Short-term Study Abroad Effect on Within-School Outcomes and Initial Career: Evidence from Random Assignment Data from a Japanese University *

Kashima, Ryohei † Maki Kato ‡

December 24, 2020

Abstract

This paper examines the effect of short-term study abroad programs on within-school outcomes (English skills, study achievement, and participation in the long-term study abroad program) and the initial employee firm's attributes of participants after graduating from university. We analyzed the effect of participation in short-term study abroad programs from randomly assigned data of applicants from a Japanese university in March and August 2014 (705 applicants leading to 300 participants). We also suggest a modification strategy for existing reapplications, which many researchers or program coordinators potentially face. We found that participation in short-term study abroad programs causally increases English test scores and long-term study abroad programs' participation rate. For the initial firms' attributes, we found that participants work for firms with a significantly higher sales and foreign stock rate than non-participants on average, although we did not find significant differences in initial monthly incomes. We provide empirical evidence to demonstrate that the average application gap to initial firms between participants and non-participants might come from the differences in employment probability from firms with top and bottom sales and foreign stock rate.

Keywords: Random Assignment, Short-term study abroad, Initial career, Higher education

*We would like to express my gratitude to Ryo Kambayashi, Masayuki Sawada, Kim Jong-Kun, Yohei Yamamoto and participants at the Hitotsubashi lunch seminar for helpful comments. Any remaining errors in this paper are entirely attributable to the authors.

†Mori Arinori Institute for Higher Education and Global Mobility, Hitotsubashi University, E-Mail: rkashi1228@gmail.com

‡International Centre, Gakushuin University

1 Introduction

Recently, the number of university students who participate in study abroad programs has increased in higher-income countries. However, despite this increase, the effect of studying abroad on students remains unclear (Savicki et al., 2015). Studying abroad has four primary objectives: improved language and communication skills, increased career selection opportunities, personal growth, and increased cross-cultural understanding (Anderson et al., 2006). With an increase in the investigation of the economic returns of higher education and the amount of related publicly available data, the careers of students who have participated in study abroad programs after graduation have attracted the attention of many researchers (Waibel et al., 2017).

Interestingly, the increase in study abroad participants is largely from short-term study abroad (hereafter referenced as SSA) programs. For example, as illustrated in Figure 1, the breakdown of the number of Japanese students and students from the United States studying abroad demonstrated a significant increase in SSA programs compared to long-term programs in both countries. In this figure, short-term is defined as a period of less than three months and long-term is defined as a period of three months or more. Lower costs, a wide variety of destinations and learning content, and the fact that the school period does not need to be extended are some of the benefits of SSA programs over long-term programs; and these are some of the reasons for the increase in the popularity of SSA programs (Kurt et al., 2013).

With the increase in the number of SSA participants, the effect of these programs has recently become a new area of academic interest (Campbell, 2016). However, empirical analyses of the effect of SSA programs are limited because of data limitations, such as the small number of participants, which is sometimes only thirty participants per program (Mapp, 2012; Kurt et al., 2013; Carley and Tudor, 2010). Although research on the impact of study abroad programs on careers has increased (Waibel et al., 2017; Rodrigues, 2012, see), it has mostly focused on long-term study abroad programs and not SSA programs. Adding

to the data limitations, research on SSA programs is limited by the programs' period. SSA programs are usually part of the standard university education period of three to four years, and so their effect has been questioned (Campbell, 2016; Gaia, 2015; Geyer et al., 2017). For instance, Kawata and Nishitani (2017) concluded that SSA programs supported by a Japanese university improved participants' English skills, even though their targeted SSA programs lasted only two weeks. Although this two-week SSA program provided evidence of statistically significant English proficiency improvements, it is unlikely that the effect was gained only through the two-week course. Instead, the effect of SSA programs is assumed to motivate language learning rather than improving language proficiency (Bodycott and Crew, 2000). Like the motivation towards language learning, study abroad programs may create new career perspectives among the participants after they graduate. Specifically, Walters et al. (2017) re-evaluated the changes in participants' beliefs and values after participating in SSA programs, which includes further study, leading to an international career after graduation.

The control of the selection bias and explanation of how study abroad programs affect careers are two major issues in the previous research on the effects of studying abroad. If the characteristics of participants of study abroad programs are different, it may not be clear whether the effect of these programs is caused by the characteristics of the participants or their participation in these programs. For instance, if the participants' wages after graduation are higher than non-participants, it is not easy to determine if this is because of their participation in a study abroad program or if their characteristics are positively evaluated in the labor market. If the selection bias is not considered, the effect of studying abroad will be overestimated. This is true for all of the effects of study abroad programs, including the improvement of language skills.

Although most previous studies face selection problems, and the results do not address the causal effects of study abroad programs, some recent studies have attempted to see causal effect. For instance, Di Pietro (2015) analyzed the effect on employability using the

instrumental variable method, while Waibel et al. (2018) used a propensity score matching method for the effect on occupational prestige by aligning the covariates. Kawata and Nishitani (2017) analyzed the ability of SSA programs on improving English test scores by estimating differences in differences (DID) using covariates based on propensity scores. However, covariates between comparison groups may not be aligned due to various constraints.

Some studies have investigated the effect on wages and concluded that international mobility and employment at large multinational companies are determinants (Liwiński, 2019b; Kratz and Netz, 2018). However, the factors involved could include individual growth obtained through the study abroad programs. One factor is foreign language proficiency, as noted by Di Pietro (2019), because it is possible to improve or become motivated to improve foreign language proficiency through study abroad programs (Sorrenti, 2017). Also, as demonstrated by Kato and Suzuki (2019), participation in SSA programs encourages long-term study abroad. Subsequently, the impact of study abroad programs on careers could be caused by both types of study abroad programs. However, as far as the authors are aware, no previous studies have analyzed the effect of SSA on career.

In response to the above issues, this study investigated the effect of SSA programs on initial-job attributes, English proficiency, and participation in long-term study abroad programs. We used randomly assigned data for participants of SSA programs lasting one month, implemented by a Japanese research university. Random assignment allows the treatment effect to be elucidated without selection bias. Subsequently, it is possible to clarify the impact of SSA programs by examining the differences between participants' post-graduate careers or specific skills before and after studying abroad. As improving language skills is often a key objective of study abroad programs, our analysis results will provide methods for improving these programs. We also treat reapplication issues which Kato and Suzuki (2019) did not include into their analysis. Consequently, the research questions are as follows:

1. Do SSA programs improve within-school outcomes such as graduation year cumulative GPA, long-term study abroad participation, and English proficiency?

2. Do SSA programs affect initial job attributes such as monthly income, corporate size based on sales or number of employees, and international relevance?

This paper proceeds as follows. Chapter 2 provides an overview of the findings and challenges of previous research. Chapter 3 explains the data, and Chapter 4 describes the methods. Chapter 5 illustrates the results, Chapter 6 discuss interpretation of the results, and the summarizes the results and conclusions.

2 Literature review and Study abroad programs

2.1 Literature review

2.1.1 Effect of study abroad programs on careers

Waibel et al. (2017) has categorized the effects of studying abroad on careers into employability, career selection, and wages or positions. As the university analyzed in this paper is a highly selective Japanese research university, and graduates have a high probability of becoming employed, this paper focuses on job attributes. Previous studies in Europe demonstrated that studying abroad increases the probability of working at a company based in a foreign country or having an internationally related job (for example, in Italy, Di Pietro (2012); in Germany, Parey and Waldinger (2010); and in the Netherlands, Oosterbeek and Webbink (2011)). A large-scale survey in Japan also demonstrated that students with an undergraduate degree from foreign universities tend to more work at foreign-affiliated companies compared to those who have shorter study abroad experiences (Shimmi et al., 2017). However, the impact of studying abroad on wages after graduation is unclear. For instance, Waibel et al. (2017) demonstrated the effect through a comprehensive review of related research on Europe and the United States. In Japan, Yokota et al. (2018) also demonstrated the effect, although the study does not control for selection bias, such as the chosen major and selection of university, which affects the wages of graduates (Kirkeboen et al.,

2016). Meanwhile, Rodrigues (2012) concluded that studying abroad effects international work content rather than careers, such as wages or employability. Recent analyses have also demonstrated that studying abroad does not affect university graduates' wages in the United States (Schmidt and Pardo, 2017). Country differences have also been reported; for example, there is little effect in Western Europe, but increased effects in Eastern or Central Europe (Rodrigues, 2013).

