# Actual Age at School Entry, Educational Outcomes, and Earnings

Daiji Kawaguchi[1]

Faculty of Economics, Hitotsubashi University

December 23, 2009

[1]Contact: Faculty of Economics, Hitotsubashi University, Naka 2-1, Kunitachi, Tokyo 186-8601, Japan, Tel: +81-42-580-8851, Fax: +81-42-580-8851, E-Mail: kawaguch@econ.hit-u.ac.jp

**Abstract**

This paper reports on the effect of actual age measured by month at school entry on test scores, eventual educational attainment, and labor market outcomes, using school test-score data and a labor- force survey of Japan. Japan is an ideal country for examining the pure effect of actual age at school entry on eventual years of education because the length of compulsory education does not vary by birth month and legal administrations assure that almost all children follow a fixed schedule of grade progress. Older children of both sexes in a school cohort obtain higher test scores and more education years than their younger counterparts. This better academic performance translates into higher annual earnings among males.
Key Words: Birth Month, Relative Age Effect, Child Development, Education, Income, Japan

JEL Classification Code: J21, J13

# 1 Introduction

Children start their schooling at different actual ages because schools accept entering students only once a year. For example, primary schools in Japan accept entering students on April 1, and children who turn age 6 on that day or before enter primary schools for that year.[1] Thus those who are born on April 1 or slightly earlier enter primary schools at an early age 6, while those who are born on April 2 or slightly after enter primary schools at a late age 6. Those who are born in April have an advantage compared with those born in March in terms of physical and mental development, and this may have favorable consequences for those born in April. Elder students at school entry perform better because of their higher absolute age or relative age than their peers.

An almost one-year age difference at age 6 could have a large impact on students' academic and physical performance, but this initial gap disappears as children age if the absolute age is an important determinant for initial performance because the fraction of age difference to absolute age disappears as children grow. In contrast, if the initial difference in performance has a causal impact on subsequent performance through feedback effects, such as stigmatization, then the birth month could have a lifetime impact on eventual educational attainment and labor market outcomes. This latter effect is called the relative age effect in the literature, and detecting it is important

[1] According to Japanese law, people officially age a day before their birthday.

because if there is such an effect, the educational system could be adjusted so that those who are relatively younger in the same school cohort could receive extra attention.

The general public recognizes the existence of a relative age effect, and, in fact, the entrance examinations of some selective primary schools (Keio and Tsukuba, for example) are given to groups of students divided by the candidates' birth months to treat younger and elder children equally. Previous studies indicate that relatively older students in the same cohort achieve more in school (Thompson (1971), Allen and Barnsley (1993), Borg and Falzon (1995), and Lien et al. (2005)), are less likely to be diagnosed as having specific learning disabilities (Martin et al. (2004)), are less likely to commit suicide (Thompson et al. (1999)), and are more likely to be class leaders in high school (Dhuey and Lipscomb (2008)). Relative age effects also are found in the field of sports (Dudink (1994) and Helsen et al. (2000)).

Economists have begun studying the relative age effect on test scores, eventual educational attainment, and labor market outcomes. Bedard and Dhuey (2006) examined the relative age effect on test scores among 4th and 8th graders in various countries and consistently found significant relative age effects. They also found that those who were born early in a school cohort are more likely to attend a 4-year college, based on data from British Columbia in Canada and the US. Using various US data, Datar (2005), Elder and Lubotsky (2008), Dobkin and Ferreira (2007) and Cascio and Schanzenbach (2007) also confirmed a significant, relative age effect on various outcome measures, such as test scores, educational attainment, and adult outcomes. Fredriksson and Ockert (2005) and McEwan and Shapiro (2008) obtained similar results for Sweden and Chile, respectively.

As many parents with relatively younger children in a cohort withhold their children because they recognize the disadvantage of being younger in a school cohort, previous studies have been limited to estimating a local average treatment effect. Dobkin and Ferreira (2007) found that parents with high socioeconomic status are more likely to postpone their children's school attendance if their children are younger in a cohort. This finding implies that parents with low socioeconomic status are more likely to be "compliers," and the estimated local average treatment effect was mainly estimated for children with low-SES parents, which could be significantly different from the population's average treatment effect.

Based on students' test score data and a large-scale labor force survey, this study reports the effect of actual age at school entry on educational attainment and labor market outcomes for both sexes in Japan. The research design based on Japanese data is advantageous for estimating the pure effect of actual age at school entry because the Japanese educational system does not induce a variation of educational attainment by the birth month as in the US (Angrist and Krueger (1991)). In addition, both delaying school attendance and repeating grades are quite rare in the Japanese educational system, and this makes it possible to estimate the population's average treatment effect of birth month on education and labor market outcomes.

An examination of test scores confirms the existing findings by Bedard and Dhuey (2006) for Japan. April-June born students score 0.11 to 0.22 standard deviation higher than January-March born students, on average, in the 4th and 8th grades. This study first finds that this difference in students' academic performance in primary and secondary education does not wash out and has life-long effects on individuals' eventual years of education. The analysis results based on the labor-force survey indicate that April-June 1968 born males in the sample had an average educational attainment of 13.20 years, whereas January-March born males had 13.03 years. This difference in educational attainment translates into an income difference between the two groups of males. Similarly, April-June 1968 born females had 12.96 years of education, whereas January-March born females had 12.88 years. Contrary to the results for males, April-June born females do not earn more than March-born females, perhaps because of a complex pattern of labor-force participation.

The rest of this paper is organized as follows. Section 2 briefly introduces the Japanese educational institutions related to this analysis. Section 3 introduces the data sets used in the study. Section 4 lays out the estimation results. Section 5 further discusses the robustness of the analysis results. The last section provides conclusions.

# 2 Primer of the Japanese School System

The Japanese school system is similar to that of the US. Compulsory education consists of 6 years of primary school and 3 years of junior high school. After junior high school, students can choose to attend 3 years of high school or 5 years of technical college (*Kosen*). After high school, the choices are 2 years of junior college or 4 years of college. Then students may choose to study in graduate school for advanced degrees. In public schools, tracking based on academic performance starts from high school. For an illustration, see Figure 1.

The school system is legally defined in the School Education Law (SEL) enacted in 1947. SEL article 22 requires parents to send their children to primary schools once their children will turn age six before the school starting day, which is April 1. According to Japanese law, people become their new age a day before their birthday; thus, children born on April 1 enter primary schools on their 6th birthday, while those born on April 2 enter primary schools on the day before their 7th birthday. So there is about a one-year maximum chronological age difference among students in the first grade.

SEL Article 23 allows a delay in school entry because of a child's illness or underdevelopment, but this exception is rarely applied. In 2004, 7,200,933 children at the primary school age (ages 6 - 12) attended primary schools, while 2,261 did not; thus the percentage of exemption is 0.03 percent. The law does not clearly prohibit students from learning in the grade above the scheduled grade, but in court cases, judges have ruled against allowing students to do so, at least in primary and junior high school. The fact that almost all children start attending school without delay or advancement contrasts with the situation in the US, where postponing school attendance has become popular among educated parents (Elder and Lubotsky (2008)).

SEL article 39 requires parents to send their children to junior high school by the end of the school year in which they turn age 15. Thus, 9 years of education is uniformly required for all children regardless of their birth month. This requirement for the school-leaving date in Japan contrasts with that of most states in the US, where the school-leaving dates often are defined by age. Those who turn 16 or 17 are allowed to leave school in many states, and those who enter primary school at an older age (measured in months) tend to have fewer years of education because they are allowed to leave school earlier. This institutional setting is exploited by Angrist and Krueger (1991) to estimate the return to education under the assumption that the quarter of birth affects earnings only through educational attainment. The Japanese institutional setting does not create a variation in the number of years of schooling by birth month. Therefore, if there is a variation in the years of education depending on the birth month in Japan, the variation is induced by individual choice.

