OSIPP Discussion Paper: DP-2013-E-007

Gender, sibling order, and differences in the quantity and quality of educational attainment: Evidence using Japanese twin data

October 3, 2013

Tien Manh Vu†

Ph.D Candidate, Osaka School of International Public Policy

Hisakazu Matsushige

Professor, Osaka School of International Public Policy

[Key words] identical twins・gender・sibling order・educational attainment・equality
[JEL Codes] I24, I25, J13, J16
[Abstract]
Using 1,045 pairs of Japanese monozygotic twins, we examine differences in educational attainment by considering both the years of schooling (quantity) and the reputation of the last attended school (quality). We find that a difference in learning performance at 15 years of age is one of the key factors connected with differences in both quantity and quality of educational attainment. We also find that when the eldest child in the family is the female twin in the 1950s and 1960s birth cohorts, she forgoes 0.542 years of schooling over her younger twin sister; but for the same birth cohorts, when the eldest child in the family is the male twin, he gains some advantage in the quality of educational attainment over his younger twin brother. However, we find that as the Japanese economy has developed, any difference between twins disappears in subsequent birth cohorts, regardless of gender and sibling order.

†Corresponding author: Tien Manh Vu.

Osaka School of International Public Policy, Osaka University,
1-31 Machikaneyama, Toyonaka, Osaka 560-0043 Japan.
Tel/Fax: +81-(0)6-6850-5656. E-mail: t-vu@osipp.osaka-u.ac.jp.

[Acknowledgments] The authors acknowledge the useful suggestions and helpful comments of Charles Yuji Horioka, Midori Wakabayashi, Wataru Kureishi, and Tsunehiro Ootsuki. The authors also gratefully acknowledge the precious supports from the members of the research project “The Assessments of the Quality and the Productivity of Non-marketable Services” (Grant-in-Aid for Scientific Research (A), No. 3243044, Research Representative: Takeshi Hiromatsu,). All remaining errors are ours.

1. Introduction

It is debatable whether there is a connection between gender, sibling birth order, educational attainment, and changes over time. More specifically, as a result of the son preference in Japan among the 1920–39 birth cohorts reported by Kureishi and Wakabayashi (2011), it is unclear whether the eldest son of that generation have received some advantage in terms of educational attainment. Further, it is unknown whether this difference has changed with the process of economic development in Japan. We argue that monozygotic twins are the best way to detect any connection between the sibling order given by parents and differences in educational attainment. In particular, monozygotic twins help to overcome several confounding problems with birth timing, genetic differences, and unobservable characteristics (Rosenzweig and Wolpin, 2000).

The purpose of this study is to examine differences in educational attainment in Japanese monozygotic twins using a specific survey detailing all the life events of the twins. Compared with previous studies, our research inspects the linkage between the sibling order and differences in educational attainment in monozygotic twins. In addition, extant studies often examine only the quantity of education using either the years of schooling or the highest grade obtained. Our analysis considers not only the difference in the years of schooling, including incomplete qualifications, but also the quality of educational attainment using the reputation of the last attended school.

We first investigate whether a difference in the years in schooling could ever exist for monozygotic twins. We measure the determinants of the difference in years of schooling using ordinary least squares (OLS) models with fixed effects to exclude the common time-invariant factors for the twins. Then, using probit models, we examine the probability of the elder twin receiving either more years of schooling or a better reputation for the last attended school than does the other twin. We separate the data into birth cohort decades and gender to obtain additional insights.

We obtain three main findings. First, the difference in educational attainment between twins is significant and substantially contributed to by the difference in learning performance when the twins are 15 years of age. Second, on average, the eldest in the 1950s and 1960s birth cohorts who is also a female twin completes 0.542 years less schooling than does her twin sibling. Meanwhile, if he is the eldest, the elder male twin gains some advantage in educational attainment once the quality of education is also considered. Finally, the apparent inequality in educational attainment by gender and sibling order in both quantity and quality disappears completely in subsequent decades.

The remainder of the paper is organized as follows. Section 2 details the arguments concerning the differences in educational attainment among siblings and between twins. Section 3 introduces the data used and Section 4 describes our empirical methodology. Section 5 presents our findings. Section 6 discusses our results and provides the conclusions.

2. Differences in educational attainment

2.1. Differences in the level of education among people in general and siblings

To the best of our knowledge, the arguments on why siblings sometimes have different levels of education focus on differences in individual cognitive abilities, schooling choices, investment in education, and differences in the family background over time. Of these, the most common explanation is differences in cognitive ability, which is largely assumed to have a substantial genetic influence. There is, of course, a transfer of human genetic code from parents to their children based on the mechanics of heredity, but in a huge variety of combinations. Thus, the intergenerational transmission of cognitive ability can come from both the paternal and maternal sides (Pronzato, 2012).

Recent research often uses weight at birth as a proxy for the differences. The reasons are twofold. First, the heavier a child is at birth, the better its physical condition. Consequently, a physical condition advantage increases the probability of learning at a higher level. Second, the weight at birth is also reflective of nutrient intake and the level of care from the parents to the child (Behrman and Rosenzweig, 2004; Royer, 2009; Rosenzweig and Zhang, 2012). However, Royer (2009) finds that the marginal benefit of birth weight in terms of education is most robust for birth weights of 2,500 grams and higher.

Another reason for differences in education is choices and/or investment. When children are young, the parents very likely influence the decision on choosing a school as a form of investment (Becker, 1991). However, when the child grows to be an adult or is at least recognized as an adult by law, the decision on schooling now becomes the choice of the individual. Tastes or the desire for schooling also influence choice. Even within the same family, tastes can vary among individuals. In addition, decisions by parents influence any differences in early life. As this affects subsequent choices, it becomes difficult to distinguish between choice (by the child) and investment (by the parents).