2.1.2 Self-selection bias

Although studying abroad has become popularized, students who participate or do not participate in these programs vary in many aspects. In particular, the socioeconomic background and other attributes of the students might correlate to participation decisions since Socioeconomic background affects students' international travel experiences (Wiers-Jenssen, 2011) and the characteristics of the universities they choose (Di Pietro, 2015; Netz, 2015; Schnepf and Colagrossi, 2020). An analysis in the United States has demonstrated that some characteristics, such as interest in diversity, are related to willingness to study abroad and affects participation alongside socioeconomic data and individual attributes (Salisbury et al., 2009, see). Additionally, living with family, understanding different cultures and foreign countries, and learning content and plans are also related to willingness to study abroad (Stroud, 2010), which has a positive relationship with participation (Luo and Jamieson-Drake, 2015). As indicated above, the factors that determine participation in studying abroad are diverse and are strongly related to learning achievement or career. However, the effects can only be determined correctly if the differences between participants and non-participants are clarified. Previous studies have addressed this issue empirically by using instrumental variable methods and propensity scores. For example, in Switzerland, Messer and Wolter (2007) demonstrates a significant effect of studying abroad on increasing the wages of the first job in OLS, but instrumental variable estimation with mothers' education as an instrument doesn't indicate any significant effect. In recent years, research has analyzed the effect of

studying abroad using causal reasoning, such as propensity scores. For example, Rodrigues (2013) targeted the effect of wages in specific European countries, Schnepf and d'Hombres (2018) examined employability in Italy and the UK, and Waibel et al. (2018) analyzed occupational status for the first three years after graduation in Germany. Kawata and Nishitani (2017) also analyzed the effect of SSA programs on English ability in Japan by estimating DID and propensity scores.

2.1.3 The effect of study abroad programs on careers

Previous studies have explained wage differences in study abroad program participants at the macro level, such as at the country income level, or the meso level, such as through wage differences by companies. For example, Liwiński (2019a) explained that the 18% higher wages of study abroad participants in Poland arose from their employment in foreign countries. Kratz and Netz (2018) explained the higher wages for graduates who study abroad through their jobs at large multinational companies. However, Sorrenti (2017) argues that foreign language proficiency is a factor that enables international employment, and this is a mechanism for the effect of studying abroad on careers. An analysis in Italy has demonstrated that studying abroad during the semester improves language proficiency leading to higher wages after graduation, although the wages varied by language (Sorrenti, 2017). Foreign language proficiency tends to increase the wages of nationals, but not the wages of foreign graduates (Ginsburgh and Prieto-Rodriguez, 2011; Saiz and Zoido, 2005; Stöhr, 2015). Therefore, international work choice leads to higher wages by improving foreign language proficiency through studying abroad.

2.2 SSA Programs

The effect of studying abroad differs depending on the length of the program. For example, participating in long-term study abroad programs often improves foreign language proficiency, learning achievement, and career selection more than SSA programs (Dwyer, 2004).

However, previous studies have not sufficiently controlled for the length of the study abroad program. For example, as described above, Di Pietro (2015) analyzed the effect of studying abroad on employability using the instrumental variables method. However, due to data constraints, the differences between the duration or content of the study abroad program were not considered. Campbell (2016) reviewed previous studies on the impact of SSA programs and focused on the effect of the global market experience of students majoring in business fields, especially on their careers. However, the research on other fields has not included these considerations. In addition to the length of the program, which is often regarded as too short to make an impact on the major objectives of higher education, it is challenging to perceive of SSA programs as a monopolistic feature because of their variety, such as programs lead by faculty members or programs with specific field characteristics, such as environmental fields surveys (Mule2018; Perry2012). The examples mentioned demonstrate that the effect of SSA programs is understudied.

2.3 Study abroad programs for Japanese university students

The number of Japanese students who study abroad had increased from around 36,000 in the 2009 academic year to around 105,000 in 2017 (MEXT, 2020). Looking at the longer trend from OECD statistics, the peaks were 83,000 students in 2004, and this decreased to about 60,000 students in 2012. The reason for this decrease has been discussed widely, including the idea that satisfaction with life in Japan and being "inwardly oriented" account for these differences (Kobayashi, 2011; Ota, 2014). Under these circumstances, in 2013, "Tobitate (Go abroad)!" was launched by the Japanese government to encourage studying abroad. This was in collaboration with industry, government, and academia, and aimed for 120,000 university students and 60,000 high school students to participate in study abroad programs by 2020 (MEXT, 2019). Universities in Japan also promote study abroad programs. For instance, since 2000, Chiba University, a large-scale national university in the Tokyo metropolitan area, has required all students to participate in a study abroad program (ChibaUniversity,

2019). The university targeted by this paper also promotes international education, including study abroad programs. In the 2018 academic year, 454 students from this university studied abroad (of which 71.4% in SSA programs), which increased by about 64% points from the 2013 academic year (of which 57% participated in SSA programs). In the 2019 academic year, about one-tenth of undergraduate students studied abroad, and the university had more than 70 international exchange agreements with higher education institutions.

3 Institutional back ground and data

The data used in this study were based on study abroad information, initial job attributes after graduation, and English proficiency test scores from a medium-sized highly selective national university in Tokyo, Japan. This article targets applicants for SSA programs described below, where the participants, who were undergraduate students majoring in social sciences and human arts, were randomly assigned.

3.1 Outline of short-term program

The SSA programs, our sample programs, were conducted in March and August 2014 (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 United Kingdom and the United States. 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 (= about 1,000 USD) in 2014, and the average cost in the 2016 to 2019 academic years was approximately 820,000 yen (= about 8,200 USD) for the equivalent program with similar length. The number of total applicants was 268 in 2013 and 437 in 2014, and the total number participants was 100 and 200, respectively. The selection method for the two targeted SSA programs was more or less 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 aims of the programs is to collect data to create future programs, the selection process does not consider foreign-language proficiency or academic achievement of applicants.

While the 2013 and 2014 school year programs were not perfectly alike, they were adequately similar in the sense of the English language study programs in native English countries. This similarity indicates that the covariates are theoretically aligned with those who participated or did not if the years and department are controlled proportionally. Meanwhile, the study institution was determined by the average language score of each institution. However, although the institutions average the language score, participating in the program is determined randomly.

3.2 Data

3.2.1 Study abroad data

As well as the SSA programs mentioned above, this study uses the information of participants who participated in long-term study abroad programs. There are several criteria to participate in long-term programs, such as high academic achievement demonstrated by a strong GPA or English proficiency demonstrated by TOEFL iBT and IELTS for English speaking countries (different language scores are required depending on the region).

3.2.2 Career Data

Career data were obtained from the university's career support office. The current data covers students who graduated from 2014 to 2018. If graduates worked at a private company listed on the Japanese stock exchange after graduation, we 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% for that period. Changes in company names, mergers, and the 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. We used information on listed firms on the stock exchange *Kaisyashikiho 2018*, which contains initial monthly salary information. We obtained information from non-listed companies from *Kaisyashikiho · Mizyozyokaisyaban 2018*, which contains financial information about sales, capital, build year, the number of workers, and initial monthly salaries. We judged a company as a foreign company if the company name is listed in *Gaishikeikigyosoran 2018*. As a limitation, data from non-listed firms are not as complete as the listed firms' data, even though *Kaisyashikiho · Mizyozyokaisyaban 2018* contains the firm's name. Also as the other limitation, reporting the company name after getting job offer is not mandatory. Subsequently, this data access imbalance and reporting error might become a source of bias in our estimation.

3.2.3 Academic data

For academic data, we used the academic achievement and attributes of the students; this included their gender, university year at the time of participation in the study abroad program, department, nationality, and whether they were 18 years old at the time of enrollment, which is the typical high school graduation age and the most common age of recent applicants. We also used the TOEFL-itp scores as the English proficiency scores. The university required first-year students to take the TOEFL test at the time of enrollment (April) and at the end of the first year (December or January). For students in their second year or later, taking a test was voluntary. The university pays the cost of taking the TOEFL-itp test in the first year, and approximately 90% of the students took the test twice.

3.3 Descriptive statistics

Of the 651 participants, 605 provided career information, of which 513 were employed at a private company at the time of graduation. Of these students, 327 (63.7%) were employed in companies listed on the Japanese stock market, and their information became the main

data source. The remaining 186 (36.3%) participants were working for firms not listed on the Japanese stock market. Table 2 illustrates the summary statistics for the SSA program applicants and participants. Some data is missing in the samples for educational outcomes and initial job attributes because of data availability. The average of some of the variables related to the program participants' initial job attributes and cultural skills tend to be slightly higher than those for non-participants. In Table 2, there are insignificant differences between participants and non-participants of SSA programs based on the t-test results on predetermined attributes, which includes the random assignment to balance the two groups. Table 2 also illustrates little difference in the gender, department, year, and nationality of the two groups as the control variables' attributes.

