Master Thesis, January 2018

Testing the Comparative Advantage Hypothesis of Demand for Higher Education in Japan

Ippei Ozawa

Graduate School of Economics

The University of Tokyo

-Abstract-

In this paper, I test the comparative advantage hypothesis of demand for higher education in Japan. The framework applied in this paper to identify selection patterns of schooling choice is the switching regression model in Wills and Rosen (1979). The results support negative selection among college graduates and positive selection among non-college graduates for the young generation. This evidence indicates a hierarchy sorting by one-dimensional unobserved heterogeneity and contradicts the comparative advantage hypothesis.

Keywords

higher education, comparative advantage, hierarchy sorting, unobserved heterogeneity

1. Introduction

One of the recent educational concerns in Japan is whether tuition fees in college and college entrance fees may cause an unfair opportunity of receiving higher education because of disparity in household incomes. The government has discussed new policies; exempting students with lower income from the tuition or enrolling fees wholly or partially, and enhancing the scholarship system more. In utility maximization problem, one`s educational choice is affected by the budget constraints of each household, then the exemption policy may relieve the budget constraint. Also, the scholarship system may also affect one`s borrowing constraint. However, improving the scholarship system in Japan has gradually made it possible for students with lower household incomes to enroll in college recently. The costs of enrolling in college may include not only the tuition fee and the enrolling fee but also a

geographic cost, a cost of effort to enroll in and graduate from college, a psychological cost, and a future cost of returning scholarship if one uses the system. These costs vary among students with different backgrounds. Moreover, one`s advantage on choosing each schooling and return to education have been thought to be important factors. Though the Japanese government seems to think that the unfair opportunity of receiving higher education is caused by tuition fees in college and college entrance fees, who enrolls in college or who does not is a problem for appropriate policy implication. We can see some possible factors related to one`s schooling choice; costs, one`s advantage on each schooling choice and so on. My motivation in this paper is to analyze what affects one`s schooling choice, selection patterns of schooling choice in another word.

Results in former researches generally imply one`s choice of whether obtaining higher education or not depends on one`s comparative advantage on each schooling choice. Under the assumption of "comparative advantage hypothesis of demand for higher education", one`s schooling choice is conducted by which type of jobs one has a comparative advantage on; type of jobs that college graduates tend to obtain and the type of jobs that non-college graduates tend to. The choice depends on one`s expected productivity as a worker for each schooling choice and the hypothesis based on standard Roy`s model is defined as;

$$\log(Y_{Ci}/Y_{Ni})_{S=1} > \log(Y_{Cj}/Y_{Nj})_{S=0} \quad \text{for all } i \neq j \tag{1}$$

where i and j are different individuals, Y_{Ci} is the lifetime earnings for the ith individual choosing college and Y_{Ni} is that for the ith individual choosing non-college. S shows the set of schooling choices. If each individual chooses college, then S=1, and S=0 if not. In this case, individual i chooses collage and individual j chooses non-college. Under equation (1), both individual i and j have a comparative advantage on their schooling choices. In qualitative interpretation of equation (1), each individual conducts a schooling choice in which the comparative productivity is higher than that of another one.

Past research in economics literatures such as Wills and Rosen, Garden (1984) generally support the comparative advantage hypothesis of schooling choice by using choice models, and Carneiro, Heckman and Vytlacil (2011) also supports it by using semiparametric framework. The research question in this paper is whether the same selection pattern is observed in Japan. Former research generally supports the selection pattern by using a database rather than in Japan and the results are not consistent with what the Japanese government seems to think. They may think disparity of opportunities for enrolling in college is due to differences of household income as we can see in

suggested policy implications such as exempting students with lower household incomes from tuition fees and entrance fees in college.

Identification strategy is applying the switching regression model[1] in Wills and Rosen; how unobserved heterogeneity determining the probability of individual schooling choice can affect future income. We can see some researches applying the similar models and supporting the comparative advantage hypothesis rather than schooling choice in labor economics. For example, Gaag,Jacques and Vijverberg (1988), Ophem and Hans (1992) and Hartog and Oosterbeek(1993) support the same hypothesis of a working choice between public sector and private sector.

The database used in this paper is taken from the Japanese General Social Survey. I conduct estimation across male samples both in groups of younger and older birth cohort after World War Ⅱ separately, with considering the possibility of structural change. Younger birth cohort is from 1941 to 1962 and older birth cohort is from 1963 to 1989.

New findings in this paper imply hierarchy sorting in a group of younger birth cohort. In this sorting pattern, those who have unobserved heterogeneity affecting positively enrolling in college tend to gain lower earnings, conditioned on observables. That is called negative selection. On the other hand, those who have unobserved heterogeneity affecting negatively enrolling in college tend to gain higher earnings, conditioned on observables. That is called positive selection. Moreover, no selection pattern is observed in a group of older birth cohort. These results contradict to the comparative advantage hypothesis supported generally by former research.

This paper mainly consists of 7 parts; Introduction part in section 1, Institutional Background part in section2, Methodology part in section 3, Data part in section 4, Results part in section 5, Conclusion part in section 6, Acknowledgement part in section 7. Appendix and References are in the end of this paper.

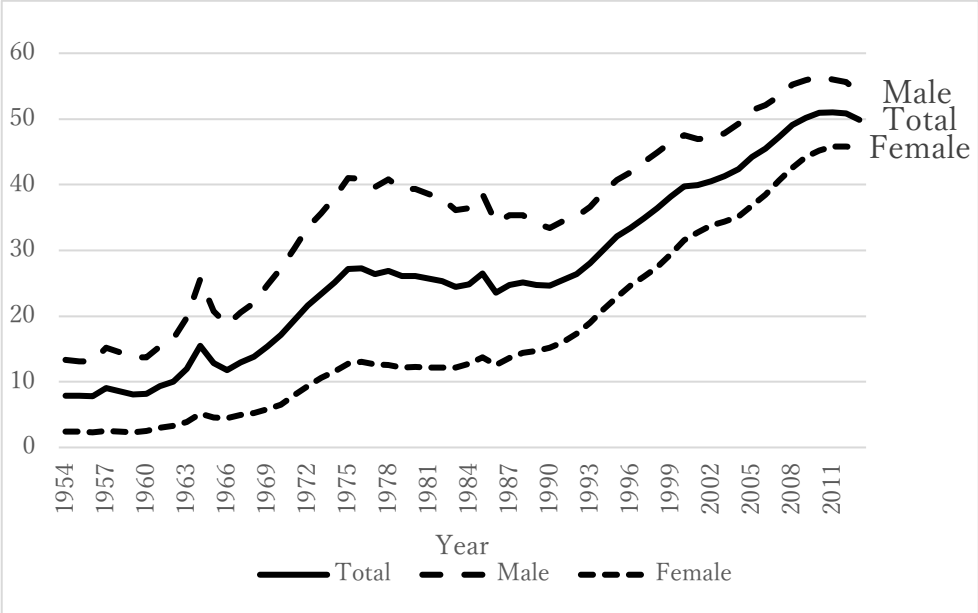
2. Institutional Background

The college entrance rate has been increasing since post ages of World War Ⅱ in Japan. In general, the trend is common in modern developed countries such as the United States and European Nations. As in figure 1, Japanese society had been facing the growth of college entrance rate by 42.6% during 1960～2012.

[1] For more efficient estimation, I conduct the switching regression by full information maximum likelihood, but it does not converge due to the complexity of models. Thus, the method in this paper is a classical approach of Heckman`s two-step method with no full information.

