2,\sigma_3)$ and $\xi_s = \frac{u^h_s}{u^h_s+e^h_{nn}} \approx \frac{(m_n(\theta_n)+m_s(\theta_s))(1-\Phi(\sigma_1))}{\eta m_s(\theta_s)(1-\Phi(\sigma_1))+m_n(\theta_n)(1-\Phi(\sigma_2))} = \xi_s(\theta_n,\theta_s,\sigma_1,\sigma_2)$. Based on vacancy values (2.12)-(2.13) and rearranged employment values (2.34)-(2.37), I get two job creation conditions which show the relationship between two labor market tightness and two cut-off abilities as follows:

$$p_n(\theta_n)\xi_n(\sigma_1,\sigma_2,\sigma_3)(1-\beta)S^l_{nn}(E[a^l_n](\sigma_2,\sigma_3))+ \\ p_n(\theta_n)(1-\xi_n(\sigma_1,\sigma_2,\sigma_3))(1-\beta)S^h_{nn}(E[a^h_n](\sigma_1,\sigma_2)]) - O_n = 0 \quad (2.38)$$

$$p_s(\theta_s)\xi_s(\theta_n,\theta_s,\sigma_1,\sigma_2)(1-\beta)S^h_{ss}(E[a^h_s](\sigma_1))+ \\ \eta p_s(\theta_s)(1-\xi_s(\theta_n,\theta_s,\sigma_1,\sigma_2))(1-\beta)S^h_{sn}(E[a^h_n](\sigma_1,\sigma_2)) - O_s = 0 \quad (2.39)$$

Based on wage equations (2.30)-(2.33) and rearranged employment value equations (2.34)-(2.37), I get matching surplus equations depending on (a,θ_n,θ_s):

$$S^l_{nn}(a,\theta_n) = \frac{y_n(a)-b}{X_1(\theta_n)} \quad (2.40)$$

$$S^h_{nn}(a,\theta_n,\theta_s) = \frac{(r+\lambda)(y_n(a)-b)+\beta\eta m_s(\theta_s)(y_s(a)-b)}{(r+\lambda)X_3(\theta_n,\theta_s)+X_5(\theta_n,\theta_s)} \quad (2.41)$$

$$S^h_{sn}(a,\theta_n,\theta_s) = \frac{X_4(\theta_s)(y_s(a)-b)-\beta m_n(\theta_n)(y_s(a)-y_n(a))}{(r+\lambda)X_3(\theta_n,\theta_s)+X_5(\theta_n,\theta_s)} \quad (2.42)$$

$$S^h_{ss}(a,\theta_s) = \frac{y_s(a)-b}{X_2(\theta_s)} \quad (2.43)$$

where $X_1(\theta_n) = r+\lambda+\beta m_n(\theta_n), X_2(\theta_s) = r+\lambda+\beta m_s(\theta_s), X_3(\theta_n,\theta_s) = r+\lambda+\eta m_s(\theta_s)+\beta m_n(\theta_n), X_4(\theta_s) = r+\lambda+\eta m_s(\theta_s)$ and $X_5(\theta_n,\theta_s) = \beta\eta m_s(\theta_s)\beta m_n(\theta_n)$.

Now I can substitute matching surplus equations (2.40)-(2.43) into job creation conditions (2.38)-(2.39). I define them as (2.38)' and (2.39)', respectively.

Self-selection for searching sector depending on individual ability

Based on Nash bargaining (2.29), matching surplus (2.40)-(2.42), the condition for search sector decision condition (2.20) can be rewritten as follows:

$$
\begin{aligned}
&m_n(\theta_n)\beta S_{nn}^h(a) - m_s(\theta_s)\beta S_{ss}^h(a) \geq 0 \\
\Leftrightarrow & m_n(\theta_n)X_2(\theta_s)(r+\lambda)(y_n(a)-b) \\
&+ (m_n(\theta_n)X_2(\theta_s)\beta\eta m_s(\theta_s) - (r+\lambda)m_s(\theta_s)X_3(\theta_n,\theta_s) - m_s(\theta_s)X_5(\theta_n,\theta_s))(y_s(a)-b) \geq 0 \\
\Leftrightarrow & m_n(\theta_n)(y_n(a)-b) - m_s(\theta_s)(y_s(a)-b) \geq 0
\end{aligned} \tag{2.44}
$$

Here, I use a property of poison probability $m_j(\theta_j)(j = n, s)$ to rearrange equation as approximated condition (2.44). If the condition of (2.44) is satisfied, some educated job seekers decide to apply for non-skilled jobs. This equation (2.44) is approximately binding for workers with the cut-off ability (σ_1) as follows:

$$m_n(\theta_n)(y_n(\sigma_1)-b) - m_s(\theta_s)(y_s(\sigma_1)-b) \approx 0 \tag{2.45}$$

(2.45) shows the relationship between endogenous cutoff σ_1 and 2 endogenous labor market tightness (θ_n,θ_s).

Self-selection for schooling

Similar to equation (2.44), condition (2.21) can be rewritten in a simpler form as follows:

$$
\begin{aligned}
&U_n^h(a) - U_n^l(a) \geq C_0 - C_1 \\
\Leftrightarrow & \beta y_s(a) - y_n(a) + (1-\beta)b \geq \frac{(C_0 - C_1)(r+\lambda)(X_3(\theta_n,\theta_s) + \beta m_n(\theta_n))}{\eta\beta m_s(\theta_s)}
\end{aligned} \tag{2.46}
$$

This equation (2.46) is approximately binding at cutoff ability $a = \sigma_2$:

$$\beta y_s(\sigma_2) - y_n(\sigma_2) + (1-\beta)b \approx \frac{(C_0 - C_1)(r+\lambda)(X_3(\theta_n, \theta_s) + \beta m_n(\theta_n))}{\eta\beta m_s(\theta_s)} \tag{2.47}$$

(2.46) shows the relationship between endogenous cutoff (σ_2) and 2 endogenous labor market tightness (θ_n,θ_s).

Self-selection for staying unemployment

Not entering the labor market might be beneficial for some low-educated workers. Whether low-educated individuals enter the labor market or not is determined by the following condition:

$$\begin{aligned} &U_n^l(a) - U_b^l(a) \geq 0 \\ \Leftrightarrow &y_n(a) \geq b \end{aligned} \tag{2.48}$$

This equation (2.48) is binding at cutoff ability $a = \sigma_3$:

$$y_n(\sigma_3) = b \tag{2.49}$$

Definitions of equilibrium

The main interest of this chapter is to obtain conditions for how cross-skilled matching occurs in equilibrium under the single application assumption (equal marginal search costs). Although there are corner solutions where no one takes education, and all workers take education, this chapter only focuses on interior equilibrium. If there exists cross-skilled matching, all matching surplus must be $S_{jk}^i \geq 0$. I define cross-skilled matching equilibrium situations as follows:

Definition 1 (Cross-skilled matching equilibrium existing 4 types workers) *The stationary equilibrium in which workers choose the education and search sector when unemployed is defined as a "cross-skilled matching equilibrium." In equilibrium, the vectors* $(\theta_n, \theta_s, \sigma_1, \sigma_2, \sigma_3) = (\theta_n^*, \theta_s^*, \sigma_1^*, \sigma_2^*, \sigma_3*)$

in equilibrium satisfies job creation conditions (2.38)' and (2.39)' and approximately bounded self-selection conditions (2.45) and (2.47) and self-selection condition for staying unemployment (2.49) are+ satisfied.

2.3.2 Features of equilibrium

In case $b = 0$, augments will be simple. Let consider $b = 0$ in this section.

Proposition 1 *For cross-skilled matching, productivity gap is necessary for increasing function of workers' ability*

Proof. *If $b = 0$ then condition (2.44) become*

$$m_n(\theta_n) - m_s(\theta_s)\tau(a) \geq 0 \tag{2.50}$$

In other words, the selection gain for all highly educated workers is the same whether they choose the non-skilled or skilled sector, and one is strictly more beneficial than the other. Suppose $\tau(a)$ is continuously increasing function for workers' ability, then $\exists a \in (0,1)$ which satisfies

$$\begin{aligned} m_n(\theta_n) - m_s(\theta_s)\tau(a) \geq 0 \quad for \quad \frac{m_n(\theta_n)}{m_s(\theta_s)} > \tau(a) > 0 \\ m_n(\theta_n) - m_s(\theta_s)\tau(a) \leq 0 \quad for \quad 1 > \tau(a) \geq \frac{m_n(\theta_n)}{m_s(\theta_s)} \end{aligned}$$

As the left-hand side (LHS) of equation (2.50) is a compact and monotone decreasing function (since $\tau(a)$ is monotone increasing for a) for ability a, this kind of ability a can exist cross point equal to zero. ■

As proposition 1, the productivity gap needs to satisfy increasing ability return. If productivity gap is linear, then there is no separation between committing initially search non-skilled sectors and skilled sectors since results of sector decision condition is identical for all ability workers.

For tractability, I define $y_n(a) = a$ and $\tau(a) = 1 + a$ here. Let $F(a) = m_n(\theta_n)y_n(a) - m_s(\theta_s)y_s(a)$, and $G(a) = \beta y_s(a) - y_n(a)$. These F and G come from equations (2.44) and (2.46).

$F(a)$ indicates the basis of a return to a cross-skilled matching route for highly educated workers. $G(a)$ indicates the basis of a return to education with cross-skilled matching. Here, $F'(a)$ and $G'(a)$ indicate their marginal return with respect to ability. Let's try to figure out equilibrium conditions graphically. The situation must be the case in Figure 2.2.

Figure 2.2 shows the schooling and sector decision conditions. The above graph shows marginal return to returns to choosing cross-skilled matching (MRC) is positive up to $a = \frac{m_n(\theta_n)-m_s(\theta_s)}{(\theta_s)}$. Educated workers with larger ability than σ_1 commit to skilled sectors jobs, and educated workers with smaller ability prefer cross-skilled matching. Also, the below figure shows the relationship between marginal return to schooling and ability. Crossing point between marginal return to schooling (MRS) = $\frac{(C_0-C_1)(r+\lambda)(X_3(\theta_n,\theta_s)+\beta m_n(\theta_n))}{\eta\beta m_s(\theta_s)}$ and $G'(a)$ become cutoff σ_2. Suppose $C_0-C_1 \to 0$ then cutoff ability become $\sigma_2 = (1-\beta)/2\beta$. Equilibrium inferiority needs to satisfy the following conditions.

Proposition 2 *Sufficient conditions of Equilibrium inferiority is following:*

$$2m_n(\theta_n^*) < m_s(\theta_s^*) \tag{2.51}$$

Proof. *Equilibrium inferiority of cross-skilled matching existing 4 types workers is satisfied if* $\sigma_1 < 1$. *Maximum point of* σ_1 *is* $\frac{m_n(\theta_n)-m_s(\theta_s)}{(\theta_s)}$. *Then, I get then I get left relation of condition (2.51).* ■

Proposition 3 *Necessary condition for cross-skilled matching existing 4 types workers is the following :*

$$m_s(\theta_s^*) < \beta m_n(\theta_n) \tag{2.52}$$

$$\beta > \frac{1}{2} \tag{2.53}$$

Proof. *Suppose there exists a cross-skilled matching, then* $\sigma_1^* > \sigma_2^*$ *must be satisfied. If schooling cost* $C_0 - C_1$ *goes to 0, then* σ_2 *takes minimum. By comparing* $\sigma_1 = \frac{m_n(\theta_n)-m_s(\theta_s)}{(\theta_s)}$ *and* σ_2*, then I*

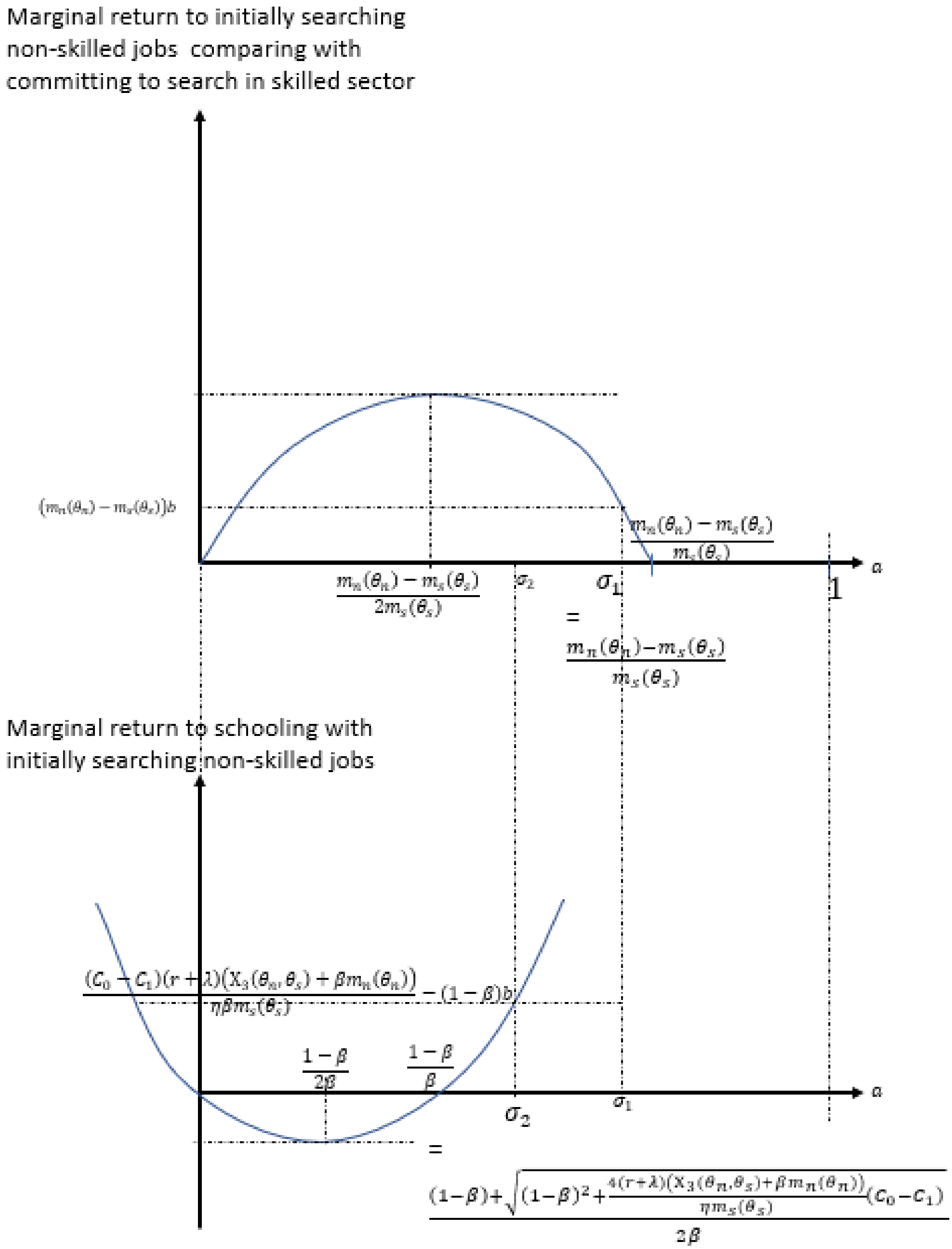

Figure 2.2: Marginal return to cross-skilled matching with respect to workers' ability

get condition (2.52). Combining condition (2.51), $beta$ should be greater than 0.5. ∎

These propositions imply a unique interior cross-skilled matching equilibrium exists for sufficiently small schooling cost $C_0 - C_1$.

Based on MRS in Figure 2.2, I get same relationship as Charlot and Decreuse (2005) that both sectors' market tightness θ_j (j=n,s) are increasing functions with respect to schooling cutoff σ_2. The composition effect (increasing expected ability with increasing schooling cost) certainly works for the cut-off. Therefore, increasing expected ability activates the matching market, and all types of matching can benefit from increasing expected ability. Sector decision cut-off has a different mechanism. Based on MRC shown at the top of Figure 2.2, an increase in σ_1 (moving to the right) leads to a decrease in the $m_s(\theta_s^*)$ and tightness θ_s.

These two cut-off mechanisms work in opposite directions. Increasing schooling cut-off leads to increasing both sector market tightness because of the composition effect. However, increasing tightness of skilled sector leads to a decrease in the sector decision cut-off. The relationship in this equilibrium is that congestion works as the main mechanism, and the more workers are educated, the lower the tightness rate in either sector, leading to an influx of highly educated workers into the less crowded, higher matching non-skilled sector and more workers with cross-skilled matching.

2.3.3 Comparative statics

Unemployment benefit

Figure 2.2 omits unemployment benefits. If I consider unemployment benefit, above figure need to appear horizontal line as $MRC = (m_n(\theta_n) - m_s(\theta_s))b$ and below figure need to subtract $(1-\beta)b$ from horizontal dot line. Then, increasing unemployment benefits b indicates that sector decision cut-off σ_1 move to 0 direction and schooling decision cut-off σ_2 also move to 0 direction. Increasing unemployment benefits leads to more education and highly educated workers who commit to skilled sectors. Both cut-offs move to 0 direction, implying that the expected ability of each type of worker decreases.