4 Empirical strategies

First, we defined each year's average treatment effect (ATE) from participating in short-term programs as an initial assignment and explained why the final assignment is not suitable for estimating the ATE. Second, we also defined the multi-year average ATE (AATE) from participating in short-term programs as an initial assignment. Third, we indicated the outcome variables using estimations, within-school outcomes, and initial firm attributes. Finally, we modified the previous AATE of participating in a short-term program as an initial assignment to the $AATE_{modi}$, which is the effect of participating in a short-term program over the treatment period. This follows Rubin's causal model (Rubin, 1974; Holland, 1986).

4.1 Short-term study abroad programs' ATE of the initial assignment

The short-term program was carried out twice in the 2013 and 2014 academic years. Ordinal RCT can be used to obtain the ATE as the treatment is usually one time, and the latent characters are the same over the treatment and control groups. However, the first assignment

results are more useful than the final assignment results.

If the final assignment is the treatment, then our estimation is as follows:

$$y_i = a_0 + a_1 LotteryWin2013_i + a_2 LotteryWin2014_i + a_3 PrgYear2014_i + \sum_{c=4}^{C} X_{ci}\alpha_c + error_i \quad (1)$$

Subscript i represents the individuals. The dependent variable y indicates within-school outcomes or initial job attributes. $LotteryWin2013/2014$ is the dummy variable indicating students who participate in the short-term program for that year. $PrgYear2014$ is the dummy variable indicating students who apply for their first application in the 2014 academic year. X represents the control variables. Under the satisfying assumption that the two 2013 and 2014 programs are independent, a_1 and a_2 become the ATE's for 2013 and 2014.

However, this is not the case as this university authority does not prohibit applicants from reapplying if they failed to gain a place in 2013. Figure 2 illustrates the application structure of short-term study abroad programs in the 2013 and 2014 academic years. Figure 3 describes the latent characteristics of Figure 2. We define each applicants' type in Table 3, state1∼6. We also defined Group1= state1,2,3,4, Group2=state3,4, and Group3=state5,6 for the convenience of notation. If we used final participation information to estimate the ATE, the gap between the "winners" (participants) state1 and "losers" (non-participants) state2 becomes 2013's ATE and 2014's ATE and is calculated by the gap between "winners" state4,6 and "losers" state3,5. In such a case, although the latent characteristics of 2014's "winners" would coincide with the "losers" latent characteristics, 2013's latent characteristics must not match, as some applicants in 2013 who do not succeed with their application reapply for the 2014 program and win at their second chance. Based on Rubin's causal model, the last estimation (1)'s a_1 is estimated ($Y_i(1)$ indicates the treated outcome, and $Y_i(0)$ is the

non-treated outcome of individual i) are as follows:

$$ATE2013 = [E(Y_i(1)–Y_i(0)|state1)P(state1) + E(Y_i(1)–Y_i(0)|state2)P(state2)]\frac{1}{P(state1,2)}$$

and a_2 is estimated as follows:

$$\begin{aligned}ATE2014 = [&E(Y_i(1)–Y_i(0)|state3)P(state3) + E(Y_i(1)–Y_i(0)|state4)P(state4)\\ &+ E(Y_i(1)–Y_i(0)|state5)P(state5) + E(Y_i(1)–Y_i(0)|state6)P(state6)]\frac{1}{P(state3,4,5,6)}\end{aligned}$$

Based on the latent characters in Figure 3, we can consider $E(Y_i(1)–Y_i(0)|state3) = E(Y_i(1)–Y_i(0)|state4)$ and $E(Y_i(1)–Y_i(0)|state5) = E(Y_i(1)–Y_i(0)|state6)$ as the ordinal counter factual consideration. This implies:

$$\begin{aligned}ATE2014 = [&E(Y_i(1)|state3)–E(Y_i(0)|state4)P(state3,4)\\ &+ E(Y_i(1)|state5)–E(Y_i(0)|state6)P(state5,6)]\frac{1}{P(state3,4,5,6)} = a_2\end{aligned} \quad (2)$$

However, state1 and state2 do not share latent covariates as the "losers" do not include individuals who apply again. Therefore, our estimation of a_1 is as follows:

$$a_1 = E(Y_i(1)|state1)–E(Y_i(0)|state2) \neq ATE2013 \quad (3)$$

If the students with apply-again characteristics have a lower treatment effect from participating in the lottery[1], this can be a source of upward bias of the ATE estimation.

To balance the latent characteristics, we considered that the first short-program assignment is the treatment in this article. If the first assignment result is the treatment, then our

[1]Students with apply-again characteristics are predetermined to participate in the other study abroad programs after their first application fails. In this case, joining a short-term study abroad has a smaller treatment effect than students who do not reapply.

estimation is as follows:

$$y_i = \alpha_0 + \alpha_1 1stLotteryWin2013_i + \alpha_2 1stLotteryWin2014_i + \alpha_3 PrgYear2014_i + \sum_{c=4}^{C} X_{ci}\alpha_c + error_i \quad (4)$$

$1stLotteryWin2013/2014$ is the dummy variable indicating students who participated in the short-term program with their first application at that year. Although some applicants are state3 or 4, our treatment initial assignment is independent between 2013 and 2014. Therefore, α_1 and α_2 become 2013's and 2014's ATE in regression form, respectively. The gap between "winners" in state1 and "losers" in state2,3,4 becomes 2013's ATE, and 2014's ATE was calculated by the gap between "winners" in state6 and "losers" in state5. Based on Rubin's causal model, the situation is as follows:

$$\begin{aligned} ATE2013 &= [E(Y_i(1)\text{–}Y_i(0)|state1)P(state1) + E(Y_i(1)\text{–}Y_i(0)|state2)P(state2) \\ &\quad + E(Y_i(1)\text{–}Y_i(0)|state3)P(state3) + E(Y_i(1)\text{–}Y_i(0)|state4)P(state4)]\frac{1}{P(state1,2,3,4)} \\ &= E(Y_i(1)|state1) - E(Y_i(0)|state2,3,4) = \alpha_1 \end{aligned} \quad (5)$$

$$\begin{aligned} ATE2014 &= [E(Y_i(1)\text{–}Y_i(0)|state5)P(state5) + E(Y_i(1)\text{–}Y_i(0)|state6)P(state6)]\frac{1}{P(state5,6)} \\ &= E(Y_i(1)|state5)\text{–}E(Y_i(0)|state6) = \alpha_2 \end{aligned} \quad (6)$$

Therefore, we can estimate the consistent ATE of initial assignment participation in short-term programs in each year.

4.2 Definition of short-term study abroad programs' AATE

We can identify each year's ATE using α_1, α_2 from equation (4). There is a possibility that the 2013 participants accidentally have characteristics to receive a higher positive treatment

effect than participants from other years. To avoid this, we mainly focused on AATE, which is the same as the weighted average of each years' ATE:

$$AATE = (\alpha_1 P(PrgYear = 2013) + \alpha_2 P(PrgYear = 2014)) \frac{1}{P(PrgYear = 2013, 2014)}$$

This AATE can be expressed as a coefficient of a simpler linear regression model:

$$y_i = \beta_0 + \beta_1 1stLotteryWin_i + \beta_2 PrgYear2014_i + \sum_{c=3}^{C} X_{ci}\beta_c + error_i \tag{7}$$

and then

$$AATE = (\alpha_1 P(PrgYear = 2013) + \alpha_2 P(PrgYear = 2014)) \frac{1}{P(PrgYear = 2013, 2014)} = \beta_1 \tag{8}$$

In this paper, we mainly report on whether AATE is significant or not. To confirm the robustness of the results, we also included the results of every single year's ATE. When the sign of the coefficient is the same, we report that the treatment effect is robust.

4.3 Outcomes of the estimation of the impact of short-term programs

4.3.1 Estimating the effect on within-school outcomes

We estimated the short-term program ATE using an increasing score of TOEFL-itp and estimated the AATE using the final year cumulative GPA and the participation rate of long-term study abroad programs. We labeled these as within-school outcomes as both outcomes are observable within the school. For calculating the effect on English skills, we only focused on first-year students in the 2014 academic year to estimate the ATE as the TOEFL-itp is optional after the second year[2]. We used the TOEFL-itp score growth from the spring to the

[2]Focusing the sample on first-year students can avoid the problem of voluntarily taking the TOEFL-itp.

winter of the first year as the outcome variable. This English score estimation is essentially the same as the estimation of DID with the propensity scores matching in Kawata and Nishitani (2017). In the linear regression model, ATE is illustrated by b_1 as follows:

$$y_i = b_0 + b_1 1stLotteryWin_i + \sum_{c=2}^{C} X_{ci} b_c + error_i \tag{9}$$

Kawata and Nishitani (2017) used a first-year students' data set, and they demonstrated a positive ATE. Therefore, we also expected a positive ATE (b_1) from our regression (9).