Figure 1

College Entrance Rate in Japan (%)



Source: Mukogawa Women`s University Institute for Education (Processed)

(http://www.mukogawa-u.ac.jp/~kyoken/data/13.pdf)

文部統計要覧昭和 31～41, 42～平成 13 年版

学校基本調査報告書昭和 40 年版

文部科学省統計要覧平成 14～25 年版

One of the factors of expanding college entrance rates is changing Japanese educational policies. With increasing demand for higher education, the Japanese Government have been expanding the capacity of college. For example, they have been admitting establishing newer private colleges and departments or faculties in each college. This trend has been mainly in the departments of human science in Japan. Senior high school students choose the course between STEM (Science, Technology, Engineering and Mathematics) and human science[2] (Non-STEM) when they are first grade students and learn in each course from second grade. This choice depends on one`s future career vision after graduating from college. For example, a student who wish to be a medical doctor in the future chooses STEM course in senior high school and applies for department of medicine in college. Another example is that a student who wishes to be a lawyer chooses a human science course in senior high

[2] Meaning of human science in this paper is non-STEM and that in general is different. STEM and human science in this paper only mean classification of different types of learning curriculum in high school, college entrance examination, and major in college.

school and applies for department of law in college generally. Expanding capacity in college is especially in departments that students who choose human science course in senior high school are the most likely to proceed to. Thus, choosing a human science course in senior high school may increase one`s probability to pass the college entrance examination. The evidence in Tanaka (2017) implies the selection pattern in choice of course in senior high school; those who have high unobserved heterogeneity in college choice are more likely to choose non-STEM majors in colleges, conditioned on observables including cognitive skills at junior and senior high schools.

Added to that, changing Japanese educational policies are curriculum and weekly school hours these policy changes compared to private school. In the background of these policy changes, cramming education and fiercely competitive college entrance examinations had been problematic in the 1900s after World War Ⅱ. Cramming education put emphasis on more knowledge in each subject of school and thus the volume of curriculum and scope of questions in college entrance examinations had been becoming larger. In Japan, school hours were set also in Saturday to catch up with this curriculum. Also, with growing college entrance rate, college entrance examinations had been becoming more competitive. Putting too much emphasis on cognitive ability and competitive college entrance examinations was considered to be a socially biased educational policy and also stressful for students. Then, more relaxed education and the importance of comprehensive abilities to live were social main interests. To address this problem, the Japanese Government had been changing educational policy from cramming education to a relaxed education gradually since 1980. In policy of relaxed education, weekly school hours and the quantity of curriculum had been reduced. For example, fully or partly no classes in Saturday. Public school basically followed these policy changes but private school was not always so. Added to that, periods for comprehensive study not limited for general curriculum have been introduced in Japanese elementary school, junior high school and senior high school. The purpose of this is to make students acquire thinking and acting subjectively through subjective activities.

In these changing social trends and policy changes, we may face the possibility of structural changes in demand for higher education among generations in Japan. Under past lower college entrance rate, those who enroll in college might have special characteristics and backgrounds; ability, household income or family backgrounds. These heterogeneities may cause disparity of schooling choice among students. However, the college entrance rate has been more than 50 % recently and increasing capacity of college and prevailing scholarship use may abbreviate the disparity. Moreover, policy change from cramming education to relaxed education may abbreviate the competition of college entrance examination and how one`s ability and family backgrounds affect possibility of

enrolling in college today may be different from that in past days. Under a competitive college entrance examination, one-dimensional ability such as cognitive skill may produce hierarchy sorting of schooling choice. On the other hand, though relaxed education may abbreviate competition and may give students more opportunities for thinking of their own career visions without too comparing their own ability with the other students.

Moreover, the structure may be different from that in other developed countries due to the uniqueness of Japanese education. Choice between STEM course and non-STEM course in senior high school is one of the characteristic educational systems in Japan. Under less capacity of STEM majors in college than non-STEM majors in Japan, students are more likely to choose non-STEM course in senior high school and that may decrease their mean earnings in the future. In this paper, I test whether the structure follows results generally supported in former researches and identify the true structure in Japan.

3. Methodology

The purpose of this model is to estimate the demand for higher education and identify the selection patterns of schooling choice. This methodology is basically followed that in Wills and Rosen. Under this framework, one`s optimal choice of whether enrolling in college or not is determined by two types of factors; future expected earnings and excluded instrument variables affecting only educational choice but not affecting future earnings directly. Then, individual discrete choice of schooling is as follows;

$$(2) \qquad S_i=1 \quad \text{if} \quad \log Y_{Ci} - \log Y_{Ni} > Z_i\beta_1 + u_{0i}$$

$$(3) \qquad S_i=0 \quad \text{if} \quad \log Y_{Ci} - \log Y_{Ni} \leqq Z_i\beta_1 + u_{0i}$$

where assume $u_{0i} \sim N(0,\sigma_u)$. In Wills and Rosen, a vector of Z_i affects one`s schooling choice through discount rate. In this paper, assume $Z_i$ be factors affecting schooling choice generally but not through life time earnings potential. This assumption is because we cannot affirm that elements of $Z_i$ affect one`s schooling choice through only one`s discount rate, and may affect through some possible factors such as budget constraint of each household. As in equation (1) of Introduction part, Y_{Ci} is the lifetime earnings for individual i choosing college, Y_{Ni} is that for choosing non-college, and S_i is an indicator of schooling choice; $S_i=1$ if choosing college and $S_i=0$ if choosing non-college.

Structural models of individual lifetime earnings are;

(4) $\log Y_{ci} = X_i\gamma_1 + \varepsilon_{Ci}$

(5) $\log Y_{Ni} = X_i\gamma_2 + \varepsilon_{Ni}$

where assume $\varepsilon_{Ci} \sim N(0,\sigma_{C\varepsilon})$ and $\varepsilon_{Ni} \sim N(0,\sigma_{N\varepsilon})$. A vector of X_i affects lifetime earnings potential for individual i. By exclusion restriction, X_i and Z_i must have elements that are not in common. Then, the reduced probit model at first stage is

(6) $S_i = 1 \text{ if } W_i\beta + u_i > 0$

(7) $S_i = 0 \text{ if } W_i\beta + u_i \leqq 0$

Denote $W_i = [X_i\ Z_i]$ and $u_i = \varepsilon_{Ci} - \varepsilon_{Ni} - u_{0i}$.

Then, one`s probability of proceeding to college is

(8) $\Pr(S_i = 1|W_i) = \Phi(W_i\beta/\sigma_u)$

Here, one`s annual wage[3], conditioned on years of working experience, for each schooling choice is

(9) $\ln w_{Ci}(exp_i) = X_i\gamma_1 + f_1(exp_i) + \varepsilon_{Ci}$

(10) $\ln w_{Ni}(exp_i) = X_i\gamma_2 + f_2(exp_i) + \varepsilon_{Ni}$

$\varepsilon_{Ci} \sim N(0,\sigma_{C\varepsilon})$ and $\varepsilon_{Ni} \sim N(0,\sigma_{N\varepsilon})$. Define exp_i as years of working experience after graduating from school one enrolled in finally. Then, one`s expected earnings, conditioned on exp_i, for each schooling choice are

(11) $E[\ln w_{ci}(exp_i)|S_i=1,W_i] = X_i\gamma_1 + f_1(exp_i) + \rho_C\sigma_{C\varepsilon}\lambda_{Ci}$

(12) $E[\ln w_{Ni}(exp_i)|S_i=0,W_i] = X_i\gamma_1 + f_2(exp_i) + \rho_N\sigma_{N\varepsilon}\lambda_{Ni}$

This method is switching regression and the inverse Mill`s ratios are

[3] In Wills and Rosen, equations of initial earnings and growth rate of annual earnings are estimated separately at second stage by using a panel dataset. However, I estimate equations of annual earnings at a cross section due to the restriction that JGSS is a cross-sectional dataset.

(13) $\lambda_{Ci}=\varphi(W_i\beta/\sigma_u)/\Phi(W_i\beta/\sigma_u)$

(14) $\lambda_{Ni}=\varphi(W_i\beta/\sigma_u)/[1-\Phi(W_i\beta/\sigma_u)]$

Denote $\rho_C=\rho(u_i/\sigma_u, \varepsilon_{Ci}/\sigma_{C\varepsilon})$ and $\rho_N=\rho(u_i/\sigma_u, \varepsilon_{Ni}/\sigma_{N\varepsilon})$. If $\rho_C\sigma_{C\varepsilon}>0$ and $\rho_N\sigma_{N\varepsilon}>0$, then that supports the comparative advantage hypothesis, conditioned on arbitrary same observable covariates among college graduates and non-college graduates. Then, one`s sorting of schooling choice depends on two-dimensional unobservable heterogeneity. Under these conditions, the definition of comparative advantage in equation (1) is satisfied, conditioned on arbitrary same observable covariates among college graduates and non-college graduates. If "$\rho_C\sigma_{C\varepsilon}>0$ and $\rho_N\sigma_{N\varepsilon}<0$" or "$\rho_C\sigma_{C\varepsilon}<0$ and $\rho_N\sigma_{N\varepsilon}>0$", then that implies hierarchy sorting, conditioned on same observable covariates between college graduates and non-college graduates. Then, one`s sorting of schooling choice depends on one-dimensional unobservable heterogeneity.