OJS efficiency

OJS efficiency does not affect sector decision cut-off σ_1 but can affect schooling cut-off σ_2 based on approximated conditions (2.45) and (2.47). X_3 contains the OJS efficiency η, so the horizontal dot line, which decides the schooling cut-off σ_2, can be moved by increasing the OJS efficiency. Suppose schooling cost $C_0 - C_1$ is sufficiently small, then increasing the OJS efficiency leads to more education. Suppose schooling cost is sufficiently high, then increasing OJS efficiency leads to less education. For educated individuals, choosing to work in the non-skilled sector will eventually lead to intersectoral mobility by OJS, which would intuitively impact the cut-off for the sector decision. The current approximation, as in equation (2.44), implies that the OJS efficiency has just a negligible impact on the search sector decision.

Job destruction rate, interest rate and schooling cost

Job destruction rate and interest rate only affect schooling cut-off. Increasing each of them leads to increasing the schooling cut-off. As a necessary condition (2.51), the matching rate is larger in the non-skilled sector than in the skilled sector in the cross-skilled matching equilibrium.

Schooling cost also only affects schooling cut-off σ_2. Increasing schooling cost leads to a higher education cut-off, implying less educated workers. In cross-skilled matching, the composition effect only affects low-educated and highly educated workers who do not commit to the skilled sector. Committed educated workers, whose ability is high enough compared with the difference in matching rate, are not affected by schooling costs. For these high-ability workers, always taking education is beneficial. Sufficiently large schooling cost leads to $\sigma_1 < \sigma_2 < 1$. In this situation, equilibrium becomes ex-post segmentation with education choice.

2.4 Socially optimal and education policy

This section basically follows Charlot and Decreuse (2005)'s ideas of social welfare and education policy. In DMP type search model, social planner maximizes stationary consumption (Hosios

(1990)). From the social planner's view, education does not directly contribute to increasing social welfare since education does not produce any consumer goods. However, education plays a role in adjusting workers' allocation and giving workers access to the specific production technology of the skilled sector. For simplicity, I assume social planners can only change workers' schooling costs ($C = C_0 - C_1$) by changing schooling subsidy C_1. Let FO indicate flow aggregate output, FV indicate flow aggregate vacancy cost and $FEdu$ indicate flow aggregate education cost. Based on Charlot and Decreuse (2005)'s setting, a social planner maximizes below stationary social welfare P in this economy:

$$arg \max_{C_1} : P = FO - FV - FEdu \tag{2.54}$$

$$s.t.(2.38)', (2.39)', (2.45), (2.47) and (2.49)$$

where

$$\begin{aligned}
FO &= e_n^l y_n(E[a_n^l]) + e_{nn}^h y_n(E[a_n^h]) + e_{sn}^h y_s(E[a_n^h]) + e_{ss}^h y_s(E[a_s^h]) \\
&= \frac{m_n(\theta_n^*)(\phi(\sigma_2^*) - \phi(\sigma_3^*))}{\lambda + m_n(\theta_n^*)} E[a_n^l] + \frac{m_n(\theta_n^*)(\phi(\sigma_1^*) - \phi(\sigma_2^*))}{\lambda + m_n(\theta_n^*)} (E[a_n^h] + E[a_n^h]^2) \\
&- \frac{m_n(\theta_n^*)\lambda(\phi(\sigma_1^*) - \phi(\sigma_2^*))}{(\lambda + m_n(\theta_n^*))(\lambda + \eta m_s(\theta_s))} E[a_n^h]^2 + \frac{m_s(\theta_s^*)(1 - \phi(\sigma_1^*))}{\lambda + m_s(\theta_s^*)} (E[a_s^h] + E[a_s^h]^2) \\
FV &= O_n v_n + O_s v_s \\
&= O_n \frac{\lambda(\phi(\sigma_1^*) - \phi(\sigma_3^*))}{\lambda + m_n(\theta_n^*)} \theta_n^* + O_s (\frac{m_n(\theta_n^*)\lambda(\phi(\sigma_1^*) - \phi(\sigma_2^*))}{(\lambda + m_n(\theta_n^*))(\lambda + \eta m_s(\theta_s))} + \frac{\lambda(1 - \phi(\sigma_1^*))}{\lambda + m_s(\theta_s^*)}) \theta_s^* \\
FEdu &= \lambda(1 - \Phi(\sigma_2))(C^0)
\end{aligned}$$

In the equilibrium existing 3 types workers, schooling subsidy can take between $C_{1min} \geq C_1 \geq C_0$, where $C_1 = C_{1min} \Leftrightarrow \sigma_1^* = \sigma_2^*$.

Based on Figure 2.2, schooling cost directly affects schooling cut-off σ_2. Increasing C_1 is identical to decreasing workers' schooling cost $C_0 - C_1$, decreasing schooling cut-off σ_2 and expanding the number of educated workers. Then, I can get the marginal return of social welfare with respect to schooling subsidy C_1 as follows:

$$
\begin{aligned}
\frac{dP}{dC_1} &= \frac{dP}{d\sigma_2^*}\frac{d\sigma_2^*}{dC_1} = (\frac{dFO}{d\sigma_2^*} - \frac{dFV}{d\sigma_2^*} - \frac{dFEdu}{d\sigma_2^*})\frac{d\sigma_2^*}{dC_1} \\
&= (\frac{m_n(\theta_n^*)}{\lambda + m_n(\theta_n^*)}(\frac{d\Phi(\sigma_2^*)}{d\sigma_2}E[a_n^l] + \Phi(\sigma_2^*)\frac{E[a_n^l]}{d\sigma_2} - \frac{d\Phi(\sigma_2^*)}{d\sigma_2}(E[a_n^h] + E[a_n^h]^2) - \Phi(\sigma_2^*)\frac{d(E[a_n^h] + E[a_n^h]^2)}{d\sigma_2} \\
&+ \frac{\lambda}{\lambda + \eta m_s(\theta_s^*)}(\frac{d\Phi(\sigma_2^*)}{d\sigma_2}E[a_n^h]^2 + \Phi(\sigma_2^*)\frac{dE[a_n^h]^2}{d\sigma_2}) + \frac{\lambda\theta_s^* O_s}{\lambda + m_s(\theta_s^*)}\frac{d\Phi(\sigma_2^*)}{d\sigma_2}) + \lambda C_0\frac{d\Phi(\sigma_2^*)}{d\sigma_2})\frac{d\sigma_2^*}{dC_1}
\end{aligned}
\tag{2.55}
$$

Here, $\frac{d\sigma_2^*}{dC_1} < 0$, $\frac{dC_0}{dC_1} > 0$. In Charlot and Decreuse (2005)'s model, the marginal return of social welfare is always monotone decreasing, and they concluded that the smallest schooling subsidy (C_{1min} in this model) is always desirable. However, (2.55) indicates that the marginal return of social welfare can be monotone, increasing if exogenous vacancy costs and schooling costs are sufficiently small. This result seems consistent with intuition.

2.5 Conclusion

The previous study assumed that in a 2×2 high and low type labor market, the highly educated unemployed could apply to both sectors at the same time. In contrast, the less educated unemployed can only apply to the low sector and consider whether a unique equilibrium exists that acts as better insurance than unemployment for the highly educated. This double application implicitly assumed that the highly educated (or skilled) unemployed would face half the marginal search cost as the low educated (unskilled) unemployed and that the highly educated workers matched with the non-skilled sector would face the same (strictly, lower by the OJS efficiency) marginal search cost as the low educated unemployed. The marginal search cost is the time available for the search. It is natural to assume that the marginal search cost is the same independent of education level since it is the length of time available for search. In this chapter, I examine under what conditions cross-skilled matching (over-education in empirical studies) occurs as an equilibrium solution when marginal search costs are identical.

I show the following. First, the productivity gap between sectors has an increasing shape with respect to worker skills, with the non-skilled sector matching rate greater than the skilled sector matching rate, but only in limited circumstances where it is not too large; some highly educated workers are cross-skilled matched and some are committed to the skilled sector and less educated workers also exist. Second, congestion works as the main equilibrium mechanism. Charlot and Decreuse (2005) concluded that education subsidies are desirable with minimal, negative subsidies for education discouragement. However, my results show that the argument differs depending on the skilled sector vacancy costs and exogenous education costs. It indicates that enrollment is encouraged when costs are sufficiently small and is discouraged when costs are sufficiently high.

This analytical model connects and extends each study by making marginal search costs identical, considering the endogenous decision structure of education, and considering OJS. However, it is still simplified and leaves out many aspects of the real world. In particular, the rate of higher education may vary with students' household budget and borrowing constraints, the growth of worker productivity as in the human capital theory, and the existence of students who go on to higher education but do not necessarily graduate drop out of school.

In addition, this chapter examined the use of a classic random matching process to clarify the difference from previous research. However, as in previous research, I rely on strong assumption that less-educated individuals do not have access to skilled jobs. Therefore, the non-skilled jobs match with less- and highly-educated workers, and the skilled job only matched with highly educated. However, in reality, some less-educated are also matched with skilled jobs. In describing this situation, it is conceivable that wage posting setting could be a better representation of reality than random matching and Nash bargaining, considering that low-educated individuals are matched with skilled jobs by luck. Incorporating such features into the model is a topic for future research.

Chapter 3

Does higher education moderate the negative effects of adverse labor market conditions at job entry?

Abstract

This chapter extends the existing literature on the hysteresis effect of labor market participation during recessions to test whether higher education can mitigate this negative effect. I include education and its interaction term in the standard estimation model of the hysteresis effect and conduct an IV estimation to handle the newly introduced ability bias. The results show that higher education mitigates part of the curse effect, but four years of education do not fully cover the curse effect of a 1% increase in the unemployment rate at entry. I also analyze whether negative shocks at entry lead to job polarization and find that the long-term rate of job polarization is not affected by a temporary recession, but the number of workers with intermediate skills declines in the short run. This suggests a relationship between the effects of the curse and the recession-induced decline in the number of middle-skilled workers.

3.1 Introduction

For most students, the situation when they enter the labor market is similar to a lottery decided only by luck of birthplace and year of birth. Evidence is accumulating that entering into adverse economic conditions has short- and long-term effects on various outcomes, such as individual earnings, employment probabilities, family formation, health status, and others (Oyer, 2006; Kahn, 2010; Oreopoulos et al., 2012; Altonji et al., 2016; Kawaguchi and Kondo, 2015). Using a sample of college graduates is the primary sample for this literature since they are less likely to return to the education sector after once employed. Recent studies have been actively analyzing the differences in the curse effect from labor market entry during a recession for various attributes, and the greatest difference has been found in differences between education levels. For example, Genda et al. (2010) analyzed education level differences in Japan and the U.S., Brunner and Kuhn (2014) in the Netherlands, Cutler et al. (2015) in multiple European countries, and Schwandt and von Wachter (2019) in the U.S. It is a reasonable explanation that less-educated workers are likely to be affected by the unemployment rate and other labor market outcomes. At the same time, it is also intuitive that less-educated workers already face sufficiently low labor market outcomes to be largely unaffected. However, it is not obvious which education level is affected more by the curse effect of initial entry labor market conditions. Previous studies, such as Genda et al. (2010), Cutler et al. (2015), and Schwandt and von Wachter (2019), conclude that the lower education group tends to be subject to larger curse effects.

The result that the negative curse effect is larger for less educated workers is not problematic in explaining the difference between groups. However, it is problematic to determine whether education has the causal power to weaken the curse effect of entry recession based on the results of previous studies. This is because the differences in individual abilities across educational backgrounds may weaken the effect of labor market entry during recessions. This type of typical ability bias problem always appears in the Mincerian specification in the return to schooling literature(Pischke and von Wachter, 2005; Card, 2001; Ashenfelter and Zimmerman, 1997; Harmon and Walker, 1995; Ashenfelter and Krueger, 1994; Card and Krueger, 1992; Angrist and Krueger, 1991; Griliches, 1977). Previous research on the curse effect field does not seriously consider this ability bias, but it is crucial to handle the endogeneity of the change in graduation timing given a certain education levelKahn (2010); Cutler et al. (2015). I thus test whether higher education itself has the power to weaken the curse effect of entering the labor market during a recession, regardless of ability. If such an effect exists, avoiding bad economic conditions through higher education might be the optimal action from an individual's perspective.

Job polarization is a growing literature stream in the field of macro and labor economics (Autor et al., 2003, 2006, 2008; Goos et al., 2009; Ikenaga, 2009; Ikenaga and Kambayashi, 2016; Cortes, 2016; Michaels et al., 2014; David, 2017; Furukawa and Toyoda, 2018; Jaimovich and Siu, 2020). In particular, Jaimovich and Siu (2020) demonstrate the relationship between recession and job loss for middle skills, and that job polarization might occur at the time of recovery from a recession. Similar to the effects of the curse of labor market entry, job polarization and job loss share the same trigger: bad economic conditions. Because of this same trigger structure, these two bodies of literature could be linked. In particular, the negative effects of a recession on labor market outcomes at the time of labor market entry might force workers to search for jobs with lower task scores than during regular periods, and the persistence of this effect might act as a mechanism to promote job polarization. Therefore, I tested how entering the labor market during a recession would change the quantity and share of each type of job task assigned to individuals in the short and long run. If polarization is accelerated, I expect the quantity and share of middle skills to decrease over the long run. Additionally, I can show how education influences such polarization.

I used the Japanese Labor Force Survey (LFS) from 2002-2017 as the data, and the results showed that higher education mitigates part of the curse effect, but even four years of education could not fully cover the curse effect from a 1% increase in the unemployment rate at entry. I also analyzed whether negative shocks at entry lead to employment polarization and found that the long-term tendency of polarization is not affected by a temporary recession, but the number of workers with middle skills declines in the short run, and this decline cannot recover in terms of job task amounts. My main finding is that education protecting the curse effect of bad labor market entry is not unique and is almost the same as in previous studies (Genda et al., 2010; Cutler et al., 2015; Schwandt and von Wachter, 2019); yet, my results are more likely to avoid the interpretation of ability differences.

The structure of this chapter is as follows. Section 2 describes the estimation methods, with some explanations from previous studies. Section 3 describes the data. Section 4 reports the estimation results. Section 5 presents the conclusions.

3.2 Empirical Strategy

3.2.1 Typical Regressions in Previous Research

The literature considering the curse effect of labor entry during recessions mainly focuses on a sample of college graduates less likely to return to education after they get a job(Kahn, 2010; Oreopoulos et al., 2012). Regarding studies that confirm the differences among educational groups, Genda et al. (2010) discussed the differences between college and high school graduates, van den Berge (2018) analyzed the differences between vocational and general education within vocational schools, and Schwandt and von Wachter (2019) considered the differences between middle school graduates, high school graduates, college graduates, and community colleges. The following Mincerian specification is typically used to confirm the curse effect:

$$
\begin{aligned}
Y_{itcr(edu)} =& \alpha_{0(edu)} + \alpha_{1(edu)} U1st_{cr(edu)} Pexp_{it(edu)} + X_{it(edu)}\theta_{(edu)} + \delta_{r(edu)} \\
&+ \delta_{t(edu)} + \delta_{c(edu)} + \epsilon_{ictr(edu)} \qquad (3.1)
\end{aligned}
$$

Essentially, estimation equation (1) is used in almost all studies in this field, although there are some differences between them, such as whether the contemporaneous unemployment term (UC) is included (Genda et al., 2010), whether some fixed effect terms are included, whether region-specific time trends are included, whether individual-level estimation (Kahn, 2010; Genda et al., 2010; van den Berge, 2018; Altonji et al., 2016; Beiler, 2017) or aggregated cell level estimation are considered(Oreopoulos et al., 2012; Schwandt and von Wachter, 2019), and whether accurate data are used. Here, Y is the labor market outcome, $U1st$ is the unemployment rate in the initial job search year, $Pexp$ refers to the years of potential experience X is a vector of controls, δ is the fixed effects term, and ϵ is the error term. Subscript i indicates individuals, t survey years, g graduation years, and r graduation regions. The curse effect of recession is illustrated by $\alpha_{1(edu)}$. Most previous studies on the differences in the initial labor market condition effect across education groups confirm the existence of differences by implicitly comparing the coefficient values, such as $\alpha_{1(edu=16)} - \alpha_{1(edu=12)}$, which are obtained by estimating (1) with the sample education level as a given (Genda et al., 2010; van den Berge, 2018; Schwandt and von Wachter, 2019). These studies conclude that the curse effect of a recession is robustly weaker for a group with a higher education level. However, while the results certainly demonstrate differences between education groups, they do not imply that receiving higher education has the causal power to ease the curse of adverse labor market entry because there is an interpretation possibility that higher-ability workers are less affected by adverse labor market entry.

3.2.2 Estimation

As mentioned in the previous subsection, the estimation of education using equation (1) is unsuitable for examining whether education mitigates the curse effect of labor market entry during a recession. As another way of examining the protective effect of education, Cutler et al. (2015) use German data to confirm whether education has a protective effect by inserting education into equation (1). I follow their estimation method to analyze whether education mitigates the negative effect of labor market entry during a recession, as in estimation (2):

$$Y_{itcrhm} = \beta_0 + \beta_1 U1st_{cr} Pexp_{it} + \beta_2 U1st_{cr} Pexp_{it} Edu_i + X_{it}\theta + \delta_r + \delta_h + \delta_m + \epsilon_{ithr} \tag{3.2}$$

Here, Edu denotes the years of education for an individual, subscript h indicates the high school graduation year (birth cohort), m indicates the survey month, and δ_m is its fixed effect. Given the lack of accurate information, I mechanically construct the potential experience and entry year, which causes the perfect linearity problem between survey year, graduation year, potential experience, high school graduation year (birth cohort), and years of education. Therefore, I choose the high school graduation year, potential experience, and education to be included in the estimation model (2) and omit the fixed effect terms of the survey year and graduation year. For simplicity, I estimate the linear probability model, even when the indicator variable is taken as an outcome.