Kato and Suzuki (2019) analyzed the effects of short-term study abroad programs on the participation rate of long-term study abroad by using the same data set. However, Kato and Suzuki (2019) implicitly assume that students who apply twice are different individuals in their estimation. As we pointed out previously, there are reapplicants, and their estimation causes a downward bias for ATE in 2013 if the treatment effect is positive. Therefore, we reanalyzed the AATE of short-term program participation in an initial assignment on the participation rate of long-term programs by taking the indicator variable for participating in long-term study abroad programs as the outcomes in equations (4) and (8).

4.3.2 Estimating the effect on the initial firm's attributes

To estimate the ATE and AATE of short-term programs on the initial firm attributes we first used the dummy variables as outcome y_i in equations (4) and (8). This indicates whether the students chose a career as a worker, which also includes public sector work, if they became an employee in a private firm, or an employee in a private firm listed in the Japanese stock market after graduation. These results of the AATE indicate whether short-term programs promote students' initial careers as employed workers. Second, we used initial monthly income, sales, capital, number of workers, and established year as the outcome variables y_i. These outcomes represent the firms' financial information. Except for monthly income and established year, the range of distribution is too large to interpret. Therefore, we interpreted

a positive statistically significant ATE as increasing the probability of receiving a job offer from a larger firm. Finally, we used foreign-related firm attributes, such as if the firm is a foreign firm and listed on the Japanese stock market, as the outcomes.

4.4 Modified estimation for the reapplication process

4.4.1 Definition of the AATE with modifications for the reapplication process

The previous AATE definition assumes that the second chance treatment for reapplicants is induced by the initial assignment result of losing the random assignment process and ATE in 2013 should be ignored. However, there is a possibility that state3 students, reapplicants that could have a second chance in 2014, certainly receive the treatment. This situation is like the always taker in non-compliance problems of experimental treatments, if we consider that the second chance treatment and first chance treatment are identical. This subsection illustrates the idea of modifying this problem of the previous AATE. We also notice that this modification provides the AATE of short-term program participation within the program periods from 2013 to 2014, which is more relaxed than the initial assignment.

In addition to the definitions in Table 3, we also defined Group1= state1,2,3,4, Group2=state3,4, and Group3=state5,6 for the convenience of notation. If we retain the assumption that a single year ATE is defined by the initial application assignment, AATE, which we want to obtain, this is defined by equation (2) without any changes. By using the states and groups, AATE can be written down as follows:

$$\begin{aligned} AATE_{modi} &= E(Y_i(1)\text{–}Y_i(0)|Group1)P(Group1) + E(Y_i(1)\text{–}Y_i(0)|Group3)P(Group3) \\ &= (E(Y_i(1)|state1) - (E(Y_i(0)|state2)P(state2) + E(Y_i(0)|state3)P(state3) \\ &\quad + E(Y_i(0)|state4)P(state4)) * \frac{1}{P(state2,3,4)})\frac{1}{P(Group1)} \\ &\quad + (E(Y_i(1)|state5) - E(Y_i(0)|state6))\frac{1}{P(Group3)} \end{aligned} \tag{10}$$

The problem with this equation (10) is the term $E(Yi(0)|state3)$ in the second line. We want to know the potential outcomes of no participation for individuals in state3 who lose in 2013 but win in the 2014 reapplication to calculate the AATE. Only $E(Yi(1)|state3)$ can be observed from the real data. If we estimate $E(Yi(0)|state3)$, we can estimate the single year ATE and AATE in the same framework as the usual RCT. Even for the second time applicants, the assignment to study abroad is random. Hence, we can consider it as follows:

$$\begin{gathered} E(Y_i(1)|state3) = E(Y_i(1)|state4) = E(Y_i(1)|Group2) \\ \wedge \\ E(Y_i(0)|state3) = E(Y_i(0)|state4) = E(Y_i(0)|Group2) \end{gathered} \tag{11}$$

Equation (11) implies that $E(Yi(0)|state3)$, which is not visible as data, can be replaced by $E(Yi(0)|state3) = E(Yi(0)|state4)$. We can express AATE by the following equation:

$$\begin{aligned} AATE_{modi} = (E(Y_i(1)|state1) - \frac{E(Y_i(0)|state2)P(state2) + E(Y_i(0)|state4)P(Group2)}{P(state2) + P(Group2)})\frac{1}{P(Group1)} \\ + (E(Y_i(1)|state5) - E(Y_i(0)|state6))\frac{1}{P(Group3)} \end{aligned} \tag{12}$$

This equation (12)'s $AATE_{modi}$ coincides with the β_1 of equation (7) in the regression form by setting 0 as the sample weight for state3 and $\frac{P(state3.4)}{P(state4)}$ for state4. We can ignore second chance winners using this weighting procedure, and the interpretation of $AATE_{modi}$ becomes the AATE of participating in the short-term program over the treatment period. Compared to (12)'s $AATE_{modi}$ expression, (8)'s AATE causes under-reporting ($AATE_{modi} > AATE$) if participation in the short-term program has a positive effect on outcome Y ($(E(Yi(1)|state3) > E(Yi(0)|state3))$)[3]. Therefore, we also report $AATE_{modi}$.

[3]This subsection argument implies that only the ATE2013 is modified by our procedure. However, because of the weight changing and other existing covariates, the newly estimated ATE2014 is also modified in our results.

5 Results

In this paper, the following reporting rules were used. When signs of the ATE of both years are the same, and the AATE is statistically significant, the programs' effect is significant and robust. In the case of different signs of ATEs within two years and a significant AATE, the effect is significant but not robust. When the AATE is not statistically significant, there is no program effect, even if it is statistically significant in a single year. All standard errors of this paper use White (1980)'s adjustment.

5.1 AATE and ATE of participating SSA programs with an initial assignment on within-school outcomes

Tables 4 presents the causal AATE and ATE of SSA program participation with the initial assignment on school outcomes. Column 1 indicates the ATE2014 on the growth score of TOEFL-itp in the first-year students in 2014. This estimation limits the sample to 2014 first-year students as the students can choose to take the TOEFL or not by paying the examination fee after their second year. The point estimation in column 1 indicates that participating in the SSA program with the initial assignment has the causal power to increase, on average, the December TOEFL score from April by 8.1 points. This result is consistent with Kawata and Nishitani (2017)'s finding[4]. Columns 2 and 4 indicate the AATE, and columns 3 and 5 illustrate each year's ATE on the cumulative GPA at the time of graduation and the probability of participation in long-term study abroad programs. Participation in SSA programs has no causal power to increase GPA. Similar to the results of Kato and Suzuki (2019), the long-term study abroad participation rate increases by around 11.5% compared to non-participants on average, and this effect is robust. Kato and Suzuki (2019)'s point estimation in table 4 at column 4 is 0.121 and ours is 0.115. This difference comes from the difference in the idea of treatment, as the initial or final assignment.

[4]Coefficient sizes are different between of our estimations and their estimations. This coefficient difference comes from the examination differences; specifically, we use TOEFL, and they use TOEIC as the outcomes.

5.2 AATE and ATE of participating in SSA programs with an initial assignment on initial employed firm's attributes

Table 5 illustrates whether students choose a career as workers in the public sector, private sector, or firms listed on the Japanese stock market. Tables 6 and 7 present whether initially employed firm characteristics are different between participants and non-participants for SSA programs. The odd number columns indicate the AATE, and the even number columns represent each single years' ATE.

Based on the results in Table 5, participating in SSA programs does not affect the probability of choosing a career in the public sector, private sector, or in firms listed on the Japanese stock market. Based on each year's ATE, the ATE2014 indicates that SSA programs decrease the likelihood of becoming a worker in listed firms. The AATE, a weighted average of ATE, concludes that SSA programs have no such power to change basic career choice. This is an example of a small sample single shot ATE mis-capturing significant causal impact. Therefore, based on the AATE, we conclude that there is no significant causal power to increase the probability of working or becoming workers in firms listed on the Japanese stock market after graduation[5].

