博士論文

論文題目　Essays on the Economics of Education and Program Evaluation

(教育経済学及びプログラム評価に関する研究)

氏名　　菊地 信義

Acknowledgments

I would like to thank my advisors, Yuji Genda, Hidehiko Ichimura, and Yasuyuki Sawada for their valuable comments and suggestions. I am especially grateful to my main advisor, Hidehiko Ichimura, for his continuous support, guidance and encouragement. I am also grateful to my dissertation committee members, Hideo Akabayashi and Hideo Owan for their fruitful comments and suggestions. I also would like to thank Naoshi Doi, Yuki Hashimoto, Daiji Kawaguchi, Ayako Kondo, Tsunao Okumura, Shinpei Sano, Shinya Sugawara, Ryuichi Tanaka, and Kengo Yasui as well as seminar participants at the University of Tokyo, and participants of the 5th Applied Econometrics Conference at Osaka University, the 2011 HiStat/TCER/Tokyo Labor Economics Workshop Joint Conference on Early Formation of Human Capital at Hitotsubashi University, Kansai Labor Seminar (*Kansai Rodo Kenkyukai*) at Osaka University, the 16th Labor Economics Conference at Nihon University, the 2011 Japanese Economic Association Autumn Meeting at the University of Tsukuba, and the 2013 Japanese Economic Association Autumn Meeting at Kanagawa University for their helpful comments and suggestions. I am also benefited from stimulating discussions with my friends and colleagues at the University of Tokyo, especially Takeshi Aida, Tomoaki Kotera, Kazuteru Takahashi, and Hirokazu Matsuyama. Finally, I thank to my parents for their unwavering support throughout my life.

The fourth chapter of this dissertation is based on my research which was financially supported by the Grant-in-Aid for Japan Society for the Promotion of Science Fellows (DC2), Grant Number: 24·10408.

Contents

List of Figures

List of Tables

Chapter 1

Overview

This dissertation consists of three essays studying the effects of an education policy and the returns to schooling in Japan as an application of the recent developments in program evaluation methods.

Policy makers and scholars broadly agree about the importance of the government's role in providing opportunities for investment in human capital. This applies to developing as well as developed countries, the latter including Japan. Japan's public investment in education in recent years appears to reflect this idea; it has increased in line with the trends seen in other member countries of the Organisation for Economic Co-operation and Development (OECD). For example, OECD (2013) notes that public expenditure on educational institutions for all levels of education in Japan increased by 5% between 2008 and 2010. However, if we were to view it as a percentage of the gross domestic product, it accounts for only 3.6% compared with the OECD average of 5.4% in 2010. This figure is the lowest among OECD countries with comparable data (OECD, 2013). How much additional public investment is needed in the education sector? Is there room to raise the efficiency of existing education policies?

To evaluate the optimality of an education policy, estimating the effects of a policy or the returns to education induced by a policy change is a central task for researchers. However, the task becomes complicated if the effects are heterogeneous and people select their treatment status based on own effects. The last two decades have seen a major development in the econometric literature on program evaluation in terms of accounting for individual level heterogeneity (Heckman and Vytlacil, 2001b).

The remainder of this chapter is organized as follows. First, I define the treatment parameters of interest. Second, I briefly describe why such parameters cannot be estimated by the conventional methods, such as ordinary least squares (OLS). Finally, I provide a short introduction of the tree essays.

1.1 Parameters of Average Treatment Effects

1.1.1 Setup

Let consider a framework that has two potential outcomes.[1]

$$
\begin{aligned}
&Y_{1i} \text{ if } D_i = 1 \\
&Y_{0i} \text{ if } D_i = 0,
\end{aligned}
\tag{1.1}
$$

where D_i denotes the treatment status for individual i that takes 1 if treated and 0 if not treated. Y_{1i} denotes a potential outcome for individual i if treated, and Y_{0i} denotes a potential outcome for individual i if not treated. The key issue of the policy evaluation is the fundamental problem of causal inference (Holland, 1986). Using the notation 1.1, it can be written as,

$$Y_i = D_i Y_{i1} + (1 - D_i) Y_{i0}, \tag{1.2}$$

where Y_i is the observed outcome. The equation 1.2 implies that we only observe either Y_{1i} or Y_{0i}, but not both states for each individual. This potential outcomes model is the Neyman-Fisher-Cox-Rubin model in statistics (Fisher, 1935; Cox, 1972; Rubin, 1974; Neyman, 1990). The model can be related to the structural model in the literature of economics, such as the switching regression model introduced by Quandt (1972) and a type of Roy (1951) model that Heckman and Honoré (1990) discuss the identification. A linear-in-the-parameters model with a selection equation is an example,

$$Y_i = X_i \beta_0 + D_i X_i (\beta_1 - \beta_0) + D(U_{1i} - U_{0i}) + U_{0i}, \tag{1.3}$$

where,

$$
\begin{aligned}
Y_{1i} &= X_i \beta_1 + U_{1i} \\
Y_{0i} &= X_i \beta_0 + U_{0i},
\end{aligned}
\tag{1.4}
$$

and,

$$
\begin{aligned}
D_i &= 1 \ \ if \ \ D_i^* \geq 0, \\
&= 0 \ \ otherwise \\
D_i^* &= \mu_D(Z_i) - V_i,
\end{aligned}
\tag{1.5}
$$

[1]Following discussion depends on Heckman, LaLonde, and Smith (1999); Heckman and Vytlacil (1999, 2005, 2007a,b); Blundell and Dias (2009).

where, X and Z are observable characteristics of an individual i and (U_1, U_0, V) are unobservables. In the following, the subscript i is omitted to simplify the notation. One interpretation of the equation 1.5 is that,

$$
\begin{aligned}
D^* &= \mu_D(Z) - V \\
&= Y_1 - Y_0 - C(W) - U_C \\
&= X(\beta_1 - \beta_0) - C(W) + (U_1 - U_0) - U_C,
\end{aligned} \tag{1.6}
$$

where, $C(W)$ and U_C indicate some observable and unobservable costs respectively. The equation 1.6 represents that individuals select into treatment if sum of the observable net benefit: $\mu_D(Z) = X(\beta_1 - \beta_0) - C(W)$ and the unobservable net benefit: $-V = U_1 - U_0 - U_C$ is positive.

1.1.2 Parameters of Interest

Now, let consider the conventional average treatment effects parameters that widely analyzed in the literature of policy evaluation. Using the notation 1.1, these parameters are defined as follows, the average treatment effect (ATE): $E(Y_1 - Y_0)$, indicates the average effect on the population, the ATE on treated (ATT): $E(Y_1 - Y_0|D = 1)$, indicates the average effect for the subpopulation who selecting into treatment, and the ATE on untreated (ATU): $E(Y_1 - Y_0|D = 0)$, indicates the average effect for the subpopulation selecting no-treatment. In the third chapter, I also consider the marginal treatment effect (MTE) introduced by Björklund and Moffitt (1987) and Heckman and Vytlacil (1999),

$$
MTE(V) = E(Y_1 - Y_0|V = v).
$$

Heckman and Vytlacil (1999, 2001b, 2005) show that the average effects parameters defined above can be recovered as weighted averages of the MTE. Heckman and Vytlacil (1999, 2001a) prove that the local instrumental variable (LIV) identifies the MTE,

$$
LIV = \frac{\partial}{\partial p} E[Y|P(Z) = p],
$$

where, $P(Z)$ denotes the probability of participation in the treatment or the propensity score. The LIV represents a limit form of the local average treatment effect (LATE) proposed by Imbens and Angrist (1994),

$$
\begin{aligned}
LATE &= \frac{E[Y|Z = z] - E[Y|Z = z^{'}]}{P(z) - P(z^{'})} \\
&= E[Y_1 - Y_0|D(z) = 1, D(z^{'}) = 0].
\end{aligned} \tag{1.7}
$$

The equation 1.7 indicates the average effect for those who are induced to change their treatment status from no-treatment: $D(z^{'}) = 0$ to treatment: $D(z) = 1$ by a discrete change in the value of Z, such that from $z^{'}$ to z.

1.2 Biases in the Conventional Method

In this section, I discuss that what type of biases the OLS estimates include when estimating the ATE parameters. Using the equations 1.1 and 1.2, I consider the following regression model,

$$Y = Y_0 + D(Y_1 - Y_0). \tag{1.8}$$

If I assume that $\beta = \beta_1 = \beta_0$ and suppress X to simplify the equation 1.3, the equation 1.8 is equivalent to the following simple regression model,

$$\begin{aligned} Y_i &= \beta + \alpha_i D_i + U_{0i} \\ &= \beta + \alpha D_i + \epsilon_i, \end{aligned} \tag{1.9}$$

where, $\alpha_i = U_{1i} - U_{0i}$ and $\epsilon_i = (\alpha_i - \alpha)D_i + U_{0i}$.

The equation 1.8 implies that the OLS estimates equal estimating the difference $E[Y|D = 1] - E[Y|D = 0]$. Due to the fundamental problem of causal inference, we can only observe $E[Y_1|D = 1]$ and $E[Y_0|D = 0]$ from the data, therefore,

$$E[Y|D = 1] - E[Y|D = 0] = E[Y_1 - Y_0|D = 1] + (E[Y_0|D = 1] - E[Y_0|D = 0]) \tag{1.10}$$

$$\begin{aligned} &= E[Y_1 - Y_0] + (E[Y_1 - Y_0|D = 1] - E[Y_1 - Y_0]) \\ &+ (E[Y_0|D = 1] - E[Y_0|D = 0]). \end{aligned} \tag{1.11}$$

The equation 1.10 indicates that if the ATT is the targeted parameter, the OLS estimate includes the bias: $E[Y_0|D = 1] - E[Y_0|D = 0]$, which is a selection bias. The equation 1.11 shows that if the ATE is the targeted parameter, the OLS estimate also includes the bias: $E[Y_1 - Y_0|D = 1] - E[Y_1 - Y_0]$, which is the sorting gain suggested in Heckman and Vytlacil (1999, 2001a, 2005). A positive sorting gain is the benefit for those who select into treatment because they expect their higher gain than the average. For example, in the equation 1.9, it implies that $\alpha_i > \alpha$ in ϵ. It is not natural to assume that the biases are zero if individuals self-select their treatment status based on own effects, such as schooling decisions. This suggests the necessity of using a program evaluation method to estimate the parameters of policy interest.

1.3 Introduction of the Essays

In this section, I introduce the three essays in the following chapters. These essays intend to estimate the causal effects of education in Japan, allowing for people self-select their treatment status based on their heterogeneous effects. The essays deal with levels of education from the compulsory schooling to the intergenerational transmission.

In the second chapter, I investigate how the reduction of instructional time affects educational attainment, using the revision of the Japanese curriculum standards in 1981 as a quasi-experiment. Although instructional time is considered an important input for the education production function, there is limited consensus on its causal effect on later outcomes. This is because of the difficulty of estimation without relying on cross-country variation or on before-and-after comparison. By using a feature of the centralized Japanese public educational system, this study estimates the effect of the revision in junior high schools as a difference-in-differences estimator. The revision is unique because it reduces the total school teaching hours by 445, which corresponds to about 13% of the previous standards, leaving the length of school weeks or the educational system unchanged. The main results show that the revision decreases schooling by about 0.5 years and the probability to enroll in high school by about 3 to 4% for women. These results are statistically significant and robust to controlling for the birth cohort or regional effects.

The third chapter examines the returns to university education in Japan using tuition, availability of universities, and labor market conditions as instruments. The magnitude of the returns to schooling has been a long lasting research interest of economists from Becker (1975), Becker and Chiswick (1966), and Mincer (1974). The returns to schooling have been estimated using data from many countries and time periods, and previous studies have found that schooling has significant impacts on a lot of grounds of later outcomes.[2] However, estimating the causal effects of education is not a simple task since the decision on an additional schooling is endogenously determined based on the individual level heterogeneity. One way to deal with the endogenous schooling decision is to use the IV approach. In this study, a set of education policy relevant instruments allows estimating the marginal effects for individuals who are induced to enroll in university by different marginal policy changes. Using the estimated MTE, this paper recovers the ATE parameters. The main empirical result shows that an additional year of university education increases the hourly wage by about 6.74% for the population. This study further finds that policies increasing the probability of university enrollment, such as free tuition and an increase in local capacities of universities, bring about positive effects of university education.

Finally, in the fourth chapter, I study the intergenerational effects of education in Japan, using a nonparametric bounds approach. The educational levels of parents are considered as key

[2]For example, Card (1999) provides a comprehensive survey for pecuniary outcomes. See also Oreopoulos and Salvanes (2011) for a survey of effects on non-pecuniary outcomes. Eide and Showalter (2011) provide a survey for the recent studies of the relation between health and education.

factors in explaining their child's educational success. Empirical studies of social science have long been interested in quantifying the magnitude of the relation, and they have often found a significant positive correlation between the parents' and child's schooling. Unfortunately, a positive intergenerational correlation does not necessarily reflect a causal relation. If the identification assumption on the exogenous selection of treatments is not valid, the OLS estimates provide a biased magnitude of the causal effect. Rather than imposing the strong assumptions required to obtain point estimates, this study derives bounds depending on relatively weak semi-monotonicity assumptions on treatment response, selection, and IVs. A combination of these three assumptions provides informative bounds on the ATE of parents' schooling on the child's schooling. The main results show that the tightest lower bounds suggest positive causal effects of parents' schooling, but the tightest upper bounds on the effects are lower than the point estimates that rely on the assumption of exogenous selection of parents' schooling. These results suggest that simple regression overestimates the true causal effect of parents' education.

References

Becker, Gary S. 1975. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research: distributed by Columbia University Press.

Becker, Gary S. and Barry R. Chiswick. 1966. "Education and the Distribution of Earnings." *American Economic Review* :358–369.

Björklund, Anders and Robert Moffitt. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *Review of Economics and Statistics* 69 (1):42–49.

Blundell, Richard and Monica Costa Dias. 2009. "Alternative Approaches to Evaluation in Empirical Microeconomics." *Journal of Human Resources* 44 (3):565–640.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, vol. 3, edited by Orley. C. Ashenfelter and David. Card, chap. 30. Amsterdam: Elsevier, 1801–1863.

Cox, David R. 1972. "Regression Models and Life Tables." *Journal of the Royal Statistical Society. Series B (Methodological)* 34 (2):187–220.

Eide, Eric R. and Mark H. Showalter. 2011. "Estimating the Relation between Health and Education: What Do We Know and What Do We Need to Know?" *Economics of Education Review* 30 (5):778–791.

Fisher, Ronald A. 1935. *The Design of Experiments.* London: Oliver & Boyd.

Heckman, James J and Bo E. Honoré. 1990. "The Empirical Content of the Roy Model." *Econometrica* :1121–1149.

Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics*, vol. 3, Part A, edited by Orley. C. Ashenfelter and David. Card, chap. 31. Amsterdam: Elsevier, 1865–2097.

Heckman, James J. and Edward Vytlacil. 1999. “Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects.” *Proceedings of the National Academy of Sciences* 96 (8):4730–4734.

———. 2001a. “Local Instrumental Variables.” In *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya*, edited by Cheng Hsiao, Kimio Morimune, and James L Powell. New York: Cambridge University Press, 1–46.

———. 2001b. “Policy-Relevant Treatment Effects.” *American Economic Review* :107–111.

———. 2005. “Structural Equations, Treatment Effects, and Econometric Policy Evaluation.” *Econometrica* 73 (3):669–738.

———. 2007a. “Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation.” In *Handbook of Econometrics*, vol. 6, Part B, chap. 70. Amsterdam: Elsevier, 4779–4874.

———. 2007b. “Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments.” In *Handbook of Econometrics*, vol. 6, Part B, chap. 71. Amsterdam: Elsevier, 4875–5143.

Holland, Paul W. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81 (396):945–960.

Imbens, Guido W. and Joshua D. Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62:467–475.

Mincer, Jacob. 1974. *Schooling, Earnings and Experience*. New York: National Bureau of Economic Research: distributed by Columbia University Press.

Neyman, Jerzy. 1990. “On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9.” *Statistical Science* 5 (4):465–472.

OECD. 2013. *Education at a Glance 2013: OECD Indicators 2013 – Country note for Japan*. Paris: OECD Publishing.

Oreopoulos, Philip and Kjell G. Salvanes. 2011. “Priceless: The Nonpecuniary Benefits of Schooling.” *Journal of Economic Perspectives* 25 (1):159–184.

Quandt, Richard E. 1972. “A New Approach to Estimating Switching Regressions.” *Journal of the American Statistical Association* 67 (338):306–310.

Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2):135–146.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5):688.

Chapter 2

The Effect of Instructional Time Reduction on Educational Attainment: Evidence from the Japanese Curriculum Standards Revision

2.1 Introduction

The publicly provided schooling at the compulsory education stage assures all children of an opportunity to learn fundamental knowledge or skills that affect their later life outcomes. As prior researches show, accumulating human capital at the earlier stages is more significant for higher educational achievement.[1] Moreover, additional schooling improves both pecuniary and non-pecuniary outcomes in adulthood.[2] Therefore, a useful educational system or policy has long been an interest of parents, educators, and policy makers, as well as economists. The classroom instructional time is one of the important educational factors with policy relevance, since it is more cost effective and simpler to change time allotment than to reform an entire educational institution.

Although there have been numerous studies about the effect of time spent in school on later achievement since the seminal paper of Card and Krueger (1992), evidence of the causal effect of instructional time itself is still limited. Because school time inputs are usually determined endogenously, it is difficult to control for an unobserved heterogeneity that correlates to the instructional time and outputs. Consequently, previous studies depend on cross-regional or

[1]See Cunha et al. (2006) as an example.

[2]For example, see Acemoglu and Angrist (2001), Harmon and Walker (1995), Oreopoulos (2006) for earnings Lochner and Moretti (2004) for criminal activities, Lleras-Muney (2005) for mortality rate, and Oreopoulos (2007) for subjective well-being.

cross-country variations of instructional time to estimate the effects.[3] Furthermore, most of these studies examine the impact of time on test scores, and less is known about the effects on later outcomes, which are arguably a greater concern.

Recent studies overcome this potential endogeneity problem of the length of school time by using educational system reform as a quasi-experiment. This study estimates the effect of a reduction of instructional time in Japan using the revision of the compulsory curriculum standards as a quasi-experiment. Among the significant relevant literature, Meghir and Palme (2005) and Pischke (2007) are the most closely related to this study, and are also valuable representatives of policy evaluation studies that directly estimate the treatment effect of compulsory schooling law reform as a difference-in-differences (DID) estimator.

Meghir and Palme (2005) use an expansion of compulsory schooling from 7 to 9 years in Sweden. The reform began from 1949 and was implemented gradually by municipality level. After the reform, the academic grades-based selection at the junior secondary school admission was abolished, enabling students to choose from one of three new courses in the same school, and the curricula were consolidated into one national curriculum. Meghir and Palme (2005) show that in the entire sample analysis, the revision significantly increases additional post-compulsory schooling by about 3% and years of schooling by 0.298 years. It is worth noting the larger effects they find for women. In the sub-sample analysis of women, the reform significantly increases additional post-compulsory schooling by about 5% and years of schooling by 0.339 years, clearly statistically significant effects.

Pischke (2007) shows the effect of short school years in the transition period for the reform that changed the start season of an academic year from spring to autumn in 1960s Western Germany. The short school years only affected the length of instructional weeks, but did not directly influence the highest grade completed or the curriculum. He reveals that about one-third reduction of instructional weeks per year for two school years significantly increases a grade repetition in primary education. In the census data analysis, the reform decreases total years of education by about 0.3 years, but he suggests that the effect on the years of education stems from the reduction of students who chose the longer academic track of secondary education.

Although a large number of researches in the literature estimated the effect of compulsory educational policy reforms, there are very few empirical studies that examine such an effect in Japan using individual data.[4] One of the exceptions is that by Kawaguchi (2011). He analyses the effect of a reduction of school days using a policy reform that changed all Saturdays to school holidays from 2002 in Japan. Using a Japanese time use survey, Kawaguchi (2011) shows that the reform reduces the studying time of junior high school students, and the effects are larger for those with less-educated parents. He also reveals this increased gap of studying

[3]For example, see Lee and Barro (2001), Wößmann (2003), Lavy (2010). Lavy (2010) also estimates the effect in Israel with the fixed effects approach. See Hanushek (2003) for a survey for the effects of school inputs.

[4]Oshio and Seno (2007) provide a survey of empirical studies in the economics of education using Japanese data.

time results in a widening gap in test scores among different parental education backgrounds. However, Kawaguchi (2011) only depends on the before-and-after policy change for the identification. Generally, a policy reform is implemented nationwide simultaneously, so that it is difficult to evaluate the effect without relying on before-and-after comparisons that may include other policy effects or macro-aggregate trends. One contribution of this study is that it estimates the effect of nationally provided educational regulation change as a DID estimator by using a feature of the centralized Japanese educational administration system.

In Japan, materials taught and teaching hours at elementary and secondary schools are regulated by the curriculum standards (*Gakushu shido yoryo*) issued by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), under the regulation of the School Educational Law (*Gakko kyoiku ho shiko kisoku*). Therefore, it is mandatory for all public schools to implement a uniform curriculum. The standards have been revised eight times from 1947 to 2012, with the fifth revision significantly reducing standard teaching hours because the previous curriculum was considered to have too much instructional time. However, private schools are not obligated to obey the standards and therefore, the effects of revision are limited to public schools.

This study analyses the greatest change in junior high school. The revision reduced school teaching hours overall by 445, which corresponds to about 13% of the previous standards. A principal reduction was implemented in the four compulsory subjects and one semi-compulsory elective subject (foreign languages), that were subjects of the entrance examinations in general. I estimate the effects of the standards revision on public school graduates through the DID method using private school graduates to capture the upward trend of educational attainment. The results show that the reduction of the instructional time decreases the years of schooling by about 0.5 years and the probability of upper secondary school enrollment by about 3 to 4% for women.

This study is the one of the first studies that estimates the causal effects of instructional time change caused by a reform of compulsory schooling regulations on later outcomes in Japan.[5] This study also contributes to the literature by evaluating a curriculum reform that changed total instructional time without changing the main source of variation in the previous studies such as school seasons, the length of school weeks, and educational systems.

The remainder of this chapter is organized as follows. Section 2 describes the 1981 Japanese curriculum standards revision. Section 3 and 4 explain the data and the estimation method, respectively. Section 5 shows and discusses the results. Section 6 presents a robustness check of the main results. Finally, Section 7 presents the conclusions.

[5]Oshio et al. (2010) analyze the subjective evaluation of a curriculum revision by parents. Nakamura (2011) also considers the same revision and data that this study uses. In the estimation, with implicit additional assumptions that private schools equally reduced expected teaching hours and that there was no transition period, he shows that the effect on educational attainment for public school graduates is smaller for women who lived in a prefecture that had a government-ordinance-designated city (*seirei shitei toshi*).

2.2 The 1981 Curriculum Standards Revision

Since 1947, the Japanese School Education Act (*Gakko Kyoiku Ho*) has implemented a nine-year compulsory schooling period,[6] consisting of a six-year elementary (primary) school period and a three-year junior high (lower secondary) school period. After completing compulsory schooling, students can attend a three-year (senior) high school or a five-year technical college (*Koto senmon gakko*) for upper academic secondary education. High school graduates can enroll into a four-year college (university) or a two-year junior college for higher education. For technical college graduates, entering into the third grade at university is another option in addition to the high school graduates' choices. A Japanese school year basically begins in April and is completed in March.

Japanese academic schools are established as public (national; *koku-ritsu* and municipal; *ko-ritsu*) and private (*shi-ritsu*) schools.[7] For admission into a post-compulsory school, students are required to pass an entrance examination, such as a paper-based test in the subjects of study or an interview. To enter national or private elementary and junior high schools, it is necessary to pass a selective examination. Entrance examinations are basically administered by each institution.[8] For municipal high schools, each municipality provides an entrance examination. In Japan, a school admission is mainly determined by performance on entrance examinations. The performance in a former school and extracurricular activities are less considered in the admission process.

The curriculum standards (the course of study; *Gakushu sido yoryo*) are curriculum guidelines for teaching issued by the MEXT. The standards specify materials taught with the amount of instructional time for all subjects for elementary, junior high and high schools. The standards are enforced under the School Education Law (*Gakko kyoiku ho shiko kisoku*). The ministry also publishes a supplemental commentary to the curriculum standards (*Gakushu sido yoryo kaisetsu*) that prescribes the materials taught, the method of teaching and the instructional time in detail.[9]

Even though the curriculum standards were revised eight times from 1947 to 2012, the strict provisions provided almost no room for discretion of teachers and schools until the time for integrated studies (*Sogo-teki na gakushu no jikan)* were regulated in the 2002 revision. Especially, the former standards of the fifth revision were considered the so-called "cramming education"

[6]In Japan, the period of compulsory schooling is defined by school years, not by age.

[7]Municipal schools are schools established by a municipality, such as city, town and village. Municipal schools also contain prefecutural schools and schools established by an agency of a local government.

[8]Public and some private universities also require applicants to take a national standardized test as a preliminary examination.

[9]The commentary is not legally binding in nominal.

(*tsumekomi kyoiku*), wherein the educational contents or the amount of instructional time were too much. This excessive quantity of teaching hours gave an opportunity for substantial instructional time reduction in the fifth standards revision in 1981, which this study analyses.

The fifth revision of the curriculum standards for elementary and junior high school was announced in 1977 after the Central Council for Education had proposed a reduction of the contents of learning in 1976. Since the fifth revision, the ministry has provided a transition period for new standards. In the transition period from the 1978 school year to the school year of full implementation, the new standards were carried out in stages for students who had already enrolled. The full implementation started from the 1980 school year in elementary schools and from the 1981 school year in junior high schools. In the revision, the biggest change was made in the junior high school standards. In the following, I describe the unchanged and changed aspects of the junior high school standards.