β_1 is the same as the ordinal effect of the persistence effect from the economic condition, and β_2 represents my main focus as the protective effect. In fact, Cutler et al. (2015) used a similar estimation and concluded that there is a protective effect because coefficient $beta_2$ is positive. The curse effect during a recession is estimated by the Mincerian specification, as in (1). When I insert education into (1) as in estimation (2), I have to face the omitted variable bias of the ability between the labor market outcomes and education, as in the literature on returns to schooling. If I do not consider this, I can interpret the highly educated group as having a lower negative effect from labor market entry during a recession but cannot conclude that higher education causally weakens the negative effects. Although Cutler et al. (2015) prove the relationship between graduation timing and the economic condition and the endogeneity of $U1st$, they do not address the ability bias between education and labor market outcomes.

Therefore, I employ the IV approach. By focusing on individuals who can change their education decisions due to changes in the accessibility to higher education institutions, I can demonstrate whether education has a protective effect regardless of individuals' abilities. To this end, I use the regional unemployment rate at the age of 18, the regional average tuition fee, and the regional capacity ratio of higher education institutions at the age of 18 as instruments for individuals' years of schooling, which are determined by their birth year and region and the decision of the higher education authorities. These instruments are standard in the literature.[1] I identify the LATE (local average treatment effect), which is different from the ATE for the entire population, in this case, all

[1]For example, Cameron and Heckman (1998), Cameron and Taber (2004), Arkes (2010), Carneiro et al. (2013), and Kikuchi (2017) use local labor market conditions, and Kane and Rouse (1993), Kikuchi (2017) tuition, Currie and Moretti (2003) Kikuchi (2017) use college capacity or new higher education institution openings.

Japanese workers.

In addition to education terms, there is concern about potential experience endogeneity, as first pointed out by Griliches (1977) since potential experience is the combination of education and age. Some Mincerian specification papers in the literature on returns to schooling use age and age squared instead of experience or potential experience and its square term (Ashenfelter and Krueger, 1994; Ashenfelter and Zimmerman, 1997). Blackburn and Neumark (1995) argued this bias problem and suggested also using instruments for potential experience to avoid misspecification, while Kling (2001) considers age as an instrument of potential experience to estimate the returns to schooling. In particular, Kahn (2010) uses age and its interaction and age squared as instruments for the potential experience and interaction with the entry year's unemployment rate. I follow this specification and use age dummies as instruments for potential experience in equation (2).

My main interests are the interaction terms in equation (2). Wooldridge (2003) suggests an IV estimation procedure with interaction terms. I follow this procedure and use the interaction age dummy and entry unemployment rate as instruments for the interaction potential experience dummy and entry unemployment rate, and the interaction age dummy and entry unemployment rate and education instrumentation terms as instruments for the interaction potential experience dummy, entry unemployment rate, and education, respectively.

In addition to the main analysis regarding education's safety function, to confirm the potential association between job polarization and economic conditions at labor market entry and the education effect on it, I run the following regression:

$$Task_{itrhm} = \gamma_0 + \gamma_1 U1st_{cr} Pexp_{it} + \gamma_2 U1st_{cr} Pexp_{it} Edu_i + X_{it}\theta + \delta_r + \delta_h + \delta_m + \epsilon_{ithr} \quad (3.3)$$

$Task_{itrhm}$ indicates the total task score and the share of task scores of individuals' occupations. Equation (3) simply substitutes task score $Task_{itrhm}$ into labor market outcomes Y_{itcrhm} in equation (2). Then, γ_1 indicates that when the unemployment rate at the time of labor market entry increases by 1%, the kind of job that an individual worker obtains subsequently and how the change in the type of job compares to a worker who can enter the labor market at a normal time in the long run. Additionally, γ_2 shows that if the unemployment rate at the time of labor market entry increases by 1%, the change in the type of occupation of workers who find jobs subsequently depends on their level of education and the magnitude of the change in the type of occupation differs in the long run from that of high school graduates.

3.3 Data

As previously mentioned, I use the 2002-2017 Japan's LFS, focusing on workers who were in their third year of high school between 1991 and 2011, as the main data for analysis.[2] To consider the education decision behavior at the age of 18, I remove from the sample workers with only junior high school education, no education information, or postgraduate education. Checking gender differences for the curse effect is interesting, but my main motivation is to shed light on educational differences. For simplicity, I focus on a sample of males. Additionally, I exclude self-employed individuals from the main analysis to target main employees. I also exclude the non-labor force.

In the literature on the curse effect of recessions, data can be classified into two categories: studies that use panel data that provide more accurate information for a small sample and studies that use repeated cross-sectional data that provide coarser information for a large sample. This chapter falls in the second category.[3] Similar to the LFS in most other countries, the Japanese have a severe limitation in that we do not know the exact values of the exact timing and region of labor market entry and the related economic conditions, which are the most important variables in this literature. As regional information, I just have access to a current residence. Therefore, I conduct two types of analyses: one assumes that people continue to live in the same place as in the year

[2]The Japanese LFS has two components, a Basic survey ("KisoChosahyo") and a Detailed Supplement survey ("TokuteiChosahyo"). The basic part has panel structures of each continuous two months in two survey years in the same month, while the detailed supplement survey is simultaneously conducted with the final fourth basic survey. I utilize the detailed supplement as the main data source and merge it with the basic survey to complement some information.

[3]In this field, even if we use panel data, the data will be used as a pooled cross-section because the main economic condition information is unique for individuals and, thus, the advantage of using panel data is lower than in other fields since we cannot take fixed effects of individuals.

of the survey and do not move as in Genda et al. (2010)'s analysis, and the other uses the highest migration patterns given a current region. In this chapter, I handle the current residential region as the main data and check result robustness using the highest transition path data. To identify the highest migration patterns, I use the Employment Trends Survey (ETS, Koyodokochousa) and the Basic School Survey (BSS, Gakkokihonchousa), both of which contain an average transition matrix of students. For the unemployment rate of each region, I use published data from the LFS on e-Stat in 0.1 point increments and the 10 regional categories from 47 provinces in the survey. These 10 categories have changed to 11 categories since 2011 when the Kyushu-Okinawa region was divided into the Kyushu and Okinawa regions. For this reason, I calculated the unemployment rate for the Kyushu-Okinawa region for 2011-2017 from the LFS basic survey.

Similar to regional information in the entry year, my data lack precise region information at the age of 18. This chapter thus uses the same assumption as for the entry year, and I conduct two estimations with the same assumption: no migration and the highest migration pattern. For instruments on education at age 18, I use the local unemployment rate, tuition fee, and capacity rate. Tuition data are based on a retail price statistics survey (Kouribukkatoukeichousa). University capacity data are from the List of Universities in Japan (Zenkokudaigakuichiran),[4] junior college capacity data are from the List of Junior Colleges in Japan (Zenkokutankidaigakuichiran), and vocational school capacity data are from the BSS.

The definition of educational differences in the effects of the recession curse is complex when a large proportion of workers return to the education sector after leaving it once. As previously mentioned, this chapter uses Japanese LFS data for analysis. As shown in Figure 1, Japan has the lowest average age of entrance students to bachelor level education institutions, which implies low re-enrollment in higher educational institutions among OECD countries since the average age must reflect the percentage of re-enrollment students. This structure is advantageous for investigating

[4]The List of Universities in Japan is the original source, the data are open, but access is obstructed for others than individual researchers generating datasets from the book. The Center for Research and Education in Program Evaluation (CREPE), University of Tokyo, offers this information as an Excel file on its web page. I am grateful to them, as I am using electronic data.

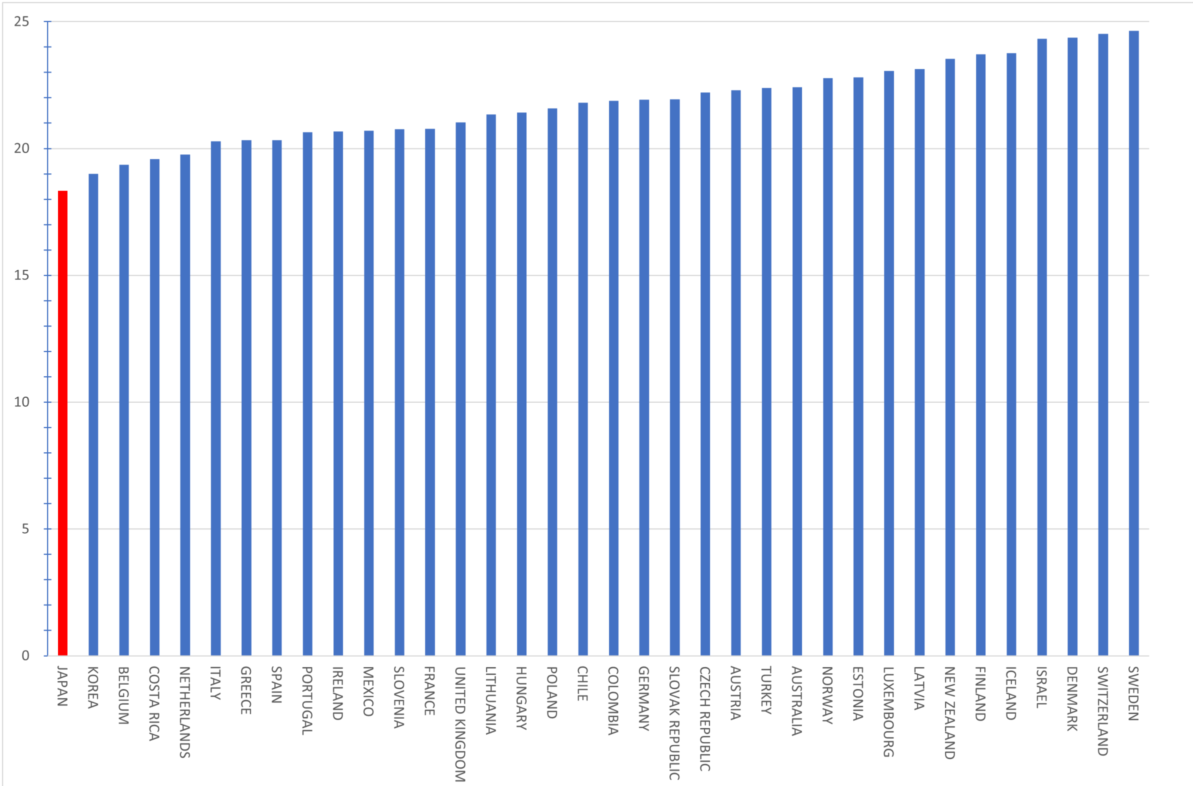

Figure 3.1: Average entry age of bachelor's or equivalent level (ISCED 2011 level 6) institution (Data extracted on 29 Aug 2021 12:48 UTC (GMT) from OECD.Stat)

the curse effect of entering the labor market during a recession.[5]

As labor market outcome variable Y_{ithr} in equation (2), I use an indicator variable for whether the worker is employed within the labor force, an indicator variable for whether the worker is a full-time employee (in Japanese "seishain," meaning higher fringe benefit workers), an indicator variable for whether the worker is a regular one, a continuous variable for working hours, and the log of annual earnings. This information comes directly from the survey questionnaires.

In Japan, the fiscal year and academic year starting in April. Further, it is usual for students who graduate from a certain level of education to go on to the next level of education or enter the labor market without a gap year, as shown in Figure 1.

footnoteIn Japan, entrance exams are required to enter universities, and around 10% of the students who fail to pass the entrance exams enter the "Rounin," a gap year unique to Japan, every year. I

[5]The Japanese government considers this feature as a problem and attempts to increase recurrent education, but the implementation is still ongoing(MHLW, 2018).

do not take this year into account.. Since job-hunting activities are conducted mainly from April to September during the academic year, I consider the year from January to December, including April, as the job-hunting year, defined as the graduation year in previous studies. Additionally, the year that includes April of the third year of high school is defined as the high school graduation year. For the unemployment rate, it is possible to assign different unemployment rates to college and high school graduates, considering that their labor markets are different. However, this chapter does not separate the unemployment rate between college and high school graduates and considers that people in the same year of entry or of age 18 and the same region share the same economic situation.

For task score data, I use the same task data in Japan as Ikenaga and Kambayashi (2016), who created a task score matrix in Japan based on the "Career matrix," a database created by The Japan Institute for Labour Policy and Training. I convert the five classifications of the task as in Autor et al. (2003) into three levels of the task class: high-skilled, middle-skilled, and low-skilled. I define non-routine analytic and non-routine interactive tasks as high-skilled tasks, routine cognitive and routine manual tasks as middle-skilled tasks, and non-routine manual tasks as low-skilled tasks.

Table 3.1 shows the summary statistics for the outcomes, entry year unemployment rate, and instruments based on the whole estimation sample and education-based sub-sample. The sample size difference comes from the features of the variables. Whether a worker is full-time, one is given by employment, and work hours are also given by employment, etc. Table 3.1 indicates that all labor market outcomes are lower in the lower education groups.

Table 3.2 shows the summary statistics for task scores and shares. To handle the case where some occupations contain a large share of the high-skilled task, but the task amount itself is not so high, I use both the task score itself and its share. Table 3.2 suggests that high-skilled and low-skilled task amounts and these shares are higher for the higher education group. The amount and share of middle-skilled tasks were lower in the higher education group.

3.4 Results

3.4.1 First-stage results

As summarized in Table 3.1, some outcomes contain different sample sizes because of their construction property. Although I obtain multiple first-stage results, I only report the first-stage result for the one with the largest number of samples, i.e., an IV estimation for employment. There are no major differences in statistics or coefficients among the multiple first-stage results. Table 3.3 presents the first-stage results when I take employment as an outcome in the second stage. Education instruments associated with the years of education and age are associated with potential experience. As the estimation uses clustered standard errors by current region and high school graduation years, I rely on Kleibergen and Paap (2006)'s test statistics because it is suitable for the IV estimation with clustered standard errors. The p-value of the under-identification test is 0.0229, and the p-value of the over-identification test is 0.6702. I conduct second-stage estimation, although the first stage is relatively weak because the F-statistics calculated by Kleibergen and Paap (2006)'s method is close to 10, which is usually used as a threshold for confirming weak instruments in the case of homoscedastic standard errors.

3.4.2 Main results: Protective effect of education

The coefficients in Tables 3.4 and 3.5 without education represent the effects of a 1% increase in the unemployment rate at labor market entry (implicitly assumed common for all education groups). In contrast, education interaction represents the protective effect (difference of the curse effect from high school graduates) of higher education.

First, I explain the ordinal curse effect using the OLS results. From Table 3.4, in column (1), a 1% increase in the unemployment rate at entry reduces the probability of employment by approximately 3% up to experience 9, and its effects will disappear by experience 13. From Table 3.4, in column (3), taking employment as a given, a 1% higher unemployment rate at entry reduces the probability of being employed as a full-time worker[6] by approximately 9% in the beginning, and this effect weakens over time to approximately 4% in experience 13, but negative effects still exist. From Tables 3.4 in (5), taking employment as given, a 1% higher unemployment rate at entry reduces the probability of being employed as a regular worker[7] by around 8% in the beginning, and this effect disappears by experience 9, the effect becoming positive at 2%. From Table 3.5, in column (1), a 1% worse unemployment rate at entry does not affect the working hour at the beginning of the entry but increases in the future by around one hour for experience 9 and two hours for experience 13. From Table 3.5, in column (3), when I focus on the sample with income, a 1% higher unemployment rate at entry results in approximately 11.5% lower annual income at the beginning, and the negative income effect remains at almost the same level by experience 13. However, the negative effects on employment-related terms in Table 3.4 are weakened or disappear.

Second, I report the main interest in the protection effect part of the OLS results. From Table 3.4, in column (1), one more year of education eased the negative effect to around 0.1% compared to 0.2% from a 1% increase in the unemployment rate at entry, which cannot cover the negative effect of a 1% increase in the unemployment rate even if workers take four more years of education; this protection effect will disappear by experience 13. From Table 3.4, in column (3), taking employment as given, one more year of education eased the negative effect to 0.5% in the beginning, and this effect slightly weakens over time to 0.3% by experience 13, which cannot cover the negative effect of a 1% increase in the unemployment rate even if workers take four more years of education, but a protective effect still exists. From Tables 3.4 in (5), taking employment as given, one more year of education eases the negative effect to 0.5% by experience 5, which cannot cover the negative effect of a 1% increase in unemployment rate even if workers take four more years of education; this protection effect slightly weakens over time until experience 9, and the protective effect share disappears with the disappearance of the ordinal negative curse effect. From Table

[6]The outcome is used here as an indicator of "seishain" in Japanese. They tend to have higher fringe benefits than other employment cases.