Table 6 shows the AATE and ATE on the initial firm's financial attributes. Income is an outcome of interest for many researchers. Column 1 in Table 6 illustrates that participating in SSA programs has no significant power to boost monthly income. Although for income, capital, number of workers, and established year, there is no significant difference between participants and non-participants, sales between participants and non-participants are significantly different, and this effect is robust. In Table 7, participants work for firms with a higher foreign stock rate than non-participants on average, although becoming a worker in affiliated foreign firms is insignificant[6].

[5]If we could conduct this random assignment process infinite times, we might capture the most precise AATE. We cannot deny the probability that our AATE is insignificant with an accidental tendency as the university conducted the random assignment process just twice, leading to a smaller sample size.

[6]A higher foreign stock rate might indicate more non-domestic branches, and Japan-oriented firms might evaluate SSA programs participation experience more than foreign origin firms. However, we cannot provide

5.3 Modified AATE and ATE results

Tables 8-11 are the results of modified point estimates of the AATE and ATE. Compared to tables 5-8, all the coefficients of tables 8-11 change slightly, and the ATE2013 is modified[7]. From Table 3, the total share of reapplicants is around 8%. Although there is no way to change this significance other than the number of workers, this modification provides precise point estimates of the AATE and ATE of SSA programs. Even if we use this modification method, we still found the same significant positive effect on the participation rate of long-term programs and initial firm's attributes, sales, and foreign stock rate[8].

This modification estimation changes the interpretation of the AATE. Without this modification, the AATE indicates the effect of participating in SSA programs at the initial application. We can interpret that the AATE with this modification indicates the effect of participating in SSA programs once within this period. The latter interpretation is more general than the previous one. Although prohibiting reapplication for applicants is the easiest way to capture precise ATEs, this tends to make the process becomes unfair for some students. This modification allows for the reapplication process and the ability to capture a precise ATE.

6 Discussion

We can easily interpret the AATE and ATE on within-school outcomes as causal impacts. However, the initial firm's attributes are somewhat complicated because participating in SSA programs does not directly differentiate firms' attributes because firms' attributes are only determined in the market. We verified how this average gap between the participants' and non-participants' financial information on initial firms is generated in Table 12. Columns 1

a clear interpretation here because of the data limitation.

[7]While the ATE2014 does not change, changing the estimation weight and other existing covariates creates a slight change in the ATE2014.

[8]This modification is not valid for the TOEFL estimation as the TOEFL estimation in table 4 uses first-year students from 2014 as the sample. This modification is only valid for the estimations, including reapplicants.

and 4 illustrate the AATE on initial firm attributes with a sample of the 10-90 percentile of firms' sales and foreign stock rate. Although tables 10 and 11 indicate significant differences of firms' attributes between the participants and non-participants, columns 1 and 4 in Table 12 illustrate that these significant differences disappear with samples in the 10-90 percentile. We also verified the effects on the employment probability from the bottom 10% in columns 2 and 5, and the top 10% of firms in columns 3 and 6. The results indicate that participation in SSA programs significantly decreases the employment probability from firms with the bottom 10% of sales and foreign stock rate and increases the employment probability from the top 10% of sales. Based on these results, initial firms' average sales differences between participants and non-participants might imply that SSA programs can increase/decrease employment probability from firms with higher/lower sales.

Additionally, as a mechanism, one possible concern is that within-school outcomes, which are predetermined before the job search period, might work as a unique mechanism path of the causal effect of participating in SSA programs on initial firms' attributes. The lottery win coefficient in Table 13 illustrates the SSA participation effect on the firms' sales and foreign stock rate through a path different from participation in long-term programs and a cumulative GPA in the graduation year. The effect on sales is significant, and the effect on the foreign stock rate is insignificant. We can learn from Tables 10 and 13 that the other paths than participation in long-term programs and cumulative GPA at graduation might exist as mechanism paths to change the initial firm's sales. However, we cannot identify the paths that increase firms' employment probability with large sales due to the data limitations. Future studies are necessary to clarify which mechanism is responsible. Table 8 illustrates that the SSA programs have the causal power to increase English scores in the first year. We do not have English score data right before graduation or in the job search period, which is an important within-school outcome for considering how the mechanism of SSA programs affects initial firms' attributes. We must rely on future studies to clarify whether English skills are the main mechanism behind the change in initial firms attributes.

Although Table 12 suggests that the differences in the initial firms' attributes between the participants and non-participants are because of the differences in employment probability from top and bottom, it is still ambiguous why employment probability is increased by participation in SSA programs. Again, this is because of the data limitations. However, there are two possible explanations for the mechanism. From the firm side preference, if students' preferences are fixed, and firms' recruiting behavior is conducted in the order of largest sales, Table 12 implies that firms have a higher priority of employing SSA participants because they would have the higher ability from the human capital accumulation in the SSA programs. From the student side preference, if firms' employment strategy is random employment, Table 12 implies that SSA participants change their attitude about jobs and start applying to firms with high sales more than non-participants. We must rely on future studies to identify which explanation is more applicable to explain why participants' employment probability from firms with top sales increases and with bottom sales decreases.

7 Conclusion

In this article, we investigated the impact of SSA programs on within-school outcomes and the initial firms' attributes after graduation. We used data from a Japanese university to randomly assign participation in SSA programs among the applicants (705 applicants and 300 study abroad participants) in the 2013 and 2014 academic years. The results demonstrate that participation in SSA programs increases some within-school outcomes (TOEFL score and the participation rate in long-term study abroad programs) and the employment probability from financially large (sales and foreign stock rate) firms compared to non-participants. However, the results demonstrate that participating in SSA programs does not increase monthly income and employment probability. The results also indicate that initial firms' attributes might change the employment probability from the bottom and top 10% of firms. Participating in SSA programs has the causal power to increase the employment probability

from firms with the top 10% sales and decrease the employment probability from the bottom 10% sales and foreign stock rate. Additionally, SSA programs still significantly affect initial firms' sales even after controlling for long-term study abroad participation and final year GPA.

Although there is a global increase in the number of SSA programs, as illustrated in Figure 1, there is a limited accumulation of previous studies that have evaluated the causal effects of SSA programs. Previous studies on the effects of studying abroad has two main challenges. The first challenge is the control of the selection bias, and the second challenge involves the unclear mechanisms that affect some labor market outcomes. Identifying the causal effects of SSA programs on some within-school outcomes and initial careers is a major contribution of this paper. Our additional contribution is offering a modification of the AATE estimation method to account for reapplication behavior. Some schools and researchers might not be able to prohibit reapplication due to these ethical restrictions, and also some researchers might not treat reapplication issues clearly into their analysis due to complicated structure. In such cases, our modification estimation might be a better option for precise point estimations for the ATE and AATE, even if there are reapplication structures.

Most of the results presented in this paper are intuitive. Regarding the initial monthly income result, it is very natural that participation in a four weeks SSA program does not affect graduates' monthly income of initial jobs. Although the treatment period of college education (four years) is 56 times larger than SSA programs (four weeks), in Japan, the ATE of college education on income is around 9% (Kikuchi, 2017). If participation in the SSA programs had the same power to increase income as college education, the ATE would become 0.16%. Such a small effect size is usually not detectable in statistical analyses.

There are some limitations to our studies. The first is external validity. The university in this paper is a research-oriented, highly selective university in Japan. To capture general results for Japan, it is necessary to use samples from diverse universities and university fixed

effects to control university specificity. To achieve this, the expansion of higher education data in Japan is necessary. On lacking external validity, the ATE and AATE are close to the lower bounds. The ATE and AATE of the other universities (at least) in Japan might be higher than our results[9]. Even though such underreporting possibility comparing to average level university, some of our results are still significants. A second limitation is the non-compliance problem after the treatment period. As lottery "losers" in this random assignment study abroad program tend to have a higher participation rate in other years or nonrandom assignment study abroad programs than "winners", this paper cannot control the other study abroad programs conducted by this university in other academic years since these participations are one of the effects of SSA programs. Additionally, the students could participate in a non-university funded study abroad program [10]. Therefore, we identified the AATE of participation in SSA programs only once within the program period. Our AATE is not purely the AATE of all students in this university, and subsequently, it is not purely the effect of participation in SSA programs only once during the university life. Compared to a purer AATE, our estimation might be biased as our sample is limited to highly motivated students and non-compliant students who take part in a study abroad treatment after this application process. Another concern is that we cannot explain why SSA programs change the initial firms' attributes. Although we offer two possible paths in preference story of firms and students. Identifying the path or mechanisms of SSA programs' effect on the initial career must be a major theme for future research on the effect of SSA programs.