The endogeneity of income is a related problem that arises when considering children's outcomes. For example, parental job loss or promotion would influence both household income and parental behaviors. Thus, it is difficult to distinguish between the effects of a variation in income and variations in other unmeasured household conditions (Dahl and Lochner, 2012). Furthermore, it is also difficult to separate differences in endowment from differences in investment (Behrman, Rosenzweig and Taubman, 1994). Griliches (1979), for example, argues that parents compensate for any inequality in innate endowment by correspondingly funding additional investment in human capital and/or adjusting the level of bequests.

Similarly, Becker (1991) asserts that parents invest more in education for a higher-ability child. In turn, the initially better-endowed child voluntarily transfers resources to the less-endowed child when they reach adulthood. In addition, suppose that parents are free to leave any volume of bequest to their children, such that the optimal school for the child will not depend on parental wealth. The child, as well as its parents, then optimizes the final level of schooling basing on its costs and benefits, perhaps even their lifetime value of schooling. However, such optimization is not always perfect, as individuals may under- or overestimate their lifetime earnings (Cunha and Heckman, 2007).

The third major factor in determining differences in education is family background. Regardless of birth order, the wealth of the family can fluctuate over time. If they have just one child, it is obvious that the parents can devote all their resources to that child. However, if there is more than one child, they may have to favor one child over another (Behrman and Taubman, 1986; de Haan, 2010). Even if the parents do not have budgetary constraints, the family background also changes over time. Research in biology and human development indicates that later born children are more likely to have birth defects and physical disadvantages (Behrman and Taubman, 1986). In addition, the level of schooling may correlate with family wealth (income), given that wealth (income) also depends on parental abilities and human capital.

In contrast, in imperfect credit markets with limited borrowing opportunities, poor parents may have to choose between their own consumption and schooling for their children (Dahl and Lochner, 2012). Thus, family income influences schooling choice. For example, Becker (1991) argues that among poor families, the marginal rate of return is higher for better-endowed children, whereas marginal utilities are higher for less-endowed children. Thus, any conclusion about what type of child poor families will invest more in is ambiguous, given that poor parents only invest in the human capital of their children. However, all earnings advantages and ancestor disadvantages disappear in three generations (Becker and Tomes, 1986). In addition, Becker (1991) argues and Haan (2010) concludes that family size could affect average investment in child education.

Moreover, abundant evidence exists of the impact of school inputs on an individual's school attainment (Krueger, 1999). In the economics of education literature, these school inputs are generally the curricula, school organization (such as class size, facilities, and administrative factors), teacher background (education level, experience, and gender), community factors (average expenditure), and school reputation. However, following Hanushek (2003), it is debatable whether these factors exert any statistically significant influence on individual educational attainment.

2.2. Twins and differences in educational attainment

Twins are the offspring of the same pregnancy and can be either dizygotic (DZ) or monozygotic (MZ) twins. DZ twins are the result of two different sperm, whereas MZ twins are the product of a single ovum dividing into two following fertilizations (Squires, 1943). Thus, MZ twins are referred to as identical twins. Parisi et al. (1983) found that twinning is inheritable through the maternal line. Thus, twins are special cases, with MZ twinning occurring at a relatively constant rate of 3.5 to 5 in every thousand births, regardless of race, with DZ twinning in Japan taking place at a rate as low as two births per thousand (Carter, 1970).

Data on twins have some advantage over household data on singletons. First, we can dismiss entirely the impact of natural birth order and differences in family background. In addition, MZ twins possess the same genetic constitution and are always of the same sex (Squires, 1943). Therefore, any difference in endowment in later life would be solely because of parental discrimination or other environmental factors. During development, there is evidence of differences in genetic aspects between identical twins. Fraga et al. (2005) find that older monozygotic twins exhibit significant differences in the overall content and genomic distribution of 5-methylcytosine DNA and histone acetylation, affecting their gene-expression portrait.

However, given that monozygotic twins have the same genetic endowment, the literature has proposed several reasons for any difference in school attainment. Using the records of twins separated at least ten years after infancy, Squires (1943) argues that birth injuries and the nurturing environment can lead to differences in educational attainment between identical twins, even though genetic constitution is a major factor in characteristics such as intelligence. In addition, noneconomic reasons could result in different nongenetic abilities. For instance, the first thing that parents must do differently is to give different first names to their twins (Ashenfelter and Rouse, 1998).

Ashenfelter and Rouse (1998) provide five main reasons for differences in educational attainment between identical twins. The first is measurement error, in that DZ twins could be mistaken for MZ twins. The family background of twins brought up differently could also contribute to differences. Environmental influences in the womb could also provide an advantage or disadvantage to one of the twins. For example, birth weights can differ between twins, and these correlate with differences in postnatal cognition (Segal, 2012), and could ultimately result in differences in educational attainment. Ashenfelter and Rouse (1998) also argue that decisions on marriage or employment could be another reason for differences. Lastly, random deviation from optimal education may also be a reason.

In addition, Isacsson (1999) tests whether differences in educational attainment are randomly determined by dividing the twin data for children aged 14–20 years into MZ twins and DZ twins, and separating males and females. The control variables are mainly height and weight of the individuals. Isacsson (1999) acknowledges that there would not be strong evidence against the assumption that differences in years of schooling are purely random, at least for MZ twins. However, we surmise that this test may underestimate educational differences, as the age for the test ranges between 14 and 20 years. This perhaps provides an insufficient number of observations for children who have completed their educational attainments, especially as the Swedish twin data in Isacsson (1999) suggest average years of schooling of 11.36–11.54 years.

As found by Kureishi and Wakabayashi (2011), the Japanese used to have a son preference, at least for the 1920–39 birth cohorts. We predict that the parents of the individuals with 1950s or 1960s birth cohorts are likely to be this 1920–39 generation. Thus, we hypothesize that the 1950s and 1960s birth cohorts are more likely to face some discrimination in educational attainment by gender and sibling order. Therefore, our study tests this hypothesis.