Pre-primary school education is quite popular in Japan. As of 2004, about 90 percent of children ages 3 to 5 attend either a day-care center (about 39 percent) or kindergarten (about 50 percent).[2] Officially, day-care facilities are non-educational institutions, and there is neither clear curriculum-based teaching nor class formation based on age cohort. In contrast, kindergartens form classes based on children's ages. Those who are ages 3, 4, and 5 by April 1 are sorted into classes corresponding to their ages. Thus, among kindergarten attendees, the relative age effect could be in motion.

# 3 Data

This study uses two data sets. The Trends in International Mathematics and Science Study (TIMSS) 2003 is used to examine the effect of birth month on academic achievement among students in the 4th and 8th grades. The Employment Status Survey (ESS) 2002 (*Shugyo Kozo Kihon Chosa*) is used to examine the effect of birth month on individuals' eventual years of education and labor market outcomes.

The TIMSS 2003 is a survey conducted by the International Association for the Evaluation of Educational Achievement that aims toward an international comparison of students' achievement. The survey takes place every 4 years and the 2003 survey targeted 4th and 8th graders. The sampling

[2] According to School Basic Statistics (*Gakkou Kihon Tokei*) by the Ministry of Education, 1,753,393 pupils attended kindergarten classes for 3, 4, and 5 year olds in 2004. The Welfare Administration Record (*Fukushi Gyosei Houkoku Rei*) by the Ministry of Welfare and Labor reported that 1,348,754 children between the ages of 3 and 5 attended day-care centers in 2004, and the population of children between the ages of 3 and 5 was 3,504,000 in 2004, according to the Annual Report of Population Estimates by the Ministry of Internal Affairs and Communications. Based on these figures, 50 percent of children between the ages of 3 and 5 attended kindergarten, and 39 percent of them attended day-care centers.

method applied for Japan is two-stage stratified random sampling. In the first stage, schools are randomly chosen with a sampling probability proportional to the number of students of each school from strata defined by region types. In the second stage, a whole class is randomly chosen from the 4th and 8th grades.[3] Japan's overall response rate was 97.4% for 4th graders and 95.9% for 8th graders ((Martin, Mullis, and Chrostowski, Martin et al., Exhibit 9.9 and Exhibit 9.10)).

Internationally standardized examinations are given in the subjects of mathematics and science. Test scores for each subject are standardized with a mean of 150 and a standard deviation of 10. About 5,000 4th graders and about 10,000 8th graders took the examination. Students who are in the 4th or 8th grade without delay from the standard schedule should be born after April 1992 or April 1988, respectively. I only kept students without delay in the sample, which resulted in 2,599 boys and 2,586 girls for the 4th graders, and 4,808 boys and 4,734 girls for the 8th graders. Reflecting the fact that grade repetition is rare in Japan, only 1.6 percent of 4th graders and 0.27 percent of 8th graders in the samples were dropped because of this sample restriction.[4]

[3]The survey structure in Japan is given on the web page of the Ministry of Education, Culture, Sports, Science and Technology - Japan. `http://www.mext.go.jp/b_menu/houdou/16/12/04121301/001.htm`

[4]Whether a student is behind the age-grade schedule or not is regressed on the birth-month dummy variables along with other explanatory variables, and the results are reported in Appendix Table 1. Only for 4th-grade girls, the January-March born are more likely to be behind than those who are born in other months. If low-ability students are more likely to be behind, estimates obtained from the sample that excludes those who are behind underestimate the effect of birth month on test scores.

The Employment Status Survey (ESS) was conducted on household members ages 15 and older in approximately 440,000 households dwelling in sampled units that cover the complete population.[5] The survey collects information as of October 1, 2002, including the labor-force status of each household member. This study utilizes micro data and extracts information on birth year, birth month, educational attainment, employment status, and annual income from the main job over the previous year.

The file contains 968,628 individuals, with 459,939 males and 508,689 females. The analysis sample is restricted to those aged 30-59 and out of school to restrict our attention to the completed years of education. Also, the inclusion of those age 59 or below in 2002 assures that people were born in 1943 or later and reached age 6 in 1949 at the earliest. Because the US occupied the Okinawa prefecture until 1973, and this prefecture was subject to a different school system from mainland Japan, the observations from Okinawa were dropped. This assures that virtually all individuals in the sample were subject to the current educational institutions adopted after World War II.

The sample is further restricted to observations with a valid birth year, birth month, educational background, and employment status. The analysis sample includes 219,207 males and 226,234 females. The survey asks respondents to indicate their educational attainment by selecting one of four

[5] Foreign diplomats, foreign military personnel and their dependents, persons dwelling in camps or ships of the Self Defense Force, and persons serving sentences in correctional institutions are excluded.

categories: 1. primary school/junior high school, 2. high school, 3. junior college/technical college(*Kosen*), and 4. college/graduate school. Nine years of education are assigned for primary school graduates, 12 years for high school graduates, 14 years for junior college or technical college graduates, and 16 years for 4-year college graduates or graduate school graduates. The survey records the previous year's annual labor earnings in ranges. The annual income ranges denominated by thousand yen are: 500 or less, 500-990, 1,000-1,490, 1,500-1,990, 2,000-2,490, 2,500-2,990, 3,000-3,990, 4,000-4,990, 5,000-5,990, 6,000-6,990, 7,000-7,990, 8,000-8,990, 9,000-9,900, 10,000-14,900, and 15,000 or above. These ranges are transformed into a continuous variable by using the center value for each range and 150,000 thousand yen for the highest range.

# 4 Estimation Results

## 4.1 Results from the Test Score Data

This subsection aims at updating the findings by Bedard and Dhuey (2006) based on the TIMSS 1995 and 1999 for Japan by using TIMSS 2003. I divide children into four groups by birth month: April-June born, July-September born, October-December born, and January-March born. The average test scores of these four groups are compared to examine the relative age effect. The April-June born are the eldest in a school cohort, and the January-March born are the youngest. The children could be divided into finer groups by birth month, but this division is chosen for an efficiency consideration, and the results reported below do not essentially change by the choice of birth-month grouping.

Figure 2 illustrates the distribution of birth months in the TIMSS sample for 4th and 8th graders. The birth months are almost uniformly distributed, and there is no clear sign of parents' manipulation of birth month around March and April.

Descriptive statistics of the TIMSS 2003 are reported in Table 1. Relatively younger students in a cohort tend to score lower than elder students. Also, younger children in a cohort are less likely to have mothers and fathers who graduate from 4-year college or above. This finding is consistent with the finding by Kureishi and Wakabayashi (2006) that working women with a higher educational background tend to deliver babies in April because they can send their children to nursery before their one-year maternity leave ends. Thus, higher test scores by older children could partly come from having a better family background. To alleviate the problem, several variables that indicate family background are controlled for in the estimation of the effect of birth month on educational attainment. Because of the low response rate for parental educational background, these controls are not perfect, however, and we should exercise caution that higher test scores among elder students are partly a result of their better family background compared with younger students.

Table 2 examines the distributional difference in background variables by birth month. Column 2 reports that relatively younger children tend to have fathers with fewer years of education. There is no other systematic difference in the distribution of other background variables by child birth month.

Table 3 reports the regression of math test scores on the dummy variables for birth month, along with other covariates. This specification allows for the non-linear effect of relative age on test scores. Results in Column 1 of Table 3 indicate that January-March born children score about 0.19 standard deviation below April-June born children among 4th-grade boys. The difference is 0.22 standard deviation for 4th-grade girls, as reported in Column 2. Columns 3 and 4 of Table 3 report the results for 8th graders. The difference between the April-June born and the January-March born in the test scores is attenuated to 0.11 standard deviation for 8th-grade boys. Similarly among 8th-grade girls, the difference is attenuated to 0.16 standard deviation. Appendix Table 2 reports the results for science test scores and exhibits qualitatively similar results.