[7]This regression uses the indicator of regular work as an outcome. Even if a worker is part-time, they are considered to be regularly working. Japanese cases of full-time work, as the other countries might be suitable for previous "seishain" or not.

3.5, in column (1), one more year of education does not have any protection effect since the curse effect does not exist, and we see a decrease of 0.06 hours (3.6 minutes) at experience 9 and 0.15 hours (9 minutes) at experience 13 against the positive curse effect, which cannot cover the positive effect of a 1% increasing unemployment rate even if workers take four more years of education. From Table 3.5, in column (3), for the sample with income, one more year of education eased the negative effect to around 0.5% in the beginning based on a 1% increase in the unemployment rate at entry, which cannot cover the negative effect of a 1% increase in the unemployment rate even if workers take four more years of education; this protection effect slightly increases from 13 to 0.8.

Then, I report the difference between the IV estimation and OLS for ordinal curse effect results. From Table 3.4, in column (2), in terms of the effect on employment probability, all coefficients are slightly larger than under the OLS, and the major difference is that there is still a negative effect for experience 13. From Table 3.4, in column (4), in terms of the effect of employment probability as a full-time position, all coefficients are slightly larger than for the OLS, and this increase is slightly larger for experiences 5 to 9. From Table 3.4, in column (6), in terms of the effect of employment probability as a regular worker, all coefficients are slightly larger than for the OLS at the beginning, and the flip-positive effect seen in OLS disappears. From Table 3.5, in column (2), in terms of the effect on working hours, the IV estimation also increased at experience 9, and its size increased slightly. From Table 3.5, in column (4), in terms of the effect on income, all coefficient sizes are larger than for the OLS, and the IV estimation might suggest that the negative effect is not stable but increases in the long run.

Finally, I report the differences between the IV estimation and OLS regarding the education protection effect. In terms of employment probability for a full-time position and as a regular worker, Table 3.4 suggests a slightly larger (0.001–0.003) protective effect from the OLS estimation. Statistical significance only occurs at employment probability with an experience of 13. From Table 3.5, regarding the effect of working hours and annual income, all coefficient sizes are larger than for the OLS, and the IV estimation suggests that the negative effect is not stable or shrinks but increases in the long run, although all of the employment-related results in Table 3.4 decrease in the long run.

For a summary with interpretation, except for working hours, my estimation results suggest there are negative effects on labor market outcomes. Even in the long run, such negative effects exist, except for regular worker status. This type of negative effect is weakened in the long run, but it is persistent or increases in income. Regarding the education protection effect, education has protective power from the negative entry condition effect, even if I exclude the ability bias of the IV estimation. However, such protective power of education cannot completely cover the negative effect of a 1% increase in the entry unemployment rate, even if workers take four years of university education. Hence, I conclude that education itself might work as insurance for adverse economic conditions at entry timing, even if I exclude the ability bias of the IV estimation. In addition, the IV estimation suggests the possibility of underestimation by the OLS.

3.4.3 Effect on job polarization

Table 3.6 shows the consequences of a 1% increase in the unemployment rate at the entry for short- and long-run changes in occupation task scores. Although the single-term coefficient of education is similar for high-skilled and low-skilled individuals between OLS and IV, the signs of the coefficients are opposite. Here, I only present the IV estimation results. The IV estimations in Table 3.6 suggest that the task amount of high- and low-skilled workers do not change with recession at the entry for both the short- and long-run. Only middle-skilled tasks decrease with recession at 1-5 potential experience around 0.8 to 1.2, and more education has the power to ease the speed of this temporal decrease but cannot completely cover a 1% shock of the unemployment rate increase in the entry year even if workers take four years more education. Table 3.6 is consistent with the findings of Jaimovich and Siu (2020) since I also confirm no positive recovery of task amounts even for potential experience 9—13, although the task amounts of high- and low-skilled workers remain constant. Table 3.7 shows the consequences of a 1% increase in the unemployment rate at the entry for short- and long-run changes in the share of occupation task scores. The IV estimation in Table 3.7 suggests that the share of middle-skilled workers shrunk during the recession.

In summary, Tables 3.6 and 3.7 suggest that the middle-skilled share shrinks with recession and due to middle-skilled task diminishing, which does not imply an increase in the high- or low-skilled tasks. This shrinking effect is just between 1-5 potential experiences and the effect disappears subsequently but does not positively recover, as Jaimovich and Siu (2020) said. Additionally, taking a higher education might slow down the speed of middle-skilled tasks to diminish speed, whose interpretation might be ambiguous. Based on these results, I suspect that middle-skilled tasks diminishing is associated with the curse effect of labor market entry under recession.

3.4.4 Robustness: The highest possibility path data usage

To confirm my previous results in Tables 3.4 and 3.5, I changed my sample construction about the entry region and age 18 regions based on the transition matrix generated by ETS and BSS. Here, I set the graduation and high school regions on the highest possible path. Table 3.8 shows the difference between the current region assumption data and the highest possible path data based on high school graduation year times in the current region. Since university graduates have a higher mobility tendency, there are 2,576 workers whose previous region information is modified. Additionally, 2 years higher education institution graduated more than high school graduates, and 311 observations were corrected, and the maximum path of high school graduates was within the same region transition for all regions. The modification rate of this procedure is as trivial as 1.35%. Tables 3.9 and 3.10 show the highest transition path data results for the robustness check in Tables 3.3 and 3.4. These results show that some coefficients change slightly, but the previous results and interpretation are robust. Using the current region does not cause a significant change in the usage of the highest possible path.

3.5 Conclusions

This chapter extends the literature on the curse effects of initial labor market participation during recessions to address whether higher education can mitigate such negative effects. I included education and its interaction terms into the standard estimation model for the hysteresis effects, as in Cutler et al. (2015), and conducted an IV estimation to handle the ability bias newly caused by inserting education in an estimation model. I addressed the ability bias with reduced-form estimation in the curse effect literature might be unique. The results demonstrate that higher education can protect against the curse effects, but the negative effect of a 1% increase in the unemployment rate at entry cannot be covered by four years of higher education in any aspect, and the protection effect itself is not as large. I also analyzed whether an entry-year negative shock is associated with job polarization and found that long-run polarization speed is not affected by a temporal recession, but the short-run decrease of middle-skilled workers, as in Jaimovich and Siu (2020), is confirmed by the data in this chapter. Considering these two relationships might be unique.

This chapter faces a couple of limitations. First, it uses the same data source as Genda et al. (2010), which only has information on current residence and lacks accurate entry year unemployment rate and age 18 information. This is an advantage of using Japanese data in that there is little movement of workers back to the education sector once they enter the labor market, but data accuracy is a major limitation. In particular, even for the case with the highest mobility rate in my robustness, the data accuracy of the highly educated is still low. When looking at the differences between education levels without ability bias, the structure must be so that high school graduates and higher education students make educational choices after observing the same situations. Second, I use the information at the age of 18, but the data accuracy is not very high. Therefore, I cannot wholly avoid the bias based on the measurement errors even in this design, where the same age of 18 information variables is used to make educational choices.

As for the implications for individual decision-making, this chapter shows that the behavior of taking temporary refuge in an institution of higher education during a recession may be a good choice in that it could weaken the curse effect of entering the labor market during a recession. However, it is worth noting that this curse effect could not be canceled out if the recession remains at the same level for the next 2-4 years. For policymakers, the results of previous studies show that individuals with higher levels of education are less damaging than those with lower levels. This implies that the negative effect on highly educated individuals may persist if the government adopts policies of zero subsidies to the highly educated. While it is crucial from the perspective of limited resource allocation to differentiate between the highly educated and the less educated, it is important to keep in mind that the benefits of education itself are not significant enough to cover all negative effects of a 1% increase in the unemployment rate.

Table 3.1: Summary statistics for outcomes, U1st, and instruments by education level

educ	sum stat	educ	employ	standard	regular	work hour	log(income)
	mean	12	0.91	0.78	0.88	45.86	5.43
	s.d.	0	0.29	0.42	0.33	13.33	0.65
educ=12	min	12	0	0	0	0	3.18
	max	12	1	1	1	120	7.32
	n	102,760	102,760	90151	93508	73669	95053
	mean	14	0.94	0.84	0.91	47.02	5.59
	s.d.	0	0.23	0.36	0.28	13.35	0.56
educ=14	min	14	0	0	0	0	3.18
	max	14	1	1	1	120	7.32
	n	33216	33216	30135	31279	23247	31551
	mean	16	0.95	0.89	0.94	47.52	5.82
	s.d.	0	0.21	0.32	0.25	13.4	0.58
educ=16	min	16	0	0	0	0	3.18
	max	16	1	1	1	120	7.32
	n	77926	77926	72269	74255	51280	74500
	mean	13.87	0.93	0.83	0.91	46.66	5.61
	s.d.	1.84	0.25	0.38	0.29	13.38	0.64
Total	min	12	0	0	0	0	3.18
	max	16	1	1	1	120	7.32
	n	213,902	213,902	192,555	199,042	148,196	201,104

educ	sum stat	U1st	U18	caprate 4 years	caprate 2 years	log(4 years tuition)	log(2 years tuition)
	mean	4.13	4.13	0.34	0.37	13.51	13.57
	s.d.	1.32	1.32	0.18	0.09	0.17	0.13
educ=12	min	1.3	1.3	0.1	0.21	13.04	13.15
	max	6.9	6.9	0.81	0.54	13.93	13.93
	n	102760	102760	102760	102760	102760	102760
	mean	4.29	3.96	0.35	0.38	13.5	13.57
	s.d.	1.16	1.32	0.18	0.1	0.17	0.13
educ=14	min	1.6	1.3	0.1	0.21	13.04	13.15
	max	6.9	6.9	0.81	0.54	13.93	13.93
	n	33216	33216	33216	33216	33216	33216
	mean	4.51	3.87	0.38	0.39	13.49	13.57
	s.d.	1.03	1.29	0.18	0.1	0.17	0.13
educ=16	min	2.3	1.3	0.1	0.21	13.04	13.15
	max	6.9	6.9	0.81	0.54	13.93	13.93
	n	77926	77926	77926	77926	77926	77926
	mean	4.3	4	0.36	0.38	13.5	13.57
	s.d.	1.2	1.31	0.18	0.1	0.17	0.13
Total	min	1.3	1.3	0.1	0.21	13.04	13.15
	max	6.9	6.9	0.81	0.54	13.93	13.93
	n	213,902	213,902	213,902	213,902	213,902	213,902

Table 3.2: Summary statistics for the task data by education level

education level	sum stat	task score			share of task score(%)		
		Highly skilled	middle	Low skilled	Highly skilled	middle	Low skilled
educ = 12	mean	51.31	37.01	15.66	49.36	35.55	15.08
	sd	4.56	3.95	1.16	1.68	1.89	0.48
	min	41.68	29.47	12.62	45.07	30.37	14.09
	max	62.29	46.78	17.84	54.16	39.69	16.38
	n	75648	75648	75648	75648	75648	75648
educ = 14	mean	53.39	37.22	16.1	50.04	34.85	15.12
	sd	4.78	3.99	1.17	1.64	1.92	0.52
	min	41.68	29.47	12.62	45.07	30.37	14.09
	max	62.29	46.78	17.84	54.16	39.69	16.38
	n	26935	26935	26935	26935	26935	26935
educ = 16	mean	54.84	36.56	16.28	50.95	33.91	15.15
	sd	4.43	3.98	1.04	1.52	1.79	0.48
	min	41.68	29.47	12.62	45.07	30.37	14.09
	max	62.29	46.78	17.84	54.16	39.69	16.38
	n	71003	71003	71003	71003	71003	71003
Total	mean	53.16	36.85	16	50.16	34.73	15.12
	sd	4.82	3.98	1.15	1.76	2	0.49
	min	41.68	29.47	12.62	45.07	30.37	14.09
	max	62.29	46.78	17.84	54.16	39.69	16.38
	n	173586	173586	173586	173586	173586	173586

Table 3.3: First-stage association between age 18 information and education decisions and potential experience and age

	(1) educ	(2) pexp
age	0.227***	0.717***
	(0.032)	(0.031)
Capacity rate of four years	21.937***	-21.674***
	(6.754)	(6.657)
Capacity rate of two years	-16.846**	16.827**
	(6.909)	(6.841)
U18	1.982***	-1.956***
	(0.37)	(0.363)
log(tuition) for four years	26.698***	-26.205***
	(4.836)	(4.762)
log(tuition) for two years	-23.353***	23.003***
	(5.184)	(5.114)
Observations	213902	213902
First stage test statistics		
Under identification test (Kleibergen-Paap rk LM statistic):		123.322**
Weak identification test (Kleibergen-Paap rk Wald F statistic):		11.987
Over identification test (Hansen J statistic)		86.481

Standard errors are in parentheses. This is the first-stage result of the association between education decisions and age 18 information and potential experience given a labor force other than self-employed. For estimation, survey month current residential regions, and high school graduation cohort fixed effects are controlled. I use the age 18 information, average tuition for university, the average tuition for junior college, capacity rate for university, capacity rate for two years higher education institution, unemployment rate at age 18 as education instruments, and use age dummies as instruments for potential experience and the interaction with entry years instruments for our main. Standard errors are clustered by region block $\times$ entry year. I rely on Kleibergen and Paap (2006)'s statistics because I use the cluster standard error with respect to current regions and high school graduation year. The p-value of the under-identification test was 0.0229, and its over identification test is 0.6702. The first-stage F-statistic is close to 10, which is relatively weak. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.4: The curse effect by labor market entry during a recession and education's protective effect on employment

	(1)	(2)	(3)	(4)	(5)	(6)
	employ		full-time ("seishain")		regular	
	OLS	IV	OLS	IV	OLS	IV
pexp=1 × U1stn	-0.027***	-0.043***	-0.090***	-0.099***	-0.078***	-0.084***
	(0.007)	(0.016)	(0.010)	(0.028)	(0.008)	(0.023)
pexp=5 × U1stn	-0.030***	-0.033*	-0.093***	-0.129***	-0.045***	-0.056**
	(0.006)	(0.017)	(0.009)	(0.034)	(0.007)	(0.026)
pexp=9 × U1stn	-0.025***	-0.053***	-0.067***	-0.130***	-0.002	-0.011
	(0.006)	(0.018)	(0.009)	(0.027)	(0.007)	(0.023)
pexp=13 × U1stn	-0.001	-0.022**	-0.039***	-0.064***	0.021***	0.012
	(0.006)	(0.010)	(0.009)	(0.020)	(0.007)	(0.015)
pexp=1 × U1stn × educ	0.001**	0.002*	0.005***	0.006**	0.005***	0.005***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
pexp=5 × U1stn × educ	0.002***	0.002*	0.005***	0.008***	0.003***	0.004**
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.002)
pexp=9 × U1stn × educ	0.002***	0.004***	0.004***	0.008***	0.001	0.001
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.002)
pexp=13 × U1stn × educ	0.000	0.002***	0.003***	0.004***	-0.001	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Observations	213902	213902	192555	192555	199042	199042

Standard errors are in parentheses. The odd-numbered columns show the results of OLS, and the even-numbered columns show the results of IV. For both estimations, survey month, current residential regions, and high school graduation cohort fixed effects are controlled. For OLS, the other controls are continuous potential experiences and continuous years of education. For years of potential experience not listed in the table, I also estimate the interaction terms for all years from 1 to 13 years. For IV estimation, I use age 18 information, the average tuition for university, the average tuition for junior college, the capacity rate for university, the capacity rate for two-year higher education institutions, unemployment rate at age 18 as education instruments, and use age dummies as instruments for potential experience and the interaction with entry year as instruments for our main. Standard errors are clustered by region block × age 18. Statistical significance is as follows: $^{*}\ p < .1$, $^{**}\ p < .05$, $^{***}\ p < .01$

Table 3.5: The curse effect by labor market entry during a recession and education's protective effect on income and working hours

	(1) work hour	(2) work hour	(3) log(income)	(4) log(income)
pexp=1 $\times$ U1st	-0.419	-0.056	-0.114***	-0.105**
	(0.398)	(1.182)	(0.016)	(0.048)
pexp=5 $\times$ U1st	0.491	1.868	-0.116***	-0.142***
	(0.354)	(1.302)	(0.013)	(0.051)
pexp=9 $\times$ U1st	1.069***	3.085**	-0.112***	-0.243***
	(0.399)	(1.224)	(0.013)	(0.050)
pexp=13 $\times$ U1st	2.129***	1.547	-0.100***	-0.163***
	(0.552)	(1.055)	(0.014)	(0.032)
pexp=1 $\times$ U1st $\times$ educ	0.030	-0.003	0.005***	0.004
	(0.028)	(0.097)	(0.001)	(0.004)
pexp=5 $\times$ U1st $\times$ educ	-0.021	-0.124	0.008***	0.011***
	(0.025)	(0.095)	(0.001)	(0.004)
pexp=9 $\times$ U1st $\times$ educ	-0.058**	-0.188**	0.009***	0.018***
	(0.029)	(0.080)	(0.001)	(0.003)
pexp=13 $\times$ U1st $\times$ educ	-0.148***	-0.104	0.008***	0.011***
	(0.039)	(0.070)	(0.001)	(0.002)
Observations	148196	148196	201104	201104