In addition, to the best of our knowledge, differences in educational attainment are most often measured by years of schooling (quantity), but not by the reputation of the last attended school (quality). Ashenfelter and Rouse (1998), using interview data, find that parents reveal that it is extremely difficult to treat identical twins in any way other than identically. Using a large sample of Swedish twins born in the period 1886–1967, Isacsson (1999) also finds little evidence for the hypothesis that differences in the years of schooling are significantly different from zero. However, we argue that we should include the quality of the last attended school when investigating any difference. We examine quality using the school’s reputation, as measured by its deviation from the mean ranking of all high schools (universities), whenever the years of schooling are identical for both twins. All other things being equal, in Japan, the higher the school ranking, the more difficult it is to gain entry, and therefore only students that are more competitive and/or a better investment will succeed in the required entrance exams.

3. Data

The data we use are from two Web-based surveys conducted in February–March 2012 and August 2012 by Rakuten Research. Rakuten is a Japanese electronic commerce and Internet company located in Tokyo, Japan. In February 2012, Rakuten administered a questionnaire to all its users, approximately 75 million users in total, of which 11.72 million users made at least one purchase on Rakuten Ichiba in each quarter in 2011. The incentive for users to join the survey was points awarded for completing the questionnaire. One point is equal to one Japanese yen. The survey first included six questions on family and siblings. The last question was about whether the respondent was a twin. The Web-based survey design then directed the respondents with a twin sibling to a survey intended for twins. All other respondents completed a separate questionnaire. All respondents received the same award points irrespective of whether they had a twin sibling. The design of the questionnaire was similar to the Princeton Twins Survey and the Employment Status Survey conducted by the Japanese Ministry of Internal Affairs and Communication.

Our data contain nonstudent respondents aged 20–60 years with a twin sibling. The sample size is 2,360 pairs of twins, comprising 1,371 pairs of MZ twins, 882 pairs of DZ twins, and 107 pairs of twins who self-identified as neither MZ nor DZ. To the best of our knowledge, this is the largest sample of Japanese twins ever examined. We found that some pairs with different genders identified themselves as MZ twins, and included these as DZ twins in the data. We used the follow-up questionnaire in August 2012 to verify the MZ twins. Based on the national population statistics, our sample comprises approximately 0.57 percent of the Japanese population of twins aged 20–60 years that are not students[1]. We employ only data relating to MZ twins in all our estimations. The actual number of observations with all the necessary information is 1,045, consisting of 545 pairs of male MZ twins and 500 pairs of female MZ twins. Table 1 provides selected descriptive statistics.

[INSERT TABLE 1 HERE]

The survey contains some specific features that can overcome the limitations of previous studies concerning educational attainment by twins. Apart from self-reported total years in school, we also identify and record years repeated in general education, years at university, and years of uncompleted school (dropouts). Therefore, our measure of years in school is more precise than indicators obtained by converting the highest diploma/degree achieved to years. Given that the Web survey allowed respondents to fill in the name of the last school attended, we combine this information with the high school and university ranking. We obtained the information on the reputational ranking of each high school and college from Kanjuku, a major provider of after-school teaching programs for elementary, junior high, and senior high school students in Japan. We further enriched the data with measures of learning

[1] Using the 2013 Japanese Statistical Yearbook.

performance at school relative to classmates and family wealth when the twins were 15 years old. We include Japanese gross domestic product (GDP) per capita at 1990 constant prices from the 1968 System of National Accounts (68SNA)[2] compiled by the Japan Statistics Bureau in the Ministry of Internal Affairs and Communications, and we specify real GDP per capita to proxy for societal wealth for the different birth cohorts.

Given that the survey respondents are all Internet users and concentrated in large cities, we acknowledge concerns that our sample is not fully random. To reflect this concern, we compare the distribution of the sample with that of the national population. In our sample, there are relatively more twins from Japan's three largest cities, namely Tokyo, Yokohama, and Osaka. The proportion of respondents in our sample aged 40–50 years is also higher than that of the national population. However, we prefer to use the data as they are rather than adjusting the weight for each individual. This is because in our analysis, we investigate past experience, and it is very likely that twins live and work in different regions/prefectures than their original hometown. Moreover, unfortunately we do not have information about the hometowns of the respondents. Therefore, weighting the current place of residence may only worsen the selection problem.

4. Empirical models and econometric specification

4.1. Empirical models

Our main hypothesis is whether the difference in educational attainment between twins is significant using information on both years of schooling completed and the reputation of the last attended school. We employ fixed effects using ordinary least squares (OLS) and probit models in our analysis. Without loss of generality, we assume that educational attainment is a function of the individuals' observable characteristics (Z_i), family background (X_i), and other unobservable characteristics for both individuals (μ_i). We assume that μ_i is uncorrelated with X_i and Z_i in a pair of twins. For example, μ_i can be the genetic endowment of MZ twins. Thus, the educational attainment functions for the elder twin[3] ($educ_{1i}$) and the younger twin ($educ_{2i}$) are

$$educ_{1i} = \alpha X_i + \alpha_1 X_i + \beta_1 Z_{1i} + \mu_i + \varepsilon_{1i}, \tag{1}$$

$$educ_{2i} = \alpha X_i + \alpha_2 X_i + \beta_1 Z_{2i} + \mu_i + \varepsilon_{2i}. \tag{2}$$

where the error terms ε_{1i} and ε_{2i} are $\mathrm{N}(0, \sigma^2)$. However, unlike previous studies, we assume that α_1 differs from α_2, as parents may generally discriminate by sibling order and gender. We then test the hypothesis $\alpha_1 - \alpha_2 = 0$ using appropriate tests. Therefore, the fixed effects (first difference) between twins is

$$deduc_i = educ_{1i} - educ_{2i} = (\alpha_1 - \alpha_2).X_i + \beta_1(Z_{1i} - Z_{2i}) + (\varepsilon_{1i} - \varepsilon_{1i}), \tag{3}$$

where all common characteristics for a pair of twins, including unobservable characteristics, μ_i,

---

[2] GDP using the 68SNA method is available for 1955–98. We convert the GDP data for 1999–2007 using 92SNA from the same source.

