Do schools form human capital?

Distributional divide and cohort-based analysis in Japan

Keiji Saito*

Graduate School of Economics, The University of Tokyo

Abstract

This paper presents empirical wage analyses in Japan on the basis of the framework of Saito(2005). Data arranged by cohorts enable analyses on the effect of the educational advancement rate with regard to the wage ratio. Empirical results in Japan suggest that a statistical trick causes most of the wage ratio changes. Further, the real effects of higher education on wages range from slightly negative to insignificant.

JEL classification: D31; J24; J31

Keywords: Wage ratio; Cohort; Human capital; Japan

* E-mail: master@keijisaito.info

1. Introduction

The object of this paper is to measure the magnitude of the fallacy demonstrated in Saito (2005) and the real effects of higher education on wages in Japan. Saito (2005) reveals the fallacy of the wage ratio using simple simulation analyses. The wage differentials and returns to education would be biased by the fallacy. However, the empirical magnitude of the fallacy remains unclear. This paper begins by confirming the considerable magnitude of the fallacy in Japan. Furthermore, the fallacy also provides a clue to measuring the real effects of higher education. I analyze the real effects of higher education on wage in Japan.

We have two major theories to analyze the effect of education or schools on economics. The first theory is the human capital theory, which considers schools to be institutions for the formation of human capital, as represented by Becker (1964). Schools increase their students' productivity and wages. The other theory, the signaling or screening theory, states that advancement to higher education eliminates asymmetric information, as suggested by Spence (1973). Students can reveal their productivity by advancing to higher education. Raising productivity in schools is not a requirement for the signaling theory.

Numerous discussions and analyses have been conducted to determine which of the two theories is more practical. In early research stages, Griliches (1977) summarizes that consistent estimators of the causal effects of education cannot be obtained by ordinary least squares. Educational status is a textbook example of an endogenous variable. Subsequent empirical studies make efforts to take into account unobserved ability and self-selection. Card (2001) reviews empirical studies in the 1990s.

However, it is difficult to distinguish the pre-school effects from the in-school ones. Moreover, one theory does not violate the other. Previous attempts to empirically distinguish between the two have been inconclusive. In contrast, this paper provides a judgement in Japan.

The paper is organized as follows: Section 2 describes the data from the Basic Surveys on Wage Structures in Japan. Section 3 analyzes the college/high-school wage ratio which is used as a primary indicator of the wage differential between educational statuses. Section 4 approaches the real effects of higher education using the ratio of quartiles. Section 5 analyzes the wage ratio on weighted averages of employees in accordance with the principles stated in Saito (2005). Section 6 presents the conclusions and provides caveats on the interpretations.

2. Data

The primary data for this paper has been obtained from 28 consecutive Basic Survey on Wage Structures (BSWS) for the survey years 1976 to 2003. BSWS provides the most comprehensive wage data in Japan. It is compiled by the Ministry of Health, Labour and Welfare. It is also the one of the most reliable and long-running wage data in the world.[1] Each survey provides wage information based on approximately 1.5 million employees.[2]

I use standard data collected from private establishments with 10 or more regular employees. BSWS has four classes of educational attainment or its equivalents: [1] graduates from junior high-schools, [2] graduates from high-schools, [3] graduates from higher professional schools and junior colleges, and [4] graduates from universities.[3] In order to correspond to the annual wage, the wage data are calculated by adding the annual special cash earnings to twelve times the contractual monthly cash earnings in June.[4] It is a prime indicator of statistics in this kind of analysis.[5]

3. The wage ratio between college and high-school graduates

This section analyzes the wage ratio between college and high-school graduates, which is considered as a primary indicator of the wage differential between educational statuses. The college/high-school wage ratio can be calculated by dividing the average annual wages of [4] graduates from universities by that of [2] graduates from high-schools in BSWS. Result-1 displays the college/high-school wage ratios of male employees as five tables. There are three ways to analyze these tables in Result-1. Firstly, the cells on the slant line going from the upper-right to the lower-left indicate the same time of the survey. Each table lines up from the upper-left line to the lower-right for every 5 surveys. Secondly, each row shows the

[1] I have conducted a detailed analysis of the US in a manner similar to Japan, using data from 35 consecutive March Current Population Surveys. The result is obscure and inconclusive, mainly because of statistical errors and sample paucity. However, it is highly beneficial to use cohorts instead of time-lines. The wage ratio of white males shows a large discontinuity in the cohort of those born in the 1950s, but not in the time of the late 1980s. Basically, the results are consistent with those of Card and DiNardo (2002).