First, I explain the unchanged points. There are three main unchanged points. First, the number of standard annual school weeks remained as 35 weeks. Second, the class unit time was kept as 50 minutes. The unchanged schooling term and class unit time guaranteed that the revision only influenced the educational contents and the instructional time, but not the fundamental educational systems or the school seasons. Third, the total number of teaching hours for art, music, and moral education was also unchanged. This suggests that the revision basically reduced the time for the main subjects that were subjects of the entrance examination and thus strongly related to later educational achievement.[10]

Second, I describe the changed parts. The fifth revision substantially reduced the amount of standard instructional time for all the junior high school years (Table 1). The total reduction in school hours over three years was 385 class units (from 3,535 to 3,150 units), that is, by about 11% of the last standards' total school hours. The school teaching hours were reduced by 445 class units (from 3,385 to 2,940 units), which corresponded to about 13% of the last standards' total teaching hours and were equivalent to more than 120 hours of reduction per grade.[11] Moreover, about 70% of the total reduction was made in the main four subjects and an elective subject (foreign language).[12]

As I explained above, there was a large reduction of the total instructional time in the fifth standards revision. However, this regulation change was only limited to public junior high schools, and private junior high schools did not have to observe the revision.[13] This quasi-

[10]The main subjects are Japanese language, mathematics, science, and social studies.

[11]The school teaching hours are school hours excluding hours for special activities. The hours of special activities are time for extracurricular activities, such as home room activities and school events.

[12]Students could choose one subject from English, French and German as a foreign language.

[13]The legal extent of the curriculum standards is unclear. Since 1958, the Minister of Education has promulgated the curriculum standards. The legal binding force of the standards to the public schools was confirmed by the Supreme Court judgment of the administrative disposition cancellation (the so-called "Denshu high school teacher disciplinary dismissal") incident in 1990. However, the judicial precedent did not provide any interpretation of the legal binding force for private schools, and thus it is possible to consider that the standards legally do not affect private schools. Therefore, many private schools are considered to be not according to the standards, but an exact

experimental situation provides an opportunity to quantify the effect of instructional time reduction on final educational achievement.

2.3 Data

2.3.1 The Japanese Panel Survey of Consumers

A data set that can be used to estimate the causal effect in this study must have two essential properties. First, it must include information on the respondent's completed schooling attainment. Second, it must include information on the establishment type of the junior high school from which the respondent graduated. This study uses the Japanese Panel Survey of Consumers (JPSC) by the Institute for Research on Household Economics. The JPSC has surveyed a nationwide representative sample of women every year since 1993. The survey began with 1,500 women aged between 24 and 34 years in 1993 (Cohort A). Samples were expanded with 500 women aged 24 to 27 years in 1997 (Cohort B), and 836 women aged between 24 to 29 years in 2003 (Cohort C).[14] Until today, the surveys are conducted between October 1 and October 31 every year. From the JPSC, I make a pooled cross-sectional data set with the first time survey of these three cohorts A, B, C, as randomly sampled three cross-sectional data sets.

In the JPSC, I can identify that the respondent graduated from a public or a private junior high school. This is an advantage that only the JPSC has, as other individual data available in Japan do not include information about the establishment type of school from which the respondent graduated. Unfortunately, in the analysis data, I am unable to differentiate between national school graduates and municipal school graduates because the data set contains only two types of responses regarding the establishment type of junior high schools: Public (national and municipal) or Private. Hence, the analysis data include both national and municipal school graduates under public junior high school graduates.

National junior high schools are subject to the curriculum standards but are also affiliated to a faculty of a national university and play an important role in the research on educational science and teaching practices. For example, a national junior high school affiliated to a faculty of education could offer a class on developing or practicing an experimental teaching method. It may affect the magnitude of effects of the revision on national junior high school graduates and bias the estimates of the treatment effects. However, I assume that the bias from the inclusion of national junior high school graduates is limited because the proportion of national junior high school students is very small. According to the School Basic Survey, an official statistical report published by the MEXT, the proportion was less than 1% of all junior high school students.

---

enforcement situation is not certain.

[14] In the JPSC, information for men is also available as husband information for women who got married. However, I do not use this data set for men since the male respondents are not randomly sampled.

The JPSC also includes information about family background including the number of siblings, the level of the last school attended by the father and mother, the father's occupation, and the prefecture where the respondent spent the longest period of time during compulsory education.[15] Other available information includes whether the respondent attended a cram school (*juku*) in addition to the regular school education.[16]

For the outcome variable, this study considers years of schooling, additional schooling as measured by binary outcomes (whether the respondent enrolls in a university, i.e. a four-year college, and whether the respondent has some schooling beyond the compulsory level). Since I am unable to obtain the years of schooling directly from the JPSC, I therefore assign a general number of education years to each degree and reduce the general years by one year for those who dropped out before finishing their last school.[17]

However, the JPSC has the following five limitations: (1) it has no information about the parent's past income; (2) the sample size is significantly small compared to the previous studies that used census or other administrative data; (3) it does not include questions about the individual's abilities, for example, academic performance or IQ score, which are included in the analysis of Meghir and Palme (2005); (4) it does not identify the respondent's exact birth month and year; and 5) the available information about paternal occupation does not always reflect the occupation at the time the respondent was in junior high school. I explain these five limitations in more detail below.

First, the income of parents has a correlation to expenditures outside the school and is also important information from the perspective of budget constraints on the decision of undertaking additional schooling. However, Meghir and Palme (2005) and Pischke (2007) also do not control for the parental income, but instead control for the number of years of parental education. I follow this substitution.

Second, I am unable to cope with the issue of small sample size because the JPSC is the only individual data set that includes information on the establishment type of the respondent's junior high school.

Third, in Japan, before the national achievement test held in 2007, there was no uniform criteria-based national academic indicator for students under the secondary educational level. In addition, students had no obligation to take an IQ test. Therefore, this analysis is unable to

[15]The father's occupation is his current occupation if he has not yet retired or his previous occupation otherwise. No information is available for maternal occupation.

[16]The cram school attendance also includes the usage of private tutoring because the data set does not differentiate using a cram school from using a private tutoring.

[17]The general number of education years is 9 years for junior high school graduates, 12 years for high school graduates, 14 years for junior college or technical college graduates, 16 years for university graduates, and 18 years for graduate school graduates. However, because of the form of the questions about parents in the JPSC, I assign 16 years for graduate school in the parental years of schooling variable. In addition, parental degree of education in Cohort A includes drop outs. The results with an alternative definition of parental education are provided in the section for the robustness check. If a respondent was a student on the survey date, I assume she completed the degree. When I assume such a respondent is a drop-out, the main results are basically same.

control for the respondent's abilities and might include the so-called ability bias. However, in Section 2.4.2, I show that there is no ability bias caused by a systematic movement of potential new students between public and private junior high schools.

Fourth, the JPSC only asks about the age on October 1 in each survey year. Therefore, I must calculate the birth year as 'the year for each survey − the answered age'.[18] To avoid misclassification by the incomplete birth year information, this study limits the sample used for analysis. I describe the sample restriction method in detail in the following section with the classification of the age (birth year) groups that were affected by the standards revision.

Fifth, the problem regarding the father's occupation information also stems from a question item of the JPSC. The JPSC asks about the current father's profession if the father is still working at the time of survey, whereas it asks about the past occupation (when he was working) if he has already retired. Thus, the available occupational information may not correspond to the father's job when the respondent was in junior high school. In particular, if an investment decision on the children's education affects the father's retirement decision, the paternal retirement timing is not random, and thus, the information on father's occupation in the JPSC is not appropriate for use in the analysis. However, since the JPSC has little information about home attributes, this study includes the father's occupation as a control variable for a robustness check and shows that there is no significant difference in the results.

2.3.2 Sample Restriction

This section explains the sample restriction. The two problems observed in sample restrictions are (1) transition period of the new curriculum standards and (2) classification of treatment cohorts by the guessed birth year. I take the following steps to restrict the sample. First, I exclude the individuals who were junior high school students during the transition period based on the nominal entrance school year. Second, I restrict the sample to adjust the misclassification caused by the guessed birth year.

This study estimates the effects of the fifth revision that was fully implemented from the 1981 school year, and thus needs two groups to analyze. One is the 'before-revision group' who have completed three years of junior high school under the last curriculum standards. Another is the 'after-revision group' who have enrolled after the fifth revision and were completely influenced by the new standards for all three years. The JPSC includes individuals who were born from 1959 to 1979. In Japan, students enter junior high school in April in the school year in which they turn 13 years old. Therefore, students who were born after April 1968 enrolled in junior high school after the full implementation of the new curriculum standards.

However, as explained in Section 2.2, the new curriculum standards were partially in effect under the transition period from the 1978 school year. The implementation during the transition

[18]Hereinafter, I call this calculated birth year as the 'guessed birth year'.

period is not clear because each prefecture, municipality, or school had different treatment situations. Based on this concern, I exclude the respondents who were born in 1963 and 1967 from the analysis sample, because these cohorts attended junior high schools during the transition period and partially received the treatment. Similarly, I drop the individuals who were born after 1975 from the analysis sample because people in these birth year groups were junior high school students during the transition period of the sixth revision of the curriculum standards announced in 1989. Ultimately, following the sample restriction, the before-revision group includes individuals born from 1959 to 1962, and the after-revision group includes those born from 1968 to 1974.

In the next step, I restrict the sample because the JPSC has no accurate information on the birth month and birth year. The JPSC only asks about the age on October 1 in each survey year but does not ask the exact birth month and year. Therefore, the guessed birth year calculated as "year for each survey − answered age" includes women who were born from October 2 to December 31 in the year previous to the guessed year in addition to those born from January 1 to October 1 in the precisely calculated year.[19]

I do not include women who guessed 1959 and 1968 as birth years in the analysis. The guessed 1959 birth year cohort also includes women who were born before March 1959. These women attended junior high school under the same standards as those who were born in the 1958 school year (April 1958 to March 1959). The 1958 birth year cohort studied for only two years under the standards that were in force from 1972, and their total instructional time was less than the amount for the before-revision group. The before-revision group studied for all three years under the 1972 standards, which had the longest instructional time. Similarly, I drop women who guessed 1968 as the birth year from the analysis sample because they were under a different treatment condition. Women from the 1967 birth year group received junior high school education only for two years after the full implementation of the 1981 standards.

Based on the above sample restrictions, the analysis sample includes women who enrolled before the revision, that is, years 1960 to 1962, and who enrolled after the revision, that is, 1969 to 1974 for the guessed birth year.

[19]For example, if a woman answered her age was 33 years on the Cohort A survey in 1993, the guessed birth year is calculated as "1993 − 33 = 1960". These 33 year-old respondents also include women who were born in 1959. For example, a woman who was born on October 2 in 1959 was still 33 years old on the survey date of October 1 in 1993.

2.4 Estimation Method

2.4.1 Estimation Model

To estimate the effect of the instructional time reduction by the curriculum standards revision, this study follows Meghir and Palme (2005) and Pischke (2007), and uses the DID method. I estimate an equation of the form:

$$y_i = \alpha + \delta D_i + \beta After_i + \theta Public_i + X_i\gamma + \omega_t + \phi_r + \epsilon_i$$

where y_i is the outcome, D_i is a dummy variable indicating the treatment status (i.e. $D_i = Public_i \times After_i$), $Public_i$ indicates those who graduated from a public junior high school, $After_i$ indicates those who enrolled in a junior high school after the standards were revised, X_i control for the observable characteristics of an individual (mother's and father's years of schooling, number of siblings, and a set of paternal job dummies),[20] and ω_t, ϕ_r are sets of dummies capturing the birth year cohort and regional effects.[21] ϵ_i is the unobservable error term. δ, the coefficient of D_i, provides the treatment effect of interest.

This model is estimated using ordinary least squares (OLS) for years of schooling, with the probit model for the discrete educational outcomes. In computing the standard errors, I allow for arbitrary correlation among public/private and birth year cohorts. To estimate the average treatment effect on treated (ATT) as the DID estimator with pooled cross-sectional data, for which the identification relies on the following two assumptions: both of them are under conditional on observed characteristics, (1) in the absence of the revision, the changes in the average outcomes of the public graduates and the private graduates would have followed a common trend before and after the revision; (2) there is no systematic movement of students between the treatment and the control group. Although, the assumption cannot be tested directly in general, this study partially tests the assumption that is most critical using an external data set.

[20]Father's job categories are defined as follows: (1) Agriculture, Forestry or Fishery, (2) Commercial, Industrial or Service business of small size, (3) Management, Freelance professional, Other professional, and Teacher, (4) Technical employee or Skilled employee, (5) Clerical employee, Sales and Service employee, and (6) Pay job at home, Without occupation, and Don't know or Missing.

[21]Based on the geographic location, I divide 47 prefectures into 11 regions: (1) Hokkaido, (2) Tohoku (Aomori, Iwate, Miyagi, Akita, Yamagata, Fukushima), (3) Kita-Kanto (Ibaraki, Tochigi, Gunma, Saitama), (4) Minami-Kanto (Chiba, Kanagawa, Yamanashi), (5) Tokyo, (6) Hokuriku and Shinetsu (Niigata, Toyama, Ishikawa, Fukui, Nagano), (7) Tokai (Gifu, Shizuoka, Aichi, Mie), (8) Kinki (Shiga, Kyoto, Osaka, Hyogo, Nara, Wakayama), (9) Chugoku (Tottori, Shimane, Okayama, Hiroshima, Yamaguchi), (10) Shikoku (Tokushima, Kagawa, Ehime, Kochi), and (11) Kyushu and Okinawa (Fukuoka, Saga, Nagasaki, Kumamoto, Oita, Miyazaki, Kagoshima, Okinawa).

2.4.2 Assumption for the Identification and Its Validity

This section discusses the validity of the assumption described in Section 2.4.1. A critical assumption for estimating the ATT parameter with pooled cross-sectional data is that there is no compositional change in the groups before and after the revision. Put it in another way, I assume that there is no intentional movement of potential students from public junior high schools to private junior high schools after the fifth revision of standards. For example, if a potential public junior high school student who had superior abilities or education-oriented parents transferred her from a public junior high school to a private junior high school, this provides an overly estimated treatment effect.

I confirm the validity of the assumption in three ways. First, I compare the parental years of schooling of public school graduates with those of private school graduates. The assumption implies that the relative achievement or abilities of private school students compared to those of public school students did not change in the later cohort. Unfortunately, data on abilities or skills, which are factors influencing potential outcomes, are not available and these are thus unobservable. Therefore, I use parental education as a proxy because, in general, parental education and children's unobservable abilities are highly correlated. If a systematic movement between public and private groups was caused by the revision, it also changed the distribution of parental education. Second, I check the time series trend of the enrollment rate of private junior high school. Third, I compare a change in the competitive ratio for private junior high school before the revision with that after the revision.[22] If education-minded families chose to enroll in a private school after the fifth revision had been implemented, there might have been a rise in the enrollment rate or in the entrance competition rate.

First, I check the distribution of parental schooling. It is clear from the descriptive statistics in Table 2.2 that the differences in parental years of schooling are not significant in both the father and mother, and that the distributions of public and private junior high school graduates are almost symmetrical. This feature is more visible when comparing histograms of the parental years of schooling (Figures 2.1 and 2.2). These histograms show no significant difference in the shape of the distribution of public and private school graduates. Moreover, both public and private school graduates only have a common increasing trend of the number of parents who graduated from high school. These results assure the validity of the assumption that children of more educated parents did not move from a public to a private school after the revision.

Second, I examine the pattern in the percentage of students who went to private junior high school around the revision.[23] If the revision of standards affects the decision to enroll in a private junior high school, the enrollment rate would have a discontinuity at the time of the revision. Figure 2.3 displays the private junior high school enrollment rates from the 1974 to

[22]The competitive ratio is the ratio of applicants to capacity at an entrance exam of a private junior high school.
[23]The number of private junior high schools in all of Japan was 551 in 1976, 547 in 1977, 551 in 1978, 552 in 1979, 548 in 1980, 550 in 1981 and 550 in 1982.

the 1985 school years. Vertical lines are placed at the school years of the announcement and the enforcement of the fifth revision. Both total and female enrollment rates have a smooth trend over the periods, and in particular, there is no consistent pattern around the vertical lines. Although there is an evidence of a drop in 1980, the change is small (less than 0.3%). Therefore, I cannot find any significant change when the revised standards were in force. Furthermore, there was no effect of the revision on private school enrollment rate even at the time of the announcement.

Finally, I investigate the change in the competitive ratio of private junior high schools. Figures 2.4 and 2.5 show the cumulative distribution of the change in the competitive ratio for private schools from the 1979 to 1980 school year (before the revision) and from the 1980 to 1981 school year (after the revision). If more children wanted to attend a private school after the revision, then more private schools would have an increase in their entrance competition rate. However, as the figure of the full sample shows, the ratio has actually declined (Figure 2.4).[24] In addition, I restrict the sample to the prestigious schools with a high deviation value in the entrance examination difficulty, and find that almost no school showed a rise in its competitive ratio (Figure 2.5). These results suggest that potential students of public junior high schools who had high academic achievement or education-oriented families did not switch to private schools after the revision.

2.5 Results

2.5.1 Effect on Total Years of Schooling

The estimated results of the effect of the curriculum standards revision on the years of schooling are shown in Table 2.3. The estimation method is OLS. In the table, "Treatment effect" corresponds to the "δ" in the estimation model above.

Column 1 shows that the revision decreases years of schooling by about 0.5 years. This result is statistically significant. Even if I control the birth year and the regional fixed effects, the coefficients show a small change (columns 2 to 4). In addition, the values are not significantly changed when I include the variable of paternal occupation; thus the results are robust to this additional control (column 5).

The revision has a significantly negative effect on the total years of education as a measure of the overall effect on educational attainment. Next, I determine how the standards revision affects the additional schooling decision in each educational level.

[24]The change in the ratio from 1979 to 1980 might reflect the lessening of competition in 1979, because there was a sharp decline in the number of elementary school graduates of the 1966 birth cohort. The fertility rate dropped sharply in 1966 due to the child avoidance during the year of the fire horse (*hinoeuma*). For further explanation of the child avoidance during the year of the fire horse, see Akabayashi (2008) and Rohlfs, Reed, and Yamada (2010).

2.5.2 University Enrollment

Table 2.4 summarizes the effect on university (four-year college) enrollment. I estimate the probit model in this analysis and report the marginal effects. Column 1 shows that the revision decreases the probability of university enrollment by about 4%. This negative effect is similar to that on years of schooling. However, the result is not statistically significant. Furthermore, when I control the fixed effects or the father's occupation, it makes the coefficients unstable (columns 2 to 5).

These results suggest that the influence of the reduction of teaching hours in the compulsory educational level does not reach the higher educational level. The negative effect on the total years of education in Table 2.3 stems from an impact in the earlier stages of education. Therefore, I estimate the effects on the academic upper secondary school enrollment in the next section.

2.5.3 Upper Secondary School Enrollment

I present the results of the upper secondary school enrollment in Tables 2.5 and 2.6. The probit model is estimated and the marginal effects are reported as well in Table 2.4.

Table 2.5 shows the 3 to 4% negative effect of the revision on high school enrollment. All estimates are statistically significant even when I control for fixed effects (columns 1 to 4). These statistical significances are different from the results of university enrollment. In addition, the statistical significance is robust when controlling for the father's occupation (column 5). These results suggest that an institutional change in junior high schools has a stronger effect on the decision of enrollment in the next educational stage. To confirm this interpretation, I show the effects on all upper secondary school enrollment decisions in Table 2.6.

In Table 2.6, the outcome variable takes unity if a respondent has some schooling in a high school or a technical college. That is, the results can be interpreted as indicating effects on decisions at all the academic post-compulsory school enrollment in Japan. The reduction of instructional time decreases upper secondary school enrollment by about 3% and the estimates are statistically significant as well, as shown in Table 2.5. Moreover, estimates are slightly smaller than the size of the coefficient of the effect on high school enrollment. These smaller effects suggest that the analysis is accurate. Therefore, the results show that the standards revision in the lower secondary school strongly influenced the enrollment in the next stage of the education such as in upper secondary school.

2.5.4 Discussion

To summarize the all estimated results, reduction in total instructional time decreases the probability of women's schooling beyond the compulsory education by about 3 to 4%. However, it has no statistically significant effect on university enrollment. Finally, all changes in educational attainment taken together translate into a decrease in total years of schooling by approximately 0.5 years.

I interpret the results of this analysis as follows. The instructional time reduction of junior high school in a compulsory education period mainly affects the formation of the academic skill attainment that is highly correlated to the individual's cognitive abilities. Therefore, the public school curriculum revision has a strong influence on the population with low motivation to study because their classroom studying time in their school occupies an important part in the overall training time for their cognitive skill development. It is also critical that the opportunities for accumulating cognitive skills decrease at a relatively early stage such as a period of the compulsory education.

This sensitive population is likely to be located at the lower tail of the education distribution.[25] A decline in accumulation of the cognitive skill at the early teens reduces the probability to enroll in an upper secondary school, which is a prerequisite for the higher education. This consequence further reduces the chance of accumulating cognitive abilities. A reduction in years of education also implies a negative effect on a chance to get more high-wage occupations.

In contrast, the top population is expected to potentially enroll in a university. The amount of time of studying in junior high school classes is not decisive of their high school enrollment. If students who are receiving the education at the time of the university enrollment decision mainly belong to the population that is not strongly affected by the reduction of time, there would no effect of the reduction on university enrollment. Based on this, the revision of curriculum standards has a statistically significant effect only on upper secondary schooling, and does not affect the university enrollment significantly.

This interpretation is similar to that suggested in the literature. For instance Pischke (2007) shows that the negative effect on the total years of education disappear when he controls the choice of secondary education track, which has different completion years. This suggests that the effect on the total years of education highly depends on the choice of courses for the secondary education. Moreover, one of the reasons for the weak instrument problem suggested in the previous studies also supports the validity of the interpretation. In the studies that use an institutional change of compulsory education system as an instrumental variable (IV), the IV has a strong impact only on students who finish their schooling at the compulsory education level. These students are at the very bottom tail of the education distribution; thus, the IV usually does not affect their college enrollment decisions.[26]

As described above, the results of this study were generally consistent with those of the previous studies. However, there is room for discussion about the size of the treatment effect.

[25]For example, Kawaguchi (2011) provides evidence that the compulsory school days reduction more significantly affects students with less-educated parents.

[26]See Black, Devereux, and Salvanes (2005), for example.

In the fifth revision of the curriculum standards used in this analysis, the total reduction of the instructional time was 445 school teaching hours over three years (about one-third year). This is a small change compared to those of previous studies that have changes totaling about one year. In spite of its relatively smaller change, this analysis shows an estimated effect larger than those of previous studies. There are two possible explanations for this result.

First, the change analyzed in this study was more influential to education attainment compared with those of previous studies. Previous studies use the variation of the nominal years of education and the length of instructional periods. It is not clear how these changes affect the instructional time and the quality or quantity of educational content. On the other hand, in the revision used in the present study, the instructional time is significantly reduced without changing the nominal years of education and instructional weeks. This had a strong negative impact on academic achievement at the post-compulsory school enrollment stage. In addition, as Meghir and Palme (2005) show, the effect of policy change has a stronger effect on women. Since the sample of this study includes only women, the estimated results can be larger than the estimates with a sample that includes men.

Second, the results of this study include biased estimates. In this analysis, there is a possibility of ability bias because I am unable to control for the abilities of individuals. However, as I show in Section 4, students with superior abilities did not move from public schools to private schools after the revision of the curriculum standards. Therefore, the first possibility is a more reasonable explanation than this one.

2.6 Robustness Check

In this section I examine the robustness of the main results presented in the results section. First, I investigate whether the revision changes cram school enrollment when an individual was a junior high school student. In addition, I show that the estimated results are not sensitive to the inclusion of an additional independent variable of cram school attendance when an individual was in an elementary school or in a junior high school. Second, I show that the estimated results are robust to the inclusion of additional controls of local labor market conditions when an individual was in a junior high school. Third, I investigate the effects of the revision in the transition period to show that the main results are not affected by the curriculum standards revision in elementary schools, which accompanied the revision in junior high schools. Fourth, I show that the main results are not sensitive to the exclusion of prefectures that experienced an exceptional high school selection system during the period under study. Fifth, I show that changing the specification of the parental education variables does not affect the previous conclusion. Finally, I show the statistical significance of the main results with different definitions of clustering of the standard errors. All the following analyses are based on the estimation model in Section 2.4.1.

First, I examine whether cram school enrollment during the junior high school period is affected by the revision. If the revision raises cram school enrollment as an additional investment in extra-school education, then studying in a cram school might compensate the effects of instructional time reduction in junior high school. Table 2.7 provides estimated results of the effect of the curriculum standards revision on cram school enrollment. The dependent variable is a dummy that takes unity if the individual attended a cram school when she was a junior high school student. The estimates suggest that the revision increases cram school enrollment, but the effect is statistically insignificant and the magnitudes of coefficient are unstable.

Although the results in Table 2.7 suggest that the revision does not significantly increase cram school enrollment rate, it is worth investigating an omitted variable bias problem of cram school enrollment. If cram school enrollment is endogenous and is affected by the revision, controlling for it captures a part of the effect of the revision, but excluding for it may provide biased estimates of the effect. Thus, I show that the additional control variable of cram school enrollment does not change the statistical significance or magnitude of the effect of the main results.

Table 2.8 reveals the effect of cram school enrollment in elementary school. The other explanatory variables are the same as those in the analysis in the previous section. For the years of schooling, cram school enrollment has a small negative effect. However, this direct effect on total education is not statistically significant. Moreover, the results of the treatment effect are similar those in Table 2.3 even if I control for the cram school enrollment (columns 1 to 3). The results from columns 4 to 6 provide evidence of the effect on the decision to enroll in high school. Cram school enrollment in elementary school has about 2% negative effect on high school enrollment. The magnitudes of the treatment effects are slightly larger than those in Table 2.5 if I include the cram school enrollment dummy, but the effects are still negative and statistically significant. These prove that the results provided in Tables 2.3 and 2.5 are robust to controls for cram school enrollment during the elementary school period.

Table 2.9 presents the results for cram school enrollment in junior high school. Columns 1 to 3 report the effect of cram school enrollment on years of schooling. Cram school enrollment in junior high school has a positive but statistically insignificant effect. The additional control of cram school enrollment does not affect coefficients of the treatment effect. As shown in columns 4 to 6, using a cram school increases high school enrollment by about 3 to 4%, and the estimates are statistically significant. Although the cram school enrollment control reduces the magnitude of the treatment effect, the changes are very small. These suggest that the main results in the previous section are robust to controls for cram school enrollment during the junior high school period.