[3] We designate the appointed elder sibling as the elder twin and the other twin as the younger twin, even though they are of the same age. Deciding upon the eldest child is common and typical among Japanese families. The survey recorded this information for all individuals.

are excluded. Equation (3) is then an unbiased estimation. This is the advantage of conducting a natural experiment using data on twins.

In order to integrate information on the difference in school reputation, we employ probit models. We assume that the elder twin obtains a relative advantage, and investigate the connections with this event. We distinguish three cases, namely the elder twin has more years of schooling, the elder twin attends a higher-ranking college given that both twins have some college attendance, and the elder twin attends a higher-ranking high school given that both twins ceased education after graduating from high school. With school ranking, we construct the dependent variable $higher$ as follows:

$$higher_i = \begin{cases} 1, if\ educ_{1i} - educ_{2i} > 0\ or\ if\ ranking_{1i} > ranking_{2i}\ |educ_{1i} = educ_{2i} \\ 0, if\ otherwise \end{cases}, \tag{4}$$

where $ranking_{1i}$ and $ranking_{2i}$ are the deviations of the school ranking from the population mean on a 0–100 scale. Only when twins have the same years of schooling do we apply either the high school ranking or university ranking based on the last school attended.

4.2. Specifications

We construct three groups of explanatory variables. The first group, ($dZ_i = Z_{1i} - Z_{2i}$), consists of the weight differences between a pair of twins at birth ($dbirth\_weight$) and differences in learning performance when the twins were 15 years old ($dlearning\_15$). The second group of variables controls for differences across families. These are the number of siblings other than the twins ($number\ of\ siblings$), the maternal age of the mother when the twins were born ($maternal\ age$), and the number of years the twins were brought up together in the same family ($together\_age$). We also include information from when the twins were 15 years of age, including relative family wealth ($familywealth\_15$) and real GDP per capita at 1990 constant prices ($real\_GDPpc\_15$). We also add an interaction term between $real\_GDPpc\_15$ and $dlearning\_15$ to examine the effect of wealth and learning performance on educational attainment. The final group of explanatory variables controls for gender and sibling order. When the data is divided into three birth cohorts, we use $BBeldest$ ($GGeldest$) to indicate MZ male (female) twins who is also the eldest child of the family. We combine gender, eldest child status, and birth cohort in all other cases. Table 1 explains the combinations. For example, we use $BBeldest70$ to designate if the twins are males from the 1970s birth cohort with no elder siblings. The reference dummy is $GGnoteldest560$ and $BBnoteldest560$ if only MZ BB twins are considered. Finally, as males and females potentially have different educational attainments by generation, we divide the data into several different scenarios by separating the data by gender, and then both by gender and across three different birth cohorts.

5. Results

5.1. Nonparametric evidence of differences between twins in years of schooling

We conduct one-sample t-tests to examine the difference in education attainment as measured by years of schooling in our sample. As shown in Table 2, we accept that the difference in years of schooling between MZ twins is zero with a 95 percent confidence level. However, similar to Isacsson (1999), we find that the probability of being correct if choosing the alternative null hypothesis (elder twin has more years in school) is 18.49 percent among MZ twins.

[INSERT TABLE 2 HERE]

We also examine whether educational attainment as measured by years of schooling changes by sibling order and gender over time. We elaborate using visual evidence from the data and illustrate the results in Graph 1. As shown, MZ GG twins in the 1950s and 1960s birth cohorts tend to have a greater variance in years of schooling than MZ BB twins of the same birth cohorts. Further, the quintile distributions of MZ BB twins exhibit less visible differences within each birth cohort, while those for MZ GG twins change noticeably over time. Clearly, MZ elder female twins tend to have fewer years of schooling when compared with their twin siblings in the 1950s and 1960s birth cohorts.

[INSERT GRAPH 1 HERE]

5.2. Differences between twins

5.2.1. Evidence from the OLS model with fixed effect

As shown in Table 3, the difference in attitude toward gender and the eldest child by birth cohort is minimal in explaining the difference in total years of schooling between the elder and younger MZ BB twins, as shown by the results of the Wald tests. However, attitudes have a significant nexus (5 percent error) with the differences in the total years of schooling between the elder and younger MZ GG twins. More specifically, the eldest female of the family who is also a twin born in the 1950s or 1960s would have approximately 0.542 years less schooling on average than her twin sister. This result is consistent with the evidence presented in Graph 1.

[INSERT TABLE 3 HERE]

5.2.2. Evidence from probit models including the difference in the reputation of the last attended school

When school rank is considered, the priority given to the elder MZ BB twin is likely to connect with the fact that he is the eldest child of the family, as evidenced in column 2 of Table 4. The priority is at least a better reputation of the last attended school. This difference is not identifiable when using only the years of schooling, as in Ashenfelter and Rouse (1998) and Isacsson (1999).

[INSERT TABLE 4 HERE]

We further examine the data by dividing the sample into three different birth cohorts. The estimated coefficient for the eldest child in column 2 of Table 5 indicates that for the 1950s and 1960s birth cohorts, there is better educational attainment by the elder twin. This result complements the findings in Kureishi and Wakabayashi (2011) on son preference in Japan. The eldest son appears to have some advantage over his siblings, even his twin brother of the same age, and the difference is more complicated than that suggested merely by years of schooling.