In Japan, exogenous shocks such as an increasing number of colleges or college entrance rate has been the main attention after World War Ⅱ. Moreover, tuition fees and college entrance fees have been different among years of enrolling in college and prefectures in Japan. Added to that, uniform distribution of colleges among prefectures may be too strong assumption in Japan because colleges are concentrated in city such as Tokyo or Osaka and not in rural area. These factors should be taken into consideration as disparity of costs of enrolling in college among birth years and prefectures living until graduating from high school. I use birth year dummies and prefecture dummies at 15 years old at first stage probit estimation. Heckman, Stixrud and Urzua (2006) and Carneiro, Heckman and Vytlacil (2011) use cohort dummies for control in estimating choice model of schooling but not reginal dummies in that model. In this paper, I try to use prefecture dummies for controlling regional heterogeneities with birth year dummies, considering Japanese backgrounds such as uniform distribution of colleges among prefectures. Also, these variables may treat for changing unemployment rate affecting one`s schooling choice.

In Wills and Rosen, number of siblings, type of high school one enrolls in, religion[4], parent`s schoolings and occupations affect one`s schooling choice through discount rate but not through lifetime earnings potential. The effect of the number of siblings on one`s schooling choice through lifetime earnings potential is limited and exclusion restriction may not be so problematic. However, exclusion restrictions for the type of high school one enrolls in, parent`s schoolings and parents`

[4] Information of one`s religion is not reported in JGSS.

occupations may be problematic. If one enrolls in a high school that vocational training is included in curriculum, then the education may affect one`s lifetime earnings potential.

For children, one of the determinants of household income may be their parents` schoolings, and the possible causalities are supported by Becker`s human capital theory (1964), Spence`s signaling theory (1973), and Thurow`s job competition model (1975). In human capital theory, higher education increases one`s productivity in labor market. In signaling theory, one`s schooling is a signaling of innate ability and productivity in labor market. In job competition model, opportunities of hiring and OJT (On the Job Training) only enhance worker`s productivity and highly educated workers can obtain the opportunities more. In each theory, parents` schoolings can be proxies for quantity of assets in each household. On the other hand, parents` schoolings can be also proxies for children`s innate cognitive ability and future productivity enhanced by discipline in household as results of parent`s schoolings. Thus, the variables may affect one`s schooling choice through lifetime earnings potential. Then, these covariates may not satisfy exclusion restrictions. Similarly, parents` occupations may not satisfy exclusion restriction. One of the determinants of budget constraints in each household and one`s discount rate may be types of parents` occupations, and they may also affect one`s occupational choice at the same time. Occupational choice is correlated with one`s lifetime earnings potential and thus, the problem of exclusion restriction may not be abbreviated enough.

In this paper, I try to use variables with respect to household income at 15 years old as the excluded instrument variables, and assume parents` schoolings and occupations may affect both one`s schooling choice and future earnings. Parents` schoolings and occupations can be proxies for quantity of assets in each household in long term and then, income at 15 years old can be proxies for temporary shocks in assets. Though parents` income is correlated with children`s future income, discussed in labor empirical works recently, temporary shocks in parents` income would not affect their schooling choice through children`s lifetime earnings potential. Whether disparity of household income may affect one`s schooling choice has been discussed in modern Japan and this problem may seriously be connected with efficiency in discussed policy of exempting students with lower income from tuition fee or college entrance fee in college. For estimating directs effect of household income on one`s schooling choice, excluding indirect effects through lifetime earnings potential, using temporary shock in household income with backgrounds of parents to abbreviate the indirect effects can be an identification strategy.

Moreover, I try to use college entrance rate of students around each sample as excluded instrument variable. This may affect one`s schooling choice because behavior of students proceeding

to higher education around each one may be one of the determinants of one`s schooling choice. This effect may be interpreted as peer group effect discussed in educational economics or bandwagon effect discussed in behavioral economics. On the other hand, it may not affect future earnings directly.

The alternative model estimated at first stage is;

$$(15) \qquad S_i=1 \quad \text{if} \quad W_i\beta+\mu_p+\xi_t+u_i>0$$

$$(16) \qquad S_i=0 \quad \text{if} \quad W_i\beta+\mu_p+\xi_t+u_i\leqq 0$$

Let μ_p and ξ_t be prefecture dummy and time dummy at 15 year`s old. Household income at 15 years old is included in X_i, thus included in W_i. Then, the models estimated at second stage are

$$(17) \qquad \ln w_{Ci}(exp_i,t)= X_i\gamma_3 + f_1(exp_i) + v_t +\rho_C\sigma_{C\varepsilon}\lambda^*_{Ci} +\eta_{Ci}$$

$$(18) \qquad \ln w_{Ni}(exp_i,t)= X_i\gamma_4 + f_2(exp_i) + v_t +\rho_N\sigma_{N\varepsilon}\lambda^*_{Ni} +\eta_{Ni}$$

Denote $\lambda^*_{Ci}=\varphi[(W_i\beta+\mu_p+\xi_t)/\sigma_u]/\Phi[(W_i\beta+\mu_p+\xi_t)/\sigma_u]$,
$\lambda^*_{Ni}=\varphi[(W_i\beta+\mu_p+\xi_t)/\sigma_u]/\{1-\Phi[(W_i\beta+\mu_p+\xi_t)/\sigma_u]\}$ and also let v_t be time fixed effect at year t and assume $\eta_{Ci}\sim N(0,\sigma_{C\eta})$, $\eta_{Ni}\sim N(0,\sigma_{N\eta})$.

4. Data

The database in this paper is from the Japanese General Social Survey[5] (JGSS) in 2001, 2002, 2005 and 2006. Samples used in this analysis are limited for male samples to prevent complex structures of working choice for women after marriage. Added to that, samples who are enrolling in school or start to work later than 3 months after graduating from school of one`s enrolling in finally are excluded to treat for attenuation bias in years of working experience. The birth cohorts of full samples used in this research is from 1941 to 1986. In this paper, I compare the results of estimations between a group of younger birth cohort and older birth cohort. Younger birth cohort is from 1962 to 1986 and older birth cohort is from 1941 to 1961. In each group of birth cohort, both samples of

[5] The Japanese General Social Surveys (JGSS) are designed and carried out at the Institute of Regional Studies at Osaka University of Commerce in collaboration with the Institute of Social Science at the University of Tokyo under the direction of Ichiro TANIOKA, Michio NITTA, Hiroki SATO and Noriko IWAI with Project Manager, Minae OSAWA. The project is financially assisted by Gakujutsu Frontier Grant from the Japanese Ministry of Education, Culture, Sports, Science and Technology for 1999-2003 academic years, and the datasets are compiled with cooperation from the SSJ Data Archive, Information Center for Social Science Research on Japan, Institute of Social Science, The University of Tokyo.

college graduates and high school graduates are included. Definition of variables used in estimations are in Table 1. Moreover, variables of real wage at price in 2015, parents` schoolings and occupations, years of working experience are processed data from row variables in JGSS. Number of observations is 1030 in a group of younger birth cohort, including 422 samples of college graduates and 608 samples of high school graduates. Number of observations is 1162 in a group of older birth cohort, including 448 samples of college graduates and 714 samples of high school graduates.

5. Results

Tables of results are in appendix part. In this paper, estimated base model in table 4 and 7 is similar to that in Wills and Rosen. Controlled base model in table 5 and 8 contains time fixed effects and prefecture dummies[6] at first stage, compared to the base model in table 4 and 7. Moreover, another set of excluded instrument variables is used in alternative model controlled by time fixed effects and prefecture dummies at first stage, in table 6 and 9. All estimations at second stage are controlled by year dummies.