Second, I show that the main results are robust to controlling for local labor market conditions. If labor market conditions changed across regions and affected parental income or raised the opportunity costs of post- compulsory schooling differently during the junior high school period, not controlling for such variations might exaggerate the negative effects of school hour reduction. To mitigate this concern, I re-estimate the main analysis while controlling for the active job opening to application ratio (*yuko kyujin bairitsu*) and the average monthly male earnings at the prefecture level. The data sources for these two variables are the report of the Employment Service Agency (*shokugyou antei gyoumu toukei*) and the Basic Survey on Wage Structure of the Ministry of Health, Labor and Welfare, respectively. These variables are averaged over the three years of junior high school. Table 2.10 shows that the main results are almost unchanged by these additional controls, except that the standard errors are a little smaller.

Third, I investigate the effects of the curriculum standards revision in junior high schools during the transition period. It shows that the main results are not affected by the accompanying curriculum standards revision in elementary schools. As described in the section 2.2, the fifth revision of the curriculum standards for public elementary schools was announced in 1977 and the full implementation started from the 1980 school year. In the analysis in the previous section, all students in the after-revision group (1969–74) were affected by the revision, therefore the estimated treatment effects might include the effects of revision in elementary schools.

In the main analysis, I assume that the effects of the revision in elementary schools are negligible for two reasons. First, the reduction in school hours was very limited. The total reduction of school hours over six years was 36 class units (from 5821 to 5785 units), that is, a reduction of about 0.6%. The reduction was equivalent to less than 6 hours per grade since the number of standard annual school weeks remained as 35 weeks and the class unit time, as 45 minutes. Second, the proportion of private elementary students was very small. According to the School Basic Survey, private elementary school students were less than 0.6% of the total elementary school students. It means that more than 99% of junior high school students attended a public elementary school under the same curriculum standards. Therefore, differences in elementary school hours across each birth year are captured by the controls for birth year effects.

Although the effects of this revision in elementary schools might be negligible, it is worth investigating the pure effects of the revision in junior high schools. However, it is difficult to estimate the pure effects directly because I am unable to define a proper control group. The data set does not indicate whether the respondents graduated from a public or a private elementary school. Furthermore, all cohorts in the after-revision group of the junior high school revision were also affected by the elementary school revision.

Therefore, I estimate the effects of the revision in the transition period instead of investigating the effects of the full implementation. To estimate the effects in the transition period, I consider individuals born from 1964 to 1965 as the new after group because they graduated from elementary school before the 1978 school year and enrolled in junior high school during the transition period of the junior high school revision. The new curriculum standards were partially in effect during the transition period, and thus, the school hours were reduced. If the reduction in school hours in public junior high schools has meaningful impact, one can expect some negative effects on educational achievement for public junior high school graduates. The magnitude of the effects is ambiguous since implementation during the transition period was not unified. The revision allowed each prefecture, municipality, or school to have different treatment situations in the transition period.

Table 2.11 reports the effects of the revision in junior high schools during the transition period. Panel A presents the effects of the revision on years of schooling. Columns 3 and 4 of the transition period show the negative treatment effects, and it can be observed that the size of the coefficients is smaller than the base results in columns 1 and 2. Panel B reports the marginal effects on high school enrollment probability.[27] As shown in columns 7 and 8, the negative treatment effects in the transition period are larger than the results from the base case specifications displayed in columns 5 and 6. To summarize the results, the revision in the transition period has stable negative effects as expected, but effects are statistically insignificant for the restricted sample size. The result in column 7 is statistically significant at the 10% level, which is an exception.

Fourth, I investigate the effect of the revision by excluding the effects of an exceptional selection system of high school entrance. One possible concern in relation to biased estimates of the effects of the revision is that some prefectures used an unusual high school entrance selection system during the period under study. The system was called the comprehensive selection (*sogo senbatsu*) system. The comprehensive selection system is one of the entrance examination systems for municipal high schools. The system was designed to ease the heated competition between students in the high school entrance examination and to correct disparities in educational levels between municipal high schools in the same school district.

In the usual selection system, junior high school graduates apply for a municipal high school in which they wish to enroll, and then, each high school chooses its students. In the comprehensive selection system, junior high school graduates cannot apply for each school in which they wish to enroll. They are assigned to a school district, or they apply for one of the groups of high schools in the school district. When applicants meet the entrance examination criteria, the system allocates applicants equally to each school in the school district or to each school in the group of schools. The system allocates applicants based on not only academic achievement, but also the distance from their residence to a school.[28]

In the period under study, in all, 15 prefectures experienced the comprehensive selection system fully or partially.[29] If the selection system directly affected high school enrollment,

[27]In Panel B, regional controls are excluded for the restricted sample size.

[28]See Minken Chuto Kyoiku Kaikaku Kenkyukai (1984) or Kurohane (1994) for a detailed description of the comprehensive selection system.

[29]The comprehensive selection system was implemented in following prefectures: Chiba, Tokyo, Fukui, Yamanashi, Gifu, Aichi, Mie, Kyoto, Hyogo, Okayama, Hiroshima, Tokushima, Nagasaki, Oita and Miyazaki.

a failure to control for the differences between systems might bias estimates of the effects of the revision. To mitigate this concern, I exclude prefectures that experienced the comprehensive selection system and re-estimate the main analysis. I do not use a dummy on the comprehensive selection system or a set of prefecture dummy variables, because the implementation of the system was different between prefectures and within each prefecture.

Table 2.12 shows the estimates without the prefectures that experienced the comprehensive selection system. Panel A presents the effects of the revision on years of schooling. Columns 4 to 6, without the comprehensive selection system, show the statistically significant negative treatment effects, and the magnitude of these effects is larger than the base results in columns 1 to 3. Panel B reports the marginal effects on high school enrollment probability. In columns 10 to 12, the treatment effects without the comprehensive selection system are similar to the base results displayed in columns 7 to 9. These results show that the statistically significant negative effects of the revision are robust to the exclusion of prefectures that experienced the comprehensive selection system. The results also suggest that the comprehensive selection system does not directly affect high school enrollment, but it slightly offsets the effects of the revision after high school enrollment.

Fifth, I examine whether the specification of the parental education variables matters with the main results. Table 2.13 shows the estimates with an alternative specification of the variables with the standard set of covariates in the previous section. In the last three columns in each panel, the parental education variable is defined as a dummy that equals one if a parent received education beyond the compulsory level. The first three columns of each panel reproduce the base case results for easy reference.

Panel A presents the effects of the revision on years of schooling. Columns 4 to 6 with the alternative specification show the significantly negative treatment effects, but the size of the coefficients is between 0.05 to 0.08, larger than the base results in columns 1 to 3. Panel B reports the marginal effects on high school enrollment probability. In columns 10 to 12, the alternative parental education controls do not change the results from the base case specifications, which are reproduced in columns 7 to 9. These results suggest that the specification change of the parental education variables does not affect the statistically significant negative effect of the treatment; thus the main results in the previous section are robust.

Finally, I discuss the standard errors. In practice, it is not always obvious how to define the clusters, and thus, it can be difficult to know the appropriate level at which to cluster, in this study. Therefore, I provide alternative definitions of clustering of the standard errors for a robustness check. Table 2.14 presents the Huber-White robust standard errors (column 1), and clustering robust standard errors clustered by the establishment type of the junior high school (column 2), by the establishment type of the junior high school and graduated from a junior high school before/after the revision (column 3), by the establishment type of the junior high school and birth year cohort (column 4), by the establishment type of the junior high school, birth year cohort and the region lived during junior high school (column 5) and by the establishment type of a junior high school, birth year cohort and the prefecture lived during junior high school (column 6). The establishment type of a junior high school is public or private. In the results section, I report the standard errors in the specification of column 4.

Panel A shows the results based on the specification of columns 1 and 4 of Table 2.3. In this case, the estimates are statistically significant from column 2 to column 4 even if I control for the fixed effects, but are not significant in the other columns. Panel B presents the results of high school enrollment with the specifications of columns 1 and 4 in Table 2.5. The estimates are almost statistically significant at the 10% level with in the Huber-White robust standard errors (column 1) and statistically significant at the 10% level under the definition of clustering that has the largest number of clustering groups (column 6) when I include fixed effect controls. As shown in the table, the statistical significance of the results is dependent on the definition of standard errors, and estimates with very small number of clusters might be biased. However, as shown in columns 2 and 3, if I drop the dimension of the clustering, the results are statistically significant at the 1% level. In addition, the standard errors are stable and small even in the specifications that cannot reject the null hypothesis. Considering the sample size is as small as about 1,100, the results are almost robust from the definition of standard errors.

2.7 Conclusion

This study studies the effects of the substantial reduction of instructional time on the educational attainment with the DID method that uses the fifth revision of curriculum standards in Japan as a quasi-experiment. The revision is unique since it reduces instructional time but does not change the main source of variation in the literature such as the length of school weeks, compulsory years of education, and fundamental educational systems.

In the main results, the 445-hour reduction of the total teaching hours over three years, which corresponds to about one-third year, decreases the years of education by about 0.5 years, and the probability of schooling beyond the compulsory education by about 3 to 4% for women. These results are statistically significant and larger compared with those of previous studies since the contents of the revision have a direct influence on the length of instructional time itself. In addition, it is necessary to note that the analysis sample of this study includes only women owing to the availability of the data set. As previous studies suggest, the revision might have a more significant effect for women.

The negative impact of the school instructional time reduction in the public education is important to later educational achievement. In particular, as the results of this study suggest, a change in compulsory education is more influential to children with families that have low educational motivation, since learning in school is critical for their academic skill development.

In this study, because the available data set has no information on parental income, I cannot analyze the role of parental income in detail. However, the heterogeneous effects of policy reform in compulsory education by degree of household poverty may be examined in further research.

References

Acemoglu, Daron and Joshua Angrist. 2001. "How Large are Human-Capital Externalities? Evidence from Compulsory-Schooling Laws." In *NBER Macroeconomics Annual 2000, Volume 15*, edited by Ben S. Bernanke and Kenneth Rogoff. Cambridge, MA: MIT Press, 9–74.

Akabayashi, Hideo. 2008. "Lives of the Firehorse Cohort: What the Statistics Show." *Japanese Economy* 35 (3):34–54.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review* 95 (1):437–449.

Card, David and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100 (1):1–40.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, vol. 1, edited by Eric A. Hanushek and Finis Welch, chap. 12. Amsterdam: Elsevier, 697–812.

Hanushek, Eric A. 2003. "The Failure of Input-based Schooling Policies." *The Economic Journal* 113 (485):F64–F98.

Harmon, Colm and Ian Walker. 1995. "Estimates of the Economic Return to Schooling for the United Kingdom." *American Economic Review* 85 (5):1278–1286.

Kawaguchi, Daiji. 2011. "Fewer School Days, More Inequality." Unpublished manuscript, Hitotsubashi University.

Kurohane, Ryoichi. 1994. "Koko Seisaku no Hensen to Senbatsu Seido (Transition of High School Policy and Selection System)." *The Journal of the Tokyo Institute for Municipal Research* 85 (3):15–25. (in Japanese).

Lavy, Victor. 2010. "Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries." Working Paper 16227, National Bureau of Economic Research, Cambridge, MA.

Lee, Jong-Wha and Robert J. Barro. 2001. "Schooling Quality in a Cross-Section of Countries." *Economica* 68 (272):465–488.

Lleras-Muney, Adriana. 2005. "The Relationship between Education and Adult Mortality in the United States." *Review of Economic Studies* 72 (1):189–221.

Lochner, Lance and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* 94 (1):155–189.

Meghir, Costas and Mårten Palme. 2005. "Educational Reform, Ability, and Family Background." *American Economic Review* 95 (1):414–424.

Minken Chuto Kyoiku Kaikaku Kenkyukai. 1984. "Koko no Sogo Senbatsu Sei Nyushi (The Comprehensive Selection System for High School Entrance Examination)." *Kokumin Kyoiku* 61:105–126. (in Japanese).

Nakamura, Ryosuke. 2011. "Yutori Kyoiku ga Kyoiku Tasseido ni Ataeta Koka no Jissho Bunseki (Empirical Analysis of the Effect of 'Yutori Education' on the Educational Attainment)." Discussion Paper 2011–14, Keio/Kyoto Global COE Discussion Paper Series. (in Japanese).

Oreopoulos, Philip. 2006. "Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter." *American Economic Review* 96 (1):152–175.

———. 2007. "Do Dropouts Drop Out Too Soon? Wealth, Health and Happiness from Compulsory Schooling." *Journal of Public Economics* 91 (11):2213–2229.

Oshio, Takashi, Shinpei Sano, Yuko Ueno, and Kouichiro Mino. 2010. "Evaluations by Parents of Education Reforms: Evidence from a Parent Survey in Japan." *Education Economics* 18 (2):229–246.

Oshio, Takashi and Wataru Seno. 2007. "The Economics of Education in Japan: A Survey of Empirical Studies and Unresolved Issues." *Japanese Economy* 34 (1):46–81.

Pischke, Jörn-Steffen. 2007. "The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years." *The Economic Journal* 117 (523):1216–1242.

Rohlfs, Chris, Alexander Reed, and Hiroyuki Yamada. 2010. "Causal Effects of Sex Preference on Sex-Blind and Sex-Selective Child Avoidance and Substitution across Birth Years: Evidence from the Japanese Year of the Fire Horse." *Journal of Development Economics* 92 (1):82–95.

Wößmann, Ludger. 2003. "Schooling Resources, Educational Institutions and Student Performance: The International Evidence." *Oxford Bulletin of Economics and Statistics* 65 (2):117–170.

Table 2.1: Number of School Hours Before and After the Fifth Curriculum Standards Revision

	(1)	(2)	(3)
Subject	1972 Standards (Before the Revision)	1981 Standards (After the Revision)	Change ((2) − (1))
Japanese Language	525	455	−70
Social Studies	455	385	−70
Mathematics	420	385	−35
Science	420	350	−70
Music	175	175	0
Art	175	175	0
Health and Physical Education	375	315	−60
Technology and Home Economics	315	245	−70
Moral Education	105	105	0
Special Activities	150	210	+60
Elective Subjects (Foreign Languages)	420	350	−70
Total School Hours	3535	3150	−385
Total School Teaching Hours (Excluding the Special Activities)	3385	2940	−445

Notes: This table reports the number of school hours (instructional time) before and after the fifth revision of the curriculum standards. The author calculated the numbers based on the Curriculum Standards Data Base of National Institute for Educational Policy Research. The standards define one school hour as 50 minutes. Column 1 shows hours of the previous standards that were in force from 1972. Column 2 shows hours of the new standards after the revision. The new standards were in force from 1981. Column 3 provides the amount of change of school hours by the revision. The numbers in the row of total school teaching hours are the sum of all school hours excluding hours for the special activities. The hours of special activities are time for extracurricular activities, such as home room activities and school events.

Table 2.2: Descriptive Statistics

Birth year:	Before the Revision 1960 - 62		After the Revision 1969 - 74	
	Public	Private	Public	Private
Years of schooling	12.87	12.19	13.13	13.00
	(1.56)	(1.92)	(1.74)	(2.15)
Father's years of schooling	10.96	11.06	11.82	11.85
	(2.36)	(2.64)	(2.44)	(2.74)
Mother's years of schooling	10.57	10.58	11.42	11.46
	(1.86)	(1.80)	(1.86)	(2.05)
Nuber of siblings	2.60	2.74	2.44	2.25
	(1.11)	(1.15)	(0.77)	(0.79)
Number of observations	323	31	694	67

Notes: This table reports the descriptive statistics. Before the revision indicates the statistics for individuals who graduated from a junior high school before the fifth revision of the curriculum standards. After the revision indicates the statistics for individuals who enrolled in a junior high school after the standards were enforced. Public and Private indicate the establishment type of the junior high school from which the individuals graduated. Standard deviations are in parentheses.

Table 2.3: The Effect of the Revision on Years of Schooling

	(1)	(2)	(3)	(4)	(5)
Treatment effect	-0.504	-0.493	-0.551	-0.538	-0.463
	(0.225)**	(0.240)*	(0.209)**	(0.221)**	(0.216)**
Birth year controls		yes		yes	yes
Regional controls			yes	yes	yes
Father's job controls					yes
Number of observations	1115	1115	1115	1115	1115

Notes: This table reports the OLS estimates of years of schooling. Treatment effect is the coefficient of a dummy variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively.

Table 2.4: The Effect of the Revision on University Enrollment

	(1)	(2)	(3)	(4)	(5)
Treatment effect	-0.037	-0.026	-0.040	-0.028	-0.017
	(0.054)	(0.058)	(0.056)	(0.058)	(0.050)
Birth year controls		yes		yes	yes
Regional controls			yes	yes	yes
Father's job controls					yes
Number of observations	1115	1115	1115	1115	1115

Notes: This table reports the marginal effects from the probit estimates of university or four year college enrollment (an indicator variable that is equal to one if an individual has ever attended a university). Treatment effect is the coefficient of a dummy variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively.

Table 2.5: The Effect of the Revision on High School Enrollment

	(1)	(2)	(3)	(4)	(5)
Treatment effect	-0.034	-0.031	-0.041	-0.038	-0.037
	(0.017)**	(0.014)**	(0.015)***	(0.016)**	(0.016)**
Birth year controls		yes		yes	yes
Regional controls			yes	yes	yes
Father's job controls					yes
Number of observations	1115	1115	1115	1115	1115

Notes: This table reports the marginal effects from the probit estimates of high school enrollment (an indicator variable that is equal to one if an individual enrolled in a high school after she graduated from a junior high school). Treatment effect is the coefficient of a dummy variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively.

Table 2.6: The Effect of the Revision on Post-Compulsory Schooling

	(1)	(2)	(3)	(4)	(5)
Treatment effect	-0.030	-0.027	-0.035	-0.032	-0.031
	(0.016)*	(0.013)**	(0.015)**	(0.015)**	(0.015)**
Birth year controls		yes		yes	yes
Regional controls			yes	yes	yes
Father's job controls					yes
Number of observations	1115	1115	1115	1115	1115

Notes: This table reports the marginal effects from the probit estimates of post compulsory school enrollment (an indicator variable that is equal to one if an individual enrolled in a high school or a technical college after she graduated from a junior high school). Treatment effect is the coefficient of a dummy variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively.

Table 2.7: The Effect of the Revision on Cram School Enrollment

	(1)	(2)	(3)
Treatment effect	0.063 (0.091)	0.034 (0.096)	0.043 (0.087)
Birth year controls		yes	yes
Regional controls		yes	yes
Father's job controls			yes
Number of observations	1101	1101	1101

Notes: This table reports the marginal effects from the probit estimates of cram school enrollment (an indicator variable that is equal to one if an individual enrolled in a cram school when she was a junior high school student). Treatment effect is the coefficient of a dummy variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively.

Table 2.8: Estimates with Cram School Control in Elementary School

Dependent Variable:	Years of Schooling (OLS)			High School Enrolment (Probit)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	-0.490	-0.530	-0.455	-0.036	-0.043	-0.042
	(0.231)**	(0.229)**	(0.221)*	(0.017)**	(0.018)**	(0.018)**
Cram school enrolment	-0.048	-0.074	-0.094	-0.020	-0.022	-0.020
in elementary school	(0.089)	(0.092)	(0.092)	(0.008)**	(0.010)**	(0.010)**
Birth year controls		yes	yes		yes	yes
Regional controls		yes	yes		yes	yes
Father's job controls			yes			yes
Number of observations	1104	1104	1104	1104	1104	1104

Notes: This table reports estimates that include an indicator for cram school enrollment in elementary school as an additional control variable. The results are estimated by OLS for years of schooling (columns 1 to 3), by probit for an indicator of high school enrollment (columns 4 to 6). The marginal effects are reported in the probit model. Treatment effect is the coefficient of a dummy variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Cram school enrollment is a dummy variable takes one if an individual used a cram school when she was an elementary school student. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively.

Table 2.9: Estimates with Cram School Control in Junior High School

Dependent Variable:	Years of Schooling (OLS)			High School Enrollment (Probit)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	-0.510	-0.544	-0.463	-0.033	-0.034	-0.034
	(0.221)**	(0.219)**	(0.216)**	(0.017)*	(0.017)**	(0.018)*
Cram school enrollment	0.119	0.068	0.008	0.030	0.031	0.034
in junior high school	(0.102)	(0.111)	(0.108)	(0.009)***	(0.009)***	(0.011)***
Birth year controls		yes	yes		yes	yes
Regional controls		yes	yes		yes	yes
Father's job controls			yes			yes
Number of observations	1101	1101	1101	1101	1101	1101

Notes: This table reports estimates that include an indicator for cram school enrollment in junior high school as an additional control variable. The results are estimated by OLS for years of schooling (columns 1 to 3), by probit for an indicator of high school enrollment (columns 4 to 6). The marginal effects are reported in the probit model. Treatment effect is the coefficient of a variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Cram school enrollment is a dummy variable takes one if an individual used a cram school when she was a junior high school student. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively.

Table 2.10: Estimates with Labor Market Conditions Controls in Junior High School

Dependent Variable:	Years of Schooling (OLS)			High School Enrolment (Probit)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	-0.559	-0.537	-0.465	-0.039	-0.037	-0.035
	(0.206)**	(0.215)**	(0.207)**	(0.014)**	(0.015)**	(0.015)**
Local labour market conditions controls	yes	yes	yes	yes	yes	yes
Birth year controls		yes	yes		yes	yes
Regional controls		yes	yes		yes	yes
Father's job controls			yes			yes
Number of observations	1115	1115	1115	1115	1115	1115

Notes: This table reports estimates that include labor market conditions as additional control variables. The results are estimated by OLS for years of schooling (columns 1 to 3), by probit for an indicator of high school enrollment (columns 4 to 6). The marginal effects are reported in the probit model. Treatment effect is the coefficient of a variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Local labor market conditions controls are the active job opening to application ratio and average monthly male earnings at the prefecture level. These variables are averaged over the three years of junior high school. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively.

Table 2.11: The Effect of the Revision in the Transition Period

Panel A				
Dependent Variable:	Years of Schooling (OLS)			
Sample Restriction:	1960-62 and 1969-74 (Base)		1960-62 and 1964-65 (Transition)	
	(1)	(2)	(3)	(4)
Treatment effect	-0.504	-0.538	-0.486	-0.488
	(0.225)**	(0.221)**	(0.325)	(0.353)
Birth year controls		yes		yes
Regional controls		yes		yes
Number of observations	1115	1115	619	619

Panel B				
Dependent Variable:	High School Enrollment (Probit)			
Sample Restriction:	1960-62 and 1969-74 (Base)		1960-62 and 1964-65 (Transition)	
	(5)	(6)	(7)	(8)
Treatment effect	-0.034	-0.031	-0.068	-0.051
	(0.017)**	(0.014)**	(0.040)*	(0.038)
Birth year controls		yes		yes
Number of observations	1115	1115	619	619

Notes: This table reports estimates of the revision in the transition period. Panel A: OLS estimates for years of schooling, Panel B: Probit estimates for an indicator of high school enrollment. Columns 1, 2, 5 and 6 reproduce the main results for easy reference. Treatment effect is the coefficient of a variable that takes one if an individual enrolled in a public junior high school after the standards were revised. In columns 3, 4, 7 and 8, the treatment variable takes one if an individual enrolled in a public junior high school during the transition period. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised or during the transition period of the revision, parental years of schooling, and number of siblings. In Panel B, regional controls are excluded. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively. In Panel B, the marginal effects are reported.

Table 2.12: The Effect of the Revision without the Comprehensive Selection System for High School

Panel A						
Dependent Variable:	Years of Schooling (OLS)					
Sample Restriction:	Base			No Comprehensive Selection System		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	-0.504	-0.538	-0.463	-0.751	-0.764	-0.539
	(0.225)**	(0.221)**	(0.216)**	(0.255)***	(0.236)***	(0.232)**
Birth year controls		yes	yes		yes	yes
Regional controls		yes	yes		yes	yes
Father's job controls			yes			yes
Number of observations	1115	1115	1115	697	697	697

Panel B						
Dependent Variable:	High School Enrollment (Probit)					
Sample Restriction	Base			No Comprehensive Selection System		
	(7)	(8)	(9)	(10)	(11)	(12)
Treatment effect	-0.034	-0.038	-0.037	-0.037	-0.039	-0.037
	(0.017)**	(0.016)**	(0.016)**	(0.018)**	(0.016)**	(0.017)**
Birth year controls		yes	yes		yes	yes
Regional controls		yes	yes		yes	yes
Father's job controls			yes			yes
Number of observations	1115	1115	1115	697	697	697

Notes: This table reports estimates with excluding prefectures that experienced the comprehensive selection system for high school entrance selection. Columns 1 to 3 and 7 to 9 reproduce the main results for easy reference. Panel A: OLS estimates for years of schooling, Panel B: Probit estimates for an indicator of high school enrollment. Treatment effect is the coefficient of a variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively. In Panel B, the marginal effects are reported.