[INSERT TABLE 5 HERE]

5.3. The disappearance of differences by gender and sibling order in recent decades

The evidence shows that the differences in educational attainment by gender and sibling order evident in the 1950s and 1960s birth cohorts largely evaporate in subsequent decades. First, the distributions of MZ BB and GG twins are much more alike among 1980s and 1990s birth cohorts, as shown in Graph 1. The distributions of the MZ elder twins and younger twins converge after the 1950s and 1960s birth cohorts. Second, as mentioned, the interaction between $rgdppc$ and $dlearning\_15$ in column 5 of Table 3 shows an improvement in equality for MZ BB twins. Similarly, as shown by the estimated coefficient for the interaction term ($\mathrm{real\_GDPpc\_15} * \mathrm{dlearning\_15}$) in column 6 of Table 4, higher real GDP per capita is associated with the elder MZ GG twin having better educational attainment, given the difference in learning performance at age 15. Thus, unlike the situation where the elder MZ GG twin would surrender some years of schooling in the past, this indicates that the elder MZ GG twin is better off with increases in wealth. Third, we witness that the elder male twin may have some advantage because he is the eldest child in the family, as shown in column 2 of Table 5. However, this advantage vanishes during the following decades. As shown in Table 5, the coefficients of eldest child in column 5 and 8 are statistically insignificant among birth cohorts in the 1970s, 1980s, and 1990s. Thus, the result supports the prediction in Vu (2013) that discrimination among children by gender due to son preference would soon disappear as a result of economic development.

5.4. Differences in learning performance at age 15 and educational attainment

Through the OLS model with fixed effects and the probit models considering both the differences in years of schooling and the reputation of the last attended school, we find that $dlearning\_15$ is the more influential factor in explaining the differences in educational attainment between MZ twins. In all of our estimations, the sign of $dlearning\_15$ is always positive when statistically significant, implying that better learning at age 15 leads to higher educational attainment. Interestingly, this challenges the conventional hypothesis that monozygotic twins are indeed strictly identical. In fact, fifteen years after their births, there are significant differences between twins in terms of their ability to learn. This finding suggests further study to investigate the development of variances between twins throughout their lifetimes.

6. Discussion and conclusion

This paper examines the differences in educational attainment between MZ twins by considering both the total years of schooling (quantity) and the reputation of the last attended school (quality). We use OLS fixed effects analysis for the former and probit models for both the former and the latter to investigate the differences in educational attainment. We find that the difference in learning performance at 15 years of age contributes significantly to the differences in educational attainment in terms of both quantity and quality. In addition, differences in years of schooling between MZ BB twins generally have minimal connection with gender, sibling order, and birth cohort. However, once we consider the reputation of the last attended school, as well as being the eldest son of the family, the MZ BB elder twin in the 1950s and 1960s birth cohorts exhibits higher educational attainment than his twin sibling. In contrast, we find the MZ female elder twin in the same birth cohort has fewer years of schooling than her twin sibling. Nevertheless, greater equality in educational attainment is in evidence in birth cohorts in recent decades. As a rule, the gap between twins, by both gender and sibling order, has disappeared as the Japanese economy has developed.

Although we do not have any direct evidence to explain the reason for the fewer years of schooling for the eldest MZ female twin in the 1950s and 1960s birth cohorts, we would argue that the notion of a role model is one possible explanation. In Japan, the elder sibling (*sempai*) assists the younger sibling (*kohai*), a custom that prevails in contemporary Japanese society. This does not conflict with the case where the eldest son of the same birth cohort is more likely to have a greater advantage than his twin. Put differently, the eldest son could reflect the lineage of the whole family, which would be more important than the relationship between twins.

We acknowledge that the difference we have found between male twins where one is the eldest child in the family provides only weak evidence, as the estimated coefficients are statistically significant at only the 90 percent level. In addition, we included an additional dummy identifying the case where the twins have no older brothers but do have one or more older sisters. In practice, Japanese families appoint the elder male twin as the eldest son even though he may not be the eldest child. However, the marginal benefit of including additional dummy variables to reflect this would be statistically small. The likelihood ratio test that we can nest the adjusted models in the original models can be accepted. The added dummy is also statistically insignificant. We predict that birth order then has a stronger negative relationship with the difference, given that the difference is statistically sensitive to change. Thus, we exclude this dummy variable from our analysis.

There are possible concerns about sample selection in our data. More specifically, given the nature of the survey, we may have inadvertently sampled only the computer literate. This may account for educated older monozygotic twins being overrepresented in the sample. However, among the 1950s and 1960s birth cohorts, the number of direct respondents is 88 elder female twins and 60 younger female twins, while the results show that the MZ GG elder twin is more likely to receive fewer years of schooling. Similarly, although the number of direct respondents includes 140 elder male twins and 78 younger male twins in the same birth cohorts, there is approximately 15 mean years of schooling for both the elder and younger twins[4]. Thus, there should be no problem of bias selection based on computer literacy. Nevertheless, we acknowledge that there could be bias as reflected in online shopping habits.

Another concern is whether we can transfer these findings to singletons within a family, given that twinning is obviously a special case. We acknowledge two other limitations of our research. First, we do not have singleton siblings with which to compare, perhaps using a similar method to that used by Behrman et al. (1994), to enable us to conclude whether the inference is valid for all kinds of siblings. Second, in our estimations, we omit the fact that it is more difficult, and possibly costly (Behrman et al., 1994), for parents to differentiate between twins than between singletons. Thus, the differences we have found may underestimate the reality among singletons. We suggest the need for further research to test a similar hypothesis for differences in educational attainment in both quantity and quality among singletons, perhaps using household fixed effects.

---

[4] The 95 percent confidence intervals for years of schooling are 15.21–15.8 and 15.01–15.6 for the elder and younger twin cohorts, respectively.

As a final point, the disappearance of the difference in educational attainments by gender and sibling order in the 1970s birth cohort and beyond implies that equality is in evidence among Japanese twins. Our estimations also suggest that monozygotic twins are not identical in educational attainment, with learning performance at age 15, sibling order, and gender being potential candidate explanations.

References

Ashenfelter, O., & Rouse, C. (1998). Income, schooling, and ability: evidence from a new sample of identical twins. *The Quarterly Journal of Economics, 113*(1), 253–284.