As shown in table 4, 5 and 6, selection terms in a group of younger birth cohort consistently imply negative selection in subsamples of college graduates and positive selection in subsamples of high school graduates. Negative selection here indicates those who have high unobservable heterogeneity in college choice equation tend to gain lower earnings than mean earnings for random samples in labor markets for college graduates, conditioned on observables. On the other hand, positive selection here indicates those who have low unobservable heterogeneity in college choice equation tend to gain higher earnings than mean earnings for random samples, conditioned on observables. However, these results are significantly robust in case of controlling prefecture dummies and time fixed effects at first stage, as shown in table 5 and 6. Comparing the results among base models in table 4, 5 and alternative model in table 6, we can see exogenous factors, such as the difference in number of colleges or unemployment rates among years and prefectures, affect selection patterns between schooling choice and future earnings in unobserved heterogeneity. In Japan, both the number of colleges and the capacity have been increasing since World War Ⅱ, and we can ignore biased distribution of colleges among prefectures, as discussed in section 2. Moreover, higher unemployment rates for non-college graduates may increase the demand for higher education to

[6] I try to estimate probit models including interaction terms of time fixed effects and prefecture dummies at first stage but they are not converged. Thus, I conduct estimations not including the interaction terms.

decrease one`s probability of facing unemployment in the future. The results may imply controlling such exogenous shocks affecting Japanese educational markets is necessarily for better estimation of selection patterns in unobserved heterogeneity. By controlling such exogenous covariates, we can see robustly negative selection in unobserved heterogeneity among subsamples of college graduates and positive selection in unobserved heterogeneity among subsamples of non-college graduates. The evidence supports hierarchy sorting hypothesis and then, the comparative advantage hypothesis is not supported.

Hierarchy sorting is caused by one-dimensional unobserved heterogeneity as discussed in Wills and Rosen. One of the interpretations to this evidence is sorting pattern by special skills high school graduates are more likely to have. In this interpretation, if one has absolute advantage on skills required more strictly in jobs that high school graduates are more likely to obtain, then he/she is more likely not to enroll in college. On the other hand, if one does not have the absolute advantage, then he/she is more likely to enroll in college. However, the skills enhance one`s productivity also in labor markets for college graduates. Though more accurate identification of components in unobserved heterogeneity may be impossible in this framework, we can guess skills causing such a sorting in schooling choice are correlated with types of non-cognitive ability.

In recent research, estimating performance of non-cognitive ability in schooling choice and the outcomes in labor market are main concerns, though effect of cognitive ability is a main concern in older researches. In Heckman, Stixrud and Urzua (2006), the evidence shows those with higher non-cognitive ability are more likely to enroll in college and gain higher wages, which seems in contradiction to the interpretation of results in this paper. However, indicator of non-cognitive ability in that paper is standardized average of person`s score on the Rotter and Rosenberg scales as a measure of non-cognitive skills; motivation, persistence, and self-esteem. Actually, non-cognitive ability includes many aspects and different types of non-cognitive ability may affect differently schooling choice and outcomes in labor market. Identifying more precisely how each type of non-cognitive ability affects differently schooling choice and outcomes in labor markets among components of non-cognitive ability is worth researching in the future for identifying selection pattern of schooling choice more precisely.

Another possible interpretation of the selection pattern is a chance factor in unobserved heterogeneity. Results in this paper imply observed factors such as grade at 15 years old may enhance one`s probability of enrolling in college. However, if a person proceeds to college by chance rather than by observed heterogeneities, then it may affect negatively on future earnings. On the other hand, if a person proceeds to college by observed heterogeneities rather than by chance, then

it may affect positively on future earnings in labor markets for college graduates. This interpretation is consistent with the evidence in this paper; negative selection in college graduates and positive selection in high school graduates.

Moreover, if one enrolls in college by a chance factor rather than observed heterogeneities in this interpretation, then it may cause over-education for the type of occupation one is assigned in labor market. Under an increasing number of college graduates, all of them are not necessarily assigned to jobs that college graduates are more likely to obtain. In Thurow`s job competition model (1975), one`s motivation to enroll in college is to obtain better opportunities of OJT enhancing one`s productivity, and more educated people have advantage on obtaining it generally. However, some of college graduates with relatively low ability among college graduates may not be able to obtain better opportunities of OJT college graduates generally obtain, due to the capacity of labor markets for college graduates under an increasing college entrance rate. On the other hand, if a person proceeds to college by observed heterogeneities rather than by chance in this interpretation, then it may cause under-education. In the same model, some of college graduates with relatively high ability among high school graduates may be able to obtain better opportunity of OJT high school graduates cannot obtain generally.

In Japan, though researches about over-education and under-education are less, Hirao (2016) estimates the effects of them on wages by using Japanese data. Due to the difficulty of obtaining Japanese data including objective indicators of over-education or under-education, he uses subjective indicators of mismatch between one`s schooling and occupation one obtains though it potentially has problem of attenuation bias. In this paper, the evidence implies over-education has negative effect on wages and under-education has positive effect on wages, and supports assumption of the job competition model. The results in my research are also consistent with job competition model in this interpretation.

Furthermore, if this selection pattern is partly or fully caused by Japanese specific characteristics, then the choice between STEM courses and human science courses in high school may be one of the factors. If one chooses a human science course, then it is easier to pass the college entrance examination due to larger capacity in college as discussed in Institutional Background in part 2. Though the difficulty in passing college entrance examinations of departments concerned with human science is different among colleges and the easiness discussed here is especially in colleges with lower rank, choosing a human science course may enhance one`s probability of passing college entrance examinations. However, the average wage for those who graduate from departments of STEM at college is higher than that of those who graduate from departments of human science at

college. These factors in Japan may cause negative selection of schooling choice for enrollers in college and the positive selection in high school graduates. This interpretation is partly consistent with the evidence in Tanaka (2017). In his paper, selection pattern between probability of choosing a STEM course and probability of enrolling in college is negative.

However, the sorting pattern discussed here is observed only in a group of younger birth cohort and we can see no robust selection pattern in a group of older birth cohort, as shown in table 7, 8 and 9. In table 7 and 8, the signs of selection terms imply negative among subsamples of college graduates and positive among non-college graduates but they are not significant. Furthermore, the signs of selection terms imply positive among subsamples of college graduates and negative among subsamples of non-college graduates. Identified selection patterns in a group of older birth cohort are not consistent.

As another part of findings in this paper, income at 15 years old has no significant effect on one`s proceeding to college in a group of younger birth cohort as shown in table 6, and has significantly positive effect in a group of older birth cohort as shown in table 9. These results imply a tradeoff between temporary shocks in household income and unobserved heterogeneity such as non-cognitive ability or a chance factor within each group of birth cohort. In this interpretation, people in the older generation are more likely to conduct one`s schooling choice by quantity of household income rather than unobserved heterogeneity. Contrarily, people in the younger generation are more likely to conduct one`s schooling choice by unobserved heterogeneity rather than quantity of household income. Behind this interpretation, improving scholarship system is a possible factor causing the structural change between younger generation and older generation. Scholarship system enables students with lower household income to enroll in college by relaxing borrowing restrictions. As a possible policy implication in this paper, the effects of exempting students with lower household income from tuition fees or entrance fees in college, on their enrolling in college, may be limited if the policy is enforced today. Moreover, the evidence implies a possibility that people not enrolling in college conduct their schooling choice by their absolute advantage on an unobserved factor related to the choice and suggested policies by the Japanese government may not necessarily enhance social welfare. However, proxies indicating income at 15 years old are based on subjective major in database used in this research, and using objective data is necessarily for abbreviating attenuation bias to identify precisely the effects of temporary shock in income on schooling choice.

6. Conclusion

By applying the switching regression model in Wills and Rosen, I identify the selection patterns of schooling choice to test the comparative advantage hypothesis of demand for higher education in Japan, both in a group of younger and older birth cohorts. In this paper, I control heterogeneities among prefectures and years, taking into account modern structure of market for higher education in Japan. Another contribution in this paper is conducting robustness checks by using another set of excluded instrument variables compared to Wills and Rosen. The evidence does not support the comparative advantage hypothesis in both groups of birth cohorts. Supported selection patterns are negative selection by unobserved heterogeneity across subsamples of college graduates and positive selection across subsamples of high school graduates in a group of younger birth cohort. These selection patterns imply hierarchy sorting by one-dimensional unobserved heterogeneity. On the other hand, no significant selection patterns are observed in groups of older birth cohort.