Overall, there is clear evidence that those who enter primary school at an older age perform better than their younger counterparts based on the TIMSS 2003. The size of coefficients for 4th graders in this study is comparable with the findings for Japan by Bedard and Dhuey (2006), who report 0.24 for 4th graders based on the TIMSS 1995 and 1999. But, the coefficient for 8th graders is smaller than the 0.23 that they found.

## 4.2 Results from the Labor Force Survey

The number of observations in the Employment Status Survey 2002 by birth year - birth month appear in Figure 3. This figure indicates that observations are more likely to be taken from winter months (i.e., between January and March) than summer months (i.e., between June and August). This seasonal pattern of births is observed among older generations because farmers tend to deliver their babies during the agricultural off-season. These figures imply that birth month might carry information about parental occupation and other socioeconomic background characteristics. In particular, those born in the winter might be more likely to have a farm background. It is notable that this seasonal pattern almost disappeared after 1968 as the share of agriculture became smaller in the Japanese economy. Because limited availability of variables in the ESS prohibits us from controlling for respondents' socioeconomic background, the analysis sample is restricted to those who were born between April 1968 and March 1972 to mitigate the contamination of analysis results from unobserved heterogeneity in socioeconomic background. The cohorts born after April 1972 are not included to avoid those who are below age 30.

Restricting our analysis sample to school cohorts that do not exhibit a clear seasonal pattern of birth month reduces the concern of unobserved heterogeneity in socioeconomic background across respondents depending on birth month. To further examine the difference of the socioeconomic background by birth month, father's years of education is regressed on the birth-month dummy variables. Father's years of education presumably captures the respondents' socioeconomic background. Thus, if birth month is correlated with socioeconomic background, then fathers' years of education will be different across groups of respondents by birth month. Fathers' years of education, however, is available only for respondents who live with their fathers. This sample restriction may cause a sample selection bias for the regression coefficients of fathers' years of education on the birth-month dummy variables.

To examine the effect of birth month on the respondent's decision to live with his/her father, the dummy variable that takes one if the respondent lives with his/her father and the father's years of education is reported, is regressed on the birth-month dummy variables. The regression results for males and females are reported in Table 4, Panel A, Column 1 and Table 4, Panel B, Column 1, respectively. These results indicate that those who were born between January and March are more likely to live with their fathers and report fathers' years of education than those who were born between April and June. For males, living with fathers between age 30 and 34 could be interpreted as a negative outcome because those who do not earn a sufficient amount to live independently are more likely to live with their fathers. In contrast, more educated women are less likely to leave their parents between age 30 and 34 because of marriage-related reasons, and thus living with fathers is considered to be a positive outcome. Indeed, more-educated male respondents are less like to live with their fathers and to report fathers' years of education in the analysis sample. I find the opposite for the female sample.

Analysis results in Table 4, Panels A and B, Columns (1) caution us a possible sample selection bias in the regression of father's years of education on the birth-month dummy variables. There could be two typical cases for the sample selection, either the respondents with higher or lower fathers' years of education are more likely to drop from the sample. Under both scenarios, the sample selection will attenuate the regression coefficients. Thus, if we assume that the observations with longer or shorter years of father's education are more likely to drop from the sample, the estimated coefficients for the regression of father's years of education on the birth-month dummy variables are the lower bound in absolute value.

Further caveat on using father's years of education as a proxy variable for respondents' unobserved ability should be emphasized because it is an imperfect proxy variable. Results in Table 4, Panels A and B, Columns (1) arguably suggest that restricting the sample to those who live with their fathers at ages 30-34 creates the samples with low unobserved ability for males and high unobserved ability for females. If the effect of birth month on school and labor market outcomes is heterogenous across unobserved ability, the effects estimated from these restricted samples could arguably be the average effects among low ability male and high ability female.

Conditional on the presence of the father in the same household, the father's years of education is regressed on the respondent's birth-month dummy variables. The estimation result for males, reported in Table 4, Panel A, Column 2, implies that those who were born between January and March have fathers with more years of education than those who were born between April and June. This result suggests that those who were born later in a school cohort tend to have a better socioeconomic background than those who were born later. This finding contradicts the previous findings based on the TIMSS, Kureishi and Wakabayashi (2006) or Kureishi and Wakabayashi (2008) based on the sample that includes younger cohorts. Strategic birth timing may well be a recent phenomena reflecting the fact that more educated mothers deliver babies between April and June for child care reasons, as originally pointed out by Kureishi and Wakabayashi (2006). The stronger socioeconomic background of the January-March born among the 1968 - 1972 cohorts creates a bias such that those who were younger in a school cohort seemingly do better in educational attainment or the labor market. If the April-June born have greater educational attainment or do better in the labor market than the January-March born, regardless of having a weaker socioeconomic background, then we can conclude that there is a causal effect of birth month on educational and labor market outcomes. Table 4, Panel B, Column 2 indicates that there is no such systematic difference in fathers' years of education by respondents' birth months among female respondents.

Figure 4, Panel A shows the month-to-month variation in the eventual years of education among males. The lower panel of the figure draws the relation between the birth month and the eventual years of education, adjusting for the difference in the years of education by school cohorts. The relation between birth month and eventual years of education is non-linear, but there is a clear discontinuity between the March born (the youngest in a school cohort) and the April born (the eldest in a school cohort). A similar but more obscure tendency is found among females, as Figure 4 Panel B illustrates. The discontinuity between the March born and the April born implies those who are born early in a cohort obtain more education than those who are born late. After adjusting for the difference in the eventual years of education across school cohorts, there is about a 0.13-year gap between those who were born between April and June and those who were born between January and March (Table 4, Panel A, Column 3). This gap is 0.08 year for females (Table 4, Panel B, Column 3).

The lower section of Figure 5, Panel A indicates the relation between birth month and cohort-adjusted log annual income. As Table 4, Panel A, Column 4 indicates, there is no systematic relation between birth month and employment status, but Column 5 indicates that January-March born people earn 3.9 percent less than April-June born people, and the October-December born earn 1.8 percent less than the April-June born. Figure 5, Panel B indicates a much less clear relation for females. April-June born people are 1.9 percentage points more likely to be employed than January-March born people but earn almost the same amount (Table 4, Panel B, Columns 4 and 5). For males, the relation between the birth month and educational outcomes carries over to the labor-market outcome, and relatively younger people in a school cohort earn less than their elder counterparts. In contrast, reading the systematic patterns of the female results is rather difficult, perhaps because of the complicated effect of birth months on labor-force participation.

The regression results for years of education indicate a discontinuity between the January-March born and the April-June born, probably because of the relative age effect. It is then interesting to examine at which margin of the years of education the discontinuity between the late born and the early born occurs. The margin of the treatment is identified by running the following regression for s=12, 14, and 16:

$$1(S \geq s) = D\beta + school\ cohort\ dummies + u, \tag{1}$$

where $S$ is the years of education, $D$ is the vector of the birth-month dummy variables, and $1(\cdot)$ is an indicator function. The results of the regression appear in Table 5. The results indicate that being April-June born uniformly increases the probability of finishing high school ($s = 12$), finishing junior college or technical college ($s = 14$), and finishing 4-year college ($s = 16$) by around 2.7 percentage points for males and by 1.2 percentage points for females. A similar amount of discontinuity across educational attainment suggests that the effect of birth month on academic outcomes does not wash out as students age. These results support relative age effects rather than absolute age effects.

# 5 Results Based On Narrower Bandwidth

Results so far indicate that elder people in a school cohort perform better than younger people both in schools and in labor markets based on the comparison between April-June born and January-March born groups. However, one may be concerned that unobserved heterogeneity between the two groups causes the gap in outcomes. One way to address this reasonable concern is to narrow the comparison window that presumably equalizes the distribution of unobserved characteristics across groups. Birth month may be correlated with parental SES through strategic birth timing but, because of the difficulty in exact timing, being born in March or April is arguably random. Based on this presumption, the gap in outcomes between March and April born is examined to see the effect of birth month on outcomes.