While these limitations are evident, this paper offers empirical evidence that participation in SSA programs increases some within-school outcomes (TOEFL score and the participation rate in long-term study abroad programs) and the employment probability with financially

[9]For example, the TOEFL score of other ordinal Japanese universities is lower than our sample university. A low start test score makes it easier for scores to increase than a high start test score. Regarding the employment probability, our sample university already has a high employment probability, and subsequently, ordinal universities might have a higher ATE and AATE from SSA programs.

[10]The *Global Study* states that a four weeks study abroad program in the Philippines costs around 180,000 yen (`https://www.global-study.jp/gashuku/school_cpils.html`). This program is similar to the 2014 programs in terms of contents (English learning) and cost.

large (sales and foreign stock rate) firms. Thus, the results will be useful for higher education and international education professionals, for considering the effect of short-term study abroad programs, and for understanding identification problems related to reapplication.

References

Anderson, Philip H., Leigh Lawton, Richard J. Rexeisen, and Ann C. Hubbard (2006) "Short-term study abroad and intercultural sensitivity: A pilot study," *International Journal of Intercultural Relations*, Vol. 30, No. 4, pp. 457–469.

Bodycott, Peter Thomas and Vernon Crew (2000) "Living the language: The value of short-term overseas English language immersion programs," *The Language Teacher*, Vol. 24, No. 9, pp. 27–34.

Campbell, Kathleen (2016) "Short-term study abroad programmes: objectives and accomplishments," *Journal of international Mobility*, No. 1, pp. 189–204.

Carley, Susan and R Keith Tudor (2010) "Assessing the impact of short-term study abroad," *Journal of Global Initiatives: Policy, Pedagogy, Perspective*, Vol. 1, No. 2, p. 5.

ChibaUniversity (2019) 「千葉大学　全員留学　Engine」, Technical report.

Di Pietro, Giorgio (2012) "Does studying abroad cause international labor mobility? Evidence from Italy," *Economics Letters*, Vol. 117, No. 3, pp. 632–635.

——— (2015) "Do study abroad programs enhance the employability of graduates?" *Education Finance and policy*, Vol. 10, No. 2, pp. 223–243.

——— (2019) "University study abroad and graduates' employability," *IZA World of Labor*, https://doi.org/10.15185/izawol.109.v2).

Dwyer, Mary M (2004) "More is better: The impact of study abroad program duration.," *Frontiers: The Interdisciplinary Journal of Study Abroad*, Vol. 10, pp. 151–163.

Gaia, A Celeste (2015) "Short-term faculty-led study abroad programs enhance cultural exchange and self-awareness.," *International Education Journal: Comparative Perspectives*, Vol. 14, No. 1, pp. 21–31.

Geyer, Alexis, Jenni Putz, and Kaustav Misra (2017) "The effect of short-term study abroad experience on American students' leadership skills and career aspirations," *International Journal of Educational Management*, Vol. 31, No. 7, pp. 1042–1053.

Ginsburgh, Victor A and Juan Prieto-Rodriguez (2011) "Returns to foreign languages of native workers in the European Union," *ILR Review*, Vol. 64, No. 3, pp. 599–618.

Holland, Paul W. (1986) "Statistics and Causal Inference," *Journal of the American Statistical Association*, Vol. 81, No. 396, pp. 945–960.

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*, Vol. 23, No. 4, pp. 411–428.

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, Mimeo,
(available at: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxrZWlzdWtla2F3YXRhN3xneDoxOTQwNDhhNDNkZGIwZWNm).

Kikuchi, Nobuyoshi (2017) "Marginal returns to schooling and education policy change in Japan."

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad (2016) "Field of Study, Earnings, and Self-Selection*," *The Quarterly Journal of Economics*, Vol. 131, No. 3, pp. 1057–1111, 05.

Kobayashi, Akira (2011) 「Factors that hinder Japanese students from studying abroad and future measures [Translated from Japanese.]」,『ウェブマガジン 「留学交流」』, 第 5 巻.

Kratz, Fabian and Nicolai Netz (2018) "Which mechanisms explain monetary returns to international student mobility?" *Studies in Higher Education*, Vol. 43, No. 2, pp. 375–400.

Kurt, Mark R, Neal H Olitsky, and Paul Geis (2013) "Assessing global awareness over short-term study abroad sequence: A factor analysis.," *Frontiers: The Interdisciplinary Journal of Study Abroad*, Vol. 23, pp. 22–41.

Liwiński, Jacek (2019a) "Does it pay to study abroad? Evidence from Poland," *International Journal of Manpower*, Vol. 40, No. 3, pp. 525–555.

——— (2019b) "Does studying abroad enhance employability?" *Economics of Transition and Institutional Change*, Vol. 27, No. 2, pp. 409–423.

Luo, Jiali and David Jamieson-Drake (2015) "Predictors of study abroad intent, participation, and college outcomes," *Research in Higher Education*, Vol. 56, No. 1, pp. 29–56.

Mapp, Susan C (2012) "Effect of short-term study abroad programs on students' cultural adaptability," *Journal of Social Work Education*, Vol. 48, No. 4, pp. 727–737.

Messer, Dolores and Stefan C Wolter (2007) "Are student exchange programs worth it?" *Higher Education*, Vol. 54, No. 5, pp. 647–663.

MEXT (2019) "About "Survey on the status of enrollment of foreign students" and "Number of Japanese students studying abroad" [Translated from Japanese.],"Technical report.

——— (2020) "Current Situation and Measures Concerning International Student Policy[Translated from Japanese.],"Technical report.

Netz, Nicolai (2015) "What deters students from studying abroad? Evidence from four European countries and its implications for higher education policy," *Higher Education Policy*, Vol. 28, No. 2, pp. 151–174.

Oosterbeek, Hessel and Dinand Webbink (2011) "Does studying abroad induce a brain drain?" *Economica*, Vol. 78, No. 310, pp. 347–366.

Ota, Hiroshi (2014) 「A Study on Inward-Oriented Japanese Students: Rethinking International-Orientedness Based on Existing Data [Translated from Japanese.]」,『留学交流』, 第 40 巻.

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*, Vol. 121, No. 551, pp. 194–222.

Rodrigues, Margarida (2012) "Determinants and impacts of student mobility: A literature review," *JRC Scientific and Technical Report JRC*, Vol. 70059.

——— (2013) "Does student mobility during higher education pay? Evidence from 16 European countries," `http://econmodels.com/upload7282/41b9a94d0d2f9bedbd271b6910121bb2.pdf`.

Rubin, Donald B (1974) "Estimating causal effects of treatments in randomized and nonrandomized studies.," *Journal of Educational Psychology*, Vol. 66, No. 5, pp. 688–701.

Saiz, Albert and Elena Zoido (2005) "Listening to what the world says: Bilingualism and earnings in the United States," *Review of Economics and Statistics*, Vol. 87, No. 3, pp. 523–538.

Salisbury, Mark H, Paul D Umbach, Michael B Paulsen, and Ernest T Pascarella (2009) "Going global: Understanding the choice process of the intent to study abroad," *Research in Higher Education*, Vol. 50, No. 2, pp. 119–143.

Savicki, V., E. Brewer, and B. Whalen (2015) *Assessing Study Abroad: Theory, Tools and Practice*: Stylus Publishing.

Schmidt, Stephen and Manuel Pardo (2017) "The contribution of study abroad to human capital formation," *The Journal of Higher Education*, Vol. 88, No. 1, pp. 135–157.

Schnepf, Sylke V and Marco Colagrossi (2020) "Is unequal uptake of Erasmus mobility really only due to students' choices? The role of selection into universities and fields of study," *Journal of European Social Policy* forthcoming.

Schnepf, Sylke V and Beatrice d'Hombres (2018) "International Mobility of Students in Italy and the UK: Does It Pay off and for Whom?," *IZA Discussion Paper*, `http://ftp.iza.org/dp12033.pdf`.

Shimmi, Yukiko, Hiroko Akiba, Ota Hiroshi, and Masahiro Yokota (2017) 「Long-term Impact of Undergraduate Study Abroad Experiences on Career: Comparative Survey Results among Degree-seeking Study Abroad, Credit-bearing Study Abroad, and Non-study Abroad Groups」,『留学交流』，第 74 巻，14–26 頁.

Sorrenti, Giuseppe (2017) "The Spanish or the German apartment? Study abroad and the acquisition of permanent skills," *Economics of Education Review*, Vol. 60, pp. 142–158.

Stöhr, Tobias (2015) "The returns to occupational foreign language use: Evidence from Germany," *Labour Economics*, Vol. 32, pp. 86–98.

Stroud, April H. (2010) "Who plans (not) to study abroad? An examination of US student intent," *Journal of Studies in International Education*, Vol. 14, No. 5, pp. 491–507.