Acknowledgement

I would like to express my gratitude to my supervisor Prof. Ryuichi Tanaka of the Institute of Social Science at the University of Tokyo, for useful comments and engagement through the research process of this master thesis. Furthermore, I would like to thank Prof. Hideo Owan, Prof. Daiji Kawaguchi, and Assistant Prof. Ayako Kondo of the Institute of Social Science at the University of Tokyo, mainly for technical advice in labor economics seminar in Graduate School of Economics at the University of Tokyo. Moreover, I would like to express my appreciation to Assistant Prof. Hitoshi Shigeoka at Simon Fraser University and the brilliant students of Graduate School of Economics at the University of Tokyo, for many comments to improve my research. All of them contribute to my future career and I never forget their kindness throughout my life.

Appendix

Table 1

Definition of Variables

Variable	Definition
S	Dummy variable; College (S=1) or non-college (S=0)
Relatively Low (Grade)	Dummy; Grade at 15 is relatively low (base is far low)
Middle (Grade)	Dummy; Grade at 15 is middle (base is far low)
Relatively Above (Grade)	Dummy; Grade at 15 is relatively above (base is far low)
Far Above (Grade)	Dummy; Grade at 15 is far above (base is far low)
NR Grade	Dummy of non-reported grade (base is far low)
Specialized High School	Dummy of specialized high school
Manager (Father)	Dummy of father`s job at 15 (Categorizing occupations in JGSS)
Security (Father)	Dummy of father`s job at 15 (Categorizing occupations in JGSS)
Clerk (Father)	Dummy of father`s job at 15 (Categorizing occupations in JGSS)
Blue Color (Father)	Dummy of father`s job at 15 (Categorizing occupations in JGSS)
NR Occupation (Father)	Dummy of non-reported father`s job (Base is the other occupations)
High School (Father)	Dummy; Father`s final schooling is high school
Junior College (Father)	Dummy; Father`s final schooling is Junior college or technical college
College (Father)	Dummy; Father`s final schooling is college or more
NR Schooling (Father)	Dummy of non-reported father`s schooling (Base is junior high school)
Part Time (Mother)	Dummy; Mother`s job at 15 was part time job
Regular Employee (Mother)	Dummy; Mother was a regular employee at 15
Side Business (Mother)	Dummy; Mother did side business at 15
Manager (Mother)	Dummy; Mother was a manager of company at 15
Died (Mother)	Dummy; Mother had been died at 15
NR Occupation (Mother)	Dummy of non-reported mother`s job
High School (Mother)	Dummy; Mother`s final schooling is high school
Junior College (Mother)	Dummy; Mother`s final schooling is Junior college or technical college
College (Mother)	Dummy; Mother`s final schooling is college or more
NR Schooling (Mother)	Dummy of non-reported mother`s schooling (Base is junior high school)
Older Brothers	Number of older brothers

Older Sisters	Number of older sisters
Younger Brothers	Number of younger brothers
Younger Sisters	Number of younger sisters
Relatively Low (Income at 15)	Dummy; Income of household at 15 is relatively low (Base is far low)
Average (Income at 15)	Dummy; Income of household at 15 is average (Base is far low)
Relatively Above (Income at 15)	Dummy; Income of household at 15 is relatively above (Base is far low)
Far Above (Income at 15)	Dummy; Income of household at 15 is far above (Base is far low)
NR Income at 15	Dummy of non-reported income at 15 (Base is far low)
College Entrance Rate	College Entrance Rate of people around oneself
NR College Entrance Rate	Dummy of non-reported college entrance rate around oneself
Wage	Real annual earnings from main job at price in 2015 (ten thousand yen) Nominal wage=35 if it is less than 70, 85 if 70～100, 115 if 100～130, 140 if 130～150, 200 if 150～250, 300 if 250～350, 400 if 350～450, 500 if 450～550, 600 if 550～650, 700 if 650～750, 800 if 750～850, 925 if 850～1000, 1100 if 1000～1200, 1300 if 1200～1400, 1500 if 1400～1600, 1725 if 1600～1850, 2075 if 1850～2300, Nominal wage=Nominal wage if Nominal wage≧2300, Wage=Nominal wage/CPI, CPI; Consumer Price Index CPI (2015 based) is taken from e-stat of Statistics Japan.
exp	Years of working experience exp = reported year (last year) – year of graduation, exp=0 if exp<0
Academy	Dummy of Academy
Junior College	Dummy of Junior college or technical college
Graduate	Dummy of graduate school
Dropout	Dummy of dropout
λ_C	Selection term of choosing college
λ_N	Selection term of choosing non-college

Table 2

Descriptive Statistics

	Younger birth cohort			
	College		Non-College	
Variable	Mean	SD	Mean	SD
Relatively Low (Grade)	0.021	0.145	0.122	0.327
Middle (Grade)	0.123	0.329	0.215	0.411
Relatively Above (Grade)	0.135	0.342	0.056	0.230
Far Above (Grade)	0.118	0.324	0.028	0.165
NR Grade	0.595	0.492	0.549	0.498
Older Brothers	0.277	0.494	0.449	0.655
Older Sisters	0.294	0.550	0.342	0.595
Younger Brothers	0.403	0.616	0.355	0.570
Younger Sisters	0.355	0.526	0.354	0.564
Specialized High School	0.012	0.108	0.127	0.333
Manager (Father)	0.166	0.372	0.051	0.220
Security (Father)	0.019	0.137	0.008	0.090
Clerk (Father)	0.405	0.492	0.289	0.454
Blue Color (Father)	0.363	0.481	0.531	0.499
NR Occupation (Father)	0.009	0.097	0.036	0.187
High School (Father)	0.408	0.492	0.405	0.491
Junior College (Father)	0.066	0.249	0.026	0.160
College (Father)	0.289	0.454	0.069	0.254
NR Schooling (Father)	0.083	0.276	0.179	0.384
Part Time (Mother)	0.223	0.417	0.252	0.434
Regular Employee (Mother)	0.363	0.481	0.429	0.495
Side Business (Mother)	0.019	0.137	0.028	0.165
Manager (Mother)	0.017	0.128	0.008	0.090
Died (Mother)	0.014	0.119	0.016	0.127
NR Occupation (Mother)	0.002	0.049	0.015	0.121
High School (Mother)	0.569	0.496	0.464	0.499
Junior College (Mother)	0.140	0.347	0.038	0.191

College (Mother)	0.069	0.253	0.031	0.174
NR Schooling (Mother)	0.095	0.293	0.169	0.375
Relatively Low (Income at 15)	0.199	0.400	0.280	0.449
Average (Income at 15)	0.521	0.500	0.493	0.500
Relatively Above (Income at 15)	0.204	0.403	0.128	0.335
Far Above (Income at 15)	0.036	0.185	0.013	0.114
NR Income at 15	0.007	0.084	0.020	0.139
College Entrance Rate	0.501	0.175	0.413	0.134
NR College Entrance Rate	0.773	0.420	0.789	0.408
wage	525.625	234.585	414.432	186.951
exp	10.123	5.485	12.901	6.206
Academy	0.031	0.173	0.110	0.313
Junior College			0.176	0.381
Graduate	0.126	0.332		
Dropout	0.038	0.191	0.044	0.206
Observations	422		608	

Table 3

Descriptive Statistics

	Older birth cohort			
	College		Non-College	
Variable	Mean	SD	Mean	SD
Relatively Low (Grade)	0.029	0.168	0.078	0.269
Middle (Grade)	0.145	0.353	0.262	0.440
Relatively Above (Grade)	0.161	0.368	0.098	0.298
Far Above (Grade)	0.214	0.411	0.055	0.227
NR Grade	0.442	0.497	0.476	0.500
Older Brothers	0.502	0.811	0.836	1.143
Older Sisters	0.563	0.833	0.838	1.025
Younger Brothers	0.435	0.656	0.506	0.764