Table 6 reports the results of this robustness check. Panel A, Columns 1 and 2 report the test score results for boys. The test score gaps for 4th and 8th graders between March born and April born groups are larger than the gaps between January-March born and April-June born reported in Table 3. These sharper results are a product of the narrower comparison window. Columns 3 to 5 report the results for eventual years of education and labor-market outcomes based on the ESS. As for the result of eventual years of education reported in Column 3, the March born group has less education than the April born group and the size of the coefficient is comparable to the gap between the January-March born and April-June born reported in Table

4, Panel A, although the estimated coefficient loses its statistical significance because of the smaller sample size. Column 4 reconfirms that birth month does not affect employment status, which was previously found in Table 4, Panel A. Column 5 indicates that the March born earn about 5 percent less than the April born and this magnitude is comparable to the gap between the January-March born and the April-June born reported in Table 4, Panel A.

Table 6, Panel B reports the robustness check for females. Test score results for 4th graders reported in Column 1 are comparable to the previous results in Table 3, but the result of 8th graders reported in Column 2 is larger than the result in Table 3. Results based on the ESS reported in Columns 3 to 5 reconfirm that elder students in a school cohort have longer eventual years of education, are more likely to be employed, and earn almost the same amount compared with younger people in the same school cohort. Overall, applying a narrower comparison window basically renders identical results and confirms the robustness of previous results based on the quarter of birth.

# 6 Conclusion and Policy Implication

This paper examined whether those who are older in a school cohort do better than their younger counterparts in terms of educational attainment and labor-market outcomes using student test-score data and a labor force survey from Japan. The phenomenon in which older children in a school cohort take advantage of their physical and mental maturity is called the relative age effect in the developmental psychology literature, and it is widely confirmed in both educational performance and sports.

This paper exploits the feature of the Japanese school system that defines the school-entering time by the child's age on April 1. Those who are born on April 1 or before enter primary schools at the beginning of age 6, while those who are born on April 2 or after start at the end of age 6. Thus those who are born in March presumably are at a disadvantage compared with those who are born in April. Exploiting this feature, we can identify the effect of age at school entry on educational attainment and labor market outcomes.

As the law requires a uniform 9 years of compulsory education irrespective of birth month, Japan is an ideal country for estimating the pure effect of age at school entry. In addition, the compliance with the de facto rule of grade progression by almost all children reduces the gap between the local average treatment effect and the average treatment effect.

The analysis results indicate that male 4th and 8th graders who were born right after the cut-off date scored about 0.2 standard deviation more than those who were born right before the cut-off date, and the gap was similar for females. This test-score gap in school by birth month translates into a difference in eventual education attainment. Both males and female who were born between April and June had more education than those who were born between January and March. The initial advantage of early-born over late-born children in primary school persists and develops into a difference in eventual educational attainment. This difference in educational attainment seem to turn into a difference in annual income among males. The persistent effect of age at school entry on outcomes suggests the importance of the relative age effect rather than the absolute age effect.

The use of school tracking during the early stage of education without giving careful consideration to the relative age effect will exacerbate the situation. The Japanese public school system usually starts tracking in high school, but in the last several years, some local governments have initiated tracking during junior high school. Extra attention should be paid to the relative age effects at the time of students' admission.

## Acknowledgement

This research was conducted as part of a project by the Economic and Social Research Institute, the Cabinet Office of Japan, and initial results were disseminated as an ESRI discussion paper (Kawaguchi (2006)). A part of the discussion paper was introduced in Kawaguchi and Mori (2007) in Japanese. I thank an anonymous referee, Akihito Asano, Orley Ashenfelter, David Card, Hidehiko Ichimura, Masahiro Kuroda, Tomohiro Machikita, Hisahiro Naito, David Neumark, Ryo Kambayashi, Ken Yamada, and seminar participants at UC Berkeley, Georgia Institute of Technology, Osaka University, and the Japanese Economic Association 2006 meeting for helpful discussions and comments. This research was completed while I was visiting UC Berkeley. I thank my colleagues there for their hospitality. I thank Yuko Ueno for helping me gain access to the government micro data. I also thank Donna Maurer for editorial assistance.

# References

Allen, J. and R. Barnsley (1993). Streams and tiers: The interaction of ability, maturity, and training in systems with age-dependent recursive selection. *The Journal of Human Resources 28*(3), 649–659.

Angrist, J. and A. B. Krueger (1991). Does compulsory schooling affect schooling and earnings? *Quarterly Journal of Economics 106*(4), 979–1014.

Bedard, K. and E. Dhuey (2006). The persistance of early childhood maturity: International evidence of long-run age effects. *Quarterly Journal of Economics 121*(4), 1437–1472.

Borg, M. G. and J. M. Falzon (1995). Birth date and sex effects on the scholastic attainment of primary schoolchildren: A cross-sectional study. *British Educational Research Journal 21*(1), 61–74.

Cascio, E. and D. W. Schanzenbach (2007). First in the class? age and the education production function. NBER working paper 13663.

Datar, A. (2005). Does delaying kindergarten entrance give children a head start? *Economics of Education Review 25*(1), 43–62.

Dhuey, E. and S. Lipscomb (2008). What makes a leader? relative age and high school leadership. *Economics of Education Review 27*(2), 173–183.

Dobkin, C. and F. Ferreira (2007). Do school entry laws affect educational attainment and labor market outcomes. Available from `http://ssrn.com/abstract=1007492`.

Dudink, A. (1994). Birth date and sporting success. *Nature 369*, 592.

Elder, T. E. and D. H. Lubotsky (2008). Kindergarten entrance age and children's achievement: Impacts of state policies, family background, and peers. Forthcoming in Journal of Human Resources.

Fredriksson, P. and B. Ockert (2005). Is early learning really more productive? the effect of school starting age on school and labor market performance. IZA DP No. 1659.

Helsen, W. F., J. L. Starkes, and J. V. Winckel (2000). Effect of a change in selection year on success in male soccer players. *American Journal of Human Biology 12*(6), 729–735.

Kawaguchi, D. (2006). The effect of age at school entry on education and income. ESRI Discussion Paper Series No. 162.

Kawaguchi, D. and H. Mori (2007). Tanjobi to gakugyo seiseki, saishu gakureki. *Nihon Rodo Kenkyu Zasshi 569*.

Kureishi, W. and M. Wakabayashi (2006). When does the stork bring the baby? shotgun babies vs non-shotgun babies. Osaka Prefecture University Discussion Paper New Series 2006-3.

Kureishi, W. and M. Wakabayashi (2008). Taxing the stork. *National Tax Journal 61*, 167–187.

Lien, L., K. Tambs, B. Oppedal, S. Heyerdahl, and E. Bjertness (2005). Is relatively young age within a school year a risk factor for mental health problems and poor school performance? a population-based cross-sectional study of adolescents in oslo, norway. *BMC Public Health 5*(102).

Martin, M., I. Mullis, and S. Chrostowski (Eds.). *TIMSS 2003 Technical Report*. TIMSS and PIRLS International Study Center, Boston College.

Martin, R. P., P. Foels, G. Clanton, and K. Moon (2004, July). Season of birth is related to child retention rates, achievement, and rate of diagnosis of specific ld. *Journal of Learning Disabilities 37*(4), 309–317.

McEwan, P. J. and J. S. Shapiro (2008). The benefits of delayed primary school enrollment: Discontinuity estimates using exact birth dates. *Journal of Human Resources 43*(1), 1–29.