Standard errors are in parentheses. The odd-numbered columns show the results of OLS and the even-numbered columns show the results of IV. For both estimations, survey month, current residential regions, and high school graduation cohort fixed effects are controlled. For OLS, the other controls are continuous potential experiences and continuous years of education. For years of potential experience not listed in the table, I also estimate the interaction terms for all years from 1 to 13 years. For the IV estimation, I use age 18 information, average tuition for university, average tuition for junior college, the capacity rate for university, capacity rate for two-year higher education institution, unemployment rate at age 18 as education instruments, and use age dummies as instruments for potential experience and the interaction with entry year as instruments for our main. Standard errors are clustered by region block $\times$ age 18. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.6: How are total task scores affected by entry labor market conditions and how does education change with OLS and IV

	(1)	(2)	(3)	(4)	(5)	(6)
	task score: high skilled		task score: middle		task score: low	
	OLS	IV	OLS	IV	OLS	IV
pexp=1 × U1st	0.739***	-0.232	0.560***	-0.802**	0.061**	-0.127
	(0.116)	(0.330)	(0.104)	(0.317)	(0.029)	(0.078)
pexp=5 × U1st	0.388***	-0.361	0.145	-1.223***	0.019	-0.066
	(0.110)	(0.393)	(0.096)	(0.380)	(0.027)	(0.093)
pexp=9 × U1st	0.623***	-0.105	0.507***	-0.387	0.092***	0.041
	(0.114)	(0.375)	(0.100)	(0.375)	(0.028)	(0.091)
pexp=13 × U1st	0.867***	0.262	0.811***	0.212	0.152***	0.099
	(0.124)	(0.318)	(0.110)	(0.322)	(0.031)	(0.076)
pexp=1 × U1st × educ	-0.056***	0.023	-0.037***	0.071***	-0.005***	0.010
	(0.008)	(0.027)	(0.007)	(0.025)	(0.002)	(0.006)
pexp=5 × U1st × educ	-0.022***	0.028	0.002	0.095***	0.000	0.004
	(0.008)	(0.030)	(0.007)	(0.028)	(0.002)	(0.007)
pexp=9 × U1st × educ	-0.035***	0.015	-0.020***	0.040	-0.004*	-0.000
	(0.008)	(0.026)	(0.007)	(0.026)	(0.002)	(0.006)
pexp=13 × U1st × educ	-0.049***	-0.011	-0.039***	-0.001	-0.007***	-0.004
	(0.009)	(0.023)	(0.008)	(0.023)	(0.002)	(0.005)
educ	1.038***	0.906***	-0.013	-0.181*	0.176***	0.164***
	(0.029)	(0.087)	(0.026)	(0.097)	(0.007)	(0.021)
Observations	173586	173586	173586	173586	173586	173586

Standard errors are in parentheses. The odd-numbered columns show the OLS results, and the even-numbered columns show IV results. For both the estimation, survey month, current residential regions, and high school graduation cohort fixed effects are controlled. For OLS, the other controls are continuous potential experiences and continuous years of education. For the IV estimation, I use age 18 information, the average tuition for university, the average tuition for junior college, the capacity rate for university, the capacity rate for two-year higher education institution, unemployment rate at age 18 as education instruments, and use age dummies as instruments for potential experience and the interaction with entry year as instruments for the main analysis. Standard errors are clustered by region block × entry year. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.7: How are the shares (percentage) of task scores affected by the entry labor market conditions, and how does education change with OLS and IV

	(1)	(2)	(3)	(4)	(5)	(6)
	task share: high skilled		task share: middle		task share: low	
	OLS	IV	OLS	IV	OLS	IV
pexp=1 × U1st	0.072*	0.326***	0.059	-0.371***	-0.130***	0.045
	(0.041)	(0.115)	(0.048)	(0.138)	(0.013)	(0.042)
pexp=5 × U1st	0.120***	0.440***	-0.064	-0.609***	-0.056***	0.169***
	(0.039)	(0.139)	(0.044)	(0.168)	(0.012)	(0.050)
pexp=9 × U1st	0.024	0.139	0.056	-0.244	-0.080***	0.105**
	(0.042)	(0.128)	(0.048)	(0.159)	(0.012)	(0.047)
pexp=13 × U1st	-0.025	-0.007	0.132**	-0.012	-0.108***	0.020
	(0.047)	(0.105)	(0.053)	(0.128)	(0.013)	(0.039)
pexp=1 × U1st × educ	-0.007**	-0.027***	-0.002	0.032***	0.009***	-0.006*
	(0.003)	(0.009)	(0.003)	(0.011)	(0.001)	(0.003)
pexp=5 × U1st × educ	-0.012***	-0.034***	0.009***	0.048***	0.003***	-0.014***
	(0.003)	(0.010)	(0.003)	(0.012)	(0.001)	(0.004)
pexp=9 × U1st × educ	-0.006**	-0.013	0.002	0.021*	0.004***	-0.008**
	(0.003)	(0.009)	(0.003)	(0.011)	(0.001)	(0.003)
pexp=13 × U1st × educ	-0.002	-0.004	-0.004	0.006	0.006***	-0.002
	(0.003)	(0.007)	(0.004)	(0.009)	(0.001)	(0.003)
educ	0.417***	0.439***	-0.415***	-0.469***	-0.002	0.030**
	(0.011)	(0.030)	(0.012)	(0.038)	(0.003)	(0.012)
Observations	173586	173586	173586	173586	173586	173586

Standard errors are in parentheses. The odd-numbered columns show the OLS results, and the even-numbered columns show the IV results. For both the estimation, survey month, current residential regions, and high school graduation cohort fixed effects are controlled. For OLS, the other controls are continuous potential experiences and continuous years of education. For IV estimation, I use age 18 information, the average tuition for university, the average tuition for junior college, the capacity rate for university, the capacity rate for two-year higher education institutions, the unemployment rate at the age of 18 as education instruments, and use age dummies as instruments for potential experience and interaction with entry year as instruments for our main. Standard errors are clustered by region block × entry year.Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.8: Differences between the current region and the highest previous region path rate

education = 16

Age 18 year \Current region	Chugoku	Hokuriku	Kita-Kanto	Shikoku	Hokkaido	Total
1991	0	255	406	223	0	884
1995	0	0	512	0	0	512
1996	405	0	0	0	0	405
2002	0	0	0	172	0	172
2003	0	0	298	0	0	298
2008	0	0	112	0	0	112
2009	0	0	65	37	0	102
2010	0	0	52	0	0	52
2011	0	0	21	0	18	39
Total	405	255	1,466	432	18	2,576

Age 18 region \Current region	Chugoku	Hokuriku	Kita-Kanto	Shikoku	Hokkaido	Total
Kinki	405	0	0	432	0	837
Minami-Kanto	0	255	1,354	0	18	1,627
Shikoku	0	0	112	0	0	112
Total	405	255	1,466	432	18	2,576

education = 14

Age 18 year \Current region	Chugoku	Kita-Kanto	Shikoku	Total
2002	148	0	0	148
2004	0	0	84	84
2007	0	79	0	79
Total	148	79	84	311

Age 18 region \Current region	Chugoku	Kita-Kanto	Shikoku	Total
Kyusyu	148	0	0	148
Tohoku	0	79	0	79
Tokai	0	0	84	84
Total	148	79	84	311

Table 3.9: The curse effect by labor market entry during a recession and education's protective effect on employment

	(1)	(2)	(3)	(4)	(5)	(6)
	employ		full-time ("seishain")		regular	
	OLS	IV	OLS	IV	OLS	IV
pexp=1 × U1st	-0.027***	-0.043***	-0.087***	-0.095***	-0.076***	-0.080***
	(0.007)	(0.016)	(0.010)	(0.027)	(0.008)	(0.023)
pexp=5 × U1st	-0.030***	-0.034**	-0.090***	-0.129***	-0.043***	-0.055**
	(0.006)	(0.017)	(0.009)	(0.033)	(0.007)	(0.026)
pexp=9 × U1st	-0.024***	-0.053***	-0.064***	-0.132***	0.000	-0.015
	(0.006)	(0.017)	(0.009)	(0.027)	(0.007)	(0.023)
pexp=13 × U1st	-0.001	-0.023**	-0.036***	-0.060***	0.022***	0.011
	(0.006)	(0.010)	(0.009)	(0.020)	(0.007)	(0.015)
pexp=1 × U1st × educ	0.001**	0.002*	0.004***	0.005**	0.004***	0.005**
	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
pexp=5 × U1st × educ	0.002***	0.002**	0.005***	0.008***	0.003***	0.004**
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.002)
pexp=9 × U1st × educ	0.002***	0.004***	0.004***	0.008***	0.000	0.002
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.002)
pexp=13 × U1st × educ	0.000	0.002***	0.002***	0.004***	-0.001**	0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Observations	213902	213902	192555	192555	199042	199042

Standard errors are in parentheses. The odd-numbered columns show the OLS results, and the even-numbered columns show the IV results. For both estimations, survey month, current residential regions, and high school graduation cohort fixed effects are controlled. For OLS, the other controls are continuous potential experiences and continuous years of education. I estimated the interaction terms for all years from 1 to 13 years for years of potential experience not listed in the table. For IV estimation, I used age 18 information, the average tuition for university, the average tuition for junior college, the capacity rate for university, the capacity rate for two-year higher education institutions, and the unemployment rate at age 18 as education instruments. I likewise used age dummies as instruments for potential experience and the interaction with entry year as instruments for our main. Standard errors are clustered by region block × age 18. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.10: The curse effect by labor market entry during a recession and education's protective effect on income and working hours with the highest transition data

	(1) work hour	(2) work hour	(3) log(income)	(4) log(income)
pexp=1 × U1st	-0.375	0.322	-0.111***	-0.089*
	(0.398)	(1.167)	(0.016)	(0.048)
pexp=5 × U1st	0.530	1.950	-0.112***	-0.127***
	(0.355)	(1.307)	(0.013)	(0.049)
pexp=9 × U1st	1.100***	3.020**	-0.109***	-0.226***
	(0.399)	(1.214)	(0.013)	(0.049)
pexp=13 × U1st	2.123***	1.954*	-0.097***	-0.146***
	(0.552)	(1.032)	(0.014)	(0.031)
pexp=1 × U1st × educ	0.028	-0.034	0.005***	0.003
	(0.028)	(0.096)	(0.001)	(0.004)
pexp=5 × U1st × educ	-0.023	-0.131	0.008***	0.009***
	(0.025)	(0.095)	(0.001)	(0.004)
pexp=9 × U1st × educ	-0.061**	-0.190**	0.009***	0.016***
	(0.029)	(0.080)	(0.001)	(0.003)
pexp=13 × U1st × educ	-0.150***	-0.137**	0.007***	0.010***
	(0.039)	(0.069)	(0.001)	(0.002)
Observations	148196	148196	201104	201104

Standard errors are in parentheses. The odd-numbered columns show the OLS results, while the even-numbered columns show the IV results. In both estimations, survey month, current residential regions, and high school graduation cohort fixed effects are controlled. For OLS, the other controls are continuous potential experiences and continuous years of education. For years of potential experience not listed in the table, I estimate the interaction terms for all years from 1 to 13 years. For IV estimation, I use age 18 information, the average tuition for university, the average tuition for junior college, the capacity rate for university, the capacity rate for two-year higher education institution, unemployment rate at age 18 as education instruments, and use age dummies as instruments for potential experience and the interaction with entry year as instruments for our main. Standard errors are clustered by region block × age 18. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$

Chapter 4

Estimating Causal Returns in a Content of College Education: Estimating the Effects of Short-Term Study Abroad

Abstract

This chapter examines the effect of participation in short-term study abroad (SSA) programs from randomly assigned data of applicants from a Japanese university in March and August 2014 (705 applicants and 300 participants). My results demonstrate that participation in SSA causally increased English test scores and long-term study abroad (LSA) participation rates. Regarding the firms' initial employment attributes, this chapter found that SSA participants tended to work for firms with significantly higher sales and foreign stock rates more than non-participants, although I did not find significant differences in initial monthly income. [1]

[1]This chapter 4 is summarized based on the co-authored paper, Kashima, Ryohei, and Maki Kato, "Short-term Study Abroad Effect on Within-School Outcomes and Initial Career: Evidence from Random Assignment Data from a Japanese University," published by Mori Arinori Institute for Higher Education and Global Mobility Working Paper Series, WP2020-01, December 2020, Tokyo, Japan.

4.1 Introduction

While higher education is structured so that all students receive the same general education through high school, higher education involves many choices for individual students, including department and class selection. When an individual choice is involved, empirical research faces the problem of endogeneity. Students who receive specific content within higher education differ from other students in general, and the potential exists for differences in their underlying attributes to cause apparent variation in the outcomes they wish to influence. In other words, it is not easy to confirm what educational content offers exist in higher education and whether there is any payoff for students. In this chapter, I examine whether specific higher education content offerings some educational returns by examining the effects of study abroad (SA), particularly short-term SA (SSA), where self-selection effects are enormous and causality has not sufficiently been identified.

Recently, there has been an increase in Japanese university students preferring to SA. During the 2009 academic year, approximately 36,000 students were studying outside Japan. However, by 2018, that number had increased by more than three times (approximately 3.19 times) (of Education Culture Sports Science and Technology Japan, 2022). Two-thirds of the increase during this period came from short-term SA (SSA, SA for less than one month), which spread faster than long-term SA (LSA, SA for longer than one month). In Japan, most university SA programs, including SSA, are subsidized by the government or the university to cover the cost of participants. The use of grants is justified when at least some benefit exists. Otherwise stated, if some causal effect is not visible, it becomes almost equivalent to taxes being spent on students' trips abroad.

If the duration is not specified, it has been shown that SA has some causal power. International mobility of labor has been a major topic in the literature on the causal effects of SA. Parey and Waldinger (2010) conducted a study using the establishment of the ERASMUS program as a natural experiment to increase the probability of SA participation within the EU, which became one benchmark study. Parey and Waldinger (2010) used the ERASMUS program as an instrumental variable and concluded that SA participation increases the probability of employment abroad in Germany, contributing in a causal sense to international labor mobility. Using a similar approach, I show that in Italy, SA participation similarly increases the probability of finding a job abroad (Di Pietro, 2012). In addition, using the cutoff in the Dutch study abroad scholarship program as an instrument, Oosterbeek and Webbink (2011) found that the number of months spent studying abroad has a positive causal effect on the likelihood of living abroad. Additionally, using the change in the language of teaching in master's programs to English as an instrument, Nocito (2021) found that being forced to use English to study at an Italian university increases the probability of subsequent work abroad by 11.3%, which suggests that improved language proficiency causally contributes to the international mobility of the labor force.

In addition to international mobility, Sorrenti (2017) showed that study abroad participation causally improves language proficiency using the ERASMUS program as an instrumental variable. Regarding early career success after college, Di Pietro (2015) in Italy and Waibel et al. (2018) in Germany showed that study abroad participation had a positive causal effect on the probability of employment three years after graduation using instrumental variable methods and propensity score matching. Using propensity score matching, Cullinan et al. (2022) showed that one semester of SA participation does not affect non-language-related academic performance.

While causal effect estimates for SA have been accumulated, studies on causal effect estimates for SSA (SA for less than one month) are limited. Kawata and Nishitani (2017) studied the causal effect of SSA participation with English examination scores in a framework of difference-in-differences analysis and propensity score matching for two-week SSA participation among first-year university applicants in Japan. They found that SSA participation has a positive impact despite the limited period. De Poli et al. (2018) studied the causal effect of SSA participation for Italian middle school students, where participation was randomly assigned to show the effects of SSA analysis over one month. Their study showed that it contributes to language proficiency and influences personality traits such as self-confidence. Kato and Suzuki (2019) showed in Japanese university data that participation in a one-month program with random SSA assignments positively affects the probability of subsequent LSA participation.

This chapter uses the same data as Kato and Suzuki (2019), university data with random SSA assignments. Here, I investigate SSA's causal effect on some learning outcomes and career choices, which Kato and Suzuki (2019) did not address. While the SA literature explores career-related causal associations such as employability and international mobility of the labor force, the SSA literature does not address career. Hence, I extend the analysis to career in addition to analysis for learning outcomes.

My findings on the causal effects of SSA participation can be summarized as follows. For learning outcomes, I found that SSA participation increased English-related outcomes but had no impact on general learning. Moreover, anent differences in career, I observed positive and significant differences in sales and foreign stock rates between SSA participation assignments. These results were robust to various modifications. The differences in careers were primarily derived from differences in the probability of employment for students who tended to work for firms in the top or bottom tiers of sales and foreign stock rates. In the remainder of this chapter, I illustrate institutional backgrounds in Section 2, discuss the data in Section 3, present my empirical strategy in Section 4, report the main results in Section 5, discuss robustness and how career differences arose in Section 6, and conclude the chapter in Section 7.

4.2 Institutional backgrounds

4.2.1 The job-hunting system for university students in Japan

As the fiscal year begins in Japan, it is common for new graduates to start job-hunting in April of their final year of study, complete job-hunting by the summer, and begin in the workplace in the new fiscal year. Accordingly, the employment rate for university graduates in Japan was 98% for the fiscal year of 2019 before the COVID-19 pandemic began (MEXT 2020), which is higher than that of EU countries (85%; Eurostat 2021). Therefore, there is not much difference among university students as to whether they were able to find a job. The main area of interest for Japanese graduate job hunters is company attributes.