Table 2.13: Estimates with Alternative Specification of Parental Education

Panel A						
Dependent Variable:	Years of Schooling (OLS)					
Specification of Parental Education:	Years of Schooling (Base)			Indicator of more than post compulsory schooling		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	-0.504	-0.538	-0.463	-0.551	-0.619	-0.514
	(0.225)**	(0.221)**	(0.216)**	(0.221)**	(0.204)***	(0.214)**
Birth year controls		yes	yes		yes	yes
Regional controls		yes	yes		yes	yes
Father's job controls			yes			yes
Number of observations	1115	1115	1115	1115	1115	1115

Panel B						
Dependent Variable:	High School Enrollment (Probit)					
Specification of Parental Education:	Years of Schooling (Base)			Indicator of more than post compulsory schooling		
	(7)	(8)	(9)	(10)	(11)	(12)
Treatment effect	-0.034	-0.038	-0.037	-0.035	-0.040	-0.039
	(0.017)**	(0.016)**	(0.016)**	(0.017)**	(0.016)**	(0.017)**
Birth year controls		yes	yes		yes	yes
Regional controls		yes	yes		yes	yes
Father's job controls			yes			yes
Number of observations	1115	1115	1115	1115	1115	1115

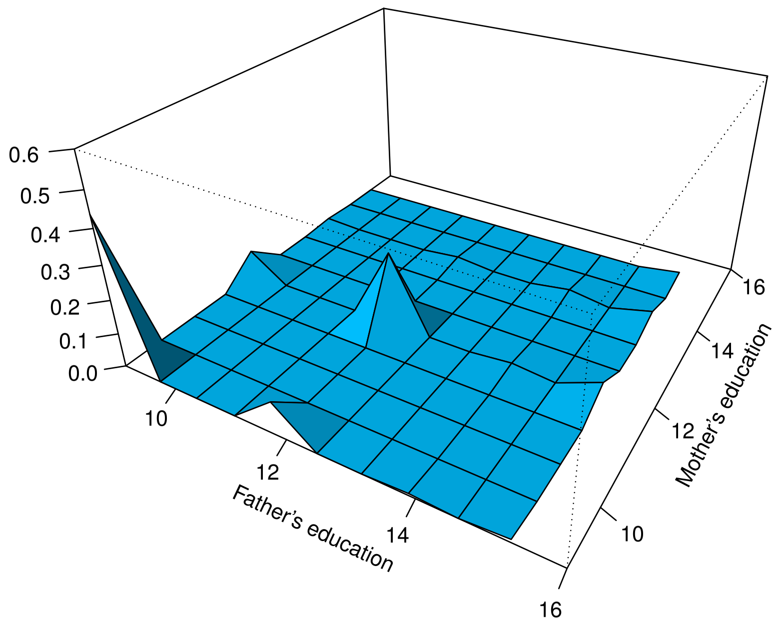
Notes: This table reports estimates with an alternative specification of mother's and father's education controls. Columns 1 to 3 and 7 to 9 reproduce the main results for easy reference. In columns 4 to 6 and 10 to 12, parental education controls are defined as an dummy variable that takes one if a parent received education beyond the compulsory level. Panel A: OLS estimates for years of schooling, Panel B: Probit estimates for an indicator of high school enrollment. Treatment effect is the coefficient of a variable that takes one if an individual enrolled in a public junior high school after the standards were revised. Birth year controls are a set of indicator variables for birth year. Regional controls include a set of indicator variables for region that an individual lived when she was a junior high school student. Father's job controls indicate a set of dummies for paternal job. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, and number of siblings. Standard errors are in parenthesis, clustered by public/private and birth year cohort. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively. In Panel B, the marginal effects are reported.

Table 2.14: Robustness to the Specification of the Standard Errors

		(1)	(2)	(3)	(4)	(5)	(6)
	Treatment Effect	Huber-White robust	Clustered in Public/Private	Clustered in (Public/Private) × (Before/After)	Clustered in (Public/Private) × Birth year	Clustered in (Public/Private) × Birth year ×Region	Clustered in (Public/Private) × Birth year ×Prefecture
Panel A: Years of schooling							
	-0.504	(0.355)	(0.019)**	(0.018)***	(0.225)**	(0.366)	(0.359)
Fixed effects controls	-0.538	(0.354)	(0.001)***	(0.035)***	(0.221)**	(0.357)	(0.356)
Panel B: High school entry							
	-0.034	(0.025)	(0.015)**	(0.014)**	(0.017)**	(0.027)	(0.025)
Fixed effects controls	-0.038	(0.024)	(0.015)***	(0.022)*	(0.016)**	(0.025)	(0.023)*

Notes: This table reports the treatment effect and its statistical significance with alternative specifications of clustering robust standard errors. Panel A: OLS estimates for years of schooling, Panel B: Probit estimates for an indicator of high school enrollment. Treatment effect is the coefficient of a dummy variable that takes one if an individual enrolled in a public junior high school after the standards were revised. The rows with fixed effects controls include sets of dummy variables that capture cohort and regional fixed effects. The following variables are also included as controls; an indicator for public junior high school graduates, an indicator for individuals who enrolled in a junior high school after the standards were revised, parental years of schooling, and number of siblings. Standard errors in parentheses are, respectively, Column (1): Huber-White robust standard.errors, Column (2): clustering robust standard errors by a public/private junior high school, Column (3): clustered by public/private junior high school and before/after the revision, Column (4): clustered by public/private junior high school and years of birth, Column (5): clustered by public/private junior high school, year of birth, and region lived at the age of junior high school, Column (6): clustered by public/private junior high school, year of birth, and prefecture lived at the age of junior high school. ***, **, and * indicate statistical significance at 1, 5, and 10% level respectively. The treatment effect in Panel B is the marginal effect.

Panel A: Public, Before the Revision



Panel B: Private, Before the Revision

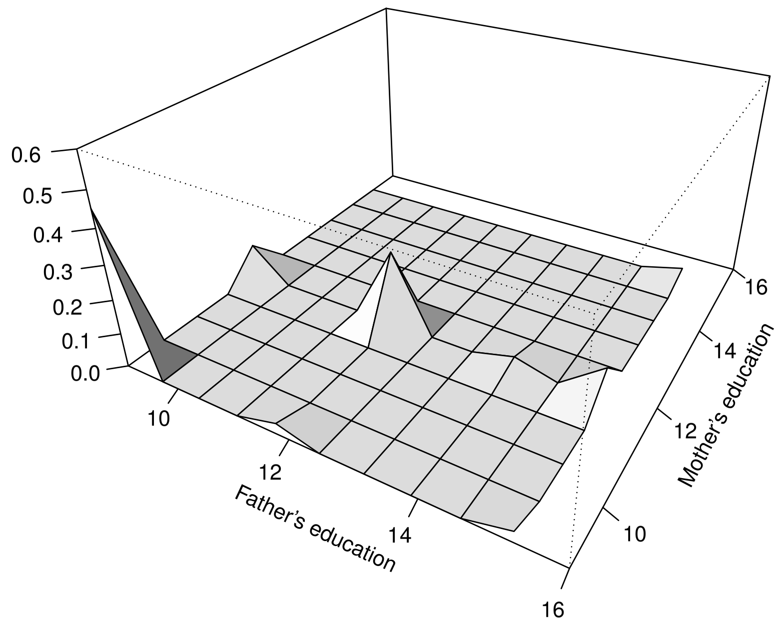


Figure 2.1: Distribution of Parental Years of Schooling of Individuals Who Graduated from Junior High School Before the Revision

Notes: These figures depict the distribution of mother's and father's years of schooling of individuals who graduated from a junior high school before the revision. Panel A is the distribution of public school graduates, Panel B is the distribution of private school graduates.

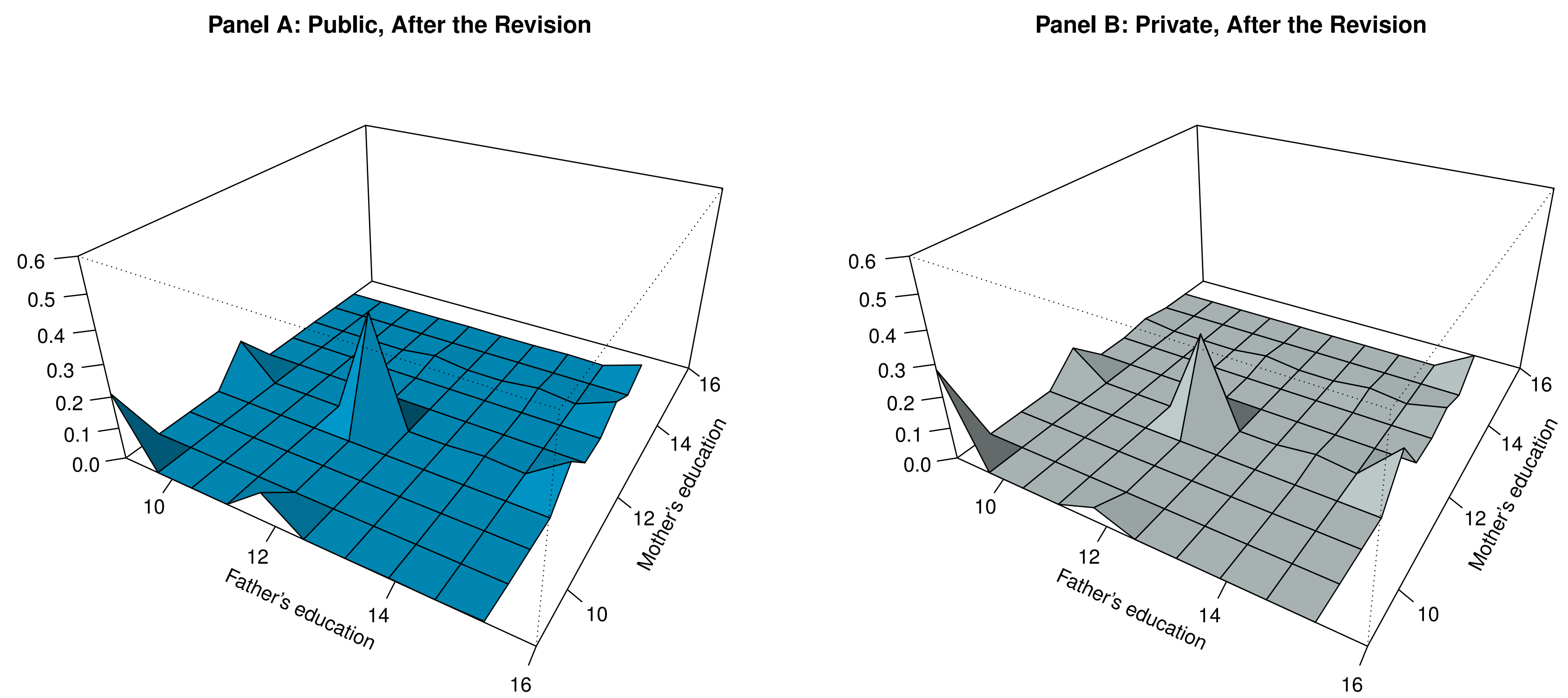


Figure 2.2: Distribution of Parental Years of Schooling of Individuals Who Graduated from Junior High School After the Revision

Notes: These figures depict the distribution of mother's and father's years of schooling of individuals who graduated from a junior high school after the revision. Panel A is the distribution of public school graduates, Panel B is the distribution of private school graduates.

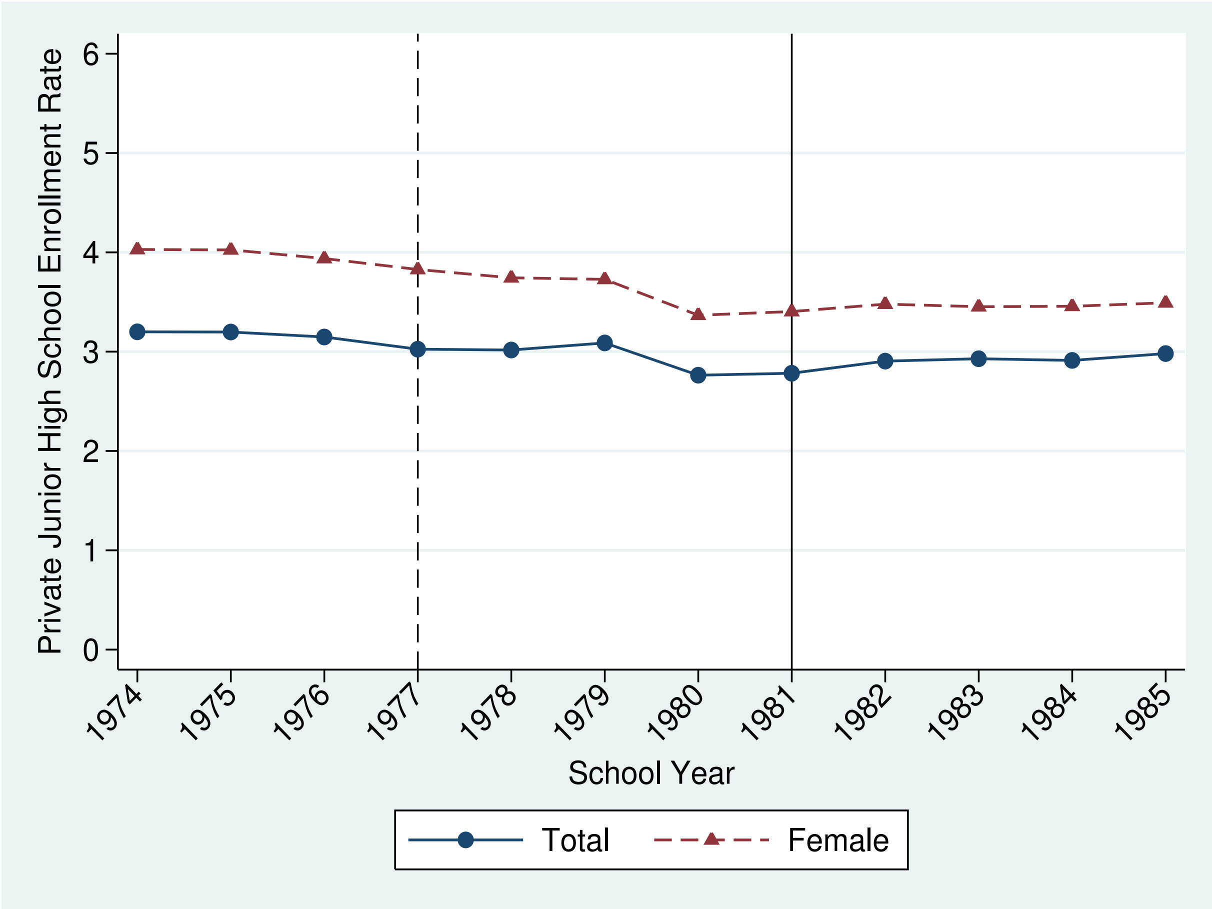


Figure 2.3: Private Junior High School Enrollment Rate

Source: *School Basic Survey*

Notes: This figure depicts the private junior high school enrollment rate from the 1974 to 1985 school year. The author calculated "The enrollment rate" as "The number of private junior high school students in the 1st grade of the school year / The number of all elementary schoolchildren in the 6th grade of the last school year" (in %) based on the *School Basic Survey*. The solid connected line is the total enrollment rate. The dashed connected line shows the female enrollment rate. The vertical dashed line indicates the year of the announcement of the fifth curriculum standards revision. The vertical solid line shows the year that the revised standards were enforced.

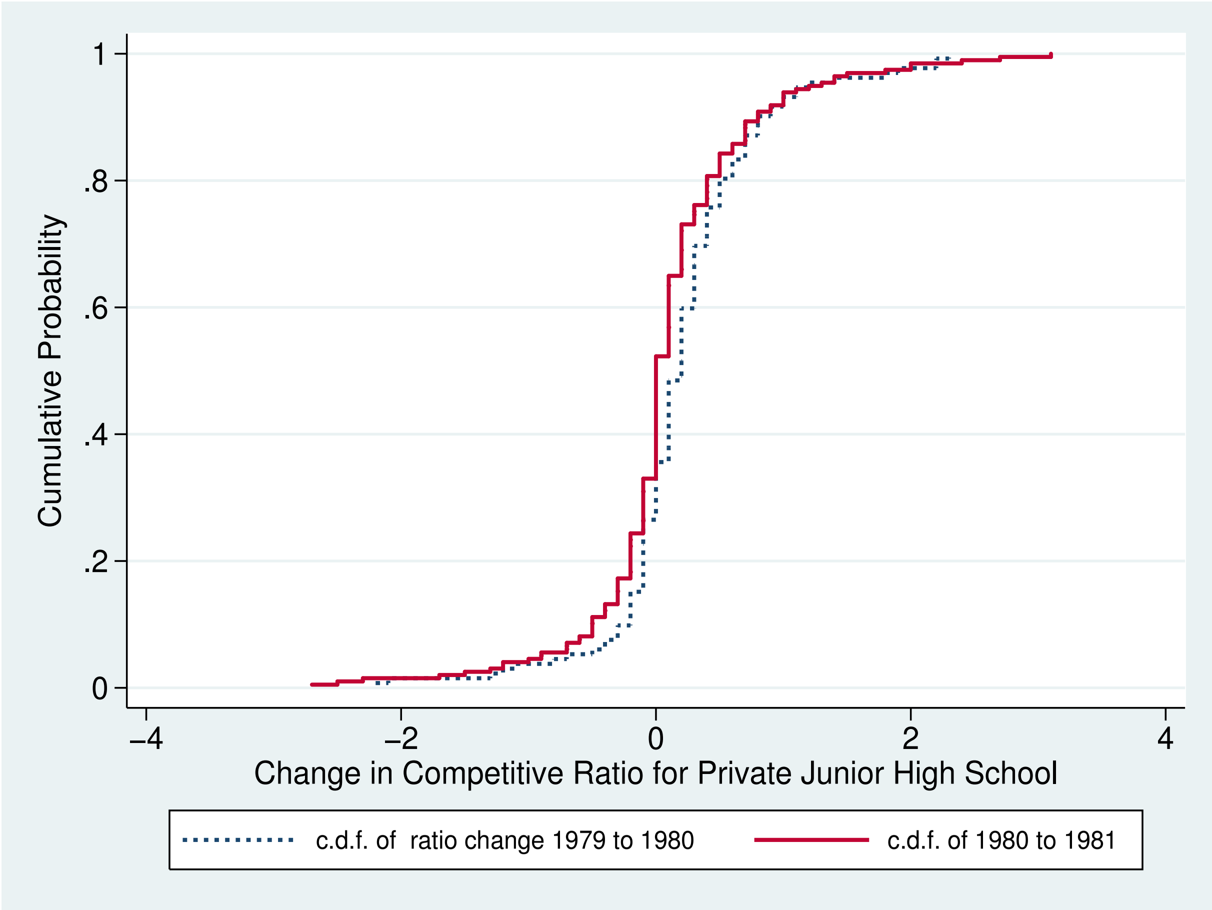


Figure 2.4: Cumulative Distribution of Change in Competitive Ratio of Private Junior High School

Source: *Shukan Sankei Rinji Zokan (Haru, Aki)* '79 - '80 and '81 - '82.

Notes: This figure depicts the cumulative distribution of the change in ratio of applicants to capacity at an entrance examination of private junior high school. The author calculated the numbers based on *Shukan Sankei Rinji Zokan*. The dashed line is the change from 1979 to 1980 (before the revision). The solid line is the change from 1980 to 1981 (after the revision).

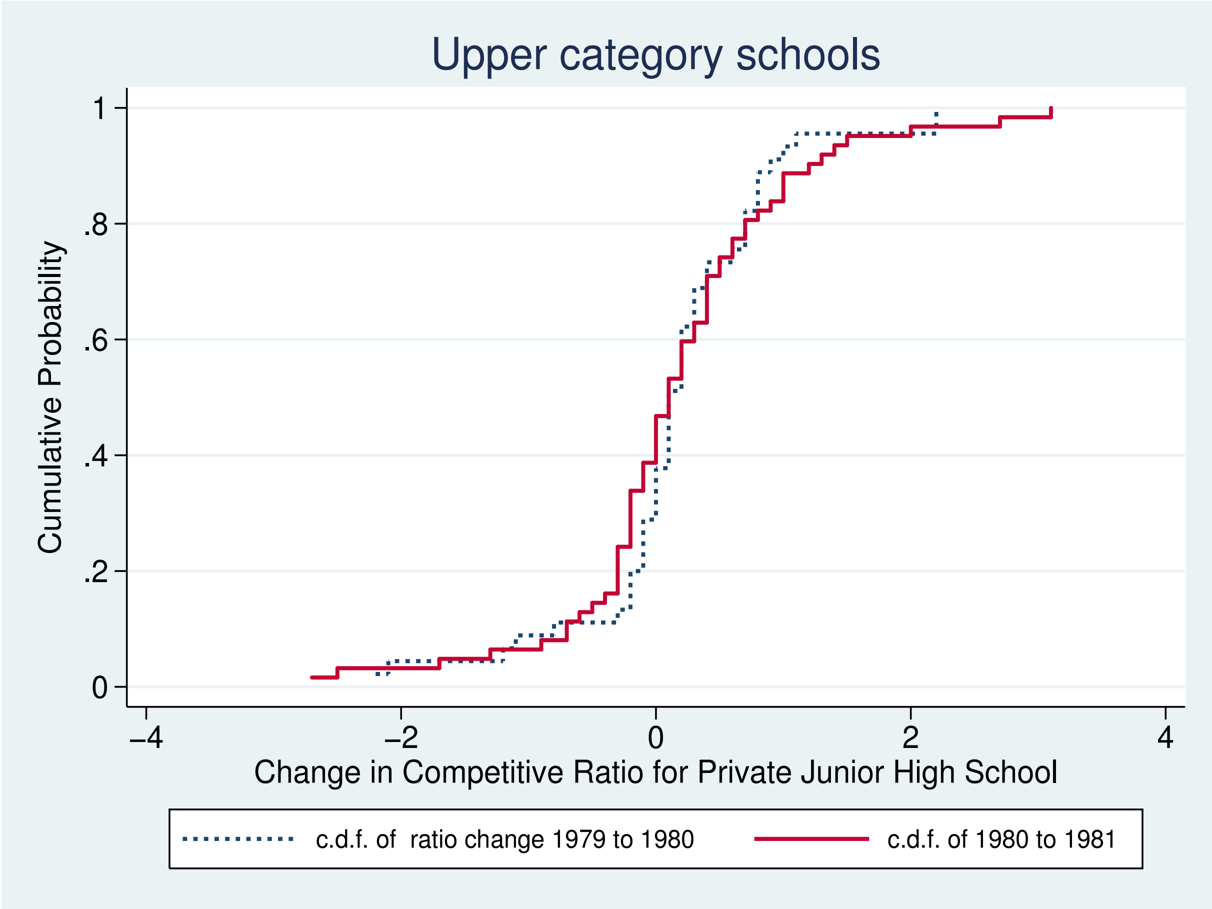


Figure 2.5: Cumulative Distribution of Change in Competitive Ratio of Private Junior High School (Upper Category Schools)

Source: *Shukan Sankei Rinji Zokan (Haru, Aki)* ’79 - ’80 and ’81 - ’82.

Notes: This figure depicts the cumulative distribution of the change in ratio of applicants to capacity at an entrance exam of private junior high school with a high deviation value in the entrance examination difficulty. The author calculated the numbers based on *Shukan Sankei Rinji Zokan*. The dashed line is the change from 1979 to 1980 (before the revision). The solid line is the change from 1980 to 1981 (after the revision).

Chapter 3

Estimating the Returns to Higher Education in Japan

3.1 Introduction

One remaining task for education policies in developed countries is providing the widespread opportunity to go to college. The trends of university enrollment rate in Japan suggest that the task is progressing satisfactorily in the aggregate. The rate has risen from 17.1% in 1970 to more than 50% in 2010 according to the School Basic Survey. By contrast, an equal opportunity to access to higher education is still a matter under discussion. In the field of sociology of education, the recent empirical studies find a regional inequality of accessibility to universities.[1] In addition, OECD (2013) points out a burden of high tuition fees in Japan: "the average annual fee to attend public tertiary institutions was USD 5019 during the academic year 2010–2011", which was "the fifth highest annual fee among (OECD) countries with available data". Therefore, researchers are expected to analyze the effects of an education policy increasing the availability of accessible, affordable higher education. To evaluate such a policy, researchers need to precisely estimate the impact of college attendance for those who are affected by the policy. However, it is hard to implement the task if effects are heterogeneous and individuals endogenously self-select into college based on their gains.

One way to deal with the endogenous schooling decision is using the instrumental variable (IV) approach. The recent literature addresses endogenous selection of college attendance by instrumenting it with a variation in local accessibility to college during adolescent years. For example, Card (1993), Kane and Rouse (1993), Currie and Moretti (2003), Cameron and Taber (2004), Carneiro, Meghir, and Parey (2013) show positive returns to schooling using an instrument of substitute for costs of college attendance, such as distance to college and college existence at the county of residence.

[1]For example, see Sasaki (2006), Kobayashi (2009), Nakazawa (2011).

Despite the growing body of the literature in this approach, it provides a limited implication for a policy intervention. The standard IV estimates capture an effect related to the local average treatment effect (LATE) of Imbens and Angrist (1994). The LATE shows the return to schooling for individuals induced to go to school by the change in the instrument. Heckman and Vytlacil (1999, 2001b, 2005), Heckman, Urzua, and Vytlacil (2006), Carneiro, Heckman, and Vytlacil (2010, 2011) point out that individuals who are induced to go to college in the LATE are not necessarily equivalent to those who are induced to go to college by a change in the policy of interest.

Heckman and Vytlacil (1999, 2001a, 2005) propose an alternative way to use instruments. They show that average treatment parameters widely analyzed in the literature of policy evaluation can be identified as weighted averages of the marginal treatment effect (MTE) introduced by Björklund and Moffitt (1987) and Heckman and Vytlacil (1999). Using this approach, recent studies find positive average effects of college education in different countries and regions, such as Chuang and Lai (2010) in Taiwan, Carneiro, Heckman, and Vytlacil (2011) in US and Nybom (2012) in Sweden.

In Japan, many studies also investigate returns to schooling. However, Oshio and Seno (2007), Yasui and Sano (2009) point out that very limited studies are motivated to estimate causal effects of schooling. Oshio and Seno (2007) also suggest that the number of empirical studies using Japanese micro data is small due to the data availability.[2] One exception is Nakamuro and Inui (2012). Using the web-based twin data, they find significant positive returns to schooling in Japan. Unfortunately, their fixed-effects approach using twins only controls for differences in family level, and thus, cannot control for unobserved individual heterogeneity that affects schooling decisions.

This study estimates returns to university education in Japan using an approach introduced by Heckman and Vytlacil (1999, 2001a, 2005), which allows self-selection of university enrollment under heterogeneous effects across individuals. To deal with the endogenous schooling decision, I use public tuition, local availability of universities, and local labor market conditions as instruments. For the local availability of universities, I collect all information of opening/closing and increasing/decreasing of accredited capacities of universities in Japan and create a measure at the prefecture level. This measure reflects cost of preparation for taking examination and probability of acceptance besides geographical moving costs or costs of not living at the home with parents.

The main results show that average effects of a year of university education in Japan are significant and positive, but highly varied across sub-group of population categorized by their treatment status. The results also suggest that policies increasing probability of university enrollment, such as free tuition and an increase in local capacities of universities, bring about positive effects of university education.

[2]Oshio and Seno (2007) comprehensively survey the empirical studies of economics of education in Japan.

The remainder of this chapter is organized as follows. Section 3.2 explains the empirical framework. Section 3.3 describes the data. Section 3.4 presents and discusses the results. Finally, I conclude in Section 3.5.