Waibel, Stine, Heiko Rüger, Andreas Ette, and Lenore Sauer (2017) "Career consequences of transnational educational mobility: A systematic literature review," *Educational Research Review*, Vol. 20, pp. 81–98.

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*, Vol. 74, pp. 45–61.

Walters, Colleen, Jennell Charles, and Sue Bingham (2017) "Impact of short-term study abroad experiences on transformative learning: A comparison of programs at 6 weeks," *Journal of Transformative Education*, Vol. 15, No. 2, pp. 103–121.

White, Halbert (1980) "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica*, Vol. 48, No. 4, pp. 817–838.

Wiers-Jenssen, Jannecke (2011) "Background and employability of mobile vs. non-mobile students," *Tertiary Education and Management*, Vol. 17, No. 2, pp. 79–100.

Yokota, Masahiro, Hiroshi Ota, and Yukiko SHIMMI (2018) *Impact of studying abroad on career and life: Measuring the effectiveness of studying abroad through a large-scale survey[Translated from Japanese.]*: GAKUBUNSHA.

Tables

Table 1: Program information

School year	2013	2014
Fees paid by students	0	10,000 yen
Number of sending schools	9	13
Cumulative total number of applicants	268	437
Cumulative total number of lottery winners	100	200
Odds of winning	2.68	2.19

100 yen is around $1 (US) in 2020. "Winners" are the study abroad participants within the applicants

Table 2: Summary statistics

		Participants(lottery_win = 1)			Non-Participants(lottery_win = 0)			
		Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	t-Value
1st job attributes	log(foreign stock rate)	126	0.31	0.13	184	0.28	0.12	
	Log(Sales)	112	14.47	1.41	170	13.99	1.97	
	Log(Capital)	126	12.01	1.92	184	11.83	2.35	
	Log(Initial monthly income)	124	6.78	0.26	170	6.74	0.25	
	Log(Number of workers)	106	8.49	1.60	154	8.32	1.77	
	Log(Build year)	126	3.84	0.90	184	3.65	0.92	
Within-school outcomes	Participation in long-term	272	0.31	0.46	379	0.16	0.37	
	GPA	264	3.12	0.49	367	3.06	0.50	
	Score difference between April and winter (only available for 2014 departure first year students)	92	19.90	27.89	97	16.04	29.89	
Pre-determined skill	TOEFL score in April of first year	258	514.20	40.85	355	513.05	39.61	0.35
	First year GPA	264	3.26	0.45	367	3.21	0.50	-1.41
	Age greater than 18	265	0.34	0.47	369	0.39	0.49	-1.47
controls (All are dummy variables)	Female	256	0.32	0.47	349	0.33	0.47	-0.14
	Second year	256	0.36	0.48	349	0.38	0.49	-0.74
	Third year	256	0.11	0.32	349	0.13	0.34	-1.12
	Major2	256	0.27	0.45	349	0.26	0.44	0.69
	Major3	256	0.13	0.34	349	0.15	0.35	-0.46
	Major4	256	0.28	0.45	349	0.24	0.43	0.49
	International students	256	0.07	0.26	349	0.11	0.32	-1.50

t-value uses the descriptive statistical attribute as an outcome to determine the difference between the groups by t-testing the coefficient values with a single regression with the 1st LotteryWin.

* Denotes the significance level, respectively, * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$.

All attributes were not statistically significant, even at the 10% level, and the treatment and control group attributes could be considered equal, suggesting that random assignment was functional.

Table 3: Type of individuals and definition of state

state	2013 result	give upcontinue	2014 second result	2014 first result	Proportion in sample
State1	Win	-	-	-	15.36%
State2	Lose	Give up	-	-	17.51%
State3	Lose	Continue	Win	-	4.30%
State4	Lose	Continue	Lose	-	3.99%
State5	-	-	-	Win	26.42%
State6	-	-	-	Lose	32.41%

Table 4: AATE and ATE of participating short-term programs on within-school outcomes

	(1) ΔTOEFL	(2) GPA	(3) GPA	(4) Long	(5) Long
lottery_win	8.089**	0.021		0.115***	
	(3.744)	(0.021)		(0.032)	
lottery_2013_win			0.002		0.184***
			(0.033)		(0.055)
lottery_2014_win			0.033		0.070*
			(0.027)		(0.039)
N	189	613	613	613	613

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method. Other controls are dummies of 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 significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$

Table 5: AATE and ATE of participating short-term program on employment probability

	(1) Working	(2) Working	(3) Private	(4) Private	(5) Market	(6) Market
lottery_win	0.033		-0.026		-0.049	
	(0.027)		(0.019)		(0.043)	
lottery_2013_win		0.060		-0.024		0.067
		(0.044)		(0.029)		(0.067)
lottery_2014_win		0.015		-0.027		-0.123**
		(0.035)		(0.025)		(0.056)
N	585	585	502	502	502	502

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method.
Other controls are dummies of female, international student, program year, entrance with age grater than 18, TOEFL-itp score at first year April, major, grade, and first year GPA.
Statistical significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$

Table 6: AATE and ATE of participating short-term program on initial firm's attributes of financial information

	(1) income	(2) income	(3) Sales	(4) Sales	(5) Capital	(6) Capital	(7) Workers	(8) Workers	(9) Build year	(10) Build year
lottery_win	0.004		0.552***		0.191		0.246		-3.319	
	(0.012)		(0.213)		(0.294)		(0.198)		(3.896)	
lottery_2013_win		-0.002		0.167		0.042		0.188		-8.189
		(0.015)		(0.306)		(0.402)		(0.287)		(6.025)
lottery_2014_win		0.008		0.838***		0.295		0.287		-0.197
		(0.016)		(0.288)		(0.414)		(0.268)		(5.087)
N	308	308	294	294	332	332	313	313	388	388

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method. Other controls are dummies of female, international student, program year, entrance with age grater than 18, TOEFL-itp score at first year April, major, grade, and first year GPA.
Depending variable worker implies number of workers. All the outcome variables take log except build year. Sample size is depending on data availability of depending variables. Data availability is very low in non-market firm and monthly income is less open than the other information.
Statistical significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$

Table 7: AATE and ATE of participating short-term program on initial firm's attributes related to international characters

	(1) Foreign firm	(2) Foreign firm	(3) Foreign stock rate	(4) Foreign stock rate
lottery_win	0.008		0.028*	
	(0.028)		(0.015)	
lottery_2013_win		0.003		0.031
		(0.043)		(0.024)
lottery_2014_win		0.011		0.026
		(0.037)		(0.019)
N	450	450	307	307

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method.
Other controls are dummies of female, international student, program year, entrance with age grater than 18, TOEFL-itp score at first year April, major, grade, and first year GPA.
Statistical significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$

Table 8: Modified AATE and ATE on within school outcomes

	(1) GPA	(2) GPA	(3) Long	(4) Long
lottery_win	0.022		0.120***	
	(0.022)		(0.033)	
lottery_2013_win		0.003		0.195***
		(0.039)		(0.057)
lottery_2014_win		0.034		0.072*
		(0.027)		(0.039)
N	585	585	585	585

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method. Other controls are dummies of female, international student, program year, entrance age grater than age 18, TOEFL-itp score at freshman April, major, grade and 1st year GPA.
Statistical significancies are following, * $p < .1$, ** $p < .05$, *** $p < .01$

Table 9: Modified AATE and ATE on employment probability

	(1) Working	(2) Working	(3) Private	(4) Private	(5) Market	(6) Market
lottery_win	0.030		-0.024		-0.053	
	(0.028)		(0.020)		(0.044)	
lottery_2013_win		0.055		-0.016		0.056
		(0.045)		(0.033)		(0.072)
lottery_2014_win		0.015		-0.028		-0.124**
		(0.035)		(0.025)		(0.056)
N	559	559	481	481	481	481

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method.
Other controls are dummies of female, international student, program year, entrance with age grater than 18, TOEFL-itp score at first year April, major, grade, and first year GPA.
Statistical significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$
Regression weight is used to handle the reapplication problem for all regression in this tables.