4.2.2 University background

The target university was built over 150 years ago and has four faculties and seven graduate schools concentrating on social sciences and human arts.[2] At the time of my study, more than 6,000 undergraduate and graduate students were enrolled. It is highly selective because it is ranked as one of the top 30 (out of 795) universities in Japan by the 2021 World University Rankings (THE, 2021) and has one of the highest average scores for the national standardized entrance exam in Japan. This selectivity implies that my results are not representative of the whole of Japan and face limited external validity.

4.2.3 Randomly assigned SSA programs

The SSA programs, my sample programs, were conducted in March and August 2014 (during the 2013 and 2014 academic years). These programs aim to improve English communication and cultural skills. The programs lasted about one month, and the destination countries were Australia, New Zealand, the UK, and the US. The differences between the programs are depicted in Table A1. The cost of participating in the program was low for students compared to the implementation costs. It was free in 2013 and cost 100,000 yen (= approximately 1,000 USD) in 2014, and the average cost in the academic years of 2016 to 2019 was approximately 820,000 yen (= approximately 8,200 USD) for equivalent programs of similar length. The number of applicants was 268 in 2013 and 437 in 2014, and the number of participants was 100 and 200, respectively. While the 2013 and 2014 school year programs were not exactly similar, they were adequately similar to English language study programs in native English countries. The selection method for the targeted SSA programs was the same. First, the applicants were recruited through posters and briefings. Next, they were selected using a random number to ensure a proportional number of participants according to university level and department. As one of the programs aims to collect data to create future SA programs, the selection process does not consider foreign-language proficiency or applicants' academic achievement, which are outcomes in my analysis. Even though the study institution was

[2]There are no language majors for whom the SA experience is an important part of university studies.

assigned by the university authority to satisfy the language requirement score of each institution, whether applicants can participate in SSA is determined randomly. Therefore, the covariates were theoretically aligned for those who did and did not participate in the years, and the department was controlled proportionally. The uniqueness of applicants' characters is vital for the external validity of the target university students. I checked university average TOEFL-ITP scores in the April of the first year with a free exam fee; most students took it. The university's average TOEFL score between 2008–2018 was 510, and my applicants' average was 513, as illustrated in Table 1. Therefore, there was no big difference between the applicants and average university students regarding initial TOEFL-ITP scores. Even though motivation differences might affect the results, my estimation results might be applicable also for non-applicants in this target university.

4.3 Data and summary statistics

4.3.1 Data

Data observed in school

I used the program applicants' data, which contained application timing, name, and the student ID of participants. I merged applicants' information with school administrative data relating to students' academic achievements and attributes; this included their gender, university year at the time of participation in the SSA, department, nationality, and whether they were 18 years old at the time of enrollment. I used the TOEFL-ITP scores and long-term exchange program participation[3], and cumulative GPA in the final year as outcomes. For TOEFL-ITP, the university requires first-year students to take the TOEFL test right after the entrance (April) and at the end of the first year (December). Taking a test was voluntary for students in their second year or later. The university pays the cost of taking the TOEFL-ITP test in the first year, and approximately 90% of students

[3] Age 18 is the typical high school graduation age and the most common age to enter university. Students older than 18 tend to experience gap years, known as "Rounin" in Japanese. Students failing the entrance exam are "Rounin" students, who might have lower ability than direct entry students of age 18. Therefore, whether students are "Rounin" might cause potential ability difference, which is why I controlled for it.

take the test twice. Therefore, the TOEFL-ITP score of first-year students almost does not contain self-selection as to whether they take the test. Information about who participated in LSA programs was also used in this chapter to reconfirm Kato and Suzuki's (2019) results. Several requirements for participation in LSA programs include a high GPA and English proficiency demonstrated by TOEFL-iBT and IELTS. First-year cumulative GPA, April TOEFL-ITP score, and whether participants' age was greater than 18 could all be predeterminants for participation in an SSA program. [4]

Career Data

I also used career data obtained from the university's career support office. The current data covers students who graduated from 2014 to 2018. I collected information on the name of the company the students worked at after graduation. Based on the foregoing information, I merged company information from several sources. If graduates worked at a private company listed on the Japanese stock exchange, I obtained the company information from the Nikkei NEEDS Financial QUEST 2.0. The percentage of graduates working for a company with an identified security code (market identifier code) was 58.4%. Changes in company names, mergers, and listing status at the time of employment were individually collected from financial reports from the end of January 2018 to the end of December 2018. I used information for firms listed on the stock exchange "Kaisyashikiho-2018," which included initial monthly salary information. I obtained information from non-listed companies from "Kaisyashikiho-Mizyozyokaisyaban-2018," which contains financial information about sales, capital, build year, number of workers, and initial monthly salaries for non-listed companies. I considered a company as a foreign company if its name was listed in "Gaishikeikigyosoran-2018". The career data used in this chapter faced some limitations as non-listed firms' information was not as complete as the data of listed firms, even if "Kaisyashikiho-Mizyozyokaisyaban-2018" contained the firm's name. Additionally, the need for graduates to report the company name after receiving a job offer is not mandatory, which poses

[4] First-year cumulative GPA could contain the effect of participation in SSA for first-year students among the 2014 applicants. This could be a source of underestimation for the effect on the final year cumulative GPA.

another limitation. Subsequently, this data access imbalance and reporting error might become a source of bias in my estimation.

4.3.2 Summary statistics

Of the 651 participants, 605 provided career information, of which 513 were employed in a private company at graduation. Of these students, 327 (63.7%) were employed in Japanese stock market-listed firms, and their information became my main data source. The remaining 186 (36.3%) participants were working for firms not listed on the Japanese stock market. Table 1 illustrates the summary statistics for the SSA program applicants and participants. Some data are missing in the samples for educational outcomes and initial job attributes owing to data availability issues. The program participants' initial job attributes and learning outcome averages tended to be slightly higher than those of non-participants. In Table 1, the t-test results demonstrated no significant differences in predeterminant skills or personal traits between participants and non-participants. Additionally, all normalized differences, suggested by Imbens and Wooldridge (2009) and Imbens and Rubin (2015) for balanced tests, are smaller than 0.2; however, some are larger than 0.1. Some variables greater than 0.1 is controlled in subsequent regressions. Although the t-tests showed no balance problems, the normalized difference results suggest that the balance may not be well balanced. Therefore, we reported results controlling for predeterminant skills and students' attributes in the main estimation results.

4.4 Empirical strategies

At the outset, I defined each year's average treatment effect (ATE) from participating in SSA with an initial assignment. I also defined the multi-year average ATE (AATE) to treat small sample problems. Subsequently, I described the outcome variables using estimations, learning outcomes, and initial firm attributes. This section follows Rubin (1974). I denoted "winners" as SSA participants with lottery assignments and "losers" as non-participants despite applying for

Table 4.1: Summary statistics

Category of Variables	Variables	Participants (winners)			Non-Participants (losers)			Balanced test	
		Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	t-Value	Norm-Dif
Learning outcomes	TOEFL overall difference	92	19.90	27.89	97	16.04	29.89		
	TOEFL listening score difference	92	1.79	3.30	97	1.94	3.23		
	TOEFL grammar score difference	92	2.42	4.56	97	2.09	5.26		
	TOEFoutcomesL reading score difference	92	1.76	3.97	97	0.81	4.90		
	Cumulative GPA at graduation	258	3.26	0.45	355	3.20	0.50		
	Participate long-term	258	3.11	0.49	355	3.05	0.50		
Choice of Career	Becoming a worker	249	0.88	0.33	336	0.84	0.37		
	Working in a private firm	219	0.94	0.25	283	0.96	0.18		
	Working in a listed firm	205	0.66	0.47	273	0.69	0.46		
Inital firm attributes, outcomes	Log(Initial monthly income)	135	5.41	0.09	173	5.41	0.11		
	Log(Sales)	116	14.40	1.54	178	13.84	2.03		
	Log(Capital)	134	11.63	2.49	198	11.41	2.75		
	Log(Number of workers)	130	8.85	1.95	183	8.58	1.97		
	Build year	163	1961.22	37.96	225	1966.12	38.51		
	Work in a foreign-affiliated firm	197	0.10	0.30	253	0.10	0.30		
	Foreign stock rate	126	0.31	0.13	181	0.28	0.12		
Controls (predeterminant skill)	April TOEFL at First year	258	514.20	40.85	355	513.05	39.61	0.35	-0.03
	First-year GPA	258	3.26	0.45	355	3.20	0.50	-1.33	-0.13
	Entrance age greater than 18	258	0.34	0.47	355	0.39	0.49	-1.31	0.10
Controls (personal attributes)	Female	258	0.33	0.47	355	0.35	0.48	-0.44	0.04
	Second year	258	0.36	0.48	355	0.39	0.49	-0.78	0.06
	Third year	258	0.09	0.29	355	0.11	0.32	-0.94	0.07
	Major2	258	0.27	0.45	355	0.25	0.43	0.73	-0.05
	Major3	258	0.14	0.34	355	0.15	0.35	-0.38	0.03
	Major4	258	0.28	0.45	355	0.25	0.43	0.89	-0.07
	International students	258	0.08	0.27	355	0.11	0.31	-1.23	0.10

The t-value is calculated by OLS regression of each of these variables on the initial lottery assignment. Norm-Dif demonstrates the normalized difference, which is suggested for balanced tests by Imbens and Wooldridge (2009) and Imbens (2015). TOEFL difference only demonstrates first-year students in 2014 applicants because of the sample selection problem. All attributes are not statistically significant, even at the 10% level of the t-test. Unfortunately, some of the absolute value of the normalized difference is greater than 0.1.

SSA.

4.4.1 Defining ATE and AATE

The SSA program was conducted twice in the 2013 and 2014 academic years. Those who did not make it in 2013 were able to reapply for the 2014 programs. Hence, the final participation status unbalanced the latent characteristics between "winners" and "losers." [5] To balance the latent characteristics, I used the initial assignment results as the treatment for individual students. Through this procedure, I calculate the program ATE each year, assuming that the subsequent behavior of

[5] I can define two types of latent characteristics as "apply again" and "give up," which are actions in case of loss in the SSA participation lottery. If I use the final participation status, then the program ATE in 2013 will be unbalanced since "winners" in 2013 completely hold both characteristics even though some "losers" in 2013 with "apply again" change their status to "winners."

reapplication and other program participation was part of a later behavior pattern for applicants if they failed. With the initial assignment result as the treatment, my estimation regression is as follows:

$$y_i = \alpha_0 + \alpha_1 1stLotteryWin2013_i + \alpha_2 1stLotteryWin2014_i + \alpha_3 PrgYear2014_i + X_i\alpha + error_i \tag{4.1}$$

where the subscript i represents the individuals. The dependent variable y_i indicates learning outcomes or initial job attributes. $1stLotteryWin2013_i$ and $1stLotteryWin2014_i$ are the dummy variables indicating students who participated in the SSA program for that year. X_i represents the control variables, and $error_i$ indicates error terms. As controls, I used dummies of major, grade, female, entrance age greater than age 18, and scores of TOEFL-ITP for first-year April and first-year GPA. In regression (1), α_1 and α_2 become 2013's and 2014's ATE, respectively. My sample size was not very large, so there is a possibility that the 2013 or 2014 participants may have accidentally experienced a treatment effect greater than in other years because of this limitation. To combat this issue, I focused mainly on AATE, which is the same as the weighted average of each year's ATE:

$$AATE = (\alpha_1 \frac{P(PrgYear = 2013)}{P(PrgYear = 2013, 2014)} + \alpha_2 \frac{P(PrgYear = 2014)}{P(PrgYear = 2013, 2014)} \tag{4.2}$$

where $P(PrgYear = T)$ is the sample weight at initial assignment in year T, and $T = 2013, 2014$ implies the full sample. This AATE can be expressed as a coefficient β_1 in a simpler regression as follows:

$$y_i = \beta_0 + \beta_1 1stLotteryWin_i + \beta_2 PrgYear2014_i + X_i\beta + error_i \tag{4.3}$$

In this chapter, I report on whether AATE was significant or not. Although reporting coefficients without any controls are sufficient for RCT estimation, Van Breukelen and Van Dijk (2007) note that including controls in RCT increases efficiency. I attached regression results with only controlling stratification variables, major and grade, as appendix Table A2, and discussed whether results with and without controls were different in the result section.

4.4.2 Description of outcomes

As learning outcomes, I used TOEFL-ITP score growth from the spring to the winter of the first year and the final year cumulative GPA and the LSA programs' participation rate. Notably, these outcomes were observable within the school and related to learning. The TOEFL-ITP score is a proxy for English proficiency. The final cumulative GPA is a proxy for overall learning attitudes, like study hours. I can understand whether English proficiency increased and overall learning attitudes changed by examining these factors as outcomes. I cannot measure ATE based on the increase in participants' TOEFL-ITP scores unless the SSA is sandwiched between the TOEFL-ITP exams. Students can take TOEFL-ITP for free in their first year in April and December and must pay an exam fee for the other grades. Hence, the other grades face self-selection problems for taking the TOEFL-ITP.[6] The sandwiched case without self-selection is that of first-year students with SSA applications in 2014. Therefore, I estimated only ATE2014 for TOEFL even though I estimated AATE on cumulative GPA and the LSA programs' participation rate. To measure the initial firm attributes, I first used dummy variables indicating whether students chose a career as a worker, became an employee in a private firm (given that they did choose a career as a worker), and became an employee in a private firm listed in the Japanese stock market (given that they did work in a private firm). Using them as outcomes in the linear probability model, as seen in (3), I estimate the AATE on the initial career choice probabilities as employed workers and workers in listed firms. Next, I used initial monthly income, sales, capital, number of workers, and established year as the outcome variables y_i. Then, I used the indicator variable of whether the firm was foreign-affiliated or not and used foreign stock rates as outcomes to check if SSA increased interest in firms' global aspects. Firms' initial attributes are only determined in the market, and SSA participation cannot

[6] Of the sample, 90% take both TOEFL-ITP. I can therefore say there is almost no self-selection for first-year TOEFL-ITP.

directly affect attributes. Therefore, AATE became the average difference in firms' initial attributes between participants and non-participants.

4.5 Results

4.5.1 Effect on learning outcomes

Table 4.2 presents the causal AATE and ATE of SSA program participation with the initial assignment on learning outcomes. Column 1 indicates the ATE2014 on the growth score of the TOEFL-ITP for first-year students in 2014. As I noted in Section 4.2, this estimation limits the sample to first-year applicants in 2014 to avoid self-selection bias regarding TOEFL-ITP scores. The coefficient in column 1 indicates that participating in the SSA program with the initial assignment has the causal power to increase TOEFL scores from December to April, on average, by 8.1 points more than by not participating. This coefficient size is around two times larger than the standard error, and the result is consistent with Kawata and Nishitani (2017) and De Poli et al. (2018), whose results indicated a significant improvement in English proficiency.[7] Column 2 indicates the AATE on the cumulative GPA at the time of graduation, and column 3 indicates the probability of participation in LSA programs. Participation in SSA programs has no causal link to increasing GPA. In line with the results from Kato and Suzuki (2019), I found that the LSA participation rate increased by 11.5% on average compared to non-participants. These results imply that participation in an SSA program increases English-related outcomes but not non-English-related outcomes.

Regarding the TOEFL result, appendix Table A2, which displays regression without any control except stratification variables, illustrates a positive but insignificant coefficient. Table 4.1 demonstrates that, overall, the difference between "winners" and "losers" was not so large, and was balanced in the whole sample. However, the appendix Figure A1 indicates a relatively large initial TOEFL score difference between "winners" and "losers"; this marks a failure at balancing

[7] Coefficient sizes are different between my estimations and the estimations of Kawata and Nishitani (2017) even considering the effect on English test scores in Japan. This could arise from the examination differences; specifically, I used TOEFL-ITP, and they used TOEIC to obtain outcomes.

Table 4.2: SSA participation effect on learning outcomes

	(1) ΔTOEFL	(2) GPA	(3) LSA
1stLotteryWin	8.089**	0.021	0.115***
	(3.744)	(0.021)	(0.032)
N	189	613	613

Standard errors are in parentheses. Standard errors are adjusted by White (1980)'s method. Other controls are dummies for female, international student, program year, entrance age greater than age 18, TOEFL-ITP score at first-year April, major, grade, and first-year GPA. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.3: SSA participation effect on the specific TOEFL score difference

	(1) Δ Listening	(2) Δ Grammar	(3) Δ Reading
1stLotteryWin	0.089	0.844	1.466**
	(0.470)	(0.702)	(0.592)
N	189	189	189

Standard errors are in parentheses. Standard errors are adjusted by White (1980)'s method. Other controls are dummies for female, international student, program year, entrance age greater than age 18, TOEFL-ITP score at first-year April, major, grade, and first-year GPA. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$. Sections scores of TOEFL-ITP are max 68 for listening and grammar and 67 for reading. Overall TOEFL-ITP score is 677, even though the total raw sections score is 203.

the sample at the stratification level sub-sample. Moreover, the appendix Figure A2 illustrates a simple linear regression fitted line between the initial TOEFL score and the increase in TOEFL score. Based on these facts, the results of Table 4.2, where students' characters are controlled for, including the initial TOEFL score, are more reliable than those displayed in Appendix Table A2 for my analysis. Therefore, I conclude that participation in SSA programs leads to a greater increase in TOEFL scores on average. Moreover, I also analyzed which English skills mainly contributed to improving the total score. Table 4.3 indicates the ATE2014 on the growth of each section of TOEFL-ITP scores for first-year students in 2014. The results demonstrate that SSA positively affects the reading section, whereas there are no specific positive effects on the grammar and listening sections. Therefore, Tables 4.2 and 4.3 indicate that the improvement of the overall score mainly comes from improving the reading score.