3.2 Empirical Framework

This section describes estimation model and method. I basically follow the model of Heckman and Vytlacil (2005); Carneiro, Heckman, and Vytlacil (2010, 2011). They consider a standard model of potential outcomes that is firstly applied to schooling in Willis and Rosen (1979).

3.2.1 Setup

Let consider a linear-in-the-parameters model with two potential outcomes:

$$\begin{aligned} Y_1 &= X\beta_1 + U_1 \qquad (3.1) \\ Y_0 &= X\beta_0 + U_0, \end{aligned}$$

where subscript 0 and 1 correspond to the untreated and treated states. X are observable and (U_1, U_0) are unobservable variables. The assumption on linearity or separable outcomes are not required for identification, but they are useful in practice for estimation and are commonly assumed in the literature. Let $D = 1$ denotes enroll in university; $D = 0$ denotes not enroll in university, the measured outcome variable Y can be written in a potential outcomes framework:

$$Y = DY_1 + (1 - D)Y_0 \qquad (3.2)$$

This equation is related to a latent variable discrete choice model that represents an individual's decision on university enrollment. I assume the following selection model:

$$\begin{aligned} D^* &= \mu_D(Z) - V \\ D &= 1 \ \ if \ \ D^* \geq 0, \\ &= 0 \ otherwise \\ or \ \ D &= \mathbf{1}[\mu_D(Z) - V \geq 0], \end{aligned}$$

where V is a unobserved continuous random variable with a strictly increasing distribution function F_V. Z is a vector of observed random variables that includes some part of X. Z also includes variables that determine the treatment decision but do not directly affect the outcome (the exclusion restriction). Formally, I assume that Z and X satisfy; $Z \not\perp\!\!\!\perp D|X$, and Z, $X \perp\!\!\!\perp (V, U_1, U_0)$. In addition, I need to assume a support condition for X on the university enrollment probability; $0 \leq Pr(D = 1|X) \leq 1$.

Let $Pr(D = 1|Z)$ denotes the probability of university enrollment conditional on Z, the assumption on V allows to rewrite the selection equation as:

$$\begin{aligned} D &= \mathbf{1}[F_v(\mu_D(Z)) \geq F_v(V)] \\ &= \mathbf{1}[P(Z) \geq U_D], \end{aligned}$$

with

$$\begin{aligned} U_D &\stackrel{\text{def}}{=} F_v(V) \sim Unif[0,1] \\ P(Z) &\stackrel{\text{def}}{=} F_v(\mu_D(Z)) = Pr[D = 1|Z]. \end{aligned}$$

Using equations 3.1 and 3.2, the observed outcome can be written as:

$$\begin{aligned} Y &= X\beta_0 + DX(\beta_1 - \beta_0) + D(U_1 - U_0) + U_0 \\ &= X\beta_0 + DX\beta + \epsilon. \end{aligned} \tag{3.3}$$

This equation indicates that the effect of university enrollment varies across individuals for differences in their X and U_1, U_0. If the enrollment decision depends on unobservable gain $U_1 - U_0$, a dummy variable D is not independent of the disturbance ϵ. In this case, neither ordinary least squares (OLS) nor simple linear IV estimates recover the standard average effect parameters, such as the average treatment effect (ATE): $E(Y_1 - Y_0)$, the ATE on treated (ATT): $E(Y_1 - Y_0|D = 1)$, and the ATE on untreated (ATU): $E(Y_1 - Y_0|D = 0)$. Heckman and Vytlacil (1999, 2001a, 2005) establish that these treatment parameters of interest can be identified as weighted averages of the MTE of Björklund and Moffitt (1987) and Heckman and Vytlacil (1999). The MTE is defined as:

$$MTE(x, u_D) \stackrel{\text{def}}{=} E(Y_1 - Y_0|X = x, U_D = u_D)$$

The MTE indicates the effects of university enrollment for individuals with $X = x$ who would be indifferent between enrollment or not under the situation that they are exogenously assigned a value of Z such as $U_D = u_D$.

3.2.2 Estimating Marginal Treatment Effect

Heckman and Vytlacil (1999, 2001a, 2005) show that the MTE can be identified by the local instrumental variables. Using the equation 3.3, the conditional expectation of Y given $X = x$, and $P(Z) = p$ is

$$\begin{aligned} E[Y|X = x, P(Z) = p] &= x\beta_0 + x(\beta_1 - \beta_0)p + K(p) \\ &= x\beta_0 + \int_0^p MTE(x, u_D) du_D, \end{aligned} \tag{3.4}$$

where,

$$\begin{aligned} K(p) &= E(U_1 - U_0 | D = 1, P(Z) = p) \\ &= \int_{-\infty}^{\infty} \int_0^p (u_1 - u_0) f(u_1 - u_0 | X = x, U_D = u_D) du_D d(u_1 - u_0), \end{aligned}$$

where $f(u_1 - u_0 | X = x, U_D = u_D)$ is the conditional density of $U_1 - U_0$. Therefore, the MTE is identified by differentiating the equation 3.4 with respect to p,

$$\frac{\partial}{\partial p} E[Y|X = x, P(Z) = p] = E[Y_1 - Y_0 | X = x, U_D = u_D]. \tag{3.5}$$

The equation 3.4 can be estimated using the model of a semi-parametric approach, such as partially linear model of Robinson (1988), and the equation 3.5 can be estimated in non-parametrically. One of the disadvantages of the semi-parametric approach is that the MTE can only be estimated over the empirical support of $P(Z)$, thus it is not possible to identify conventional treatment parameters that require full unit interval of $P(Z)$. An alternative way of estimating the MTE is a parametric approach assuming on the joint normal distribution of the unobservables (U_0, U_1, V), and estimating the outcome and selection equations together using the method of maximum likelihood.[3] This parametric approach is a conventional way of estimating the equations and is related to the switching regression models (Björklund and Moffitt, 1987; Willis and Rosen, 1979).[4] An advantage of specifying the normality assumption is that it helps to estimate the MTE over full unit interval of $P(Z)$ and to recover the treatment effect parameters of interest.

3.2.3 Policy Relevant Treatment Effects

Once the MTE is estimated, the parameters that are directly relevant to the policy questions can also be estimated as weighted averages of it. I compute the policy relevant treatment effects

[3] In the estimation, I normalize the variance of V to 1.

[4] See also Heckman, Tobias, and Vytlacil (2001) for further description.

(PRTE) introduced by Heckman and Vytlacil (2001b) and marginal version of PRTE (MPRTE) proposed by Carneiro, Heckman, and Vytlacil (2010). Let D^*, Y^* and P^* denote the treatment state, outcome and probability of university enrollment after the policy change, Heckman and Vytlacil (2005, 2007) define the PRTE when $E(D^*) \neq E(D)$ as,

$$\frac{E(Y^*) - E(Y)}{E(D^*) - E(D)} \stackrel{\text{def}}{=} \int_0^1 MTE(u_D)\omega_{PRTE}(u_D)du_D,$$

where,

$$\omega_{PRTE}(u_D) = \frac{F_p(u_D) - F_{p^*}(u_D)}{E_{F_{p^*}}(P) - E_{F_p}(P)},$$

where F_p^* and F_p are the distribution of P^* and P, respectively.[5]

The MPRTE is defined as the limit of PRTE with a sequence of alternative policies indexed by a scalar variable α such that $lim_{\alpha\to 0}P_\alpha^*(Z) = P(Z)$. I consider two policy sequences as defined in Carneiro, Heckman, and Vytlacil (2010, 2011): (1) a policy that increases the probability of university enrollment by α so that $P_\alpha^* = P + \alpha$ and (2) a policy that changes each individual's probability of university enrollment by the proportion $(1+\alpha)$, so that $P_\alpha^* = (1+\alpha)P$

3.3 Data

3.3.1 The Japanese General Social Survey

The main analysis data are the Japanese General Social Survey(JGSS).[6] JGSS are repeated cross-section data for men and women aged 20–89 on 1st September of each survey year. This study uses surveys conducted in 2000, 2001, 2002, 2005, 2006, 2008, and 2010, and pools male respondents from all waves. From the pooled original data, I exclude the observations in four ways.

First, I limit the sample by their age. I exclude individuals who were younger than 28 years old on the survey year because they might not complete their academic schooling. Second, I drop the individuals who answered that their father or mother was absent at the age of 15.

[5] To simplify the notation, I suppress control variables.

[6] The Japanese General Social Surveys (JGSS) are designed and carried out by the JGSS Research Center at Osaka University of Commerce (Joint Usage / Research Center for Japanese General Social Surveys accredited by Minister of Education, Culture, Sports, Science and Technology), in collaboration with the Institute of Social Science at the University of Tokyo. The data for this secondary analysis, the JGSS, the JGSS Research Center, was provided by the Social Science Japan Data Archive, Center for Social Research and Data Archives, Institute of Social Science, The University of Tokyo.

Because their single parental structure might substantially differ from the families with couple parents, unobservable effects cannot be controlled. Third, I only use the observations who reached their first university enrollment decision after 1972 school year due to the availability of tuition data. Finally, I can use the respondents whose observational characteristics are available to match the comparable information on the instruments and covariates for the estimation explained below. The remained sample after restrictions contains male workers were born in 1953–1979 and were 28–54 years old on the date of survey.

The JGSS have the advantage of including the information on workers' annual income and working hours per week. For the outcome variable, I compute the worker's personal hourly wage. Unfortunately, the JGSS only ask income measures by 19 categories. Following Oshio and Kobayashi (2009) and Sano and Yasui (2009), I assign the median value of each category and evaluate it at the 2005 consumer prices and transform it in logarithm.[7]

One of the disadvantages of the JGSS is that the data set only contains limited information on the residence at the university enrollment decision. I use the information on the prefecture where individuals resided at the age of 15 and assume that in the year of university enrollment, their home (at least their parental residence) was in the prefecture. The control variables from the JGSS are mother's and father's education (and these squares), number of siblings (and its square), dummy variables indicating urban residence and rural (farm or fishing village) residence at the age of 15, a set of dummies for prefecture resided at the age of 15, and cohort dummies. For the educational attainment of the respondents and their parents, I use the information on the level of the last school attended and assign the standard years of schooling in Japan.

I also control for long-term trends of the active job opening to application ratio (*yuko kyujin bairitsu*) from the report of the Employment Service Agency (*shokugyou antei gyoumu toukei*) and the annual average monthly total cash earnings of the Monthly Labor Survey (*Maigetsu Kinro Tokei Chosa*), and local estimated population size at age 15–19 of the Population Census (and its square) at the prefecture where individuals resided when they were 15 years old. In addition, I include following three controls for current conditions: years of current job experience (and its square), unemployment rate in the region of residence on the survey year, and the average monthly male earnings in the prefecture of residence on the survey year. The data sources for unemployment rate and the average monthly male earnings are the Monthly Labor Survey and the Basic Survey on Wage Structure.

In the analysis, I consider the binary treatment decision for university enrollment at the completion of upper secondary education. Therefore, I separate individuals into two groups: (1) individuals who graduated from high school or completed an upper secondary education,

[7]For the lowest category (less than 700,000 yen), I assign 700,000. For the highest category (over 23,000,000 yen), I assign 23,000,000. When I exclude the people in the lowest and the highest categories, the estimated results are basically same.

and (2) individuals who had some university education or more.[8]

3.3.2 Instrumental Variables

This section discusses the IVs for the university enrollment decision. I use differential changes in the direct and opportunity costs of university attendance across prefectures and cohorts, while controlling for both permanent differences and aggregate trends. The IVs are local university availability measures containing all universities, average tuition of public university in the freshman year, and local labor market conditions in high school years, in the prefecture where the individual resided when he was 15 years old.[9]

3.3.2.1 Capacity of Universities

First, I describe a measure of the availability of local universities. Local college availability measures are firstly used by Card (1993) and Kane and Rouse (1993) as a proxy of direct costs of college attendance, and widely used in the literature. Kane and Rouse (1993) uses a distance to college measure as an instrument for schooling, and followed by Carneiro et al. (2011), Nybom (2012) in recent years. An indicator of presence of college in the county of residence is used by Card (1993) as a substitute of distance to college, and it is commonly used in the literature, for example, Kling (2001), Cameron and Taber (2004), Carneiro, Heckman, and Vytlacil (2010, 2011), Carneiro, Meghir, and Parey (2013). Unfortunately, the JGSS only have information on the prefecture of residence at the age of 15. Because all prefectures have a university during the analyzed years in Japan, I am unable to use this local presence measure. Currie and Moretti (2003) use the number of two- and four-year colleges as an instrument for college attendance. This measure is superior to the indicator definition because its continuous variations across residential areas and years allow to control for permanent differences across counties. Although the number of colleges measure partly takes into account the quantitative differences in college availability across residential areas, it is too rough to capture the exact differences in local opportunities of college education. Because the size of the colleges are different among the region, each college has different effects on local availability of colleges. In Japanese empirical studies in sociology of education, the number of enrollments is widely used as a substitute of capacities of colleges offered.[10] However, as Currie and Moretti (2003) point

[8]For the respondents, completion of their last schooling is available. When I construct the variable of years of schooling, I reduce the number by one year from standard years for those who dropped out before finishing their last school. However, I am unable to know whether dropouts of technical college completed their upper secondary education. In the analysis, I assume that they completed their upper secondary education and included them in the analysis sample. If I exclude them from the analysis sample, the main results are basically same.

[9]The construction of these data is amply described in the Data Appendix.

[10]For example, see Sasaki (2006), Kobayashi (2009).

out, not only the supply of college places, but also the demand for these places determine the number of enrollments, and thus it is not a valid instrument.

In this study, I construct a unique data set that contains the information on all accredited capacities for new enrollments of universities by prefecture and year in Japan. I collect the information at the department level and total up capacities of national, prefectural, municipal, and private co-ed universities by prefecture. This measure is merged to the individual data based on high school graduate's standard college examination year and the residence at the age of 15.[11] I assume that this measure of local availability of universities is a proxy of easy access to a local university; i.e., costs of geographical moving or costs of not living at their home with parents. The quantitative characteristics of this measure also reflects costs of preparation for taking examination and probability of acceptance.[12]

One potential problem of this capacity measure is that changes in cohort size are likely to have an impact on the availability of universities given any fixed number of capacities (Card and Lemieux, 2001; Currie and Moretti, 2003). To avoid this problem, I control for local cohort size at age 15–19 when the individual was at age 18 in both selection and outcome equations.

Another concern on using this measure is that it is affected by new openings and an increase in size of a university reflecting an expected increase of local demand for university education, as well as the number of colleges measure of Currie and Moretti (2003). Although I am unable to completely rule out these possibilities, the Japanese centralized educational system partly mitigates this concern. The Japanese School Education Act (*Gakko Kyoiku Ho*) prior to the revision in 2003 prescribes that all openings and closings of department of university are required to be approved by the national government in advance. Private universities also need an approval for changes in capacities at the department level in advance.[13] Therefore, all universities cannot freely control their capacities in response to the expected local demand of university education.

3.3.2.2 Average Tuition in Public Universities

Second, I explain the variable of tuition for new enrollments. The tuition measure is created as accredited capacities weighted averages over all public co-ed universities in a prefecture, or at the regional level if there is no public university in the prefecture. Kane and Rouse (1993),

[11]When I merge the year at age of 18, the acquired results are basically same.

[12]Although I integrate all capacities of universities into one measure, changes in capacities might have heterogeneous effects on college attendance by student's major field of study. In this study, because the available data set has no information on the major field of study, I cannot analyze the impacts of college major choice in detail, which I leave to future research.

[13]Public universities need to notice their changes in capacities in advance. Before 1974, private universities are allowed to increase their capacities with a notification to the government in advance, therefore capacities of private universities might not be a valid instrument in these periods. For a robustness check, I exclude individuals who were born before 1955 and re-estimate the analysis. The results are almost similar but more imprecise for the smaller sample size.

Cameron and Heckman (1998, 2001), Carneiro, Heckman, and Vytlacil (2010, 2011), and Carneiro, Meghir, and Parey (2013) use tuition as an instrument to predict college attendance.

One concern to use tuition as an instrument is that the variable is highly correlated with quality of the university (Cameron and Heckman, 2001). If the measure captures both costs of college attendance and college qualities, it directly affects differences in wage. To mitigate this concern, I only include entrance fees and course fees of prefectural and municipal universities in the tuition measure. Because these fees are specified by the national and local governments, each university is not permitted to change the amount of these fees. In Japan, new enrollments of university need to pay entrance fees, course fees, and other fees for school expenses.[14] Before 2003, entrance fees and course fees of national universities are specified by the national government.[15] Based on the amount of the fees of national universities, the fees of prefectural and municipal universities are regulated by a local government or a governmental agency. Therefore, entrance and course fees of public universities are determined at local government level and basically reflect a variation at the prefecture level. I do not include other fees in public tuition because each university is allowed to determine the amount of such fees at the department level. The same thing applies to the tuition of private universities. It varies from university to university, and thus might reflect the quality of the university. I exclude the tuition of national universities from the public tuition measure because it was unified over the country until 2003. The differences in tuition of national universities are captured by the cohort dummies.

Using the local tuition at prefecture level presume that the variable matters for the schooling choice of the individual. One could argue that individuals might move to a different prefecture for their university education to avoid high tuition costs of local public universities (Carneiro, Meghir, and Parey, 2013). However, it appears reasonable to consider that prefectural variation matters in Japan. Because prefectural and municipal universities usually set lower price of tuition for intra-regional students, movers are not only prevented from the option of living at home, but also disadvantaged for a higher tuition for extra-regional students.

3.3.2.3 Local Labor Market Conditions

Third, I use labor market conditions as instruments for university enrollment decision. I use the active job opening to application ratio and the annual average monthly total cash earnings in the prefecture of residence at age 15. Local labor market conditions are used as instruments by Cameron and Heckman (1998), and followed by Cameron and Taber (2004), Arkes (2010), Carneiro, Heckman, and Vytlacil (2010, 2011), Carneiro, Meghir, and Parey (2013), among others.

I construct these measures as three years averaged over high school period and merge them

[14]Other fees contain, for example, training fees, fees for facilities and equipment.

[15]In 2004, national universities are incorporated by the the National University Corporation Act (*Kokuritsu Daigaku Hojin Ho*).

to the individual data at the year of individual was 18 years old. I presume that local earnings capture temporary shocks to family income. Local job openings reflect the speed of job transitions or of finding a job if unemployed, and thus it also is related to temporary variation in family resources. In addition, local earnings might capture foregone earnings as opportunity costs of an additional schooling. A potential problem of using local labor market conditions is that long-run trends of labor market conditions might affect both these measures and residential choice at age 15. If local active job opening to application ratio and local earnings in high school are correlated with the unobservables in the outcome equations, would not be a valid instrument. To avoid this concern, I include averaged over 6 years trends of local labor market conditions at age 13–18 in both selection and outcome equations to control for residential choice at the age of 15 and long-run differences in labor market conditions of prefecture of residence in adolescent years. In addition, I control for a set of dummy variables of the prefecture of residence at age 15, allowing for permanent or aggregate differences in characteristics of the prefecture.

Table 3.1 shows summary statistics for instruments with the outcomes and covariates. It shows that individuals with some university education have, on average, higher wages than those without university education. They have less years of experience in current job since they have longer years of schooling. The difference between the two groups is about 3.81 years of schooling. Using this figure, all estimates of treatment effects reported below are annualized. Individuals with some university education have more-educated parents, have smaller number of siblings, and have lived in a prefecture where better labor market conditions in both adolescents and the survey year. Their residence at the age of 15 was more likely to be in urban areas, to be in a prefecture where has larger number of adolescents. They are less likely to come from rural areas than those without university education. For instruments, individuals with some university lived in a prefecture where has larger number of capacities of universities and better labor market conditions when they were in high school or upper secondary education. However, individuals with some university education were in a prefecture where costs, on average, higher tuition of public universities than those without university education.

3.4 Results

3.4.1 First Stage Results

Table 3.2 presents the estimates of the selection model of university enrollment. I estimate the logit model where the dependent variable is a dummy variable that equals one if the individual has ever attended university, and I report the marginal effects at the mean value of each variable. All controls reported in the table perform well, and estimates show expected signs reported in the previous studies. For example, individuals who have parents with longer years of schooling are more likely to enroll in university than those whose parents with lower levels of education. The instruments are jointly strong predictors of university enrollment, though local active job opening to application ratio (local job openings) is not individually significant. Local capacity of universities is important determinant of university enrollment. If local capacities increase by 100 places, probability of university enrollment increases by about 1%. Local tuition in public university also has statistically significant effects on university enrollment. If local tuition rises by 10000 yen, it decreases university enrollment by about 0.6%. Local earnings play a role of an opportunity cost variable of university enrollment. If local earnings averaged over high school years increase by 1%, university enrollment decreases by about 2%. A better local active job opening to application ratio at university enrollment increases probability of university enrollment. The 0.1 point improvement of local active job opening to application ratio in high school period increases probability of university enrollment by about 1.1%, but it has not statistically significant effect.

3.4.2 OLS and IV Results

In Table 3.3, I present standard OLS and IV estimates to compare with the estimates in the previous studies that use these methods. The OLS estimate shows that annualized returns to university education is about 5.32%. One could argue that the magnitude of the OLS estimate in this study is much smaller than the estimates derived in the previous studies using Japanese micro-data. These studies report about 7 to 11% of OLS estimates of returns to schooling (e.g. Tachibanaki, 1988; Trostel, Walker, and Woolley, 2002; Ono, 2004; Sano and Yasui, 2009; Yasui and Sano, 2009; Nakamuro and Inui, 2012). The differences in magnitude of the estimates between their results and my result might be explained by differences in sample restriction and in control variables. There are three main differences. First, I only use male observations. Trostel, Walker, and Woolley (2002), Sano and Yasui (2009), and Yasui and Sano (2009) suggest larger magnitude of the returns for female. Including female observations may provide larger estimates of returns to schooling. I exclude female from the analysis to avoid the selection bias in labor market participation. Second, I consider binary treatment of university enrollment, thus exclude the individuals with less than upper secondary education. It is possible to consider a model with multiple levels of treatment or with a continuous treatment of schooling, but it needs more restrictions and additional instruments to identify the treatment parameters of interest.[16] Finally, I include variables of local labor market conditions in the survey year when estimating the model. Because the analysis data are based on pooled cross-section data sets, controlling for current labor market conditions relieve sampling biases. These controls are not included in the previous studies. Controlling for these variables might weaken the magnitude of the coefficient

[16] See, for example, Heckman, Urzua, and Vytlacil (2006), Heckman and Vytlacil (2007), and Florens et al. (2008).

of university enrollment.

Table 3.3 shows that the IV estimate of returns to university education is 10.67%. In the line with the literature, the IV estimate is larger than the OLS estimate. Card (2001) suggests that such a finding indicates returns to schooling are heterogeneous and higher for the individuals who are induced to enroll in university by the changes in the instruments than those who have average returns. This interpretation is related to the LATE parameter of Imbens and Angrist (1994). However, the IV estimates do not necessary reflect the original LATE parameter if instruments are multiple and the model includes a set of controls, which is the case of this study.[17] Interpreting an IV estimate is not always straightforward. Even if it shows LATE, it indicates the policy effect of interest only if the variation of the instrument corresponds exactly to the policy variation (Heckman, Urzua, and Vytlacil, 2006; Carneiro, Heckman, and Vytlacil, 2011).

3.4.3 Marginal, Average and Policy Relevant Effects

In Figure 3.1, I estimate the MTE assuming joint normality of (U_1, U_0, V) and plot it with 90% confidence interval bands.[18] The MTE is monotonically declining as U_D increases, and it suggests substantial heterogeneity of marginal effects of university education. When the U_D is particularly low, for individuals who are more likely to enroll in university, the marginal effects are high: about 31.1%, but for those who have low values of U_D, the effects are substantially low: −12.6%. They incur negative gain from attending university. Carneiro, Heckman, and Vytlacil (2011) note that these results imply individuals self-select university enrollment based on their comparative advantage with respect to their gains.

The IV and MTE estimates suggest that university education has substantial and heterogeneous effects on future pecuniary outcome. Yet, it is not clear how large the average impacts for the different subpopulations, and how these effects are related to educational policies. To examine these issues, I show the treatment effects parameters in Table 3.3, which are constructed from the MTE using the weights presented in Heckman and Vytlacil (2005) and Carneiro, Heckman, and Vytlacil (2010, 2011).

Average effects are substantially different by the groups of population. The ATE shows an additional year of university education increases hourly wage by 6.74%. The ATT is larger than the ATE. It suggests that a return to one year of university education is 11.59% for those who enrolled in university. The ATU is much smaller than the ATE and the ATT. It shows the effect is 2.79% for those who did not enroll in university, which they would gain if they had enrolled in university.

Conventional average treatment parameters are important themselves, but these only ad-

[17] See Angrist and Imbens (1995) and Heckman and Urzua (2010) for their discussion about identification and interpretation of IV estimates.

[18] To depict the MTE, I evaluate it at mean values of control variables.

dress policy questions in extreme cases. For example, the ATU indicates the effects of a policy forcing entire population to receive university education. In contrast, the MPRTE parameter of Carneiro, Heckman, and Vytlacil (2010, 2011) answers questions about the marginal gains of specific policies in more general case. Table 3.3 presents estimates of different definitions of the MPRTE, where the policy is considered as a marginal change in probability of university enrollment. The MPRTE with a policy that increases the probability of university enrollment by an amount α, is 7.18%. A policy that changes the probability of university enrollment by the proportion $(1 + \alpha)$ provides slightly smaller effects. The MPRTE of such a policy is 5.39%, but standard errors are large.

Finally, I calculate the average returns for those who are induced to enroll in university by a particular policy shift. Table 3.3 reports the PRTE of two counter-factual policies: (1) a policy of free tuition in public universities, (2) a policy that increases capacities of university by 1000 places if the prefecture has less than 5000 places. The PRTE of free public tuition shows the effect is 6.47%, which is similar to the magnitude of the ATE. The PRTE of increasing capacities suggests a larger impact, 7.78%.