Thompson, A. H., R. H. Barnsley, and R. J. Dyck (1999). A new factor in youth suicide: The relative age effect. *Canadian Journal of Psychiatry 44*(1), 82–85.

Thompson, D. (1971). Season of birth and success in the secondary school. *Educational Research 14*(1), 56–60.

Table 1: Descriptive Statistics of TIMSS 2003 Sample

| | April-June (Oldest) | July-September | October-December | January-March (Youngest) | Total |
|---|---|---|---|---|---|
| $4^{th}$ Grade Boys Math Score | 151.66 | 151.31 | 150.08* | 149.70* | 150.75 |
| $4^{th}$ Grade Girls Math Score | 150.74 | 150.61 | 149.40* | 148.70* | 149.91 |
| $8^{th}$ Grade Boys Math Score | 150.29 | 150.68 | 150.34 | 149.20* | 150.17 |
| $8^{th}$ Grade Girls Math Score | 150.43 | 150.38 | 149.70 | 148.42* | 149.77 |
| Fraternal Education | | | | | |
| Junior High School | 3.25 | 3.73 | 3.92 | 4.33 | 3.79 |
| High School | 26.93 | 27.31 | 24.13 | 28.49 | 26.7 |
| Junior College / Technical College | 5.96 | 4.86 | 5.43 | 5.29 | 5.38 |
| Four-year college | 29.65 | 28.12 | 29.03 | 26.56 | 28.37 |
| Graduate School | 1.93 | 1.86 | 2.32 | 1.25 | 1.85 |
| Missing | 32.28 | 34.12 | 35.17 | 34.07 | 33.91 |
| Maternal Education | | | | | |
| Junior High School | 3.25 | 2.11 | 2.85 | 3.08 | 2.8 |
| High School | 34.91 | 31.36 | 30.01* | 34.17 | 32.56 |
| Junior College / Technical College | 17.11 | 18.48 | 19.32 | 17.52 | 18.12 |
| Four-year college | 16.67 | 16.94 | 17.63 | 15.21 | 16.64 |
| Graduate School | 0.96 | 0.49 | 0.8 | 0.67 | 0.73 |
| Missing | 27.11 | 30.63* | 29.39 | 29.36 | 29.14 |
| Community Type | | | | | |
| More than 500,000 | 23.3 | 23.28 | 24.85 | 24.16 | 23.87 |
| 100,001 to 500,000 | 36.3 | 36.75 | 37.59 | 35.82 | 36.63 |
| 50,001 to 100,000 | 10.27 | 11.4 | 10.83 | 11.84 | 11.08 |
| 15,001 to 50,000 | 18.34 | 17.94 | 17.32 | 18.69 | 18.06 |
| 3,001 to 15,000 | 11.62 | 10.25 | 9.3* | 9.17* | 10.12 |
| -3,000 | 0.17 | 0.37 | 0.12 | 0.31 | 0.24 |
| Computer possession at home | 78.87 | 79.78 | 80.86 | 80.27 | 79.93 |
| Numbers of books at home | | | | | |
| 0-10 | 13.57 | 13.39 | 11.64 | 13.1 | 12.94 |
| 11-25 | 23.47 | 22.94 | 24.21 | 24.44 | 23.73 |

| | | | | | |
|---|---|---|---|---|---|
| 26-100 | 33.88 | 35 | 35.07 | 33.77 | 34.45 |
| 101-200 | 15.57 | 14.86 | 16.71 | 15.58 | 15.66 |
| 200- | 13.51 | 13.81 | 12.37 | 13.1 | 13.22 |
| Number of people at home | 4.85 | 4.86 | 4.82 | 4.79 | 4.83 |

Note: * indicates that the mean is different from the mean for April-June born at the 5-percent significance level.

Table 2: Age at School Entry and Student Background

Sample: TIMSS 2003, 4th and 8th Graders, Boys and Girls Pooled

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | Mother's Years of Education | Father's Years of Education | Community Type | Computer | Number of Books at Home | Number of People at Home |
| Estimation Method | Ordered Probit | Ordered Probit | Ordered Probit | Probit | Ordered Probit | OLS |
| July-September | 0.07 | -0.05 | -0.02 | 0.01 | 0.01 | 0.01 |
| | (0.05) | (0.06) | (0.03) | (0.01) | (0.04) | (0.04) |
| October-December | 0.09 | 0.02 | -0.08 | 0.02 | 0.02 | -0.03 |
| | (0.06) | (0.05) | (0.03) | (0.01) | (0.03) | (0.04) |
| January-March | -0.03 | -0.13 | -0.04 | 0.01 | -0.01 | -0.06 |
| (Youngest) | (0.06) | (0.05) | (0.03) | (0.01) | (0.04) | (0.04) |
| Constant | - | - | - | - | - | 4.85 |
| | | | | | | (0.04) |
| Observations | 6428 | 5996 | 14004 | 14004 | 14004 | 14004 |
| R-squared | - | - | - | - | - | 0.00 |

Note: Standard errors robust against school-level clustering are reported in parentheses. Mothers' and fathers' years of education are only available for 8th graders and they are categorical variables: 1: primary school, 2: junior high school, 3: high school, 4: junior college / technical college, 5: four-year college, 6: graduate school. Community type is a categorical variable: 1: more than 500,000 people, 2: 100,001 to 500,000 people, 3: 50,001 to 100,000 people, 4: 15,001 to 50,000 people, 5: 3,001 to 15,000 people, 6: fewer than 3,000 people. Computer is a dummy variable that takes one if a student has a computer at home. The coefficient is the marginal effect. Number of books at home is a categorical variable: 1. 0-10 books, 2: 11-25 books, 3: 26-100 books, 4: 101-200 books, 5: more than 200 books.

Table 3: Age at School Entry and Math Test Scores (Mean = 150, SD = 10)

Sample: TIMSS 2003, 4th and 8th Graders

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Grade | 4th | | 8th | |
| Sex | Boys | Girls | Boys | Girls |
| July-September | -0.49 | -0.33 | 0.36 | -0.02 |
| | (0.62) | (0.59) | (0.53) | (0.55) |
| October-December | -1.56 | -1.52 | -0.34 | -0.65 |
| | (0.68) | (0.62) | (0.50) | (0.58) |
| January-March | -1.86 | -2.22 | -1.13 | -1.59 |
| (Youngest) | (0.60) | (0.63) | (0.55) | (0.54) |
| N | 2453 | 2479 | 4558 | 4514 |
| R-squared | 0.09 | 0.07 | 0.11 | 0.16 |

Note: Standard errors robust against school-level clustering are reported in parentheses. The following are included as explanatory variables: categorical dummy variables for mothers' and fathers' years of education (including missing category), community type, and number of books at home, a dummy variable for computer possession at home, and a continuous variable for the number of people. For 4th graders, the examination was given in February 2003. The sample includes those who were born in April 1992 or after. Observations from 150 schools are included. For 8th graders, the examination was given in February or March 2003. The sample includes those who were born April 1988 or after. Observations from 144 schools are included.