Table 4.4: SSA participation effect on employment probability

	(1) Working	(2) Private firm	(3) Market firm
1stLotteryWin	0.033	-0.026	-0.031
	(0.027)	(0.019)	(0.043)
N	585	502	478

Parentheses denote standard errors. Standard errors are adjusted by White (1980)'s method. Other controls are dummies for female, international student, program year, entrance age greater than age 18, TOEFL-ITP score at first-year April, major, grade, and first-year GPA. Statistical significance is as follows: * p <.1, ** p< .05, *** p < .01. The outcome variable "Working" indicates choosing a career as a worker, "Private firm" indicates choosing a career as non-public officers given a working career, and "Market firm" indicates choosing a career as a worker in a firm listed on Japanese stock market. We dropped some students from our sample in (1) since some students did not report their initial career path to the subject university (n=28). The regression in (2) focuses the sample on the case where the outcome in (1) takes 1 (drop n=83), and the regression in (3) focuses the analysis on the case where the outcome in (2) takes 1 (drop n=24).

4.5.2 Effect on the initial firm's attributes

Table 4.4 illustrates the AATE of SSA regarding each employment probability. Based on the results in Table 4.4, participating in SSA programs does not affect employment probability or whether graduates will choose a career in the public sector, private sector, or at a firm listed on the Japanese stock market. Therefore, I conclude that SSA programs have no power to change basic initial career choices at the level of employment probability. Table 4.4 presents the information acquired, that is, no difference between participants and non-participants, which could work as a balanced test in subsequent regression analysis for identifying the various attributes of firms where the graduates are employed.

Table 4.5 presents the differences in the attributes of firms at which graduates were initially employed and outlines the different trends between SSA program participants and non-participants. Column 1 in Table 4.5 indicates that participation in SSA programs has no power to boost monthly income. There is also no significant difference regarding capital, the number of workers, or established years and whether the firm is foreign-affiliated or not. Only the sales between participants and non-participants were significantly different at the 1% level, and the foreign stock rate was also significant at the 10% level. The results regarding the number of workers and established years might indicate that SSA has no power to increase working in venture capital companies since the

Table 4.5: SSA participation effect on the initial firm's attributes

	(1) log(Initial Income)	(2) log Sales) (Sales)	(3) log (Capital)	(4) log(Number of Workers)	(5) Build year	(6) Foreign firm	(7) Foreign stock rate
1stLotteryWin	0.004	0.552***	0.191	0.246	-3.319	0.008	0.028*
	(0.012)	(0.213)	(0.294)	(0.198)	(3.896)	(0.028)	(0.015)
N	308	294	332	313	388	450	307

Parentheses denote standard errors. Standard errors are adjusted by White (1980)'s method. Other controls are dummies for female, international student, program year, entrance age greater than age 18, TOEFL-ITP score at first-year April, major, grade, and first-year GPA. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$. The sample size depends on data availability of outcome variables. The sample includes non-listed firms in the Japanese stock market, whereas in the case of non-listed firms, the number of samples is limited due to not a few missing values.

number of workers and years of establishment tends to be small. As initial firms' attributes are only determined in the market, I can interpret that SSA program participation might have the power to increase employment probability from the point of view of relatively large sales and foreign stock rates, even though most initial firms' attributes were the same.

4.6 Discussion

4.6.1 Robustness check for treatment

The SSA programs and my analysis faced some problems with the randomness of assignments. To confirm the robustness of the results in Section 5, the issues of reapplication mechanisms and observational data attrition were considered. These issues were considered because of the potential sources of bias in the estimation.

Serious consideration of reapplication structure

By considering the presence of loser reapplication behavior as part of the program effect, I estimated AATE in Sections 4 and 5. Given that the two programs are homogeneous, the situation in which some of the losers of the initial allocation receive treatment is similar to the noncompliance problem. As only 4.3% of the sample is a winner in the reapplication, the effect in this regard is weak, but the earlier results may be subject to downward bias. Owing to the random assignment, the potential outcomes of not receiving the treatment of the "Winners" who reapplied could be considered substitutable for the "Losers" outcomes. Therefore, in this section, I reconfirm the AATE effect by removing "Winners" from the sample and increasing the weight of the "Losers" sample by that amount to fill the difference for the initial placement discussed in Section 5. As illustrated in Appendix Table A3, accounting for the reapplication structure did not significantly change the results.

Serious consideration of data attrition

In the estimations of the employment attributes in Tables 4.4 and 4.5, the sample size was limited due to data access constraints on the outcome variables. Even if the university authorities' initial SSA assignment was random, the properties of the random assignment and balancing of the distribution might have been broken. If the data attrition method creates more missing outcome variables for the losers, this could be a biased source. Moreover, data attrition may shift the ratio of winners to losers from the random assignment made by the university authorities, and alter the standard errors. Subsequently, I checked whether the data attrition patterns differ between winners and losers. I created a dummy variable that takes 1 if the outcome variable was observable, and 0 if it was not and performed the same estimation as in Section 5. Any difference in data attrition was likely to be statistically significant. Appendix Table A4 demonstrates no significant difference in the amount of data attrition for each allocation result. The sample proportion discrepancies between random assignment and data attrition were corrected by adjusting the weights and reanalyzing. As illustrated in Appendix Table A5, the results did not change significantly after adjusting for the random assignment rate of winning.

4.6.2 How are differences in firms' initial attributes generated?

Here, I examined how the difference in mean between the participants' and non-participants' sales and foreign stock rates at initial firms was generated. I verified whose employment probability

Table 4.6: Conditional quantile regression coefficients of 1stlottery win for row outcomes

Outcome \Quantile	Q.10	Q.20	Q.30	Q.40	Q.50	Q.60	Q.70	Q.80	Q.90
log(Sales)	1.560***	0.930	0.648	0.463	0.178	0.063	0.171	0.204	0.422**
	(0.548)	(0.601)	(0.427)	(0.285)	(0.239)	(0.166)	(0.118)	(0.136)	(0.189)
N	294	294	294	294	294	294	294	294	294
Foreign stock rate	0.060*	0.038*	0.026	0.022	0.018	0.014	0.030**	0.034*	0.042
	(0.035)	(0.022)	(0.018)	(0.020)	(0.020)	(0.016)	(0.015)	(0.020)	(0.030)
N	307	307	307	307	307	307	307	307	307

Parentheses are standard errors. Other controls are dummies for female, international student, program year, entrance age greater than age 18, TOEFL-ITP score at first-year April, major, grade, and first-year GPA. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$. Each column corresponds to a separate regression coefficient of row outcomes on 1stLottery Win.

was changed by participation in SSA programs. Figure A3 illustrates the distributions of log sales and foreign stock rates for participants and non-participants. Figure A3 suggests the possibility of a difference in employment probability between the firms with top and bottom sales or foreign stock rates. To confirm this case, I ran a conditional quantile regression with the assumption of rank preservation for firms' characters. My method shares the same spirit as Cullinan et al. (2022) quantile analysis to consider heterogeneity. Table 4.6 confirms whether the tendency appeared in Figure A3 by each 10% conditional quantile regression for sales and foreign stock rate. Participants had higher sales in the ten percentiles with a 1% significance level, in the 90 percentiles with a 5% significance level, and the foreign stock rate in the 10 and 20 percentiles with a 10% significance level, 70 percentiles with a 5% significance level, and 80 percentiles with a 10% significance level.

I could infer that SSA participation does not necessarily increase all student's employment probability in firms with higher sales and foreign stock rates, but enhances the probability of obtaining employment at such firms for students who tend to work at firms with relatively larger or lower quantile in terms of sales and foreign stock rate.

4.6.3 Potential explanations for how SSA could change firms' initial attributes

Tables 4.5 and 4.6 suggest that the differences in the initial firms' attributes between the participants and non-participants were derived from the differences in employment probability from relatively larger or smaller firms. However, it is unclear why employment probability changes due to participation in SSA programs. There are two possible explanations for the mechanism based on labor demand- and supply-side preferences. From the firm's perspective, if students' preferences are fixed, and firms' recruiting behavior is conducted in the order of the largest sales, Tables 4.5 and 4.6 imply that firms prioritize employing SSA participants because they tend to have higher ability than the non-participants, based on human capital accumulation in SSA programs. From the student's perspective, if firms' employment strategies are random, Tables 4.5 and 4.6 imply that SSA participants change their attitude about job hunting and apply to firms with high sales more than non-participants. Given data limitations, I could not analyze the details of mechanisms based on my results and dataset. Petzold (2017) found that in Germany, including SA experience on a curriculum vitae for an internship application leads to quicker response rates and slightly increases the probability of a successful application. Petzold (2017) suggested employers' preferences might explain the results illustrated in Tables 4.5 and 4.6 if Japanese employers' tendencies are the same as German employers. I must rely on future studies to identify which explanation is more applicable to why the probability of obtaining employment at relatively large firms increases, and that for relatively small firms decreases.

4.7 Conclusion

This chapter considers whether differences in the content of higher education change educational returns, using SSA as an example. In addition, this chapter investigated the impact of SSA programs on learning outcomes and firms' initial attributes after graduation. To control for self-selection bias, I used data from a Japanese university to randomly assign participation in SSA programs among the applicants for the 2013 and 2014 academic years. My results demonstrate that participation in SSA programs increases English-related learning outcomes and the probability of obtaining employment at financially large (in terms of sales and foreign stock rates) firms compared with non-participants. Accordingly, these results were robust after accounting for various potential sources of bias. The average difference in the initial firm's attributes might arise because participation in the SSA programs increased the applicant's probability of employment at a relatively large firm and decreased the applicant's probability of employment at a relatively small firm.

Even with a global increase in the number of SSA programs, there is limited research on the causal effects of SSA programs. The major contribution of this chapter is the derivation of an approximation of the causal effects of SSA. I derived the AATE of the initial assignment effect using the random assignment and LATE for participation in SA programs more than once. Moreover, I provide causal inferences about the effects of participation in SSA. I focus on the individual labor market outcomes through human capital accumulation and learning outcomes as the effect of SSA, while previous studies have not yet investigated this aspect of SSA. Additionally, I demonstrate causal evidence for participation in SSA in Asia, whereas most previous studies were either in European or North American countries.

However, I faced some limitations. The first is external validity. The university examined in this chapter is a research-oriented, highly selective university in Japan. To capture general results in Japan, I would need to use samples from a diverse sample of universities. A second limitation is that my estimated effects were not pure ATE of "only once participation in SSA." There was participation in the other SA programs, and excluding all the other effects of randomized SSA programs was difficult. The third limitation is the skepticism of randomness. Although I demonstrated in Section 6.1 that the results in Section 5 are robust and not significantly different after accounting for these issues, I cannot completely deny the design problems with the AATE estimation. Additionally, I could not have explained why SSA programs change the attributes of initially employed firms. I offered two possible explanations for firms' and students' preferences in Section 6.3. Identifying the mechanisms behind the effects of SSA program participation on initial career characteristics must be a focus of future research.

While these limitations are worth bearing in mind, my results are still meaningful and demonstrate that SSA program participation increases some learning outcomes and influences firms' initial attributes. While my sample university was highly selective and lacked representativeness, my AATE was close to the lower bounds results, and the AATE of other universities in Japan might be higher than mine. Although there is, in this chapter, a higher possibility of under-reporting than if my sample was from an average university, some of my results are still significant.

Thus, these results are informative for higher education as well as international education professionals and contribute to my understanding of the effect of SSA programs and identifying problems related to reapplication. These findings might also be beneficial to students considering applying to SSA programs.

From the perspective of human capital accumulation, this chapter suggests that educational content within higher education has an effect on some specific human capital accumulation (competency) and that the returns depend on whether it is valued in the market. While this chapter has illustrated the effect with the example of SSA, It is important for future research to identify what the other educational subjects and faculty education encourage human capital accumulation and what kind of returns each one has.

Bibliography

Acemoglu, Daron and Joshua Angrist (2000) "How large are human-capital externalities? Evidence from compulsory schooling laws," *NBER macroeconomics annual*, 15, 9–59.

Albrecht, James and Susan Vroman (2002) "A Matching Model with Endogenous Skill Requirements," *International Economic Review*, 43 (1), 283–305, 10.1111/1468-2354.t01-1-00012.

Altonji, Joseph G, Lisa B Kahn, and Jamin D Speer (2016) "Cashier or consultant? Entry labor market conditions, field of study, and career success," *Journal of Labor Economics*, 34 (S1), S361–S401.

Angrist, Joshua D. and Alan B. Krueger (1991) "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics*, 106 (4), 979–1014, `http://www.jstor.org/stable/2937954`.

Arkes, Jeremy (2010) "Using Unemployment Rates as Instruments to Estimate Returns to Schooling," *Southern Economic Journal*, 76 (3), 711–722, https://doi.org/10.4284/sej.2010.76.3.711.

Ashenfelter, Orley and Alan Krueger (1994) "Estimates of the Economic Return to Schooling from a New Sample of Twins," *The American Economic Review*, 84 (5), 1157–1173, `http://www.jstor.org/stable/2117766`.

Ashenfelter, Orley and David J. Zimmerman (1997) "Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons, and Brothers," *The Review of Economics and Statistics*, 79 (1), 1–9, `http://www.jstor.org/stable/2951427`.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2006) "The Polarization of the U.S. Labor Market," *American Economic Review*, 96 (2), 189–194, 10.1257/000282806777212620.

——— (2008) "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 90 (2), 300–323, 10.1162/rest.90.2.300.

Autor, David H, Frank Levy, and Richard J Murnane (2003) "The skill content of recent technological change: An empirical exploration," *The Quarterly journal of economics*, 118 (4), 1279–1333.

Beiler, Hendrik (2017) "Do you dare? The effect of economic conditions on entrepreneurship among college graduates," *Labour Economics*, 47, 64–74, https://doi.org/10.1016/j.labeco.2017.05.003, EALE conference issue 2016.

van den Berge, Wiljan (2018) "Bad start, bad match? The early career effects of graduating in a recession for vocational and academic graduates," *Labour Economics*, 53, 75 – 96, https://doi.org/10.1016/j.labeco.2018.05.011.

Blackburn, McKinley L. and David Neumark (1995) "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look," *The Review of Economics and Statistics*, 77 (2), 217–230, http://www.jstor.org/stable/2109861.

Brunner, Beatrice and Andreas Kuhn (2014) "The impact of labor market entry conditions on initial job assignment and wages," *Journal of Population Economics*, 27 (3), 705–738.

Cameron, Stephen V. and James J. Heckman (1998) "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106 (2), 262–333, 10.1086/250010.

Cameron, Stephen V. and Christopher Taber (2004) "Estimation of Educational Borrowing Constraints Using Returns to Schooling," *Journal of Political Economy*, 112 (1), 132–182, 10.1086/379937.

Card, David (2001) "Estimating the return to schooling: Progress on some persistent econometric problems," *Econometrica*, 69 (5), 1127–1160.

Card, David and Alan B Krueger (1992) "Does school quality matter? Returns to education and the characteristics of public schools in the United States," *Journal of political Economy*, 100 (1), 1–40.

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil (2011) "Estimating marginal returns to education," *American Economic Review*, 101 (6), 2754–81.

Carneiro, Pedro and Sokbae Lee (2009) "Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality," *Journal of Econometrics*, 149 (2), 191–208, https://doi.org/10.1016/j.jeconom.2009.01.011.

Carneiro, Pedro, Costas Meghir, and Matthias Parey (2013) "Maternal Education, Home Environments, and the Development of Children and Adolescents," *Journal of the European Economic Association*, 11 (1), 123–160, 10.1111/j.1542-4774.2012.01096.x.

Charlot, Olivier and Bruno Decreuse (2005) "Self-selection in education with matching frictions," *Labour Economics*, 12 (2), 251–267, 10.1016/j.labeco.2003.11.004.

——— (2010) "Over-education for the rich, under-education for the poor: A search-theoretic microfoundation," *Labour Economics*, 17 (6), 886 – 896, 10.1016/j.labeco.2010.05.004.

Cortes, Guido Matias (2016) "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data," *Journal of Labor Economics*, 34 (1), 63–105, 10.1086/682289.

Cullinan, John, Darragh Flannery, and Donal Palcic (2022) "Study abroad programme participation and subsequent academic performance: evidence from administrative data," *Education Economics*, 30 (3), 251–269, 10.1080/09645292.2021.1978936.

Cunha, Flavio, James J Heckman, Lance Lochner, and Dimitriy V Masterov (2006) "Interpreting the evidence on life cycle skill formation," *Handbook of the Economics of Education*, 1, 697–812.