Becker, G. S. (1991). *A treatise on the family*. Cambridge, Mass., Harvard University Press.

Becker, G. S., & Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics, 4*(3), S1–39.

Behrman, J. R., & Rosenzweig, M. R. (2004). Returns to birth weight. *Review of Economics and Statistics, 86*(2), 586–601.

Behrman, J. R., Rosenzweig, M. R., & Taubman, P. (1994). Endowments and the allocation of schooling in the family and in the marriage market—the twins experiment. *Journal of Political Economy, 102*(6), 1131–1174.

Behrman, J. R., & Taubman, P. (1986). Birth-order, schooling, and earnings. *Journal of Labor Economics, 4*(3), S121–S145.

Carter, C. O. (1970). Twin births. *Nature, 226*, 1271–1271.

Cunha, F., & Heckman, J. J. (2007). Identifying and estimating the distributions of ex post and ex ante returns to schooling. *Labour Economics, 14*(6), 870–893.

Dahl, G. B., & Lochner, L. (2012). The impact of family income on child achievement: evidence from the earned income tax credit. *American Economic Review, 102*(5), 1927–1956.

de Haan, M. (2010). Birth order, family size and educational attainment. *Economics of Education Review, 29*(4), 576–588.

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestart, M. L., et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America, 102*(30), 10604–10609.

Griliches, Z. (1979). Sibling models and data in economics: beginnings of a survey. *Journal of Political Economy, 87*(5), S37–64.

Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal, 113*, F64–F98.

Isacsson, G. (1999). Estimates of the return to schooling in Sweden from a large sample of twins. *Labour Economics, 6*(4), 471–489.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics, 114*(2), 497–532.

Kureishi, W., & Wakabayashi, M. (2011). Son preference in Japan. *Journal of Population Economics*, *24*(3), 873–893.

Parisi, P., Gatti, M., Prinzi, G., & Caperna, G. (1983). Familial incidence of twinning. *Nature,* *304*(5927), 626–628.

Pronzato, C. (2012). An examination of paternal and maternal intergenerational transmission of schooling. *Journal of Population Economics, 25*(2), 591–608.

Rosenzweig, M. R., & Wolpin, K. I. (2000). Natural "natural experiments" in economics. *Journal of Economic Literature, 48*(4), 827–847.

Rosenzweig, M. R., & Zhang, J. (2012). Economic growth, comparative advantage, and gender differences in schooling outcomes: evidence from the birth weight differences of Chinese twins. Yale University Discussion Paper.

Royer, H. (2009). Separated at girth: US twin estimates of the effects of birth weight. *Applied Economics, 1*(1), 49–85.

Segal, N. L. (2012). Twins, birth weight, cognition, and handedness. *Proceedings of the National Academy of Sciences of the United States of America, 109*(48), E3293–E3293.

Squires, J. C. (1943). Studies of twins in the social sciences. *Nature, 152*, 437–440.

Vu, T. M. (2013). One male offspring preference: evidence from Vietnam using a split-population model. *Review of Economics of the Household*. In press. DOI: 10.1007/s11150–013–9183–z.

Graph 1
Quintile distributions of educational attainment by sex, sibling order, and birth cohorts among MZ twins

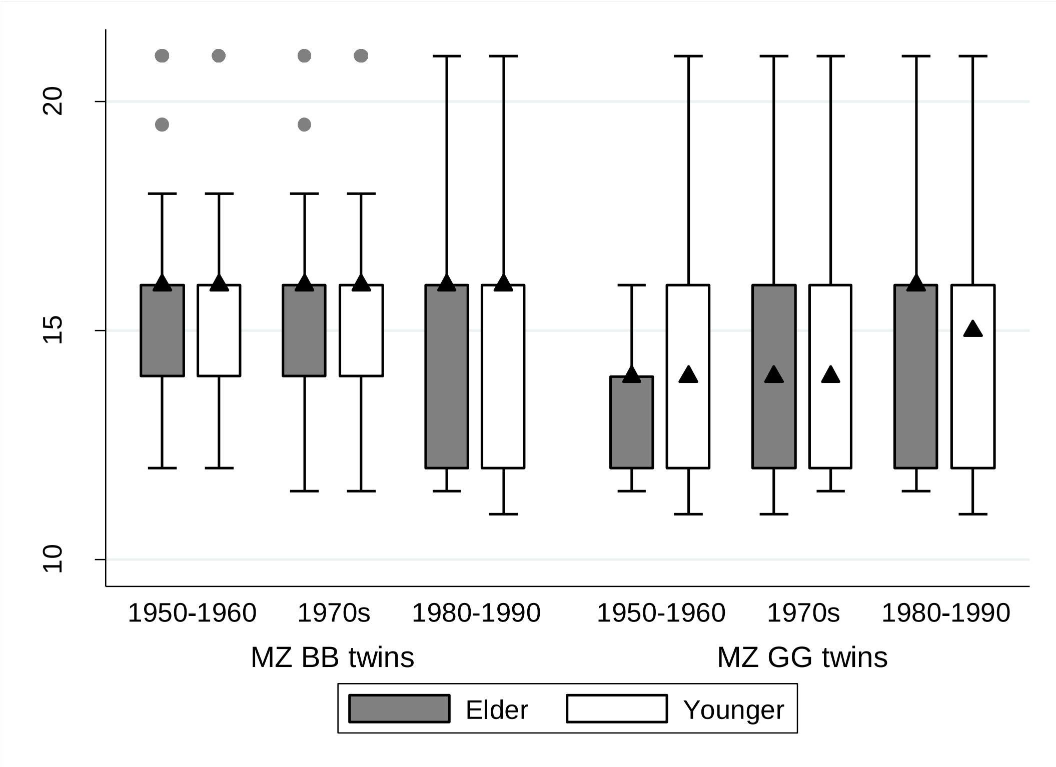

Table 1
Descriptive statistics.