Table 4: The Effect of Age at School Entry on Final Education and Labor Market Outcomes, Ages 30-34 in 2002

Panel A: Male Sample, born between April 1968 and March 1972

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dependent Variable | Father's Years of Education Available | Father's Years of Education | Years of Education | Employed | Log (Income) |
| July-September | 0.010 | 0.14 | 0.05 | -0.000 | -0.005 |
| | (0.008) | (0.07) | (0.04) | (0.004) | (0.010) |
| October-December | 0.005 | 0.09 | -0.03 | 0.001 | -0.018 |
| | (0.008) | (0.07) | (0.04) | (0.004) | (0.010) |
| January-March | 0.022 | 0.18 | -0.13 | -0.002 | -0.039 |
| (Youngest) | (0.008) | (0.07) | (0.04) | (0.004) | (0.010) |
| Constant | 0.291 | 10.81 | 13.20 | 0.935 | 5.945 |
| | (0.008) | (0.06) | (0.04) | (0.004) | (0.009) |
| Observations | 26716 | 8875 | 26716 | 26716 | 24627 |
| R-squared | 0.004 | 0.00 | 0.00 | 0.000 | 0.008 |

Panel B: Female Sample, born between April 1968 and March 1972

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dependent Variable | Father's Years of Education Available | Father's Years of Education | Years of Education | Employed | Log (Income) |
| July-September | 0.004 | 0.02 | -0.05 | 0.021 | -0.034 |
| | (0.007) | (0.09) | (0.03) | (0.008) | (0.019) |
| October-December | 0.009 | 0.05 | -0.03 | 0.010 | -0.036 |
| | (0.007) | (0.09) | (0.03) | (0.008) | (0.019) |
| January-March | 0.012 | 0.02 | -0.08 | 0.019 | -0.006 |
| (Youngest) | (0.007) | (0.09) | (0.03) | (0.008) | (0.018) |
| Constant | 0.148 | 11.19 | 12.96 | 0.571 | 5.062 |
| | (0.006) | (0.10) | (0.03) | (0.008) | (0.018) |
| Observations | 27801 | 5369 | 27801 | 27801 | 16310 |
| R-squared | 0.006 | 0.00 | 0.00 | 0.001 | 0.003 |

Note: All specifications include school cohort dummy variables for 1969-1972. Heteroskedasticity robust standard errors are in parentheses.

Table 5: The Effect of Birth Month on the Years of Education, Ages 30-34 in 2002

Male and Female Sample, born between April 1968 and March 1972

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | 1(Educ≥12) | 1(Educ≥14) | 1(Educ≥16) | 1(Educ≥12) | 1(Educ≥14) | 1(Educ≥16) |
| Sample | Male | Male | Male | Female | Female | Female |
| July-September | 0.000 | 0.012 | 0.013 | -0.008 | -0.006 | -0.007 |
| | (0.005) | (0.009) | (0.008) | (0.004) | (0.009) | (0.006) |
| October-December | 0.001 | -0.004 | -0.011 | -0.004 | -0.007 | -0.005 |
| | (0.005) | (0.009) | (0.008) | (0.004) | (0.009) | (0.006) |
| January-March | -0.008 | -0.023 | -0.027 | -0.005 | -0.019 | -0.012 |
| (Youngest) | (0.005) | (0.008) | (0.008) | (0.004) | (0.008) | (0.006) |
| Constant | 0.908 | 0.429 | 0.309 | 0.943 | 0.442 | 0.123 |
| | (0.005) | (0.008) | (0.008) | (0.004) | (0.008) | (0.005) |
| Observations | 26716 | 26716 | 26716 | 27801 | 27801 | 27801 |
| R-squared | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | 0.001 |

Note: Heteroskedasticity robust standard errors are in parentheses. All specifications include school cohort dummy variables.

Table 6: Robustness Check using Narrower Birth-Month Window using Only April- and March-Born for the Analysis Sample

Panel A: Males

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Sample | TIMSS | | | ESS | |
| Dependent Variable | 4th Grade Math | 8th Grade Math | Years of Education | Employed | Log (Income) |
| Born in March | -3.98 | -1.55 | -0.09 | -0.009 | -0.054 |
| | (1.15) | (0.97) | (0.06) | (0.007) | (0.016) |
| Constant | 148.86 | 146.13 | 13.13 | 0.940 | 5.979 |
| | (2.74) | (3.76) | (0.07) | (0.008) | (0.018) |
| N | 396 | 692 | 4678 | 4678 | 4334 |
| R-squared | 0.16 | 0.11 | 0.00 | 0.001 | 0.013 |

Panel B: Females

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Sample | TIMSS | | | ESS | |
| Dependent Variable | 4th Grade Math | 8th Grade Math | Years of Education | Employed | Log (Income) |
| Born in March | -2.22 | -2.59 | -0.10 | 0.039 | -0.030 |
| | (1.24) | (0.78) | (0.05) | (0.014) | (0.031) |
| Constant | 149.15 | 155.88 | 12.95 | 0.572 | 5.105 |
| | (2.65) | (6.47) | (0.05) | (0.016) | (0.035) |
| N | 379 | 722 | 4820 | 4820 | 2820 |
| R-squared | 0.10 | 0.21 | 0.00 | 0.002 | 0.002 |

Note: Standard errors robust against school-level clustering are in parentheses for columns (1) and (2). Heteroskedasticity robust standard errors are in parentheses for columns (3) to (5). All specifications include school cohort dummy variables. The ESS sample includes those who were born between April 1968 and March 1972.

Appendix Table 1: Age at School Entry and Behind Age-Grade Schedule

Sample: TIMSS 2003, 4th and 8th Graders

Dependent Variable: =1 if the Student is behind Age-Grade Schedule

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Grade | 4th | | 8th | |
| Sex | Boys | Girls | Boys | Girls |
| July-September | -0.00 | 0.01 | 0.00 | -0.00 |
| | (0.00) | (0.01) | (0.00) | (0.00) |
| October-December | -0.01 | -0.00 | 0.01 | 0.00 |
| | (0.00) | (0.01) | (0.00) | (0.00) |
| January-March | 0.00 | 0.06 | 0.00 | -0.00 |
| (Youngest) | (0.01) | (0.02) | (0.00) | (0.00) |
| Constant | 0.02 | 0.00 | 0.01 | 0.01 |
| | (0.01) | (0.02) | (0.01) | (0.01) |
| N | 2525 | 2558 | 4558 | 4646 |
| R-squared | 0.01 | 0.03 | 0.01 | 0.00 |

Note: The linear probability model is estimated. Standard errors robust against school-level clustering are reported in parentheses. The following variables are included as explanatory variables: categorical dummy variables for mothers' and fathers' years of education (including missing category), community type, and number of books at home, a dummy variable for computer possession at home, and a continuous variable for the number of people.

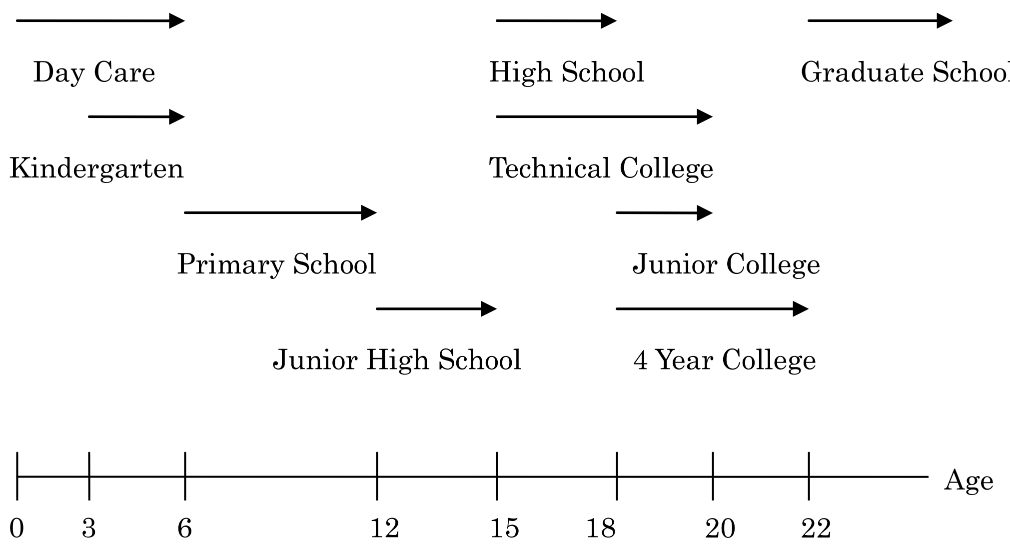
Appendix Table 2: Age at School Entry and Science Test Scores (Mean = 150, SD = 10)