Younger Sisters	0.371	0.576	0.405	0.620
Specialized High School	0.020	0.140	0.174	0.379
Manager (Father)	0.152	0.359	0.035	0.184
Security (Father)	0.020	0.140	0.008	0.091
Clerk (Father)	0.328	0.470	0.223	0.416
Blue Color (Father)	0.408	0.492	0.611	0.488
NR Occupation (Father)	0.011	0.105	0.011	0.105
High School (Father)	0.239	0.427	0.185	0.388
Junior College (Father)	0.127	0.334	0.052	0.222
College (Father)	0.150	0.357	0.025	0.157
NR Schooling (Father)	0.109	0.312	0.210	0.408
Part Time (Mother)	0.083	0.276	0.109	0.312
Regular Employee (Mother)	0.397	0.490	0.550	0.498
Side Business (Mother)	0.018	0.133	0.032	0.177
Manager (Mother)	0.002	0.047	0.003	0.053
Died (Mother)	0.016	0.124	0.027	0.161
NR Occupation (Mother)	0.013	0.115	0.013	0.112
High School (Mother)	0.415	0.493	0.204	0.404
Junior College (Mother)	0.098	0.298	0.018	0.134
College (Mother)	0.033	0.180	0.004	0.065
NR Schooling (Mother)	0.094	0.292	0.209	0.407
Relatively Low (Income at 15)	0.281	0.450	0.346	0.476
Average (Income at 15)	0.442	0.497	0.398	0.490
Relatively Above (Income at 15)	0.219	0.414	0.092	0.290
Far Above (Income at 15)	0.022	0.148	0.010	0.099
NR (Income at 15)	0.007	0.082	0.017	0.129
College Entrance Rate	0.514	0.190	0.396	0.161
NR College Entrance Rate	0.708	0.455	0.752	0.432
wage	788.838	352.458	610.984	549.858
exp	27.850	5.385	32.366	5.357
Academy	0.011	0.105	0.057	0.233
Junior College			0.102	0.303

Graduate	0.060	0.238		
Dropout	0.042	0.202	0.032	0.177
Observations	448		714	

Table 4

Base Model

Younger birth cohort

	(1)	(2)	(3)
	Probit	**College**	**Non-College**
Relatively Low (Grade)	-0.150	0.092	0.027
	(0.419)	(0.104)	(0.099)
Middle (Grade)	0.510	-0.051	0.054
	(0.380)	(0.097)	(0.092)
Relatively Above (Grade)	1.321***	-0.040	-0.012
	(0.395)	(0.098)	(0.102)
Far Above (Grade)	1.563***	0.060	0.008
	(0.409)	(0.108)	(0.144)
NR Grade	0.674*	-0.543***	-0.200
	(0.366)	(0.140)	(0.197)
Older Brothers	-0.322***		
	(0.091)		
Older Sisters	-0.134		
	(0.094)		
Younger Brothers	-0.078		
	(0.085)		
Younger Sisters	-0.149		
	(0.097)		
Specialized High School	-1.335***		
	(0.240)		
Manager (Father)	0.509**		

	(0.247)
Security (Father)	0.750*
	(0.431)
Clerk (Father)	0.594***
	(0.202)
Blue Color (Father)	0.355*
	(0.199)
NR Occupation (Father)	-0.202
	(0.416)
High School (Father)	0.103
	(0.144)
Junior College (Father)	0.475*
	(0.256)
College (Father)	0.759***
	(0.198)
NR Schooling (Father)	-0.066
	(0.243)
Part Time (Mother)	-0.182
	(0.119)
Regular Employee (Mother)	-0.203*
	(0.108)
Side Business (Mother)	-0.392
	(0.269)
Manager (Mother)	0.161
	(0.427)
Died (Mother)	-0.024
	(0.335)
NR Occupation (Mother)	-0.952*
	(0.506)
High School (Mother)	0.360**
	(0.151)
Junior College (Mother)	0.719***

	(0.222)		
College (Mother)	0.307		
	(0.271)		
NR Schooling (Mother)	0.181		
	(0.245)		
exp		0.142***	0.067***
		(0.023)	(0.015)
exp²		-0.004***	-0.001**
		(0.001)	(0.001)
Academy		-0.235	0.103
		(0.205)	(0.066)
Junior College			0.108**
			(0.054)
Graduate		0.240***	
		(0.052)	
Dropout		-0.200	-0.054
		(0.146)	(0.098)
λ_C		-0.079	
		(0.071)	
λ_N			0.127*
			(0.069)
Constant	-1.370***	5.360***	5.381***
	(0.432)	(0.144)	(0.208)
Control	No	Yes	Yes
N	1030	422	608
R-sq		0.376	0.231

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Control variables include both time fixed effects and prefecture dummies.

Table 5

Controlled Base Model

Younger birth cohort

	(1)	(2)	(3)
	Probit	**College**	**Non-College**
Relatively Low (Grade)	-0.306	0.107	0.033
	(0.433)	(0.118)	(0.098)
Middle (Grade)	0.375	-0.054	0.049
	(0.390)	(0.111)	(0.091)
Relatively Above (Grade)	1.290***	-0.077	-0.033
	(0.411)	(0.111)	(0.099)
Far Above (Grade)	1.557***	0.014	-0.009
	(0.418)	(0.120)	(0.138)
NR Grade	0.553	-0.542***	-0.212
	(0.374)	(0.137)	(0.196)
Older Brothers	-0.324***		
	(0.100)		
Older Sisters	-0.116		
	(0.101)		
Younger Brothers	-0.034		
	(0.091)		
Younger Sisters	-0.128		
	(0.106)		
Specialized High School	-1.405***		
	(0.240)		
Manager (Father)	0.549**		
	(0.248)		
Security (Father)	0.644		
	(0.516)		
Clerk (Father)	0.656***		
	(0.206)		
Blue Color (Father)	0.428**		

	(0.198)		
NR Occupation (Father)	-0.366		
	(0.419)		
High School (Father)	0.106		
	(0.152)		
Junior College (Father)	0.500*		
	(0.265)		
College (Father)	0.745***		
	(0.211)		
NR Schooling (Father)	-0.080		
	(0.240)		
Part Time (Mother)	-0.184		
	(0.126)		
Regular Employee (Mother)	-0.193*		
	(0.116)		
Side Business (Mother)	-0.437		
	(0.299)		
Manager (Mother)	0.064		
	(0.458)		
Died (Mother)	-0.099		
	(0.330)		
NR Occupation (Mother)	-0.863		
	(0.588)		
High School (Mother)	0.382**		
	(0.154)		
Junior College (Mother)	0.754***		
	(0.238)		
College (Mother)	0.316		
	(0.290)		
NR Schooling (Mother)	0.160		
	(0.244)		
exp		0.140***	0.064***

		(0.022)	(0.014)
exp²		-0.004***	-0.001*
		(0.001)	(0.001)
Academy		-0.260	0.088
		(0.206)	(0.066)
Junior College			0.092*
			(0.055)
Graduate		0.234***	
		(0.051)	
Dropout		-0.199	-0.050
		(0.144)	(0.097)
λ_C		-0.146**	
		(0.067)	
λ_N			0.191***
			(0.061)
Constant	-1.324	5.433***	5.385***
	(0.807)	(0.146)	(0.203)
Control	Yes	Yes	Yes
N	1030	422	608
R-sq		0.382	0.239

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Control variables include both time fixed effects and prefecture dummies.