3.5 Conclusion

This study investigates the returns to schooling of university education in Japan. I create instruments reflecting the direct costs of college attendance: total accredited capacities of all universities and public tuition in the prefecture of residence at the age of 15. This measure captures cross-time and cross-prefecture variation, because I also control for a set of birth cohorts and prefecture dummies. Using local capacities of universities, public tuition, and local labor market conditions as instruments, this study recovers the average effects of university education as weighted averages of the MTE. The result shows that additional university schooling increases the hourly wage, on average, 6.74% for population. In addition, this study finds heterogeneous marginal effect by individual and heterogeneous average effects by the groups of subpopulation. Average effect for those enrolled in university is 11.59%, but the effect is less than 3% for those who did not enroll in.

This study further investigates the effect of university education for those who are induced to enroll in university by a specific policy change. I find that policies increasing probability of university enrollment provide positive effects of university education.

Data Appendix

This appendix describes the construction of the instrumental variables. The data source for capacity of universities is *Zenkoku Daigaku Ichiran*. This book is published with the list of all accredited national, prefectural, municipal and private universities in each academic year by the *Bunkyo Kyokai*. It contains the detailed information on the accredited capacities, the location, and the date of opening and closing by the department of the university. I collected the total quota for new enrollment offered in each prefecture of an academic year in the department level. If the department is located in more than one prefecture, I take the prefecture where students of the department stay in longer. If general education courses are collectively offered in the other prefecture, I assign each department to the prefecture where they offer an upper-level or a specialized course. I exclude the universities that offer only in correspondence, Internet learning and a graduate school. I also exclude following categories of departments: art, music, religious, home economics, and physical education. An important problem is the conversion of female to co-ed universities. *Zenkoku Daigaku Ichiran* provides the information on conversion of single-sex to co-ed in university or college level, but not in department level. I search the history of the university in the official web cite or in an official report published by the university, and I identify the department started to offer courses to male. If I am unable to identify which department was changed to co-ed, I assume that all departments offer co-ed courses based on the information of *Zenkoku Daigaku Ichiran*.

Tuition data are based on *Keisetsu Jidai* in 1972-2000 published by Obunsha. I define tuition is sum of entrance fees and course fees. Some prefectural and municipal universities have price discrimination by residential area of students. I assign minimum price of tuition for intra-regional students if prefectural or municipal universities are available and assign maximum price of tuition for extra-regional students if there is no prefectural or municipal universities in the prefecture of residence at the age of 15. I construct the measure as accredited capacities weighted averages over prefectural and municipal universities in a prefecture, or at the regional level if prefectural and municipal universities are not available. The region is based on the definition of region code of the Labor Force Survey (*Rodo-Ryoku Chosa*) of the Statistics Bureau of the Ministry of Internal Affairs and Communications.

The active job opening to application ratio (*yuko kyujin bairitsu*) is based on the report on employment service (*shokugyou antei gyoumu toukei*) of the Bureau of Employment Security of the Ministry of Health, Labour and Welfare. It is defined as number of active job openings per number of active applications. I use the ratio that excludes new graduates and part-timers. I construct average local earnings in high school years from the annual average monthly total cash earnings (establishments with 30 employees or more) of the Monthly Labor Survey (*Maigetsu Kinro Tokei Chosa*). Local earning is evaluated at the 2005 consumer prices and transform it in logarithm.

References

Angrist, Joshua D. and Guido W. Imbens. 1995. “Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity.” *Journal of the American Statistical Association* 90 (430):431–442.

Arkes, Jeremy. 2010. “Using Unemployment Rates as Instruments to Estimate Returns to Schooling.” *Southern Economic Journal* 76 (3):711–722.

Björklund, Anders and Robert Moffitt. 1987. “The Estimation of Wage Gains and Welfare Gains in Self-Selection Models.” *Review of Economics and Statistics* 69 (1):42–49.

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

———. 2001. “The Dynamics of Educational Attainment for Black, Hispanic, and White Males.” *Journal of Political Economy* 109 (3):455–499.

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

Card, David. 1993. “Using Geographic Variation in College Proximity to Estimate the Return to Schooling.” Working Paper 4483, National Bureau of Economic Research, Cambridge, MA.

———. 2001. “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems.” *Econometrica* 69 (5):1127–1160.

Card, David and Thomas Lemieux. 2001. “Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis.” *Quarterly Journal of Economics* 116 (2):705–746.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. 2010. “Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin.” *Econometrica* 78 (1):377–394.

———. 2011. “Estimating Marginal Returns to Education.” *American Economic Review* 101 (6):2754–2781.

Carneiro, Pedro, Michael Lokshin, Cristobal Ridao-Cano, and Nithin Umapathi. 2011. “Average and Marginal Returns to Upper Secondary Schooling in Indonesia.” IZA Discussion Papers 6162, Institute for the Study of Labor (IZA).

Carneiro, Pedro, Costas Meghir, and Matthias Parey. 2013. “Maternal Wducation, Home Environments, and the Development of Children and Adolescents.” *Journal of the European Economic Association* 11 (s1):123–160.

Chuang, Yih-chyi and Wei-wen Lai. 2010. “Heterogeneity, Comparative Advantage, and Return to Education: The Case of Taiwan.” *Economics of Education Review* 29 (5):804–812.

Currie, Janet and Enrico Moretti. 2003. “Mother’s Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings.” *Quarterly Journal of Economics* 118 (4):1495–1532.

Florens, Jean-Pierre, James J. Heckman, Costas Meghir, and Edward Vytlacil. 2008. “Identification of Treatment Effects Using Control Functions in Models with Continuous, Endogenous Treatment and Heterogeneous Effects.” *Econometrica* 76 (5):1191–1206.

Heckman, James J., Justin L. Tobias, and Edward Vytlacil. 2001. “Four Parameters of Interest in the Evaluation of Social Programs.” *Southern Economic Journal* 68 (2):210–223.

Heckman, James J. and Sergio Urzua. 2010. “Comparing IV with Structural Models: What Simple IV Can and Cannot Identify.” *Journal of Econometrics* 156 (1):27–37.

Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. “Understanding Instrumental Variables in Models with Essential Heterogeneity.” *Review of Economics and Statistics* 88 (3):389–432.

Heckman, James J. and Edward Vytlacil. 1999. “Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects.” *Proceedings of the National Academy of Sciences* 96 (8):4730–4734.

———. 2001a. “Local Instrumental Variables.” In *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya*, edited by Cheng Hsiao, Kimio Morimune, and James L Powell. New York: Cambridge University Press, 1–46.

———. 2001b. “Policy-Relevant Treatment Effects.” *American Economic Review* :107–111.

———. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73 (3):669–738.

———. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments." In *Handbook of Econometrics*, vol. 6, Part B, chap. 71. Amsterdam: Elsevier, 4875–5143.

Imbens, Guido W. and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62:467–475.

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

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

Kobayashi, Masayuki. 2009. *Daigaku Shingaku no Kikai (An Opportunity of University Enrollment)*. Tokyo, Japan: the University of Tokyo Press. (in Japanese).

Nakamuro, Makiko and Tomohiko Inui. 2012. "Estimating the Returns to Education Using the Sample of Twins: The Case of Japan." Discussion Paper 12-E-076, RIETI Discussion Paper Series.

Nakazawa, Wataru. 2011. "Shussin Chiiki ni yoru Kosotsugo Shingaku Kikai no Hubyodo (Inequality of Opportunity for the Progression to Higher Education on the Basis of Residential Areas)." Discussion Paper 43, Panel Survey Project Discussion Paper Series, Institute of Social Science, the University of Tokyo. (in Japanese).

Nybom, Martin. 2012. "College Choice, Abilities and Lifetime Earnings: A Local IV Approach with Swedish Registry Data." Unpublished manuscript, Stockholm University.

OECD. 2013. *Education at a Glance 2013: OECD Indicators 2013 – Country note for Japan*. Paris: OECD Publishing.

Ono, Hiroshi. 2004. "College Quality and Earnings in the Japanese Labor Market." *Industrial Relations: A Journal of Economy and Society* 43 (3):595–617.

Oshio, Takashi and Miki Kobayashi. 2009. "Happiness, Self-Rated Health, and Income Inequality: Evidence from Nationwide Surveys in Japan." PIE/CIS Discussion Paper 451, Center for Intergenerational Studies, Institute of Economic Research, Hitotsubashi University.

Oshio, Takashi and Wataru Seno. 2007. “The Economics of Education in Japan: A Survey of Empirical Studies and Unresolved Issues.” *Japanese Economy* 34 (1):46–81.

Robinson, Peter M. 1988. “Root-N-Consistent Semiparametric Regression.” *Econometrica* 56 (4):931–54.

Sano, Shinpei and Kengo Yasui. 2009. “Nihon ni okeru Kyoiku no Return no Suikei (Estimating the Returns to Education in Japan).” *Kokumin Keizai Zassi* 200 (5):71–86. (in Japanese).

Sasaki, Yosei. 2006. “Kyoiku Kikaino Chiikikan Kakusa (Regional Gaps in Educational Opportunities).” *The Journal of Educational Sociology* 78:303–320. (in Japanese).

Tachibanaki, Toshiaki. 1988. “Education, Occupation, Hierarchy and Earnings.” *Economics of Education Review* 7 (2):221–229.

Trostel, Philip, Ian Walker, and Paul Woolley. 2002. “Estimates of the Economic Return to Schooling for 28 Countries.” *Labour Economics* 9 (1):1–16.

Willis, Robert J and Sherwin Rosen. 1979. “Education and Self-Selection.” *Journal of Political Economy* 87 (5, Part 2):S7–S36.

Yasui, Kengo and Shinpei Sano. 2009. “Kyoiku ga Chingin ni Motarasu Ingateki na Koka ni tsuite (Causal Effects of Education on Wages).” *Nihon Rodo Kenkyu Zasshi* 588:16–33. (in Japanese).

Table 3.1: Summary Statistics

Variables	High school (D = 0)	University (D = 1)
Years of scholing	12.3299	16.1439
	(0.7384)	(0.6304)
Log hourly wage	7.4813	7.7428
	(0.5436)	(0.5369)
Control Variables		
Years of current job experience	13.3465	12.5917
	(9.1970)	(7.9454)
Local log earnings in survey year	12.7634	12.8029
	(0.1032)	(0.1064)
Local unemployment rate	4.6543	4.6229
in survey year (in %)	(0.9872)	(0.9839)
Mother's years of schooling	10.0539	11.4722
	(2.2403)	(2.2391)
Father's years of schooling	10.1556	12.2285
	(2.5825)	(3.0751)
Number of siblings	1.6836	1.3527
	(1.0304)	(0.8185)
Urban residence at age 15	0.0871	0.1937
	(0.2822)	(0.3955)
Rural residence at age 15	0.4346	0.2575
	(0.4960)	(0.4375)
Local population of age 15–19	28.5065	34.4926
(in 10000)	(22.3170)	(25.8994)
Local log earnings at age 13–18	12.5964	12.6493
	(0.2311)	(0.2265)
Local job openings at age 13–18	1.0370	1.0819
	(0.7304)	(0.7093)
Instrumental Variables		
Capacity of universities	11.0961	17.3315
(in 1000)	(18.4653)	(26.8027)
Tuition in public universities	25.0574	25.9429
(in 10000)	(18.3435)	(17.7490)
Local log earnings in high school	12.6435	12.6926
	(0.1996)	(0.1971)
Local job openings in high school	1.0296	1.0402
	(0.7328)	(0.7067)
Number of observations	964	862

Notes: This table reports summary statistics of the analysis data. Local job openings indicate the active job opening to application ratio. Standard deviations are in parentheses.

Table 3.2: University Enrollment Decision

Dependent variable:	University Enrollment
Control Variables	
Father's years of schooling	0.0332
	(0.0074)
Mother's years of schooling	0.0421
	(0.0091)
Number of siblings	-0.0819
	(0.0163)
Urban residence at age 15	0.1167
	(0.0486)
Rural residence at age 15	-0.0825
	(0.0272)
Local population of age 15-19	-0.0113
(in 10000)	(0.0047)
Local log earnings at age 13-18	1.7652
	(0.9886)
Local job openings at age 13-18	-0.1306
	(0.0845)
Instrumental Variables	
Capacity of universities	0.1078
(in 1000)	(0.0044)
Tuition in public universities	-0.0060
(in 10000 yen)	(0.0020)
Local log earnings in high school	-2.0409
	(0.8869)
Local job openings in high school	0.1120
	(0.0814)
Test for joint significance of IVs	
Chi-square	18.61
p-value	0.0009

Notes: This table reports the marginal effects evaluated at the mean value of each variable from a logit regression of university enrollment (a dummy variable that is equal to one if an individual has ever attended university, and equal to zero if he has never attended university but has completed upper secondary education). Local job openings indicate the active job opening to application ratio. Cohort dummies and a set of dummies for prefecture of residence at the age of 15 are also controlled in the model but not reported. Robust standard errors are in parentheses. Chi-square and p-values indicate the results of the test of joint significance of coefficients on the instrumental variables.

Table 3.3: Estimates of Returns to a Year of University Education

Parameters	
OLS	0.0532
	(0.0059)
IV	0.1067
	(0.0499)
ATE	0.0674
	(0.0370)
ATT	0.1159
	(0.0637)
ATU	0.0279
	(0.0453)
MPRTE	
$P_\alpha = P + \alpha$	0.0718
	(0.0393)
$P_\alpha = (1 + \alpha)P$	0.0539
	(0.0367)
PRTE	
Free tuition	0.0647
	(0.0373)
Increase in capacities of universities	0.0778
	(0.0403)

Notes: This table reports estimates of returns to university education for the normal selection model: average treatment effect (ATE), average treatment on the treated (ATT), average treatment on the untreated (ATU), the marginal policy relevant treatment effect (MPRTE), and policy relevant treatment effect (PRTE). The PRTE corresponds to the two counter-factual policies: (1) Free tuition: a policy of free tuition in public universities, (2) Increase in capacities of universities: a policy that increases capacities of universities by 1000 places if the prefecture has less than 5000 places. The IV estimate uses $P(Z)$ as the instrument (Logit model for first stage with all instruments). Standard errors (in parentheses) are obtained by the bootstrap method (250 replications). All estimates are annualized (divided by 3.81 years).

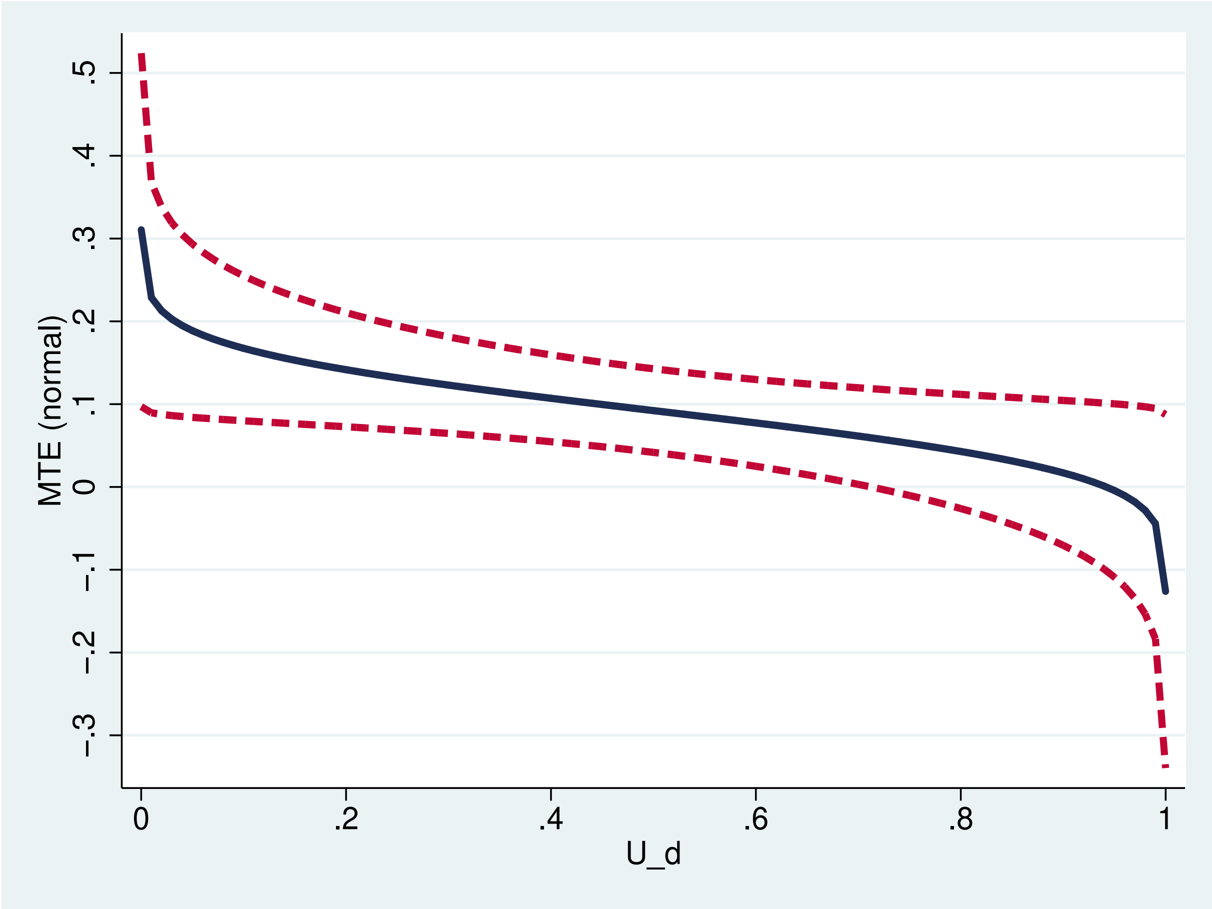


Figure 3.1: Marginal Treatment Effect

Notes: This figure depicts marginal treatment effect with assumptions on normally distributed unobservables. A solid line indicates the estimated effect. Dashed lines indicate 90% confidence intervals.

Chapter 4

Intergenerational Transmission of Education in Japan

4.1 Introduction

The education of parents is one of the fundamental factors in explaining the child's educational success. Empirical studies of social science pay much attention to the intergenerational correlation between parents' education and child's education. The recent literature advances the analysis to quantify the causal effects of parents' schooling in addition to obtaining precise estimates of the correlation.

To estimate the causal impact, the literature depends on the identification strategies that use identical twins, adopted children, and instrumental variables (IV). However, as Holmlund, Lindahl, and Plug (2011) point out, different strategies provide different results. Hence, these studies do not reach a consensus on the magnitude of the effects.[1]

In Japan, previous empirical studies also find a significant positive correlation between parents' schooling and child's schooling.[2] Unfortunately, these studies have paid limited attention to the causal relationship. Researchers require micro-data sets that contain twins, adopted children or plenty of information about individual characteristics to estimate the causal effects. However, it is difficult to access such data sets in Japan.[3] This limited accessibility to informative micro-data sets prevents researchers from controlling individual heterogeneity, and it magnifies the inevitable concern for point identification under strong assumptions on treatment selection. For this reason, I derive sets of nonparametric bounds under relatively weak and testable assumptions that appear to share broad consensus in the literature.

In the empirical field of social science, the credibility of mean-independence or exogenous

[1]See also Black and Devereux (2011) for a survey.

[2]For example, see Tachibanaki (1988) and Yamada (2011). See also Kariya (2001) as an example of sociological research on the relationship between educational attainment and family background.

[3]Nakamuro and Inui (2012) are one of the exceptions. They collect twin data through a web-based survey.

treatment conditions is usually a matter of considerable disagreement. Furthermore, it is difficult to find a variable that satisfies the IV assumption. If the identification assumption on the exogenous treatment is not valid, a simple regression coefficient provides a biased estimate of the causal effect. Manski (1997) and Manski and Pepper (1998, 2000) introduce the nonparametric bound analysis under a set of monotonicity assumptions. They show that even if the strong assumptions for point identification are not satisfied, relatively weak assumptions can effectively tighten the bounds on the causal effect of the treatment. In recent years, the number of studies applying a nonparametric bounding method has been growing (e.g. Blundell et al., 2007; Gonzalez, 2005; Gundersen and Kreider, 2009; Kreider and Hill, 2009; Kreider and Pepper, 2007; Lee and Wilke, 2009; Okumra and Usui, 2010).[4] De Haan (2011) uses the nonparametric bound method to examine the effects of parents' schooling on child's schooling. With the U.S. data, she shows that mother's or father's college degree has a positive effect on child's schooling. Moreover, the estimated bounds are significantly different from zero but substantially lower than the ordinary least squares (OLS) estimates.

Most of the studies on causal effect of parents' education on the child's education, including De Haan (2011), have focused on the effect of a single treatment of each of the parents. However, the estimated effect might not be the treatment effect of interest in the single treatment framework. The estimated effect of a parental schooling includes both the direct transfer from the given parent and the indirect transfer from the other parent due to assortative mating effects. On the other hand, the studies using the regression method control a variable of spousal education as an additional regressor, but usually ignore the endogeneity of the variable.[5]

As Holmlund, Lindahl, and Plug (2011) note, it is not clear whether spousal education should be controlled in the analysis, and if so, how this may be done, because the strong positive correlation of father's and mother's schooling makes it difficult to interpret the coefficient for each parent separately. Oreopoulos, Page, and Stevens (2006) propose an alternative way to control spousal education. They use the sum of mother's and father's completed education as the explanatory variable instead of including each parent's education separately. It allows directly estimating the effect of a one-year increase in either parent's schooling. However, the specification assumes that the effects of father's and mother's education are the same. It is also not clear that the assumption is satisfied in general.

In practice, children are exposed to a combination of multiple treatments of mother's and father's educational level rather than exposed to a single treatment of each parent's educational

[4]See Manski (2003) and Manski and Pepper (2009) for a review.

[5]An advantage of using the regression method is that parental income can also be included as a control. This helps to understand the mechanisms by which parent's education affects the child's education, because parental education has both a direct impacts through generating differences in the quality of child's endowments and an indirect impacts of higher quantity or quality of inputs due to the positive effect of parent's education on the family income (Chevalier et al., 2013). Because the data set does not contain precise information on parental income during the adolescent years, I am unable to control parental income in the analysis. Therefore, the analyzed effects include the indirect impact of parent's education on the child's education.

level, at least biological meanings. Therefore, it is natural to consider a set of treatments of both parents' schooling simultaneously. A specification of multiple treatments avoids the discussion about the proper way to control the spousal effects, but directly evaluates the effects of a combination of both parents' schooling. Furthermore, it allows the heterogeneous effects of each parent's schooling. Unfortunately, very little is known about the effects of a combination of father's and mother's schooling as an application of the evaluation method with multiple treatments.

This study contributes to the literature by assessing the treatment effect of a combination of parents' schooling on the child's schooling in Japan using a nonparametric bounds approach. To do so, I consider a set of treatment vectors of mother's and father's education and assume that the treatment takes semi-ordered multiple values. In the analysis, rather than imposing the strong assumptions required to obtain point estimates, I rely on a set of monotonicity and semi-monotonicity assumptions to acquire informative bounds. The assumptions are the monotone treatment response (MTR) of Manski (1997), the monotone selection (MTS), and the monotone instrumental variables (MIV) of Manski and Pepper (1998, 2000).

The main results show that the tightest bounds on the average effects of parents' schooling are positive, but lower than the point estimates that rely on the assumption on exogenous selection of parents' schooling. These results suggest that the regression estimates that are mainly provided in the previous Japanese studies have an upward bias from the true causal effect of parents' schooling.

The remainder of this chapter is organized as follows. Section 4.2 defines the parameter of interest and describes the identification assumptions. Section 4.3 explains the data set. Section 4.4 shows and discusses the results. Finally, Section 4.5 presents the conclusions.

4.2 Method

This study employs a nonparametric technique to obtain the bounds on the causal effect of parents' education on child's education. I basically follow the setup of Manski and Pepper (2000). There is a probability space (J, Ω, P) of individuals. Each member j of population J has observable covariates $x_j \in X$ and a response function $y_j(\cdot) : T \to Y$ mapping the mutually exclusive and exhaustive treatments $t \in T$ into real-valued outcomes $y_j(t) \in Y$, where the treatment T is the level of schooling that the parent completed, thus assumed to be an ordered set. Y is years of schooling or university graduation of the child j, and $x = (w, v)$, $X = W \times V$. Each value of (w, v) defines an observable sub-population of children. I assume that the outcome space Y has the greatest lower bound $Y_{min} \equiv min\ Y$ and the least upper bound

$Y_{max} \equiv max\, Y$.[6]

Child j has a realized treatment $z_j \in T$ and a realized outcome $y_j \equiv y_j(z_j)$, both of which are observable. The latent outcomes $y_j(t), t \neq z_j$ are not observable. A researcher learns the distribution P(x, z, y) of covariates, realized treatments and realized outcomes by observing a random sample of the population. To simplify the notation, the subscript j will be dropped except when it required.

The parameter of interest is the average treatment effect (ATE) between parental levels of education $t \in T$ and $t' \in T$, $s.t.\ t > t'$:

$$ATE(t,t') \equiv E[y(t) - y(t')] = E[y(t)] - E[y(t')]. \tag{4.1}$$

The linearity of expectations provides the second equality. The empirical problem comes down to learning the $E[y(\cdot)]$ from the empirical evidence of the the distribution $P(x, z, y)$ of covariates, realized treatments and realized outcomes with assumptions. The fact that $E[y|z = t] = E[y(t)|z = t]$ and the law of iterated expectations gives,

$$E[y(t)] = E[y|z = t] \cdot Pr(z = t) + E[y(t)|z \neq t] \cdot Pr(z \neq t). \tag{4.2}$$

With a data set, researchers can estimate $E[y|z = t], Pr(z = t), Pr(z \neq t)$, but they cannot estimate $E[y(t)|z \neq t]$ without adding any assumptions. Manski (1989, 1990) shows that no assumption (worst case) bounds on $E[y(t)]$ can be identified if the support of the dependent variable is bounded:

$$\begin{aligned} E[y|z = t] \cdot Pr(z = t) + Y_{min} \cdot Pr(z \neq t) \\ \leq E[y(t)] \leq \\ E[y|z = t] \cdot Pr(z = t) + Y_{max} \cdot Pr(z \neq t). \end{aligned} \tag{4.3}$$

The width of bounds is $(Y_{max} - Y_{min}) \cdot Pr(z \neq t)$, thus the bounds are informative if $Pr(z \neq t) < 1$. These bounds on $E[y(\cdot)]$ provide the lower (upper) bound on the ATE(t,t') by subtracting the upper (lower) bound on $E[y(t')]$ from the lower (upper) bound on $E[y(t)]$. Hence the identification region $H\{\cdot\}$ for the ATE is the interval:

[6] When I analyze the effects on child's years of schooling, I assume $Y_{min} = 9$ and $Y_{max} = 18$ years of schooling from the construction of the data set.

$$
\begin{aligned}
&H\{E[y(t)] - E[y(t')]\}\\
&= \{E[y|z=t] \cdot Pr(z=t) + Y_{min} \cdot Pr(z \neq t) - E[y|z=t'] \cdot Pr(y=t') + Y_{max} \cdot Pr(z \neq t'),\\
&E[y|z=t] \cdot Pr(y=t) + Y_{max} \cdot Pr(z \neq t) - E[y|z=t'] \cdot Pr(z=t') + Y_{min} \cdot Pr(z \neq t')\}.
\end{aligned}
$$

This interval contains the value zero. Its width is $(Y_{max} - Y_{min}) \cdot [Pr(z \neq t) + Pr(z \neq t')]$. The no assumption bounds are important as a beginning because these bounds include all results that depend on different assumptions about $E[y(t)|z \neq t]$. However, the no assumption bounds are too wide to obtain an informative result in this study. Therefore, some additional assumptions are necessary to tighten the bounds. In the remainder of this section, I explain how assumptions tighten the bounds. I consider the exogenous treatment selection (ETS) assumption and the following three assumptions: (1) the MTR, (2) the MTS and (3) the MIV. In addition, I apply the semi-monotonicity version of these three assumptions.