[2] The definition of an employee in the data is one of the following: [1] Employees hired for an indefinite period. [2] Employees hired for longer than one month. [3] Employees hired for less than one month or by the day and who were hired for 18 days or more in April and May.

[3] Universities include graduate schools.

[4] BSWS is based on pre-tax wage.

[5] The annual wage or one-twelfth of the annual wage is used in Katz and Revenga (1992), Ohtake and Inoki (1997), and Genda (1997). Hourly wage is not common because the working time is hard to measure, especially among highly-educated employees. Brunello and Ariga (1997) use hourly wages to correspond with the literature on earning functions. Although I have used hourly wages instead of the annual wages, the principal results given in this paper are robust.

same age bracket. Information regarding change in wages within the same age bracket can be obtained by checking along a row. Thirdly, the columns show the same birth bracket data. Each column provides cohort wage information. The effect of the birth bracket is the most important aspect to be considered in this paper.

As a matter of course, the cohorts that were between 20 and 24 years old in 1991 became between 25 and 29 years old in 1996, and between 30 and 34 years old in 2001. This cohort represents the birth bracket from 1966 to 1971.[6] Published BSWS records the average wages of five-year age brackets from 20 to 64 years old. Due to the unavailability of one-year age bracket data, I cannot trace one-year period cohort transitions completely. These five lined-up tables in Result-1 are based on five-year period cohorts. Owing to this five-year arrangement, each table displays cohort transition of the wage ratio without overlapping. On the other hand, the oldest group in the each birth bracket moves to the prior birth bracket in the next survey. The transitions of table-to-table display the changes year by year with cohort overlapping.

The educational advancement rate in Japan as well as other developed countries has increased substantially for the 20th century. Schooling during youth virtually decides the share of educational status in each birth bracket. Fixing the birth bracket is almost equivalent to fixing the share of educational status regardless of the age bracket. Therefore, comparing the birth bracket columns enables analysis of the effect of educational status on the wage ratio.

I have marked the comparison to its immediate left column in Result-1 if there are two or more comparable cells from 20-24 to 55-59 years old. With the generational change, the wage ratio increases after a decrease, except for two irregular shifts in about 1953 and 1973. While Result-1 shows the entire male college/high-school wage ratio information recorded in BSWS, the effect of each of the birth brackets is unclear. The blank area of the upper-left and the lower-right in the tables prevent the standard estimation. Therefore, I adopt the following estimation equation considering each bracket variable as a dummy[7]:

$$[\text{Wage Ratio}_{ij}] = \alpha + \beta_i \sum_{i=1}^{n} [\text{age bracket}_i] + \gamma_j \sum_{j=1}^{k} [\text{birth bracket}_j] + \varepsilon_{ij} \tag{1}$$

If a cell contains a value, the corresponding age and birth bracket variables equal one, and the other variables equal zero. On considering that the intercept α as the estimator of a dummy constantly takes the value one, all variables in the equation are composed of dummy variables.[8] Throughout this paper, I use the observations of the age brackets from 20-24 to 55-59 years old and the birth brackets from

[6] As noted in the Result-1, the recorded age in BSWS is counted on June 30. Although it does not exactly correspond with birth year, I have used the term for the sake of simplicity.

[7] The method is used in Deaton and Paxson (1994) and Ohtake and Inoki (1997).

[8] This estimation method is a special case of fixed effects estimation.

1921-1926 to 1978-1983. The number of observations is 219. To avoid multicollinearity, each bracket variable has to take one reference group. I set the youngest and the earliest brackets from among the observations as the reference groups. The reference brackets are the age bracket of 20-24 years old and the birth bracket of 1921-1926. In addition, this dummy variable regression leads to neither the customary standard error nor the t-value [Appendix 1.A].