Variable	Description	Obs.	Mean	Std. Dev.	Min.	Max.
deduc	Difference in years of schooling between the elder and younger twin	1,045	0.042	1.517	–9	9
higher	=1 if the elder twin has more years of schooling or better reputation for the last attended school	1,045	0.553	0.497	0	1
BBeldest	MZ male–male twins who do not have elder siblings	1045	0.260	0.439	0	1
GGeldest	MZ female–female twins who do not have elder siblings	1045	0.234	0.424	0	1
BBeldest560	BB twins who are the eldest children of the family and born in 1950s or 1960s	1,045	0.105	0.307	0	1
BBeldest70	BB twins who are the eldest children of the family and born in 1970s	1,045	0.096	0.294	0	1
BBeldest890	BB twins who are the eldest children of the family and born in 1980s or 1990s	1,045	0.059	0.236	0	1
BBnoteldest560	BB twins who have elder sibling and born in 1950s or 1960s	1,045	0.103	0.305	0	1
BBnoteldest70	BB twins who have elder sibling and born in 1970s	1,045	0.101	0.302	0	1
BBnoteldest890	BB twins who have elder sibling and born in 1980s or 1990s	1,045	0.056	0.231	0	1
GGeldest560	GG twins who are the eldest children of the family and born in 1950s or 1960s	1,045	0.066	0.248	0	1
GGeldest70	GG twins who are the eldest children of the family and born in 1970s	1,045	0.098	0.297	0	1
GGeldest890	GG twins who are the eldest children of the family and born in 1980s or 1990s	1,045	0.071	0.257	0	1
GGnoteldest560	GG twins who have elder sibling and born in 1950s or 1960s	1,045	0.076	0.264	0	1
GGnoteldest70	GG twins who have elder sibling and born in 1970s	1,045	0.102	0.303	0	1
GGnoteldest890	GG twins who have elder sibling and born in 1980s or 1990s	1,045	0.066	0.248	0	1
number of siblings	Number of siblings other than the twins	1,045	1.099	1.071	0	8
together_age	Total years the twins were raised together	1,045	16.015	10.653	0	40
familywealth_15	Family wealth when twins were 15 years old (=1 if wealth was average or above)	1,045	0.294	0.456	0	1
real_GDPpc_15	Real GDP per capita of Japan when twins were 15 years old (68SNA)	1,045	309.363	70.865	134.18	425.96
maternal age	Maternal age of the mother when twins were born	1,045	27.502	4.079	18	48
dbirth_weight	Difference in weight at birth between the elder and younger twin	1,045	41.874	292.181	–3001	1500
dlearning_15	Difference in learning performance when twins were 15 years old	1,045	0.063	0.814	–4	4
real_GDPpc_15 * dlearning_15	Interaction term between real_GDPpc_15 and dlearning_15	1,045	19.856	264.225	–1639.55	1548.948

Table 2

Tests for differences in educational attainment between twins by years in school.

H0: The years of schooling are the same for both twins			
	MZ twins	MZ BB twins	MZ GG twins
Pr(T<t)	0.8151	0.9779	0.2111
Pr(T>t)	0.1849	0.0221	0.7889
Pr(|T|>|t|)	0.3699	0.0442	0.4222
N (pairs of twins)	1,045	545	500

Table 3

OLS with fixed effects estimates of differences in years of schooling between MZ twins.

Twins sample	Pooled MZ	MZ BB	MZ GG	Pooled MZ	MZ BB	MZ GG
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	deduc	deduc	deduc	deduc	deduc	deduc
BBeldest560	0.1045	0.0346		0.0825	0.0385	
	(0.1954)	(0.2170)		(0.1940)	(0.2167)	
BBeldest70	0.4038	0.2451		0.3935	0.2602	
	(0.2876)	(0.3243)		(0.2881)	(0.3230)	
BBeldest890	0.1004	–0.1240		0.0914	–0.1129	
	(0.3504)	(0.4111)		(0.3501)	(0.4095)	
BBnoteldest70	0.2629	0.0273		0.2498	0.0296	
	(0.2552)	(0.2825)		(0.2564)	(0.2824)	
BBnoteldest890	0.5400	0.2299		0.5154	0.2042	
	(0.3504)	(0.4030)		(0.3523)	(0.4037)	
BBnoteldest560	0.1075			0.0851		
	(0.2043)			(0.2047)		
GGeldest560	–0.4444**		–0.5463**	–0.4521**		–0.5415**
	(0.2131)		(0.2377)	(0.2148)		(0.2362)
GGeldest70	0.1846		0.1959	0.1720		0.1998
	(0.3022)		(0.3837)	(0.3029)		(0.3899)
GGeldest890	0.5416		0.6370	0.5294		0.6398
	(0.3317)		(0.4499)	(0.3329)		(0.4535)
GGnoteldest70	0.0185		0.1297	0.0003		0.1349
	(0.2820)		(0.3502)	(0.2834)		(0.3565)
GGnoteldest890	0.3404		0.5305	0.3244		0.5335
	(0.3399)		(0.4734)	(0.3412)		(0.4782)
number of siblings	–0.0167	0.0475	–0.1041	–0.0137	0.0556	–0.1046
	(0.0530)	(0.0631)	(0.0962)	(0.0529)	(0.0619)	(0.0967)
together_age	0.0060	0.0011	0.0105	0.0061	0.0013	0.0105
	(0.0047)	(0.0062)	(0.0069)	(0.0047)	(0.0062)	(0.0069)
familywealth_15	–0.0779	–0.1808	0.0387	–0.0741	–0.1665	0.0396
	(0.0944)	(0.1192)	(0.1474)	(0.0944)	(0.1182)	(0.1473)
real_GDPpc_15	–0.0019	–0.0012	–0.0029	–0.0018	–0.0009	–0.0029
	(0.0017)	(0.0023)	(0.0026)	(0.0018)	(0.0024)	(0.0027)
maternal age	–0.0056	–0.0101	0.0005	–0.0052	–0.0095	0.0005
	(0.0106)	(0.0143)	(0.0162)	(0.0106)	(0.0143)	(0.0163)
dbirth_weight	–0.0002	0.0001	–0.0004*	–0.0002	0.0001	–0.0004**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0002)
dlearning_15	0.4960***	0.5168***	0.4648***	0.8444***	1.1295***	0.3317
	(0.0733)	(0.0892)	(0.1222)	(0.2951)	(0.2881)	(0.7095)
real_GDPpc_15 * dlearning_ 15				–0.0011	–0.0020**	0.0004
				(0.0009)	(0.0009)	(0.0019)
Constant	0.5339	0.6389	0.6303	0.4997	0.5095	0.6292
	(0.5666)	(0.7744)	(0.8515)	(0.5720)	(0.7813)	(0.8514)
N (pairs of twins)	1,045	545	500	1,045	545	500
R-squared	0.092	0.094	0.095	0.093	0.100	0.095
Wald test with H0: All dummies combining gender, eldest and birth cohort are zero						
F-statistic	1.80	0.78	2.24	1.73	0.74	2.27
Prob>F	0.0493	0.5679	0.049	0.0627	0.5913	0.0468