Table 6

Alternative Model

Younger birth cohort

	(1)	(2)	(3)
	Probit	**College**	**Non-College**
Relatively Low (Grade)	-0.162	0.113	0.018
	(0.444)	(0.154)	(0.105)
Middle (Grade)	0.438	-0.098	0.039
	(0.403)	(0.147)	(0.101)
Relatively Above (Grade)	1.264***	-0.152	-0.031
	(0.426)	(0.152)	(0.108)
Far Above (Grade)	1.504***	-0.110	-0.053
	(0.441)	(0.162)	(0.138)
NR Grade	0.866**	-0.685***	-0.251
	(0.397)	(0.187)	(0.199)
Older Brothers	-0.310***		
	(0.097)		
Older Sisters	-0.125		
	(0.099)		
Younger Brothers	0.002		
	(0.090)		
Younger Sisters	-0.127		
	(0.103)		
Specilaized High School	-1.471***	0.174	0.227***
	(0.303)	(0.152)	(0.075)
Manager (Father)	0.529**	-0.046	-0.158
	(0.265)	(0.105)	(0.104)
Security (Father)	0.669	-0.237	-0.181
	(0.499)	(0.168)	(0.196)
Clerk (Father)	0.692***	-0.044	-0.140**
	(0.220)	(0.088)	(0.069)
Blue Color (Father)	0.439**	-0.028	-0.123*

	(0.212)	(0.089)	(0.065)
NR Occupation (Father)	-0.287	0.306**	-0.026
	(0.406)	(0.149)	(0.144)
High School (Father)	0.116	0.055	0.002
	(0.152)	(0.060)	(0.057)
Junior College (Father)	0.475*	0.084	-0.047
	(0.272)	(0.094)	(0.100)
College (Father)	0.699***	-0.085	0.039
	(0.215)	(0.081)	(0.093)
NR Schooling (Father)	-0.090	0.228	0.009
	(0.244)	(0.141)	(0.088)
Part Time (Mother)	-0.194	0.085	0.021
	(0.127)	(0.074)	(0.051)
Regular Employee (Mother)	-0.216*	0.038	-0.104**
	(0.117)	(0.057)	(0.048)
Side Business (Mother)	-0.442	0.150	-0.083
	(0.322)	(0.139)	(0.153)
Manager (Mother)	-0.012	-0.099	0.143
	(0.461)	(0.141)	(0.196)
Died (Mother)	-0.145	-0.310	-0.138
	(0.330)	(0.342)	(0.148)
NR Occupation (Mother)	-0.884	0.525***	-0.122
	(0.580)	(0.092)	(0.320)
High School (Mother)	0.364**	0.079	-0.017
	(0.154)	(0.068)	(0.055)
Junior College (Mother)	0.691***	0.096	-0.143
	(0.243)	(0.093)	(0.112)
College (Mother)	0.316	0.134	-0.077
	(0.287)	(0.160)	(0.112)
NR Schooling (Mother)	0.145	-0.029	-0.009
	(0.251)	(0.156)	(0.095)
Relatively Low (Income at 15)	-0.032		

	(0.248)		
Average (Income at 15)	0.049		
	(0.241)		
Relatively Above (Income at 15)	0.225		
	(0.270)		
Far Above (Income at 15)	0.167		
	(0.388)		
NR Income at 15	-0.797		
	(0.557)		
College Entrance Rate	1.048**		
	(0.425)		
NR College Entrance Rate	-0.559***		
	(0.175)		
exp		0.136***	0.063***
		(0.023)	(0.015)
exp2		-0.004***	-0.001*
		(0.001)	(0.001)
Academy		-0.241	0.104
		(0.208)	(0.068)
Junior College			0.112*
			(0.058)
Graduate		0.236***	
		(0.055)	
Dropout		-0.163	-0.056
		(0.140)	(0.106)
λ_C		-0.273***	
		(0.097)	
λ_N			0.268***
			(0.073)
Constant	-1.447*		
	(0.825)		

Control	Yes	Yes	Yes
N	1030	422	608
R-sq		0.415	0.264

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Control variables include both time fixed effects and prefecture dummies.

Table 7

Base Model

Older birth cohort

	(1)	(2)	(3)
	Probit	**College**	**Non-College**
Relatively Low (Grade)	0.222	-0.077	-0.026
	(0.347)	(0.436)	(0.140)
Middle (Grade)	0.514*	0.105	0.017
	(0.305)	(0.330)	(0.130)
Relatively Above (Grade)	1.057***	0.255	0.230
	(0.314)	(0.326)	(0.142)
Far Above (Grade)	1.465***	0.271	0.338*
	(0.317)	(0.335)	(0.190)
NR Grade	0.561*	-0.184	0.148
	(0.296)	(0.522)	(0.176)
Older Brothers	-0.201***		
	(0.048)		
Older Sisters	-0.154***		
	(0.049)		
Younger Brothers	-0.204***		
	(0.061)		
Younger Sisters	-0.188**		
	(0.078)		
Specialized High School	-1.472***		

	(0.202)
Manager (Father)	0.511**
	(0.226)
Security (Father)	-0.034
	(0.402)
Clerk (Father)	0.275*
	(0.166)
Blue Color (Father)	-0.037
	(0.154)
NR Occupation (Father)	0.289
	(0.376)
High School (Father)	0.017
	(0.132)
Junior College (Father)	-0.052
	(0.187)
College (Father)	0.357
	(0.219)
NR Schooling (Father)	-0.069
	(0.209)
Part Time (Mother)	-0.390**
	(0.153)
Regular Employee (Mother)	-0.326***
	(0.099)
Side Business (Mother)	-0.835***
	(0.268)
Manager (Mother)	-1.018*
	(0.546)
Died (Mother)	-0.317
	(0.318)
NR Occupation (Mother)	-0.138
	(0.411)
High School (Mother)	0.486***

	(0.123)		
Junior College (Mother)	1.045***		
	(0.237)		
College (Mother)	0.980**		
	(0.418)		
NR Schooling (Mother)	-0.162		
	(0.212)		
exp		0.114*	0.136***
		(0.069)	(0.048)
exp²		-0.002	-0.002***
		(0.001)	(0.001)
Academy		-0.779	-0.250*
		(0.689)	(0.137)
Junior College			0.129
			(0.080)
Graduate		-0.085	
		(0.228)	
Dropout		0.039	-0.207
		(0.089)	(0.175)
λ_C		-0.125	
		(0.102)	
λ_N			0.017
			(0.085)
Constant	-0.535	4.984***	4.218***
	(0.341)	(1.017)	(0.783)
Control	No	Yes	Yes
N	1162	448	714
R-sq		0.068	0.059

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Control variables include both time fixed effects and prefecture dummies.

Table 8

Controlled Base Model

Older birth cohort

	(1)	(2)	(3)
	Probit	**College**	**Non-College**
Relatively Low (Grade)	0.318	-0.087	-0.031
	(0.373)	(0.431)	(0.139)
Middle (Grade)	0.700**	0.091	0.009
	(0.331)	(0.325)	(0.129)
Relatively Above (Grade)	1.321***	0.260	0.214
	(0.337)	(0.322)	(0.141)
Far Above (Grade)	1.674***	0.285	0.310*
	(0.342)	(0.330)	(0.183)
NR Grade	0.689**	-0.197	0.142
	(0.320)	(0.527)	(0.176)
Older Brothers	-0.181***		
	(0.049)		
Older Sisters	-0.131**		
	(0.052)		
Younger Brothers	-0.171***		
	(0.066)		
Younger Sisters	-0.138*		
	(0.081)		
Specialized High School	-1.552***		
	(0.202)		
Manager (Father)	0.530**		
	(0.235)		
Security (Father)	0.038		
	(0.398)		
Clerk (Father)	0.319*		
	(0.174)		
Blue Color (Father)	-0.041		

	(0.160)		
NR Occupation (Father)	0.248		
	(0.381)		
High School (Father)	-0.034		
	(0.137)		
Junior College (Father)	-0.004		
	(0.191)		
College (Father)	0.258		
	(0.239)		
NR Schooling (Father)	-0.111		
	(0.216)		
Part Time (Mother)	-0.416**		
	(0.168)		
Regular Employee (Mother)	-0.278***		
	(0.105)		
Side Business (Mother)	-0.816***		
	(0.276)		
Manager (Mother)	-0.893		
	(0.582)		
Died (Mother)	-0.355		
	(0.320)		
NR Occupation (Mother)	-0.149		
	(0.449)		
High School (Mother)	0.456***		
	(0.128)		
Junior College (Mother)	1.070***		
	(0.265)		
College (Mother)	1.063**		
	(0.448)		
NR Schooling (Mother)	-0.163		
	(0.215)		
exp		0.114*	0.139***

		(0.069)	(0.047)
exp²		-0.002	-0.002***
		(0.001)	(0.001)
Academy		-0.786	-0.253*
		(0.689)	(0.136)
Junior College			0.127
			(0.080)
Graduate		-0.083	
		(0.226)	
Dropout		0.027	-0.205
		(0.088)	(0.174)
λ_C		-0.091	
		(0.085)	
λ_N			0.061
			(0.074)
Constant	-0.596	4.950***	4.159***
	(0.604)	(1.025)	(0.779)
Control	Yes	Yes	Yes
N	1162	448	714
R-sq		0.066	0.060

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Control variables include both time fixed effects and prefecture dummies.