4.2.1 Exogenous Treatment Selection

The ETS assumes that the realized treatment z is statistically independent of the latent outcomes $y(\cdot)$, and can be expressed as:

$For\ each\ t \in T, and\ all\ (u_1, u_2) \in T \times T$

$$E[y(t)|z = u_1] = E[y(t)|z = u_2]. \tag{4.4}$$

or

$$E[y(t)] = E[y|z = t]. \tag{4.5}$$

Under the ETS assumption, the parameter of interest becomes: $E[y|z = t] - E[y|z = t']$. The ETS assumption or covariates x conditional version of it, $E[y(t)|x] = E[y|z = t, x]$, is implicitly imposed in much of the empirical literature. This assumption implies that treatments are assigned randomly to the population and excludes the existence of heterogeneous treatment effects with essential heterogeneity (Heckman, Urzua, and Vytlacil, 2006). It does not seem to be appropriate in the intergenerational transmission of education. For example, one could consider a situation that parents received different educational levels because they differ substantially from one another, but this heterogeneity is unobservable by the researchers. If the heterogeneity also determines child's schooling, the realized treatment is endogenous in the analysis. Moreover, it is difficult to use micro-data sets with plenty of information on individual characteristics in Japan. In that case, the researchers cannot fully control the heterogeneity, thus endogenous selection of parents' education violates the ETS assumption.

However, one advantage of assuming the ETS is that it yields point identification. The ETS estimates are equivalent to the coefficients obtained by regression child's education on dummy variables for each parental educational level and no other covariates. This study shows the point estimates under the ETS assumption for comparison with bounds estimates.

4.2.2 Monotone Treatment Response

Manski (1997) proposes the MTR assumption that states response functions are weakly increasing. That is,

For each $j \in J$ *and all* $(t_1, t_2) \in T \times T$ *s.t.* $t_1 \leq t_2$,

$$t_1 \leq t_2 \Rightarrow y_j(t_1) \leq y_j(t_2). \tag{4.6}$$

The MTR assumes that an increasing parents' schooling does not decrease the child's schooling. The validity of this assumption is suggested in the literature. For instance, the human capital theory (Becker, 1975; Becker and Tomes, 1979; Solon, 1999) suggests a positive impact of increasing parents' schooling on child's schooling. Empirical studies of returns to schooling show that an additional year of schooling has positive effects on both pecuniary and non-pecuniary outcomes (Card, 1999; Oreopoulos and Salvanes, 2011). The higher average income of highly educated parents makes investment in education easier with credit constraints. More educated parents more likely use their time for their children (Guryan, Hurst, and Kearney, 2008). It is worth noting that the MTR assumption does not exclude zero effect of parental education. It allows the case that the positive correlation between parents' schooling and child's schooling has no causal relationship.

Manski (1997) shows that the MTR assumption tightens the no assumption bound,[7]

$$\begin{aligned} t < z_j &\Rightarrow Y_{min} \leq y_j(t) \leq y_j, \\ t = z_j &\Rightarrow y_j(t) = y_j, \\ t > z_j &\Rightarrow y_j \leq y_j(t) \leq Y_{max}. \end{aligned} \tag{4.7}$$

[7]For a full derivation of the MTR bounds, see Manski (1997).

thus,

$$
\begin{aligned}
E[y|z \leq t] \cdot Pr(z \leq t) + Y_{min} \cdot Pr(z > t)& \\
\leq E[y(t)] \leq& \\
Y_{max} \cdot Pr(z < t) + E[y|z \geq t] \cdot Pr(z \geq t)&.
\end{aligned} \tag{4.8}
$$

4.2.3 Monotone Treatment Selection

Manski and Pepper (1998, 2000) weaken equation 4.4 to an inequality and yield the MTS assumption:

$For\ each\ t \in T\ and\ all\ (u_1, u_2) \in T \times T\ s.t.\ u_1 \leq u_2$

$$u_1 \leq u_2 \Rightarrow E[y(t)|z = u_1] \leq E[y(t)|z = u_2]. \tag{4.9}$$

If the MTS assumption holds, then,

$$
\begin{aligned}
u < t &\Rightarrow Y_{min} \leq E[y(t)|z = u] \leq E[y|z = t] \\
u = t &\Rightarrow E[y(t)|z = u] = E[y|z = t] \\
u > t &\Rightarrow E[y|z = t] \leq E[y(t)|z = u] \leq Y_{max}.
\end{aligned} \tag{4.10}
$$

It follows that

$$
\begin{aligned}
Y_{min} \cdot Pr(z < t) + E[y|z = t] \cdot Pr(z \geq t)& \\
\leq E[y(t)] \leq& \\
E[y|z = t] \cdot Pr(y \leq t) + Y_{max} \cdot Pr(z > t)&.
\end{aligned} \tag{4.11}
$$

Manski and Pepper (1998, 2000) also show that combining the MTR and the MTS assumptions can have substantial identifying power.[8] Let the MTR and the MTS assumptions 4.6 and 4.9 hold. Then, bounds reduce to

[8] For a full derivation of the MTS and the MTR-MTS bounds, see Manski and Pepper (1998, 2000).

$$E[y|z < t] \cdot Pr(z < t) + E[y|z = t] \cdot Pr(z \geq t)$$
$$\leq E[y(t)] \leq \tag{4.12}$$
$$E[y|z = t] \cdot Pr(z \leq t) + E[y|z > t] \cdot Pr(z > t).$$

If I invoke the interpretation of Manski and Pepper (1998, 2000), the MTS assumption indicates that parents who select higher levels of schooling have a weakly higher mean response function than those who select lower levels. It implies that a genetic transmission of unobserved ability from parents to children provides the correlation of educational levels. Researchers commonly assume that unmeasured inherent differences are key source of endogeneity of the education. Therefore, the MTS assumption is consistent with the arguments in the literature. Furthermore, Manski and Pepper (1998, 2000) suggest a simple test for the validity of the joint MTR-MTS assumption. Under the assumption,

$For\ all\ (u1, u2) \in T \times T,\ s.t.\ u_1 \leq u_2,$

$$E[y|z = u_1] = E[y(u_1)|z = u_1] \leq E[y(u_1)|z = u_2] \leq E[y(u_2)|z = u_2] = E[y|z = u_2]. \tag{4.13}$$

The MTS assumption implies the first inequality and the MTR assumption gives the second inequality. Under this hypothesis, $E(y|z = u)$ must be a weakly increasing function of u. Hence the researcher should reject the hypothesis if the mean outcomes of the child's education are not weakly increasing in the realized level of schooling of the parent.

4.2.4 Monotone Instrumental Variables

The recent literature of causal intergenerational schooling effects has moved away from estimating OLS under the ETS assumption, and relies on alternative identification assumptions (Holmlund, Lindahl, and Plug, 2011). One of the assumptions is the IV. Using the notation in the present paper, the familiar mean-independence form of the IV assumption can be written as,

$For\ each\ t \in T\ and\ all\ (u_1, u_2) \in V \times V,$

$$E[y(t)|v = u_1] = E[y(t)|v = u_2]. \tag{4.14}$$

If the treatment effects are homogeneous, the IV estimator recovers the parameter of interest. Even if the effects are heterogeneous, researchers can recover the parameter of interest or a related interpretable parameter using an IV with additional assumptions on a selection model.[9] Unfortunately, it is hard to find an IV that satisfies the mean-independence conditions. Instead of assuming mean-independence, Manski and Pepper (1998, 2000) introduce a weaker assumption to replace the equality in 4.14 by an inequality, yielding a weakly monotone relation between the instrumental variable and the mean response function.

For each $t \in T$ *and all* $(u_1, u_2) \in V \times V$,

$$u_1 \leq u_2 \Rightarrow E[y(t)|v = u_1] \leq E[y(t)|v = u_2]. \tag{4.15}$$

In the case that the MIV assumption holds, the bounds to be:

$$\begin{aligned}
&\sum_{u \in V} Pr(v = u) \cdot \Big\{max_{u_1 \leq u}[E[y|z = t, v = u_1] \cdot Pr(z = t|v = u_1) + Y_{min} \cdot Pr(z \neq t|v = u_1)]\Big\} \\
&\leq E[y(t)] \leq \\
&\sum_{u \in V} Pr(v = u) \cdot \Big\{min_{u_2 \geq u}[E[y|z = t, v = u_2] \cdot Pr(z = t|v = u_2) + Y_{max} \cdot Pr(z \neq t|v = u_2)]\Big\}.
\end{aligned} \tag{4.16}$$

The MIV and the MTR-MTS assumptions make distinct contributions to the identification. Manski and Pepper (1998, 2000) suggest that combining the MIV and the MTR-MTS assumptions yields particularly interesting bounds.[10]

$$\begin{aligned}
&\sum_{u \in V} Pr(v = u) \cdot \Big\{max_{u_1 \leq u}[E[y|z < t, v = u_1] \cdot Pr(z < t|v = u_1) \\
&\qquad + E[y|z = t, v = u_1] \cdot Pr(z \geq t|v = u_1)]\Big\} \\
&\leq E[y(t)] \leq \\
&\sum_{u \in V} Pr(v = u) \cdot \Big\{min_{u_2 \geq u}[E[y|z = t, v = u_2] \cdot Pr(z \leq t|v = u_2) \\
&\qquad + E[y|z > t, v = u_2] \cdot Pr(z > t|v = u_2)]\Big\}.
\end{aligned} \tag{4.17}$$

[9]See Imbens and Angrist (1994) and Heckman and Vytlacil (2007) for more details.
[10]For a full derivation of the MIV and the MTR-MTS-MIV bounds, see Manski and Pepper (1998, 2000).

4.2.5 Semi-Monotone Instrumental Variables

Manski and Pepper (1998) also suggest the ways to combine multiple scalars of IV or MIV assumptions. In this study, I use two MIVs simultaneously to tighten the bounds. To do so, I assume that the pair of MIVs, (v^a, v^b) is a two-dimensional semi-monotone instrumental variable (SMIV):

$$For\ each\ t \in T\ and\ all\ [(u_1^a, u_1^b), (u_2^a, u_2^b)] \in (V^a \times V^b) \times (V^a \times V^b)\ s.t.\ u_1^a \leq u_2^a\ and\ u_1^b \leq u_2^b$$

$$\Rightarrow E[y(t)|(v^a, v^b) = (u_1^a, u_1^b)] \leq E[y(t)|(v^a, v^b) = (u_2^a, u_2^b)]. \tag{4.18}$$

In this assumption, a pair of MIVs is assumed to be semi-ordered rather than ordered because it includes some pairs of values that are not ordered.[11] The MTR-MTS-MIV bounds in 4.17 can be extended to the MTR-MTS-SMIV bounds if the maxima and minima operations are taken over pairs of values that are ordered.

4.2.6 Semi-Monotone Treatment Response

The previous studies of the effect of parent's schooling on child's schooling have focused on the identification of a single treatment of one parent and have estimated the effect of mother's and father's schooling severally. By constrast, children are affected by a combination of both parents' levels of education in practice. This indicates that the treatment of parents' schooling are multiple for the child. In this study, I analyze effects of both parents' schooling simultaneously in the framework of the nonparametric bounding approach.[12]

Manski (1997) proposes the assumption of semi-monotone treatment response (SMTR) where a semi-ordered vector of multi-valued treatments $\mathbf{T}$ exists. It suggests that the discussion based on the single ordered treatment of a parent's schooling can be extended to a semi-ordered vector of multiple treatments, which is a combination of each parent's schooling. This study considers two-dimensional treatment vectors of mother's and father's schooling, $(t^M, t^F) = \mathbf{t} \in \mathbf{T}$. In this case, the SMTR assumption states:

$$For\ each\ j \in J\ and\ all\ (\mathbf{t}_1, \mathbf{t}_2) \in \mathbf{T} \times \mathbf{T}\ s.t.\ \mathbf{t}_1 \leq \mathbf{t}_2,$$

$$\mathbf{t}_1 \leq \mathbf{t}_2 \Rightarrow y_j(\mathbf{t}_1) \leq y_j(\mathbf{t}_2). \tag{4.19}$$

[11] For example, consider a case that $u_1^a < u_2^a$ *and* $u_1^b > u_2^b$. It does not predict the ordering of $E[y(t)|(v^a, v^b) = (u_k^a, u_k^b)]$, $k = 1, 2$.

[12] The matching approach is an alternative method to analyze multiple treatments. For example, see Imbens (2000); Lechner (2001). See also Flores and Mitnik (2013) as an example applying the difference-in-difference method for multiple treatments. Frölich (2004) is a comprehensive survey of the analysis on multiple treatments.

where I define $\mathbf{t}_1 \leq \mathbf{t}_2$ $if\ and\ only\ if\ t_1^M \leq t_2^M\ and\ t_1^F \leq t_2^F$.

The analysis under the MTR assumption still holds under the assumption of SMTR except the case where treatments are not ordered:[13]

$$
\begin{aligned}
&\mathbf{t} < \mathbf{z}_j \Rightarrow Y_{min} \leq y_j(\mathbf{t}) \leq y_j, \\
&\mathbf{t} = \mathbf{z}_j \Rightarrow y_j(\mathbf{t}) = y_j, \\
&\mathbf{t} > \mathbf{z}_j \Rightarrow y_j \leq y_j(\mathbf{t}) \leq Y_{max}. \\
&\mathbf{t} \varnothing \mathbf{z}_j \Rightarrow Y_{min} \leq y_j(\mathbf{t}) \leq Y_{max}.
\end{aligned}
\tag{4.20}
$$

where $\varnothing$ denotes that $\mathbf{t}$ and $\mathbf{z}$ are not ordered. Then, I have the following the SMTR bounds:

$$
\begin{aligned}
E[y|\mathbf{z} \leq \mathbf{t}] \cdot Pr(\mathbf{z} \leq \mathbf{t}) + Y_{min} \cdot Pr(\mathbf{z} > \mathbf{t}) + Y_{min} \cdot Pr(\mathbf{z} \varnothing \mathbf{t}) \\
\leq E[y(\mathbf{t})] \leq \\
Y_{max} \cdot Pr(\mathbf{z} < \mathbf{t}) + E[y|\mathbf{z} \geq \mathbf{t}] \cdot Pr(\mathbf{z} \geq \mathbf{t}) + Y_{max} \cdot Pr(\mathbf{z} \varnothing \mathbf{t}).
\end{aligned}
\tag{4.21}
$$

4.2.7 Semi-Monotone Treatment Selection

The SMIV assumption of Manski and Pepper (1998) suggests an assumption on selection with a semi-ordered vector of multi-valued treatment. The semi-monotone treatment selection (SMTS) assumption is the special SMIV assumption that the MIVs, $\mathbf{v} = (v^a, v^b)$ are the realized treatments $\mathbf{z}$. With the SMTS assumption, the analysis under the MTS assumption is also available except the case where treatments are not ordered. Therefore, I can get the SMTR-SMTS bounds by combining the SMTR assumption and the SMTS assumption as well as the analysis on the single treatment. Let the SMTR and the SMTS assumption hold.[14] Then,

$$
\begin{aligned}
&\mathbf{u} < \mathbf{t} \Rightarrow Y_{min} \leq E[y|\mathbf{z} = \mathbf{u}] \leq E[y(\mathbf{t})|\mathbf{z} = \mathbf{u}] \leq E[y|\mathbf{z} = \mathbf{t}] \\
&\mathbf{u} = \mathbf{t} \Rightarrow E[y(\mathbf{t})|\mathbf{z} = \mathbf{u}] = E[y|\mathbf{z} = \mathbf{t}] \\
&\mathbf{u} > \mathbf{t} \Rightarrow E[y|\mathbf{z} = \mathbf{t}] \leq E[y(\mathbf{t})|\mathbf{z} = \mathbf{u}] \leq E[y|\mathbf{z} = \mathbf{u}] \leq Y_{max}. \\
&\mathbf{u} \varnothing \mathbf{t} \Rightarrow Y_{min} \leq y(\mathbf{t}) \leq Y_{max}.
\end{aligned}
\tag{4.22}
$$

[13]For a full derivation of the SMTR bounds, see Manski (1997).

[14]For a full derivation MTR-MTS bounds and a further discussion on the semi-monotonicity assumptions, see Manski (1997) and Manski and Pepper (1998).

It follows that

$$
\begin{aligned}
E[y|\mathbf{z} < \mathbf{t}] \cdot Pr(\mathbf{z} < \mathbf{t}) + E[y|\mathbf{z} = \mathbf{t}] \cdot Pr(\mathbf{z} \geq \mathbf{t}) + Y_{min} \cdot Pr(\mathbf{z} \varnothing \mathbf{t}) \\
\leq E[y(\mathbf{t})] \leq \\
E[y|\mathbf{z} = \mathbf{t}] \cdot Pr(\mathbf{z} \leq \mathbf{t}) + E[y|\mathbf{z} > \mathbf{t}] \cdot Pr(\mathbf{z} > \mathbf{t}) + Y_{max} \cdot Pr(\mathbf{z} \varnothing \mathbf{t}).
\end{aligned}
\tag{4.23}
$$

These bounds can be viewed as a natural extension of the MTR-MTS bounds. The bounds can be combined with the MIV and the SMIV assumptions as well as the single treatment bounds since the assumptions are not mutually exclusive.

4.3 Data

The main analysis data are the Japanese General Social Surveys (JGSS).[15] The JGSS are repeated cross-section data for men and women aged 20-89 on each survey date. This study uses the surveys conducted in 2000, 2001, 2002, 2003, 2005, 2006, 2008, and 2010, and pools the respondents from all waves. From the pooled original data, I exclude the observations by following four steps.

First, I restrict the sample by their age. I exclude the people who were younger than 25 years old on the survey year because they might not complete their academic schooling. Second, I drop the individuals who answered that their father or mother was absent at the age of 15, because I am unable to control unobservable characteristics if parental structures are substantially different between those with single parent and those with couple parents. Third, I only use the observations who were born after 1940 because the current educational system was enacted after 1947. Finally, I use the individuals whose observational characteristics are available. So the remained sample after the restrictions contains 13669 individuals were born in 1940–1984 and were at the age of 25–69 on the date of survey.

For the educational attainment of the respondents (children) and their parents, I use the information about the level of the last school attended and assign the standard years of schooling in Japan. For children, completion of their last schooling is available. I follow Tanaka (2008), reduce the number of years of education by one year from standard years for those who dropped

[15]The Japanese General Social Surveys (JGSS) are designed and carried out by the JGSS Research Center at Osaka University of Commerce (Joint Usage / Research Center for Japanese General Social Surveys accredited by Minister of Education, Culture, Sports, Science and Technology), in collaboration with the Institute of Social Science at the University of Tokyo.

out before finishing their last school. To examine the robustness of the definition of the outcome variable, I also use an indicator of university graduation.

In the analysis, I consider the following four schooling levels for parents: (1) Less than High School denotes that the parent completes compulsory schooling or lower secondary education, (2) High School denotes that the parent completes high school or upper secondary education, (3) Some College denotes that the parent completes a college or a higher education, (4) University denotes that the parent has a bachelor's degree. I also use schooling of parent as an MIV. When I obtain bounds on the effect of mother's schooling, I use the indicator of father's university graduation as an MIV. For father's schooling, the indicator of mother's college graduation takes on a role of an MIV.

Under the SMIV assumption, I use two additional dummy variables as MIVs. The first additional MIV is a dummy variable of father's "regular" worker status that takes one if the father was an executive of a company or was an employee with non-terminable contract (*joji-koyo no rodosha*) when the child was 15 years old. A father's more stable working status implies more stable parental income in the child's adolescent years. It is natural to assume that a child with more parental income is less likely to face credit constraints at the decision on an additional schooling. The second additional MIV is a dummy variable that takes one if the child was born after 1975. Taking the birth cohort as an MIV is equivalent to assuming that children who were born in more recent years have weakly higher mean outcome functions than those who were born in older years. This assumption can be attributed to the facts of decreasing trends of number of children, non-decreasing trends of number of schools per child, and stable non-decreasing trends of post compulsory school or university enrollment rate in recent years in Japan. Table 4.1 gives some summary statistics.

4.4 Results

Before presenting the results of bounds on the ATE, I display that the combined assumption of the MTR and the MTS cannot be rejected in the analysis data. Table 4.2 reports mean schooling outcomes of children by educational level of parents. All variables show that the MTR-MTS assumption is not rejected since average values of outcome variables are weakly increasing both in the level of mother's and father's schooling.

4.4.1 Single Treatment Effects on Child's Years of Schooling

I show the results of analyses on child's years of schooling. Table 4.3 provides bounds on the ATE of mother's and father's schooling on child' years of schooling. Following Manski and Pepper (1998, 2000), Gonzalez (2005), and Giustinelli (2011) all tables below also report the

95% confidence intervals that obtained by percentile bootstrap of 3000 replications.

The column of the ETS assumption reports point estimates of the effects. The effect of the change of mother's schooling from junior high school to high school shows an increase in child's schooling by about 1.50 years. The corresponding effect of father's schooling increases child's schooling by about 1.30 years. The change of mother's schooling from high school to some college increases child's schooling by about 1.02 years. When increasing father's schooling from high school to some college, it increases child's schooling by about 0.74 years. These results suggest stronger effects of maternal education than paternal education on child's years of schooling. However, the ETS results of bachelor's degree seem to be mixed. Compared to college degree, mother's bachelor degree has smaller effects than father's bachelor degree. When increasing mother's schooling from junior high school to bachelor's degree, it increases child's schooling by about 2.95 years. It is larger than the corresponding effect of father. When increasing father's schooling from junior high school to bachelor's degree, it increases child's schooling by about 2.65 years.

It is clear from the columns (2) and (7), the no assumption bounds are extremely wide. Imposing the MTR assumption increases significantly the lower bounds by the definition of the assumption that implies non-negative effects (columns (3) and (8)). The MTS assumption allows tightening upper and lower bounds compared to the no assumption bounds, but this is not powerful enough to acquire information on the treatment effect (columns (4) and (9)). The combination of the MTR and the MTS assumptions leads substantially tighter bounds than imposing each assumption alone. Unfortunately, the MTR-MTS assumption cannot provide informative results itself (columns (5) and (10)).

Table 4.3 also shows the bounds under the MIV assumption and combination of it. The MIV assumption tightens upper and lower bounds compared to the no assumption bounds, but this is not powerful enough to obtain informative bounds as well as the MTS assumption (columns (11) and (15)). These results do not depend on the definition of the MIV (columns (12) and (16)). The columns of (13), (14), (17) and (18) show the bounds under the MTR-MTS-MIV assumption. The combination of these three assumptions leads to informative bounds. I get upper bounds that are lower than the ETS results in some cases. The ETS point estimates fall outside the confidence intervals when increasing parent's schooling from junior high school to university.

4.4.2 Single Treatment Effects on Child's University Graduation

To examine the robustness of the results of Table 4.3, I estimate bounds on an alternative definition of outcomes. Table 4.4 displays the results of the effects on child's university graduation. The MTR, the MTS bounds are rather wide and not very informative, these are not shown in Table 4.4.

In all levels of parental education, the ETS estimates show that additional years of schooling of a parent have significant impacts on probability of the child's university graduation (columns (1) and (2)). However, the MTR-MTS-MIV upper bounds suggest that the ETS estimates overstate the ATE in some levels of parental education (columns (4), (5), (9), and (10)). The upper bounds under the MTR-MTS-MIV assumption suggest much smaller effects of parent's bachelor's degree on child's university graduation when I compare with the parent who completes only compulsory education or lower secondary education. Unfortunately, the ETS estimates of lower categories of parent's schooling report very small effects. Even I impose the MTR-MTS-MIV assumption, the upper bounds on the effects cannot exclude the small effects of the point estimates in these levels of parents' education.

4.4.3 Single Treatment Effects on Child's Education under the SMIV assumption

The results under the MTR-MTS-MIV assumption provide informative upper bounds since they are substantially smaller than the point estimates obtained under the ETS assumption, but lower bounds do not exclude zero effect of parent's schooling on both years of schooling and probability of university graduation of the child. In this section, I use the two MIVs simultaneously under the SMIV assumption instead of using the schooling of the other parent and father's regular job variable separately.

Panel A of Table 4.5 provides the effects on child's years of schooling. It shows that SMIV assumption gives very informative bounds (columns (5)). The upper bounds are lower than the MTR-MTS-MIV bounds and indicate that both mother's and father's college degree increases child's schooling by about 1.41 years. The upper bounds results of a change of parent's schooling from junior high school to bachelor's degree show that the mother's university graduation increases child's schooling by about 2.09 years and the father's university graduation increases child's schooling by about 2.12 years. These upper bounds are substantially lower than the estimates under the ETS assumption.

Combining the two MIVs also plays an important role to tighten the lower bounds. For both mother's and father's schooling, the results show that the effect of an increase in parent's schooling to a college degree or more has an impact on child's schooling that is significantly different from zero. For example, a change of mother's schooling from junior high school to bachelor's degree provides an increase in child's schooling by about 0.39 years. The father's bachelor's degree increases child's schooling by about 0.43 years.