Result-2A is the estimation result of the college/high-school wage ratio of male employees. While Result-2A indicates that the wage ratio increases almost linearly with aging, the birth bracket effects decrease from the high level of the 1920s.[9] Then, the birth bracket of the 1940s hits the bottom. The wage ratio takes an upward turn after the irregular fluctuation of the 1950s. The figure of birth bracket effects in Result-2A corresponds to the apparent essence of Result-1. This transition of the birth bracket effects is observed not only for male employees but also for female employees. Result-2B is the estimation result of female employees. Because of marital status and child-rearing, the decision to work of female employees has been more complicated than those of male employees. However, Result-2B shows that the female college/high-school wage ratio is similar in transition to that of male. On the whole, the birth bracket effects increase after the decrease. In addition, Result-2A and Result-2B show that the wage ratio for the births in about 1973 has an irregular decrease regardless of gender. These irregular shifts imply the initial effect of new graduates [Appendix 2].

Despite the disturbance caused by the initial effects, on the whole, the birth bracket effects increase after they decrease along with the generational progress. This transition matches the simulation result of mountain-shaped distributions shown in Saito (2005). However, Saito (2005) demonstrates the wage ratio transition based on monotonically increasing advancement rate. The college advancement rate of male has not risen monotonically in Japan. Both the Ministry of Education statistics and the share of employees in BSWS report that the college advancement rate of male for the 1960s birth bracket had decreased slightly. I estimate the share of college graduates among employees in the same manner as (1).

$$[\text{Share of Educational Status}_{ij}] = \alpha + \beta_i \sum_{i=1}^{n} [\text{age bracket}_i] + \gamma_j \sum_{j=1}^{k} [\text{birth bracket}_j] + \varepsilon_{ij} \qquad (2)$$

Result-2C indicates the share of college graduates among male employees. Similarly, Result-2D indicates the share of college graduates among female employees. Although linear estimations do not provide accurate estimates especially in the old cohort, the birth bracket effects show the approximate transition of the share of college graduates.[10] Result-2C shows that the share of college graduates among male

[9] Even though the productivity does not correspond to the spot wage such as the efficiency wage, the age brackets control them.

[10] Some estimated values of the 20-24 age bracket in the old cohort are negative. However, as far as the birth bracket

employees had decreased slightly at the 1960s birth bracket. There seems to be a discrepancy between simulations in Saito (2005) and data in Japan.

The reason for the discrepancy is that the wage ratio between [4] graduates from universities and [2] graduates from high-schools in BSWS does not satisfy one of the assumptions in the simulations. The simulations in Saito (2005) demonstrate the relationship between one dividing line and the wage ratio. On the other hand, the wage ratio taken from BSWS considers the two educational statuses of high-school and university graduates, excluding the other two statuses. Using the same estimation as Result-2C and Result-2D, I have estimated the share of other three educational statuses by gender. The graphs in Result-3 highlight the estimated share of educational statuses among 25-29 and 35-39 years old.[11] Result-3A and Result-3B show that the share of graduates from higher professional schools and junior colleges in male employees increases on the 1960s birth bracket.

An amendment to the School Education Law in 1975 curbed the increase in the college advancement rate. Since the amendment erected legal frameworks for higher professional schools, the share of higher professional school graduates increased rapidly.[12] As a result, the share of college graduates among male employees had decreased slightly on the cohorts. As mentioned in Saito (2005), different rates of decrease of the higher and lower average cause the change in the wage ratio. In the simulated situation, both the numerator and the denominator decrease. However, if the intermediate educational status between the higher and the lower expands, the average capability of higher educational status increases and the average capability of lower educational status decreases. Wherever the dividing line is, the wage ratio between the higher and the lower educational status would increase.

4. The ratio of quartiles and net human capital loss

The simulation of four educational categories require strong assumptions about distribution and parameters. Even though I assume a mountain-shaped distribution, there is no general result without specifying parameters. A simple approach without specifying parameters is integration by weighted average. For example, the average wage of lower educational status is calculated from the weighted average of [1] graduates from junior high-schools and [2] graduates from high-schools. Similarly, the average wage of higher educational status is calculated from the weighted average of [3] graduates from higher professional

effects show an approximate share transition, it need not be taken into consideration. Even if I estimate without observation of the 20-24 age bracket, the main results are robust.

[11] The dummy variable regression necessarily makes the sum of the share estimates to be 1 [Appendix 1.B].