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 4

Marginal effects using probit estimation for the differences in both years of schooling and the reputation of the last attended school between MZ twins.

Twins sample	Pooled MZ	MZ BB	MZ GG	Pooled MZ	MZ BB	MZ GG
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	mfx	mfx	mfx	mfx	mfx	mfx
BBeldest	0.0631	0.0953*		0.0630	0.0993*	
	(0.0451)	(0.0555)		(0.0451)	(0.0556)	
GGeldest	0.0403		0.0240	0.0401		0.0254
	(0.0468)		(0.0609)	(0.0469)		(0.0609)
number of siblings	0.0027	0.0029	0.0132	0.0025	0.0052	0.0117
	(0.0185)	(0.0242)	(0.0297)	(0.0185)	(0.0243)	(0.0298)
together_age	0.0032**	0.0025	0.0034	0.0032**	0.0025	0.0033
	(0.0016)	(0.0023)	(0.0022)	(0.0016)	(0.0023)	(0.0022)
familywealth_15	–0.0042	0.0355	–0.0429	–0.0043	0.0378	–0.0410
	(0.0351)	(0.0484)	(0.0519)	(0.0352)	(0.0484)	(0.0520)
real_GDPpc_15	0.0007***	0.0006**	0.0006*	0.0007***	0.0007**	0.0006*
	(0.0002)	(0.0003)	(0.0004)	(0.0002)	(0.0003)	(0.0004)
maternal age	0.0061	0.0015	0.0135**	0.0061	0.0017	0.0135**
	(0.0040)	(0.0054)	(0.0061)	(0.0040)	(0.0054)	(0.0061)
dbirth_weight	–0.0000	0.0000	–0.0001	–0.0000	0.0000	–0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
dlearning_15	0.1021***	0.1203***	0.0831***	0.0817	0.2734**	–0.2120
	(0.0203)	(0.0288)	(0.0288)	(0.0947)	(0.1283)	(0.1566)
real_GDPpc_15 * dlearning_ 15				0.0001	–0.0005	0.0009*
				(0.0003)	(0.0004)	(0.0005)
N (pairs of twins)	1,045	545	500	1,045	545	500

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 5

Marginal effects using probit estimation for the differences in both years of schooling and the reputation of the last attended school between MZ twins by birth cohort.

Birth cohort	1950s–1960s			1970s			1980s–1990s		
Sample	Pooled MZ	MZ BB	MZ GG	Pooled MZ	MZ BB	MZ GG	Pooled MZ	MZ BB	MZ GG
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	mfx	mfx	mfx	mfx	mfx	mfx	mfx	mfx	mfx
BBeldest	0.0988	0.1672*		0.0656	0.0844		0.0156	0.0035	
	(0.0738)	(0.0868)		(0.0736)	(0.0935)		(0.0925)	(0.1211)	
GGeldest	–0.0056		0.0155	0.0369		0.0102	0.0998		0.1472
	(0.0884)		(0.1291)	(0.0721)		(0.0893)	(0.0857)		(0.1122)
number of siblings	0.0648*	0.0580	0.1348*	–0.0015	0.0183	–0.0241	–0.0525	–0.0743	–0.0121
	(0.0356)	(0.0426)	(0.0708)	(0.0306)	(0.0421)	(0.0454)	(0.0330)	(0.0478)	(0.0486)
together_age	0.0011	0.0019	–0.0010	0.0035	0.0021	0.0053	0.0068**	0.0044	0.0089**
	(0.0031)	(0.0041)	(0.0050)	(0.0025)	(0.0037)	(0.0035)	(0.0030)	(0.0047)	(0.0038)
familywealth_15	–0.1094*	–0.0469	–0.1436	0.0480	0.0390	0.0651	0.0486	0.1968**	–0.0948
	(0.0587)	(0.0759)	(0.0988)	(0.0600)	(0.0878)	(0.0846)	(0.0634)	(0.0925)	(0.0901)
maternal age	0.0049	0.0043	0.0136	0.0046	–0.0061	0.0185*	0.0092	0.0111	0.0088
	(0.0064)	(0.0082)	(0.0109)	(0.0069)	(0.0094)	(0.0103)	(0.0080)	(0.0127)	(0.0105)
dbirth_weight	0.0001	0.0001	–0.0000	–0.0000	0.0000	–0.0000	–0.0001	–0.0002	–0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)
dlearning_15	0.0857**	0.1544***	0.0207	0.1252***	0.1289***	0.1102*	0.0944***	0.0804	0.1011**
	(0.0354)	(0.0499)	(0.0560)	(0.0362)	(0.0471)	(0.0569)	(0.0347)	(0.0618)	(0.0418)
Observations	366	218	148	415	206	209	264	121	143

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.