Sample: TIMSS 2003, 4th and 8th Graders

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Grade | 4th | | 8th | |
| Sex | Boys | Girls | Boys | Girls |
| July-September | -0.70 | -0.38 | 0.25 | 0.17 |
| | (0.67) | (0.57) | (0.60) | (0.53) |
| October-December | -1.54 | -1.38 | -0.60 | -0.72 |
| | (0.66) | (0.63) | (0.57) | (0.58) |
| January-March | -1.75 | -1.92 | -1.72 | -1.41 |
| (Youngest) | (0.63) | (0.53) | (0.59) | (0.57) |
| N | 2453 | 2479 | 4558 | 4514 |
| R-squared | 0.05 | 0.06 | 0.09 | 0.14 |

Note: Standard errors robust against school-level clustering are reported in parentheses. The following variables are included as explanatory variables: categorical dummy variables for mothers' and fathers' years of education, community type, and number of books at home, a dummy variable for computer possession at home, and a continuous variable for the number of people. For 4th graders, the examination was given in February 2003. The sample includes those who were born in April 1992 or after. Observations from 150 schools are included. For 8th graders, the examination was given in February or March 2003. The sample includes those who were born April 1988 or after. Observations from 144 schools are included.

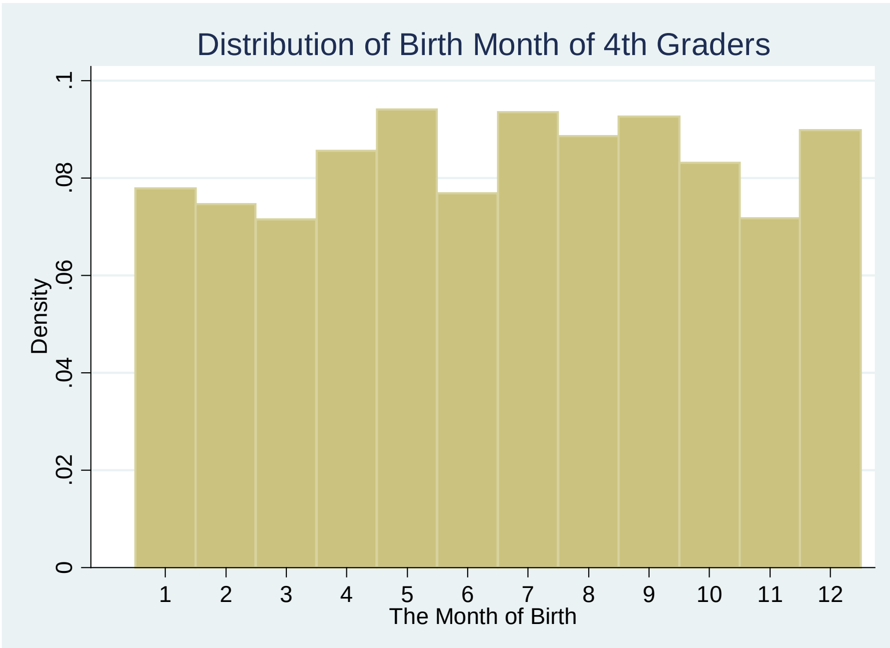
Figure 1: The Japanese School System




Note: Primary school and junior high school are compulsory. Major tracking starts from age 15.

Figure 2: Distribution of Birth Months in the TIMSS

Panel A: Distribution among 4th Graders




Panel B: Distribution among 8th Graders

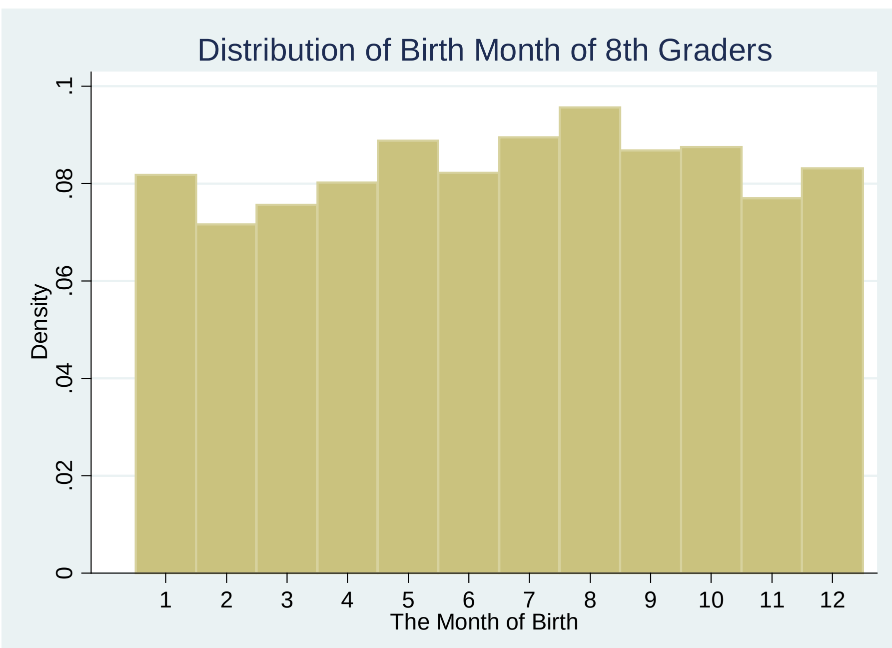
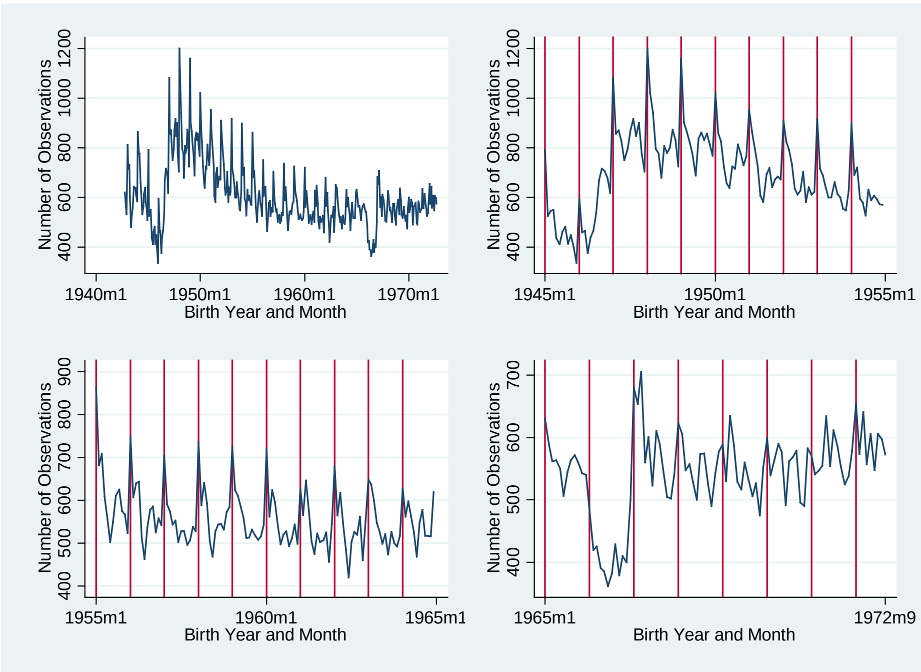