Currie, Janet and Enrico Moretti (2003) "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings," *The Quarterly Journal of Economics*, 118 (4), 1495–1532, http://www.jstor.org/stable/25053945.

Cutler, David M, Wei Huang, and Adriana Lleras-Muney (2015) "When does education matter? The protective effect of education for cohorts graduating in bad times," *Social Science & Medicine*, 127, 63–73.

David, Benjamin (2017) "Computer technology and probable job destructions in Japan: An evaluation," *Journal of the Japanese and International Economies*, 43, 77–87, https://doi.org/10.

1016/j.jjie.2017.01.001.

De Poli, Silvia, Loris Vergolini, and Nadir Zanini (2018) "The impact of a study abroad programme on learning abilities and personality traits: evidence from a randomization," *Applied Economics Letters*, 25 (8), 562–566, 10.1080/13504851.2017.1346354.

Di Pietro, Giorgio (2012) "Does studying abroad cause international labor mobility? Evidence from Italy," *Economics Letters*, 117 (3), 632–635, https://doi.org/10.1016/j.econlet.2012.08.007.

——— (2015) "Do study abroad programs enhance the employability of graduates?" *Education Finance and policy*, 10 (2), 223–243.

Diamond, Peter A (1982) "Aggregate demand management in search equilibrium," *Journal of political Economy*, 90 (5), 881–894.

Dolado, Juan J., Marcel Jansen, and Juan F. Jimeno (2009) "On-the-Job Search in a Matching Model with Heterogeneous Jobs and Workers*," *The Economic Journal*, 119 (534), 200–228, 10.1111/j.1468-0297.2008.02210.x.

Doyle, William R. and Benjamin T. Skinner (2016) "Estimating the education-earnings equation using geographic variation," *Economics of Education Review*, 53, 254–267, https://doi.org/10.1016/j.econedurev.2016.03.010.

of Education Culture Sports Science, Ministry and (MEXT) Technology Japan (2022) "Survey of International Student Enrollment and Number of Japanese Students Studying Abroad [Translated from Japanese.] Accessed 13 September 2022.,"Technical report, `https://www.mext.go.jp/content/20220603-mxt_gakushi02-100001342_2.pdf`.

Flinn, Christopher and Joseph Mullins (2015) "LABOR MARKET SEARCH AND SCHOOLING INVESTMENT," *International Economic Review*, 56 (2), 359–398, https://doi.org/10.1111/iere.12108.

Furukawa, Yosuke and Hiroki Toyoda (2018) "The recent process of job polarization in Japan: evidence from 1980 to 2010," *Applied Economics Letters*, 25 (7), 456–460, 10.1080/13504851.2017.1332741.

Gautier, Pieter A. (2002) "Unemployment and Search Externalities in a Model with Heterogeneous

Jobs and Workers," *Economica*, 69 (273), 21–40, 10.1111/1468-0335.00269.

Genda, Yuji, Ayako Kondo, and Souichi Ohta (2010) "Long-term effects of a recession at labor market entry in Japan and the United States," *Journal of Human resources*, 45 (1), 157–196.

Goos, Maarten, Alan Manning, and Anna Salomons (2009) "Job Polarization in Europe," *American Economic Review*, 99 (2), 58–63, 10.1257/aer.99.2.58.

Grenet, Julien (2013) "Is extending compulsory schooling alone enough to raise earnings? Evidence from French and British compulsory schooling laws," *The Scandinavian Journal of Economics*, 115 (1), 176–210.

Griliches, Zvi (1977) "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45 (1), 1–22, http://www.jstor.org/stable/1913285.

Harmon, Colm and Ian Walker (1995) "Estimates of the economic return to schooling for the United Kingdom," *The American Economic Review*, 85 (5), 1278–1286.

Hartog, Joop (2000) "Over-education and earnings: where are we, where should we go?" *Economics of Education Review*, 19 (2), 131 – 147, https://doi.org/10.1016/S0272-7757(99)00050-3.

Hosios, Arthur J. (1990) "On The Efficiency of Matching and Related Models of Search and Unemployment," *The Review of Economic Studies*, 57 (2), 279–298, 10.2307/2297382.

Ikenaga, Toshie (2009) "Polarization of the Japanese Labor Market-Adoption of ICT and changes in tasks required," *Japanese Journal of Labour Studies*, 584, 73–90.

Ikenaga, Toshie and Ryo Kambayashi (2016) "Task Polarization in the Japanese Labor Market: Evidence of a Long-Term Trend," *Industrial Relations: A Journal of Economy and Society*, 55 (2), 267–293.

Imbens, Guido W and Donald B Rubin (2015) *Causal inference in statistics, social, and biomedical sciences*: Cambridge University Press.

Imbens, Guido W and Jeffrey M Wooldridge (2009) "Recent developments in the econometrics of program evaluation," *Journal of economic literature*, 47 (1), 5–86.

Jaimovich, Nir and Henry E. Siu (2020) "Job Polarization and Jobless Recoveries," *The Review of*

Economics and Statistics, 102 (1), 129–147, 10.1162/rest_a_00875.

Kahn, Lisa B. (2010) "The long-term labor market consequences of graduating from college in a bad economy," *Labour Economics*, 17 (2), 303 – 316, https://doi.org/10.1016/j.labeco.2009.09.002.

Kamhöfer, Daniel A. and Hendrik Schmitz (2016) "Reanalyzing Zero Returns to Education in Germany," *Journal of Applied Econometrics*, 31 (5), 912–919, 10.1002/jae.2461, jae.2461.

Kane, Thomas J and Cecilia E Rouse (1993) "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?" Working Paper 4268, National Bureau of Economic Research, 10.3386/w4268.

Kato, Maki and Ken Suzuki (2019) "Effective or Self-Selective: Random Assignment Demonstrates Short-Term Study Abroad Effectively Encourages Further Study Abroad," *Journal of Studies in International Education*, 23 (4), 411–428, 10.1177/1028315318803713.

Kawaguchi, Daiji and Ayako Kondo (2015) "The Effects of Graduating from College during a Recession on Consumption and Asset Holding," *Research Institute of Economy, Trade and Industry (RIETI)*.

Kawata, Keisuke and Hajime Nishitani (2017) "Impact of short-term study abroad programs on language achievement: Evidence from the case of Hiroshima University START Program [Translated from Japanese.]," May, Accssed `https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxrZWlzdWtla2F3YXRhN3xneDoxOTQwNDhhNDNkZGIwZWNm`).

Kikuchi, Nobuyoshi (2017) "Marginal returns to schooling and education policy change in Japan," ISER Discussion Paper 0996r, Institute of Social and Economic Research, Osaka University.

Kleibergen, Frank and Richard Paap (2006) "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 133 (1), 97–126, https://doi.org/10.1016/j.jeconom.2005.02.011.

Kling, Jeffrey R (2001) "Interpreting Instrumental Variables Estimates of the Returns to Schooling," *Journal of Business & Economic Statistics*, 19 (3), 358–364, 10.1198/073500101681020006.

Leuven, Edwin and Hessel Oosterbeek (2011) *Overeducation and Mismatch in the Labor*

Market, 4 of Handbook of the Economics of Education, Chap. 3, 283–326: Elsevier, 10.1016/B978-0-444-53444-.

McGuinness, Seamus (2006) “Overeducation in the Labour Market.,” *Journal of Economic Surveys*, 20 (3), 387 – 418, `http://search.ebscohost.com/login.aspx?direct=true&db=eoh&AN=0867306&lang=ja&site=ehost-live`.

MHLW (2018) “Analysis of the Labor Economics in 2008 - Human Resource Development in Response to Diversification of Work Styles[Translated from Japanese.],” `https://www.mhlw.go.jp/wp/hakusyo/roudou/18/dl/18-1.pdf`.

Michaels, Guy, Ashwini Natraj, and John Van Reenen (2014) “Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years,” *Review of Economics and Statistics*, 96 (1), 60–77.

Mortensen, Dale T (1986) “Job search and labor market analysis,” *Handbook of labor economics*, 2, 849–919.

Mortensen, Dale T and Christopher A Pissarides (1994) “Job creation and job destruction in the theory of unemployment,” *The review of economic studies*, 61 (3), 397–415.

Mortensen, Dale T. and Christopher A. Pissarides (1999) “Chapter 39 New developments in models of search in the labor market,” 3 of Handbook of Labor Economics, 2567 – 2627: Elsevier, https://doi.org/10.1016/S1573-4463(99)30025-0.

Nocito, Samuel (2021) “The effect of a university degree in english on international labor mobility,” *Labour Economics*, 68, 101943, https://doi.org/10.1016/j.labeco.2020.101943.

Oosterbeek, Hessel and Dinand Webbink (2011) “Does Studying Abroad Induce a Brain Drain?” *Economica*, 78 (310), 347–366, https://doi.org/10.1111/j.1468-0335.2009.00818.x.

Oreopoulos, Philip (2005) “Canadian Compulsory School Laws and their Impact on Educational,” *Analytical Studies Branch Research Paper Series, Statistics Canada, catalogue* (11F0019MIE).

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz (2012) “The Short- and Long-Term Career Effects of Graduating in a Recession,” *American Economic Journal: Applied Economics*, 4 (1), 1–29, 10.1257/app.4.1.1.

Oyer, Paul (2006) "Initial labor market conditions and long-term outcomes for economists," *Journal of Economic Perspectives*, 20 (3), 143–160.

Parey, Matthias and Fabian Waldinger (2010) "Studying Abroad and the Effect on International Labour Market Mobility: Evidence from the Introduction of ERASMUS," *The Economic Journal*, 121 (551), 194–222, 10.1111/j.1468-0297.2010.02369.x.

Petrongolo, Barbara and Christopher A. Pissarides (2000) "Looking Into the Black Box: A Survey of the Matching Function," CEP Discussion Papers dp0470, Centre for Economic Performance, LSE, `https://ideas.repec.org/p/cep/cepdps/dp0470.html`.

Petzold, Knut (2017) "Studying abroad as a sorting criterion in the recruitment process: a field experiment among German employers," *Journal of Studies in International Education*, 21 (5), 412–430.

Pischke, Jörn-Steffen and Till von Wachter (2005) "Zero Returns to Compulsory Schooling In Germany: Evidence and Interpretation," Working Paper 11414, National Bureau of Economic Research, 10.3386/w11414.

——— (2008) "Zero returns to compulsory schooling in Germany: Evidence and interpretation," *The Review of Economics and Statistics*, 90 (3), 592–598.

Rubin, Donald B (1974) "Estimating causal effects of treatments in randomized and nonrandomized studies.," *Journal of Educational Psychology*, 66 (5), 688–701.

Schwandt, Hannes and Till von Wachter (2019) "Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets," *Journal of Labor Economics*, 37 (S1), S161–S198, 10.1086/701046.

Sorrenti, Giuseppe (2017) "The Spanish or the German apartment? Study abroad and the acquisition of permanent skills," *Economics of Education Review*, 60, 142–158, https://doi.org/10.1016/j.econedurev.2017.07.001.

THE, (Times Higher Education) (2021) "THE World University Rankings Japan Edition,"Technical report, `https://japanuniversityrankings.jp/rankings/`.

Van Breukelen, Gerard JP and Koene RA Van Dijk (2007) "Use of covariates in randomized controlled trials," *Journal of the International Neuropsychological Society*, 13 (5), 903–904.

Verdugo, Richard R. and Naomi Turner Verdugo (1989) "The Impact of Surplus Schooling on Earnings: Some Additional Findings," *Journal of Human Resources*, 24 (4), 629–643, https://EconPapers.repec.org/RePEc:uwp:jhriss:v:24:y:1989:i:4:p:629-643.

Waibel, Stine, Knut Petzold, and Heiko Rüger (2018) "Occupational status benefits of studying abroad and the role of occupational specificity-A propensity score matching approach," *Social Science Research*, 74, 45–61.

White, Halbert (1980) "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica*, 48 (4), 817–838, http://www.jstor.org/stable/1912934.

Wooldridge, Jeffrey M. (2003) "Further results on instrumental variables estimation of average treatment effects in the correlated random coefficient model," *Economics Letters*, 79 (2), 185 – 191, https://doi.org/10.1016/S0165-1765(02)00318-X.

Appendix A

Table A1: Program detail for each year

School year	2013	2104
Fees paid by students	0	100,000 yen
Number of sending schools	9	13
Cumulative total number of applicants	269	437
Cumulative total number of lottery “winners”	100	200
Odds of winning	2.68	2.19

Notes: 100 yen was around $1 (US) in 2020. “Winners” are the study abroad participants among the applicants.

Table A2: AATE without controlling predeterminant characters

	(1) ΔTOEFL	(2) GPA	(3) Long	(4) Working	(5) Private	(6) Market	(7) log(Initial Income)
1stLotteryWin	4.291	0.060	0.130***	0.032	-0.024	-0.026	0.007
	(4.176)	(0.040)	(0.035)	(0.028)	(0.019)	(0.043)	(0.011)
N	189	613	613	585	502	478	308
Significance difference	Yes	No	No	No	No	No	No

	(8) log(Sales)	(9) log(Capital)	(10) log(Number of Workers)	(11) Build year	(12) Foreign firm	(13) Foreign stock rate
1stLotteryWin	0.568***	0.206	0.275	-3.708	0.005	0.030**
	(0.209)	(0.291)	(0.195)	(3.854)	(0.028)	(0.015)
N	294	332	313	388	450	307
Significance difference	No	No	No	No	No	No

Parentheses denote standard errors. Standard errors are adjusted by White (1980)’s method. Other controls are dummies for major, grade, and initial application year. Statistical significance is as follows: * $p < .1$, ** $p < .05$, *** $p < .01$. Significance difference is Yes if the statistical significances of without control (only control stratifications) and with control regression are different.

Table A3: AATE considering reapplication structure

	(1) GPA	(2) Long	(3) Working	(4) Private	(5) Market	(6) log(Initial Income)
1stLotteryWin	0.022	0.120***	0.030	-0.024	-0.037	0.005
	(0.022)	(0.033)	(0.028)	(0.020)	(0.044)	(0.013)
N	585	585	559	481	457	298
	(7) log(Sales)	(8) log(Capital)	(9) log(Number of Workers)	(10) Build year	(11) Foreign firm	(12) Foreign stock rate
1stLotteryWin	0.534**	0.214	0.268	-2.741	0.007	0.027*
	(0.216)	(0.299)	(0.198)	(4.049)	(0.029)	(0.015)
N	283	320	301	375	434	297

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method. Other controls are dummies for female, international student, program year, entrance age greater than age 18, TOEFL-ITP score at first-year April, major, grade, and first-year GPA. Statistical significance is as follows:* $p < .1$, ** $p < .05$, *** $p < .01$. In the current estimation, to account for the behavior of reapplying to the random assignment programs, we set the weight of reapplied "Winners" (28 applicants) to zero (excluded from the sample). The weight of the "losers" (26 applicants) among reapplicants was adjusted ((28+26)/26 times). We omit TOEFL results since the TOEFL outcome does not have a reapplication problem.

Table A4: Randomization check for data attrition problem

	(1)	(2)	(3)	(4)	(5)
Regression Outcome	Indicator variables of non-missing data with respect to columns				
Data attrition variables	Working	Private	Market	log(Initial Income)	log(Sales)
1stLotteryWin	-0.019	-0.052	-0.026	-0.036	0.052
	(0.017)	(0.031)	(0.034)	(0.041)	(0.041)
N	613	613	613	613	613
	(6) log(Capital)	(7) log(Number of Workers)	(8) Build year	(9) Foreign firm	(10) Foreign stock rate
1stLotteryWin	0.038	0.012	0.002	-0.051	0.021
	(0.041)	(0.041)	(0.039)	(0.036)	(0.041)
N	613	613	613	613	613

Parentheses are standard errors, which are adjusted by White (1980)'s method. Other controls are dummies for female, international student, program year, entrance age greater than age 18, TOEFL-ITP score at first-year April, major, grade, and first-year GPA. Statistical significance is as follows:* $p < .1$, ** $p < .05$, *** $p < .01$. The coefficients in this table are calculated by the regression adjusted by the sample randomization rate ('# of winners in sample') as weight.

Table A5: AATE with weight to adjust as randomization “winners” and “losers” rate

	(1) Working	(2) Private firm	(3) Market firm	(4) log(Initial Income)	(5) log(Sales)
1stLotteryWin	0.033	-0.026	-0.031	0.004	0.553***
	(0.027)	(0.019)	(0.043)	(0.012)	(0.213)
N	585	502	478	308	294

	(6) log(Capital)	(7) log(Number of Workers)	(8) Build year	(9) Foreign firm	(10) Foreign stock rate
1stLotteryWin	0.191	0.246	-3.314	0.008	0.028*
	(0.294)	(0.198)	(3.896)	(0.028)	(0.015)
N	332	313	388	450	307

Parentheses are standard errors. Standard errors are adjusted by White (1980)’s method. Other controls are dummies for female, international student, program year, entrance age greater than age 18, TOEFL-ITP score at first-year April, major, grade, and first-year GPA. Statistical significance is as follows:* $p < .1$, ** $p < .05$, *** $p < .01$. This estimation was adjusted by weight to ensure that the “win” ratio within the analysis sample is consistent with the “win” ratio of the random assignment of the overall applicants, as the number of samples used varies depending on the outcome variable.