Panel B of Table 4.5 shows the effects on the probability of child's university graduation. The results using this outcome variable are very similar to the results using child's years of schooling as the outcome variable. The MTR-MTS-SMIV bounds give not only more informative upper bounds but also more informative lower bounds (column (10)). The results of increasing father's or mother's schooling to a college degree or more has a positive effect on child's university graduation that is statistically significantly different from zero, but substantially lower than the point estimates under the assumption of ETS.

To summarize the estimated results, the combination of the MTR-MTS-MIV assumptions provides informative bounds on the ATE of parent's schooling on child's schooling. The main results show that the tightest upper bounds on the effects are lower than the point estimates that rely on the ETS assumption. These results suggest that the regression estimates have an upward bias from the true causal effect of parent's schooling. The combination of the MTR-MTS-SMIV assumptions also gives informative lower bounds, which suggest positive effects of parental education on child's education.

4.4.4 Multiple Treatment Effects on Child's Years of Schooling and University Graduation

In this section, I provide the results of analyses on multiple treatments of both parents' schooling. The partial identification approach discussed in the method section gives bounds on the ATE of the multiple treatments under different semi-monotonicity assumptions as well as the single treatment case. Table 4.6 shows the multiple treatments effects of parents' schooling on child's schooling. The SMTR and the SMTS bounds are too wide to provide informative results, these are not shown in the table.[16]

Panel A of Table 4.6 provides the effects on child's years of schooling. I begin by examining the estimated ATE under the assumption of ETS for a useful benchmark. Children whose parents complete college or more education have about 2.27 longer years of schooling than those whose parents complete high school or less education (column (1)). Compared to the point estimates under the ETS assumption, the conservative no assumption bounds are [-4.3261 5.8333], which are too wide to acquire information on the ATE of parents' schooling (column (2)).

By combining different sets of assumptions, the results clearly illustrate more informative bounds on the ATE. The estimated bounds are [0, 2.2697] under the SMTR-SMTS assumption and [0, 2.1060] under the SMTR-SMTS-MIV assumption with an MIV of father's regular job or [0, 2.2388] with an MIV of the indicator of child was born after 1975 (columns (3), (4), and (5)). These upper bounds are smaller than the estimates under the ETS assumption. While both of these MIV assumptions substantially reduce the upper bounds, there still remains much uncertainty about the lower bounds on the ATE. Under the combined SMTR-SMTS assumption and SMIV assumption, the bounds narrow to [0.2397, 1.2077] and the confidence intervals are (0.1913, 1.3243), which exclude zero (column (6)). Thus, a set of semi-monotonicity assumptions provides the non-negative impact of parents' schooling on child's schooling.

Panel B of Table 4.6 shows the effects on the probability of child's university graduation.

[16]In this table, I only use the two levels of parents' schooling (college degree or more and high school or less) because the number of observations is too small to estimate with four levels of parents' schooling.

The results of this outcome variable are similar to the results of child's years of schooling. Under the ETS assumption, the point estimates indicate that children whose parents complete college or more education are more likely graduate university than those whose parents complete high school or less education (column (7)). The difference in the probability is 42%. In comparison with the point estimates under the ETS assumption, the conservative no assumption bounds are [-30.17, 82.71]%, which include negative effects of parents' schooling (column (8)).

The estimated bounds are [0, 42.00]% under the SMTR-SMTS assumption and [0, 38.39]% under the SMTR-SMTS-MIV assumption with an MIV of father's regular job or [0, 41.43]% with an MIV of the indicator of the child was born after 1975 (columns ((9), (10), and (11)). These upper bounds are smaller than the estimates under the ETS assumption. Under the SMTR-SMTS-SMIV assumption, the bounds narrow to [4.54, 19.84]% and the confidence intervals are (3.53, 22.69), which exclude zero (column (12)). It suggests that parents' college degrees are at least somewhat beneficial on child's university graduation. However, this estimated bound is strictly smaller than the point estimates under the assumption of the ETS, which is sometimes implicitly assumed in the literature.

To summarize the estimated results, the combination of the semi-monotonicity assumptions provides informative bounds on the ATE of parents' schooling on child's schooling in the multiple treatments framework. The main results show that the tightest upper bounds on the effects are substantially lower than the point estimates that assume the ETS. These results suggest that the regression estimates have an upward bias from the true causal effect of parents' education. The combination of the SMTR-SMTS-SMIV assumptions also gives informative lower bounds, which suggest non-zero positive effects of parents' education on child's education.

4.5 Conclusion

Quantifying the causal effects of parents' schooling on child's schooling is a goal of the recent empirical literature on intergenerational relation of education. Previous Japanese studies give large, positive and statistically significant estimates by the OLS method. These studies implicitly assume that parents' schooling is assigned randomly to the child. It does not seem to be appropriate if unobserved heterogeneity determines both parents' and child's educational levels.

The main contributions of this study are twofold. First, this study estimates the causal effects of parents' schooling on child's schooling in Japan using a nonparametric bounds analysis. It provides the ATE bounds without relying on an invalid exogenous selection assumption. Second, this study shows results of multiple treatments of both parents' schooling as an application of the nonparametric bounds method of multiple treatments evaluation. To obtain informative bounds, this study imposes relatively weak and partially testable monotonicity and semi-monotonicity assumptions on treatment response, selection, and instrumental variables.

The tightest bounds under the combination of these three assumptions show that obtained lower bounds on the ATE are positive and significantly different from zero, but upper bounds on the ATE are lower than the point estimates that rely on the assumption on exogenous selection of parents' schooling. These results suggest that the previous studies using OLS overestimate the true causal effect of parents' education.

References

Becker, Gary S. 1975. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research: distributed by Columbia University Press.

Becker, Gary S. and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87 (6):1153–1189.

Black, Sandra E. and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*, vol. 4, Part B, edited by David Card and Orley C. Ashenfelter, chap. 16. Amsterdam: Elsevier, 1487–1541.

Blundell, Richard, Amanda Gosling, Hidehiko Ichimura, and Costas Meghir. 2007. "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds." *Econometrica* 75 (2):323–363.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, vol. 3, edited by Orley. C. Ashenfelter and David. Card, chap. 30. Amsterdam: Elsevier, 1801–1863.

Chevalier, Arnaud, Colm Harmon, Vincent O'Sullivan, and Ian Walker. 2013. "The Impact of Parental Income and Education on the Schooling of Their Children." *IZA Journal of Labor Economics* 2 (1):8.

De Haan, Monique. 2011. "The Effect of Parents' Schooling on Child's Schooling: A Nonparametric Bounds Analysis." *Journal of Labor Economics* 29 (4):859–892.

Flores, Carlos A. and Oscar A. Mitnik. 2013. "Comparing Treatments across Labor Markets: An Assessment of Nonexperimental Multiple-Treatment Strategies." *Review of Economics and Statistics* 95 (5):1691–1707.

Frölich, Markus. 2004. "Programme Evaluation with Multiple Treatments." *Journal of Economic Surveys* 18 (2):181–224.

Giustinelli, Pamela. 2011. "Non-Parametric Bounds on Quantiles under Monotonicity Assumptions: With an Application to the Italian Education Returns." *Journal of Applied Econometrics* 26 (5):783–824.

Gonzalez, Libertad. 2005. "Nonparametric Bounds on the Returns to Language Skills." *Journal of Applied Econometrics* 20 (6):771–795.

Gundersen, Craig and Brent Kreider. 2009. "Bounding the Effects of Food Insecurity on Children's Health Outcomes." *Journal of Health Economics* 28 (5):971–983.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. "Parental Education and Parental Time with Children." *Journal of Economic Perspectives* 22 (3):23–46.

Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *Review of Economics and Statistics* 88 (3):389–432.

Heckman, James J. and Edward Vytlacil. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments." In *Handbook of Econometrics*, vol. 6, Part B, chap. 71. Amsterdam: Elsevier, 4875–5143.

Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2011. "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods." *Journal of Economic Literature* 49 (3):615–651.

Imbens, Guido W. 2000. "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika* 87 (3):706–710.

Imbens, Guido W. and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62:467–475.

Kariya, Takehiko. 2001. *Kaisoka Nihon to Kyoiku Kiki (Education in Crisis in Stratified Japan)*. Tokyo, Japan: Yushindo Kobusha. (in Japanese).

Kreider, Brent and Steven C. Hill. 2009. "Partially Identifying Treatment Effects with an Application to Covering the Uninsured." *Journal of Human Resources* 44 (2):409–449.

Kreider, Brent and John V. Pepper. 2007. "Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors." *Journal of the American Statistical Association* 102 (478):432–441.

Lechner, Michael. 2001. "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption." In *Econometric Evaluation of Labour Market Policies*, edited by Michael Lechner and Friedhelm Pfeiffer. Heidelberg: Physica-Verlag/Springer, 43–58.

Lee, Sokbae and Ralf A. Wilke. 2009. "Reform of Unemployment Compensation in Germany: A Nonparametric Bounds Analysis Using Register Data." *Journal of Business & Economic Statistics* 27 (2):193–205.

Manski, Charles F. 1989. "Anatomy of the Selection Problem." *Journal of Human Resources* 24 (3).

———. 1990. "Nonparametric Bounds on Treatment Effects." *American Economic Review* 80 (2):319–23.

———. 1997. "Monotone Treatment Response." *Econometrica* 65 (6):1311–1334.

———. 2003. *Partial Identification of Probability Distributions*. New York: Springer.

Manski, Charles F. and John V. Pepper. 1998. "Monotone Instrumental Variables: With an Application to the Returns to Schooling." Working Paper 224, National Bureau of Economic Research, Cambridge, MA.

———. 2000. "Monotone Instrumental Variables: With an Application to the Returns to Schooling." *Econometrica* 68 (4):997–1010.

———. 2009. "More on Monotone Instrumental Variables." *The Econometrics Journal* 12 (s1):S200–S216.

Nakamuro, Makiko and Tomohiko Inui. 2012. "Estimating the Returns to Education Using the Sample of Twins: The Case of Japan." Discussion Paper 12-E-076, RIETI Discussion Paper Series.

Okumra, Tsunao and Emiko Usui. 2010. "Concave-Monotone Treatment Response and Monotone Treatment Selection: With an Application to the Returns to Schooling." IZA Discussion Paper 4986, The Institute for the Study of Labor (IZA).

Oreopoulos, Philip, Marianne E. Page, and Ann H. Stevens. 2006. "The Intergenerational Effects of Compulsory Schooling." *Journal of Labor Economics* 24 (4):729–760.

Oreopoulos, Philip and Kjell G. Salvanes. 2011. "Priceless: The Nonpecuniary Benefits of Schooling." *Journal of Economic Perspectives* 25 (1):159–184.

Solon, Gary. 1999. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics*, vol. 3, Part A, edited by Orley C. Ashenfelter and David Card, chap. 29. Amsterdam: Elsevier, 1761 – 1800.

Tachibanaki, Toshiaki. 1988. "Education, Occupation, Hierarchy and Earnings." *Economics of Education Review* 7 (2):221–229.

Tanaka, Ryuichi. 2008. "The Gender-Asymmetric Effect of Working Mothers on Children's Education: Evidence from Japan." *Journal of the Japanese and International Economies* 22 (4):586–604.

Yamada, Ken. 2011. "Family Background and Economic Outcomes in Japan." Working Paper 1246, Singapore Management University.

Table 4.1: Summary Statistics

	mean	sd	min	max
Chlid's years of schooling	13.0893	2.1713	9	18
Chlid's university graduation	0.2452	0.4302	0	1
Mother's years of schooling	9.7783	2.5953	6	18
Mother's some college	0.0750	0.2634	0	1
Mother's university graduation	0.0258	0.1584	0	1
Father's years of schooling	10.2526	3.1694	6	18
Father's some college	0.0681	0.2519	0	1
Father's university graduation	0.1231	0.3286	0	1
Father had a regular job	0.5872	0.4924	0	1
Child was born after 1975	0.0941	0.2920	0	1
Female	0.5406	0.4984	0	1

Notes: This table reports summary statistics of the analysis data. Number of observations is 13,669. Father had a regular job indicates that the father was an executive of a company or was a regular employee (*joji-koyo no rodosha*) when the child was 15 years old.

Table 4.2: Mean of Child's Schooling Outcomes by Parent's Education Level

	Mother		Father	
Education level of parent	Chlid's years of schooling	Child's university	Chlid's years of schooling	Child's university
Junior high school	12.1753	0.1184	12.1779	0.1179
High school	13.6801	0.3095	13.4729	0.2691
Some college	14.6966	0.5220	14.2095	0.4135
University	15.1222	0.6648	14.8247	0.5639

Notes: This table reports mean schooling outcomes of children by educational level of parents for the test of the MTR-MTS assumption.

Table 4.3: Bounds of ATE of Parent's Schooling on Child's Years of Schooling

Assumption	ETS		No Assumption		MTR		MTS		MTR MTS	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
	Effect of Mother's Schooling									
	(1)		(2)		(3)		(4)		(5)	
E[y(t=HS)] - E[y(t=JH)]	1.5049		-4.2558	5.6509	0	5.3291	-3.7843	1.9400	0	1.6182
	(1.4365	1.5743)	(-4.2673	5.6327)	(0	5.3626)	(-3.8194	2.0043)	(0	1.6873)
E[y(t=Coll)) - E(y(t=HS)]	1.0165		-6.7095	6.7336	0	5.1737	-4.5414	3.2915	0	1.7316
	(0.8969	1.1355)	(-6.7224	6.7343)	0	5.2075)	(-4.5855	3.4028)	(0	1.8479)
E[y(t=Univ)) - E(y(t=Coll)]	0.4256		-8.5946	8.4987	0	4.9942	-5.6240	5.5483	0	2.0438
	(0.1902	0.6548)	(-8.6021	8.4932)	(0	5.0297)	(-5.7323	5.7433)	(0	2.2469)
E[y(t=Univ)) - E(y(t=JH)]	2.9469		-6.1168	7.4401	0	7.4401	-6.1168	2.9469	0	2.9469
	(2.7452	3.1443)	(-6.1874	7.4581)	(0	7.4631)	(-6.1402	3.1480)	(0	3.1562)
	Effect of Father's Schooling									
	(6)		(7)		(8)		(9)		(10)	
E[y(t=HS)] - E[y(t=JH)]	1.2950		-4.7572	5.9639	0	5.3148	-3.9018	2.1607	0	1.5116
	(1.2222	1.3728)	(-4.7708	5.9647)	(0	5.3485)	(-3.9380	2.2265)	(0	1.5806)
E[y(t=Coll)) - E(y(t=HS)]	0.7366		-7.0819	7.1973	0	5.3335	-4.3424	3.2763	0	1.4125
	(0.6015	0.8681)	(-7.0943	7.1987)	(0	5.3671)	(-4.3933	3.3848)	(0	1.5240)
E[y(t=Univ)) - E(y(t=Coll)]	0.6153		-8.0247	8.2542	0	5.2369	-4.9590	4.8285	0	1.8111
	(0.4633	0.7619)	(-8.0270	8.2639)	(0	5.2702)	(-5.0624	4.9196)	(0	1.9110)
E[y(t=Univ)) - E(y(t=JH)]	2.6468		-5.5845	7.1362	0	7.1362	-5.5845	2.6468	0	2.6468
	(2.5472	2.7423)	(-5.5978	7.1416)	(0	7.1612)	(-5.6104	2.7476)	(0	2.7480)

Notes: This table reports average treatment effects of parent's education on child's years of schooling. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are (1) JH = less than high school, (2) HS = high school, (3) Coll = some college, (4) Univ = bachelor's degree or more. Number of observations is 13,669.

Table 4.3: Bounds of ATE of Parent's Schooling on Child's Years of Schooling (cont.)

Assumption	MIV		MIV		MTR MTS MIV		MTR MTS MIV	
MIV	Other parent's univ / college grad		Father's regular job		Other parent's univ / college grad		Father's regular job	
	LB	UB	LB	UB	LB	UB	LB	UB
	Effect of Mother's Schooling							
	(11)		(12)		(13)		(14)	
E[y(t=HS)] - E[y(t=JH)]	-4.2558	5.4734	-4.2558	5.2105	0	1.4162	0	1.4958
	(-4.2877	5.5076)	(-4.2862	5.2594)	(0	1.5011)	(0	1.5648)
E[y(t=Coll)) - E(y(t=HS)]	-6.7095	6.1976	-6.5188	6.6702	0	1.4125	0	1.6587
	(-6.7327	6.2405)	(-6.5492	6.6945)	(0	1.5514)	(0	1.7822)
E[y(t=Univ)) - E(y(t=Coll)]	-8.0586	8.0606	-8.5312	8.4728	0	1.4178	0	1.8613
	(-8.0974	8.0996)	(-8.5442	8.4839)	(0	1.9128)	(0	2.1019)
E[y(t=Univ)) - E(y(t=JH)]	-6.1168	6.8245	-6.1168	7.1645	0	2.1948	0	2.6893
	(-6.1405	6.8711)	(-6.1405	7.2055)	(0	2.7014)	(0	2.9373)
	Effect of Father's Schooling							
	(15)		(16)		(17)		(18)	
E[y(t=HS)] - E[y(t=JH)]	-4.6679	5.8243	-4.7572	5.4669	0	1.3649	0	1.4038
	(-4.6989	5.8570)	(-4.7869	5.5166)	(0	1.4448)	(0	1.4715)
E[y(t=Coll)) - E(y(t=HS)]	-7.0819	6.4940	-6.9310	7.1391	0	1.2180	0	1.2909
	(-7.1026	6.5455)	(-6.9591	7.1615)	(0	1.3513)	(0	1.4148)
E[y(t=Univ)) - E(y(t=Coll)]	-7.4106	7.0613	-7.9665	8.1252	0	1.5775	0	1.6869
	(-7.4599	7.1286)	-7.9825	8.1426)	(0	1.6966)	(0	1.8045)
E[y(t=Univ)) - E(y(t=JH)]	-5.5845	5.8037	-5.5845	6.6612	0	2.2836	0	2.4542
	(-5.6098	5.8780)	(-5.6097	6.7041)	(0	2.4119)	(0	2.5719)

Notes: This table reports average treatment effects of parent's education on child's years of schooling. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are (1) JH = less than high school, (2) HS = high school, (3) Coll = some college, (4) Univ = bachelor's degree or more. MIVs are the indicator of other parent's university / college graduation for mother's / father's education and father's regular job status. Number of observations is 13,669.

Table 4.4: Bounds of ATE of Parent's Schooling on Child's University Graduation

Assumption	ETS		No Assumption		MTR MTS		MTR MTS MIV		MTR MTS MIV	
MIV	No		No		No		Other parent's univ / college grad		Father's regular job	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
	Effect of Mother's Schooling									
	(1)		(2)		(3)		(4)		(5)	
E[y(t=HS)] - E[y(t=JH)]	0.1912		-0.4539	0.6468	0	0.2163	0	0.1815	0	0.2036
	(0.1772	0.2054)	(-0.4669	0.6471)	(0	0.2302)	(0	0.1987)	(0	0.2178)
E[y(t=Coll)) - E(y(t=HS)]	0.2124		-0.6630	0.8306	0	0.3056	0	0.2326	0	0.2881
	(0.1801	0.2446)	(-0.6652	0.8323)	(0	0.3373)	(0	0.2687)	(0	0.3200)
E[y(t=Univ)) - E(y(t=Coll)]	0.1428		-0.9470	0.9522	0	0.4233	0	0.2795	0	0.3896
	(0.0829	0.2024)	-0.9487	0.9536)	(0	0.4736)	(0	0.4072)	(0	0.4491)
E[y(t=Univ)) - E(y(t=JH)]	0.5464		-0.5703	0.9360	0	0.5464	0	0.3838	0	0.5079
	(0.4977	0.5967)	(-0.5764	0.9365)	(0	0.5977)	(0	0.5128)	(0	0.5678)
	Effect of Father's Schooling									
	(6)		(7)		(8)		(9)		(10)	
E[y(t=HS)] - E[y(t=JH)]	0.1512		-0.4983	0.6930	0	0.1973	0	0.1785	0	0.1889
	(0.1371	0.1663)	(-0.5038	0.6987)	(0	0.2103)	(0	0.1948)	(0	0.2028)
E[y(t=Coll)) - E(y(t=HS)]	0.1445		-0.7194	0.8671	0	0.2330	0	0.1917	0	0.2036
	(0.1098	0.1779)	(-0.7242	0.8721)	(0	0.2611)	(0	0.2268)	(0	0.2346)
E[y(t=Univ)) - E(y(t=Coll)]	0.1503		-0.8906	0.9181	0	0.3372	0	0.2794	0	0.3153
	(0.1100	0.1904)	(-0.8943	0.9218)	(0	0.3633)	(0	0.3096)	(0	0.3436)
E[y(t=Univ)) - E(y(t=JH)]	0.4460		-0.5218	0.8917	0	0.4460	0	0.3728	0	0.4218
	(0.4209	0.4702)	(-0.5263	0.8965)	(0	0.4720)	(0	0.4056)	(0	0.4498)

Notes: This table reports average treatment effects of parent's education on child's university graduation. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are (1) JH = less than high school, (2) HS = high school, (3) Coll = some college, (4) Univ = bachelor's degree or more. MIVs are the indicator of other parent's university / college graduation for mother's / father's education and father's regular job status. Number of observations is 13,669.

Table 4.5: Bounds of ATE of Parent's Schooling on Child's Years of Schooling and University Graduation

Assumption	ETS		No Assumption		MTR MTS MIV		MTR MTS MIV		MTR MTS SMIV	
MIV	No		No		Other parent's univ / college grad		Father's regular job		Other parent's univ/college and Father's regular job	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
	Panel A: Effect on Chlid's Years of Schooling									
Mother's Schooling	(1)		(2)		(3)		(4)		(5)	
E[y(t=Coll+)) - E(y(t=HS-)]	1.9083		-3.8263	5.1737	0	1.4588	0	1.7959	0.0475	1.4133
	(1.8040	2.0119)	(-3.8609	5.2100)	(0	1.5945)	(0	1.9147)	(0.0028	1.5528)
E[y(t=Univ)) - E(y(t=JH)]	2.9469		-6.1168	7.4401	0	2.1948	0	2.6893	0.3890	2.0886
	(2.7452	3.1443)	(-6.1874	7.4581)	(0	2.7014)	(0	2.9373)	(0.3175	2.4566)
Father's Schooling										
E[y(t=Coll+)) - E(y(t=HS-)]	1.8748		-3.6665	5.3335	0	1.5601	0	1.7141	0.1308	1.4145
	(1.7930	1.9562)	-(3.7019	5.3685)	0	1.6600)	(0	1.8095)	(0.0867	1.5182)
E[y(t=Univ)) - E(y(t=JH)]	2.6468		-5.5845	7.1362	0	2.2836	0	2.4542	0.4321	2.1167
	(2.5472	2.7423)	(-5.5978	7.1416)	(0	2.4119)	(0	2.5719)	(0.4221	2.2564)

	Panel B: Effect on Probability of Child University Graduation									
Mother's Schooling	(6)		(7)		(8)		(9)		(10)	
E[y(t=Coll+)) - E(y(t=HS-)]	0.3484		-0.2334	0.7666	0	0.2429	0	0.3245	0.0092	0.2317
	(0.3214	0.3755)	(-0.2402	0.7738)	(0	0.2797)	(0	0.3533)	(0.0007	0.2689)
E[y(t=Univ)) - E(y(t=JH)]	0.5464		-0.5703	0.9360	0	0.3838	0	0.5079	0.0567	0.3773
	(0.4977	0.5967)	(-0.5764	0.9365)	(0	0.5128)	(0	0.5678)	(0.0413	0.4734)
Father's Schooling										
E[y(t=Coll+)) - E(y(t=HS-)]	0.3279		-0.2412	0.7588	0	0.2618	0	0.2981	0.0246	0.2338
	(0.3072	0.3486)	(-0.2478	0.7656)	(0	0.2863)	(0	0.3200)	(0.0153	0.2596)
E[y(t=Univ)) - E(y(t=JH)]	0.4460		-0.5218	0.8917	0	0.3728	0	0.4218	0.0665	0.3499
	(0.4209	0.4702)	(-0.5263	0.8965)	(0	0.4056)	(0	0.4498)	(0.0557	0.3849)

Notes: This table reports average treatment effects of parent's education on child's years of schooling and child's university graduation. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are JH = less than high school, Univ = bachelor's degree or more, HS− = high school or less, Coll+ = college or more. MIVs are the indicator of father's regular job status and the indicator that denotes the child was born after 1975. Number of observations is 13,669.

Table 4.6: Bounds of ATE of Parents' Schooling on Child's Years of Schooling and University Graduation

Assumption	ETS		No Assumption		SMTR SMTS		SMTR SMTS MIV		SMTR SMTS MIV		SMTR SMTS SMIV	
MIV	No		No		No		Father's regular job		Born After 1975		Father's regular job and Born After 1975	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
	Panel A: Effect on Chlid's Years of Schooling											
	(1)		(2)		(3)		(4)		(5)		(6)	
E[y[(t^M,t^F)=(Coll+, Coll+)]	2.2697		-4.3261	5.8333	0	2.2697	0	2.1060	0	2.2388	0.2397	1.2077
- E[y[(t^M,t^F)=(HS-, HS-)]]	(2.1610	2.3856)	(-4.3591	5.8669)	(0	2.3864)	(0	2.2344)	(0	2.3499)	(0.1913	1.3243)
	Panel B: Effect on Probability of Child University Graduation											
	(7)		(8)		(9)		(10)		(11)		(12)	
E[y[(t^M,t^F)=(Coll+, Coll+)]	0.4200		-0.3017	0.8271	0	0.4200	0	0.3839	0	0.4143	0.0454	0.1984
- E[y[(t^M,t^F)=(HS-, HS-)]]	(0.3893	0.4491)	(-0.3080	0.8332)	(0	0.4510)	(0	0.4177)	(0	0.4443)	(0.0353	0.2269)

Notes: This table reports average treatment effects of a set of vector of mother's and father's education on child's years of schooling and child's university graduation. t^M and t^F denotes the treatment level of mother's and father's education respectively. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are Coll+ = college or more. MIVs are the indicator of father's regular job status and the indicator that denotes the child was born after 1975. Number of observations is 13,669.