Figure 3: Number of Observations

Panel A: Males 30-59 in 2002




Panel B: Females 30-59 in 2002

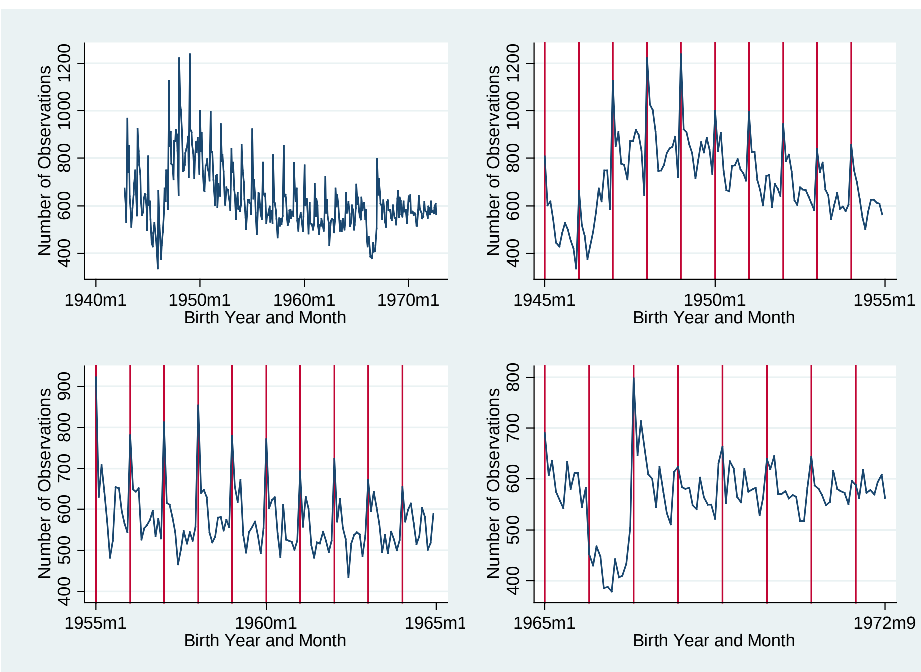
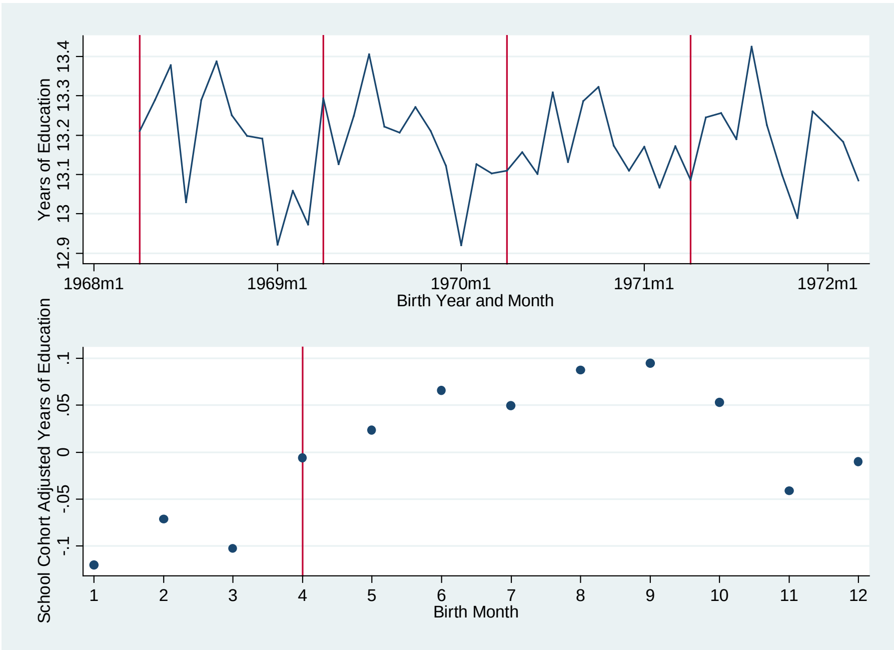


Figure 4: Years of Education

Panel A: Males 30-34 in 2002, born between April 1968 and March 1972




Panel B: Females 30-34 in 2002, born between April 1968 and March 1972

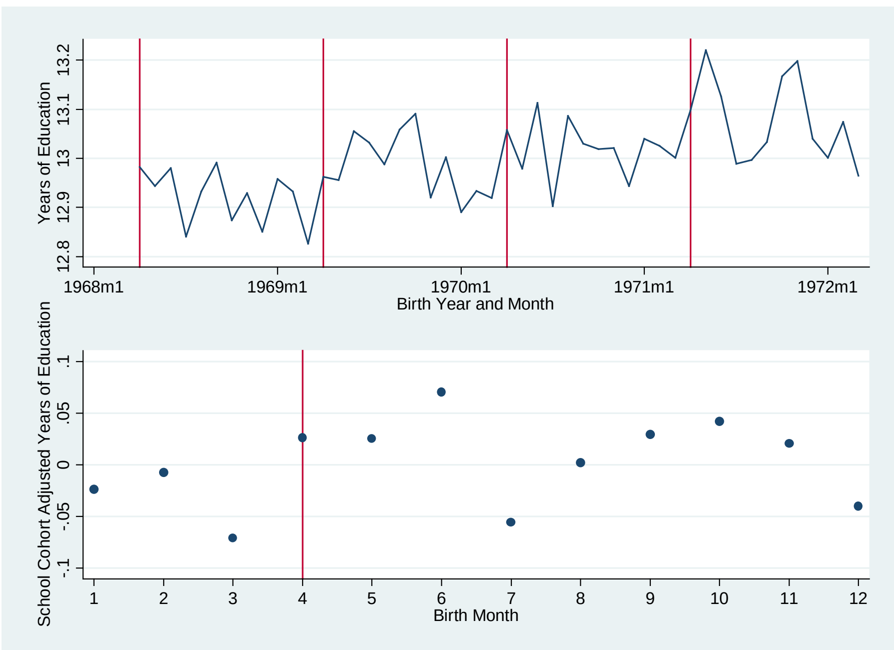
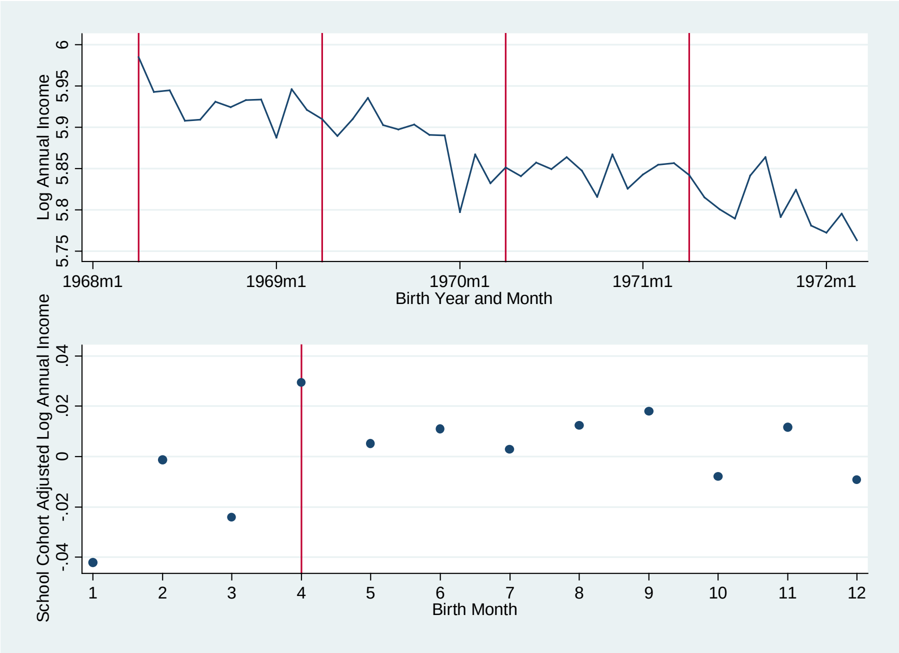


Figure 5: Log Earnings

Panel A: Males 30-34 in 2002, born between April 1968 and March 1972




Panel B: Females 30-34 in 2002, born between April 1968 and March 1972