Table 9

Alternative Model

Older birth cohort

	(1)	(2)	(3)
	Probit	**College**	**Non-College**
Relatively Low (Grade)	0.357	-0.175	-0.039
	(0.384)	(0.469)	(0.145)
Middle (Grade)	0.722**	0.125	0.015
	(0.340)	(0.365)	(0.135)
Relatively Above (Grade)	1.224***	0.344	0.239
	(0.346)	(0.361)	(0.146)
Far Above (Grade)	1.592***	0.370	0.368**
	(0.352)	(0.367)	(0.187)
NR Grade	0.888***	-0.170	0.157
	(0.332)	(0.559)	(0.187)
Older Brothers	-0.175***	-0.175***	
	(0.052)	(0.052)	
Older Sisters	-0.126**	-0.126**	
	(0.053)	(0.053)	
Younger Brothers	-0.147**	-0.147**	
	(0.069)	(0.069)	
Younger Sisters	-0.126	-0.126	
	(0.085)	(0.085)	
Specilaized High School	-1.625***	-0.246	0.013
	(0.255)	(0.179)	(0.115)
Manager (Father)	0.479**	0.186	0.063
	(0.244)	(0.169)	(0.118)
Security (Father)	0.054	-0.386	0.136
	(0.429)	(0.577)	(0.133)
Clerk (Father)	0.200	0.155	-0.081
	(0.186)	(0.140)	(0.087)
Blue Color (Father)	-0.118	0.037	-0.012

	(0.170)	(0.137)	(0.076)
NR Occupation (Father)	0.275	0.564**	-0.049
	(0.385)	(0.223)	(0.170)
High School (Father)	-0.082	-0.119	0.027
	(0.140)	(0.104)	(0.085)
Junior College (Father)	-0.051	-0.009	-0.051
	(0.192)	(0.153)	(0.116)
College (Father)	0.210	0.019	0.187
	(0.245)	(0.144)	(0.119)
NR Schooling (Father)	-0.099	-0.124	-0.138
	(0.218)	(0.120)	(0.089)
Part Time (Mother)	-0.370**	0.012	-0.131
	(0.171)	(0.099)	(0.091)
Regular Employee (Mother)	-0.253**	-0.080	-0.055
	(0.109)	(0.080)	(0.059)
Side Business (Mother)	-0.710**	-0.596	-0.065
	(0.306)	(0.457)	(0.191)
Manager (Mother)	-0.611	0.418**	-0.537
	(0.535)	(0.198)	(0.454)
Died (Mother)	-0.331	0.181	-0.108
	(0.336)	(0.221)	(0.150)
NR Occupation (Mother)	-0.171	0.187	0.108
	(0.465)	(0.140)	(0.105)
High School (Mother)	0.422***	0.079	0.115
	(0.131)	(0.101)	(0.086)
Junior College (Mother)	0.909***	0.044	-0.050
	(0.259)	(0.142)	(0.209)
College (Mother)	1.051**	-0.054	0.374
	(0.440)	(0.277)	(0.237)
NR Schooling (Mother)	-0.182	0.060	0.144*
	(0.217)	(0.160)	(0.081)
Relatively Low (Income at 15)	0.757***		

	(0.187)		
Average (Income at 15)	0.713***		
	(0.187)		
Relatively Above (Income at 15)	1.176***		
	(0.212)		
Far Above (Income at 15)	1.210***		
	(0.426)		
NR Income at 15	0.029		
	(0.433)		
College Entrance Rate	0.898**		
	(0.361)		
NR College Entrance Rate	-0.510***		
	(0.151)		
exp		0.118*	0.138***
		(0.071)	(0.047)
exp2		-0.002	-0.002***
		(0.001)	(0.001)
Academy		-0.744	-0.267*
		(0.667)	(0.143)
Junior College			0.116
			(0.084)
Graduate		-0.085	
		(0.224)	
Dropout		0.000	-0.164
		(0.098)	(0.177)
λ_C		0.079	
		(0.101)	
λ_N			-0.042
			(0.091)
Constant	-1.276*	4.709***	4.249***
	(0.686)	(1.187)	(0.791)

Control	Yes	Yes	Yes
N	1162	448	714
R-sq		0.113	0.079

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Control variables include both time fixed effects and prefecture dummies.

References

[1] Robert J.Wills; Sherwin Rosen. "Education and Self-Selection." The Journal of Political Economy, Vol.87, No.5.Part 2: Education and Income Distribution (Oct., 1979), S7-S36.

[2] John Garden. "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable." Econometrica, Vol. 52, No.5 (Sep.,1984), pp. 1199-1218

[3] Pedro Carneiro, James J. Heckman and Edward Vytlacil. "Estimating Marginal Returns to Education." American Economic Review. 2011 Oct; 101(6): 2754-2781.

[4] David Card. "Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems" Econometrica, Vol. 69, No.5 (September, 2001), 1127-1160.

[5] James J. Heckman, Jora Stixrud, Sergio Urzua. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." Journal of Labor Economics, 2006, vol.24, no.3.

[6] James J. Heckman. "Sample Selection Bias as a Specification Error." Econometrica. Vol.47, No.1 (Jan., 1979), pp. 153-161

[7] Joop Hartog and Hessel Oosterbeek. "Public and private sector wages in the Netherlands." European Economic Review 37 (1993) 97-114. North-Holland.

[8] Hans van Ophem. "A Modified Switchng Regression Model for Earning Differentials Between the Public and Private Sectors in the Netherlands." The Review of Economics and Statistics, Vol.75, No.2 (May, 1993), pp. 215-224.

[9] Monazza Aslam, Greeta Kingdon. "Public-Private sector segmentation in the Pakistani labor market." Journal of Asian Economics 20 (2009) 34-49

[10] Jacques van der Gaag and Wim Vijverberg. "A Switching Regression Model for Wage

Determinants of a Developing Country." The Review of Economics and Statistics, Vol.70, No.2 (May, 1988), pp. 244-255.

[11] Arega D. Alene, V. M. Manyong. The effects of education on agricultural productivity under traditional and improved technology in northern Nigeria: an endogenous switching regression analysis. Empirical Economics (2007) 32:141-159.

[12] Michael Spence. "Job Market Signialing." Quartely Journal of Economics, Volume 87, Issue (Aug., 1973), 355-374.

[13] Jennie E. Brand and Yu Xie. "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." American Sociological Review 75(2) 273-302, American Sociological Association 2010.

[14] 田中隆一. "大学教育需要を考える", 日本労働研究雑誌 第 687 号（2017 年 10 月号）

[15] 平尾智隆. "若年労働市場における学歴ミスマッチ–教育過剰の発生とその賃金への影響–. 日本労務学会誌 Vol.17, No.2:4 – 18 (2016)

[16] 安井健吾, 佐野晋平. "教育が賃金にもたらす因果的な効果について–手法のサーヴェイと新たな推定–. 日本労働研究雑誌, No.588/July 2009.

[17] 馬欣欣. "正規と非正規の就業形態およびその賃金格差に関する日中比較." 大原社会問題研究所雑誌, No.601/2008.12

[18] 濱中淳子. "検証・学歴の効用." 勁草書房, 2013

[19] 松繁寿和. "大学教育効果の実証分析–ある国立大学卒業生たちのその後–."日本評論社,2004

[20] 有賀健, 林文夫（編）. "経済制度の実証分析と設計 第 1 巻 経済停滞の原因と制度 第 8 章 新規高卒者の労働市場." 勁草書房, 2007