Table 10: Modified AATE and ATE on initial firm's attributes of financial information

	(1) income	(2) income	(3) Sales	(4) Sales	(5) Capital	(6) Capital	(7) Workers	(8) Workers	(9) Build year	(10) Build year
lottery_win	0.005		0.534**		0.214		0.268		-2.741	
	(0.013)		(0.216)		(0.299)		(0.198)		(4.049)	
lottery_2013_win		0.000		0.121		0.099		0.241		-7.367
		(0.019)		(0.321)		(0.436)		(0.292)		(6.568)
lottery_2014_win		0.008		0.841***		0.294		0.288		0.279
		(0.017)		(0.288)		(0.414)		(0.269)		(5.105)
N	298	298	283	283	320	320	301	301	375	375

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method. Other controls are dummies of female, international student, program year, entrance with age grater than 18, TOEFL-itp score at first year April, major, grade, and first year GPA.
Depending variable worker implies the number of workers. All the outcome variables take log except build year. Sample size is dependent on data availability of depending variables. Data availability is very low in non-market firm and monthly income is less open than the other information.
Statistical significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$
Regression weight is used to handle the reapplication problem for all regression in this tables.

Table 11: Modified AATE and ATE on initial firm's attributes related to international

	(1) Foreign firm	(2) Foreign firm	(3) Foreign stock rate	(4) Foreign stock rate
lottery_win	0.007		0.027*	
	(0.029)		(0.015)	
lottery_2013_win		-0.002		0.029
		(0.045)		(0.024)
lottery_2014_win		0.012		0.026
		(0.037)		(0.019)
N	434	434	297	297

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method.
Other controls are dummies of female, international student, program year, entrance with age grater than 18, TOEFL-itp score at first year April, major, grade, and first year GPA.
Depending variable foreign firm indicate a student's company is listed in *Gaishikeikigyosoran 2018*.
Data about foreign stock rate is only available for firms listed in the Japanese stock market.
Statistical significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$
Regression weight is used to handle the reapplication problem for all regression in this tables.

Table 12: AATE on employment probability for firms with top and bottom 10

	(1) Sales	(2) 0-10 Sales firm	(3) 90-100 Sales firm	(4) Foreign stock rate	(5) 0-10 Foreign stock rate firm	(6) 90-100 Foreign stock rate firm
lottery_win	-0.209	-0.113***	0.081**	0.006	-0.056*	0.044
	(0.151)	(0.034)	(0.039)	(0.011)	(0.033)	(0.037)
N	225	283	283	236	297	297

Parentheses are standard errors. Standard errors are adjusted by White (1980)'s method. Other controls are dummies of female, international student, program year, entrance with age grater than 18, TOEFL-itp score at first year April, major, grade, and first year GPA.
Column (1) and (4) are AATE on average log sales and foreign stock rate with 10-90 percentile sample. Column (2),(3),(5),(6) show AATE on employment probability from bottom or top 10% firms. Statistical significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$
Regression weight is used to handle the reapplication problem for all regression in this tables.

Table 13: Controlling long-term study abroad participation and check significance in short-term participation

	(1) Sales	(2) Foreign stock rate
long_go	-0.100	0.003
	(0.310)	(0.013)
final_GPA	0.593	0.020
	(0.493)	(0.028)
lottery_win	0.533**	0.026
	(0.219)	(0.016)
N	283	297

Parentheses are standard errors. Standard errors are adjusted
by White (1980)'s method. Other controls are dummies of female,
international student, program year, entrance with age grater than 18,,
TOEFL-itp score at first year April, major, grade, and first year GPA.
Statistical significances are as follows, * $p < .1$, ** $p < .05$, *** $p < .01$
Regression weight is used to handle the reapplication problem for all regression in this tables.

Figures

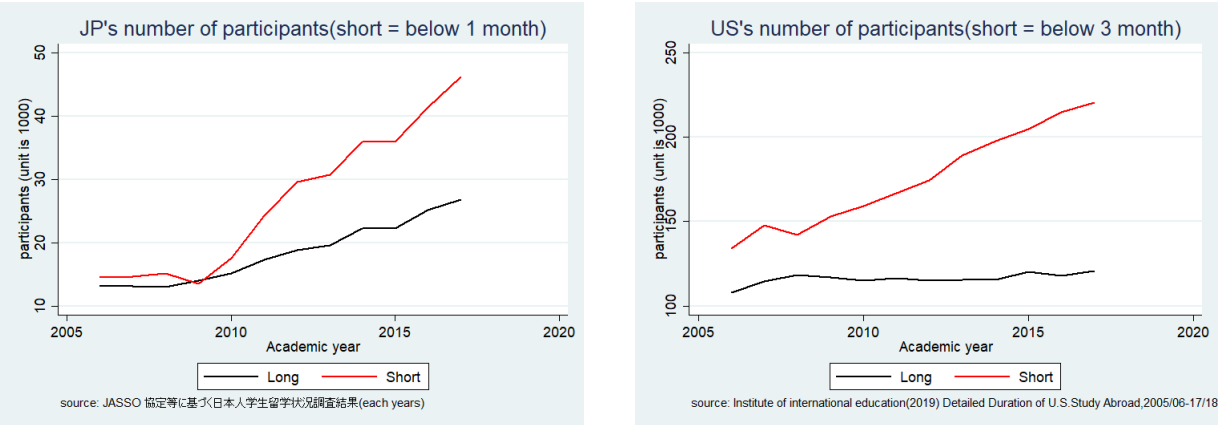

Figure 1: The number of participants in study abroad programs in Japan and the United States

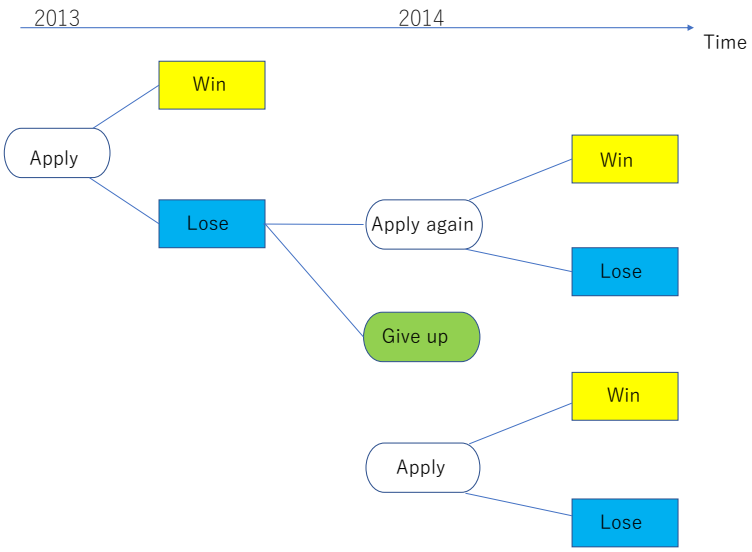

Figure 2: Reapplication structure of short-term program

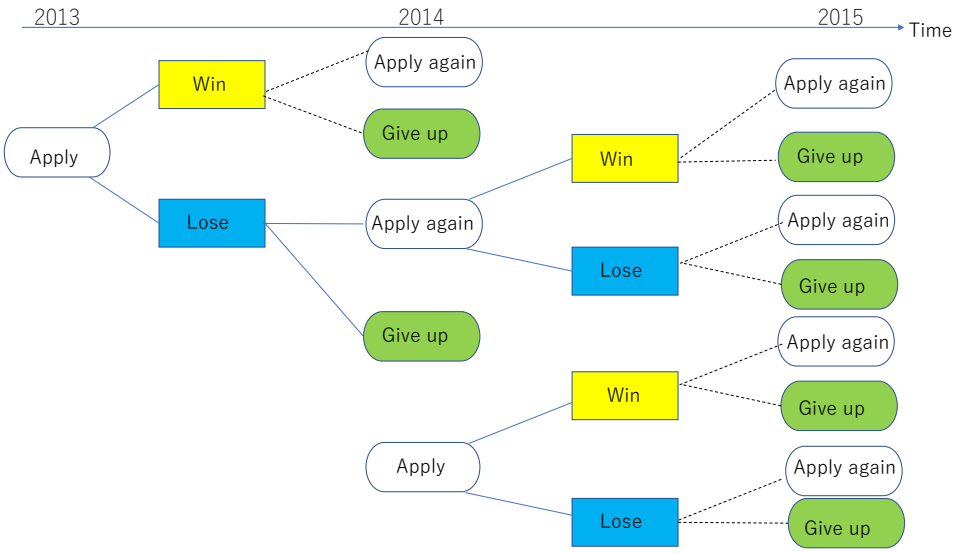

Figure 3: Reapplication structure of short-term program and latent unobservable characters

Win/Lose indicates the success of individual applications for short-term programs in the random assignment process. "Apply again" or "Give up" are the two characters of the applicants, which are the individual ' s available choices of action if they are rejected after their application: they can choose to reapply or to not reapply The dot-lines indicate the unobservable actions; the solid-lines indicate the applicants ' choices as observable actions leading to the action's result.