[12] The share of graduates from higher professional schools and junior colleges among male employees was about 6% in the birth brackets of late 1950s. This share had increased to 12% in the birth brackets of late 1960s. According to the Ministry of Education statistics, the share of male junior college graduates had decreased at that time.

schools and junior colleges and [4] graduates from universities. However, the weighted average method requires that the human capital formation effects of schools be negligible.

As discussed in Saito (2005), regardless of human capital formation, the wage ratio of the mountain-shaped distribution increases after it decreases once along with the increase in the educational advancement rate. Therefore, the trends of college/high-school wage ratio do not imply that human capital formation in higher education is small. The wage ratios of averages cannot identify the effects of human capital formation. Therefore, I describe a primitive approach to identify the effects of human capital formation.

Let me present the primitive approach with A, B and C, as in the beginning of Section 2 in Saito (2005). Generally, the wage ratios $\frac{A}{\frac{B+C}{2}}$ and $\frac{\frac{A+B}{2}}{C}$ do not balance even if higher education has no effect. However, if higher education has no effect on wages, there are ratios that always balance regardless of the position of the dividing line. The ratios are simply $\{\frac{A}{B}, \frac{B}{C}, \frac{A}{C}\}$. If higher education has no effect, these ratios always remain constant. Conversely, let me assume that higher education forms human capital. Then, as the phase advances, B proceeds to higher education. In this assumption, the simple result is that the raised denominator decreases $\frac{A}{B}$, and the raised numerator increases $\frac{B}{C}$. As shown in Figure 1, I can approach the effects of higher education without specifying the distribution.

The empirical approach is similar to that of the wage ratio between averages. Without separation on the basis of educational statuses, I have tabulated quartile wages by gender. Then, the higher quartile is divided by the lower quartile based on cells in a similar manner to Result-1. If higher education has effects on wages, the ratio changes when the educational advancement rate exceeds 25%, 50%, and 75%. On the contrary, if higher education has no effect on wages, the ratio remains constant in the same situation.

Published BSWS does not record quartiles that are convertible to annual wages. However, BSWS records quartiles of scheduled cash earnings in June. There is a strong correlation between annual wages and scheduled cash earnings.[13] Therefore, the quartile analyses use scheduled cash earnings as a proxy for annual wages. In addition, the share of each educational status among employees is required to analyze the effect of higher education. I use the estimated values in Result-3 as an indicator of share of educational statuses among employees.

Result-4 contains four results of the wage ratio of quartiles. With reference to Result-3B and Result-3D, I draw dotted lines at the points where educational partitions come across 25%, 50%, and 75% in the birth bracket graphs in Result-4. Result-4A indicates the ratio between the median and first quartile of male employees. The upper side of the birth bracket graph represents the widening of the gap. The quartile

[13] In each survey, the correlation coefficient between annual wages and scheduled cash earnings is approximately 0.995, based on the age brackets from 20-24 to 55-59 years old and by gender and educational status.

ratio decreases on the birth bracket that the high-school graduate line cut across at 50%. Afterwards, the ratio increases on the birth bracket that the high-school graduate line cut across at 75%. This transition implies that the high-school effect is smaller than that of On-the-Job Training (OJT). In other words, high-schools have net negative effects on the formation of human capital, which is reflected in the wages. It is a *net human capital loss*. High-schools repress the human capital that is supposed to be formed in the labor market.

In Result-4A, the quartile ratios change just before the educational partitions pass over the quartiles. This is consistent if schools repress human capital formation. I present a simple univariate example in Figure 2. There are nine people numbered 1 through 9 sequentially. Each of them begins with the number of capabilities equal to his/her assigned number. If the dividing line has a negative effect (-2), when does the wage ratio between the medium and the third highest change? When the second most-capable person proceeds to higher education, the fractional number changes from 1.4 to 1.2. This is not the time for the third most-capable person. The point is not phase 3 but phase 2 in Figure 2. If higher education has net negative effects on wages, the prior change is consistent with the inference.

Result-4B indicates the ratio between the third quartile and the median of male employees. On one hand, the effects of the decrease in the 1930s birth bracket might appear to precede the net human capital loss in colleges as mentioned above. On the other hand, when the higher professional schools and junior colleges partition cut across the median line in the 1970s birth bracket, the ratio decreased. Among Result-4, this is the only observation that appears to indicate that higher education has positive effects on wages. However, Result-4A seems to indicate that the higher professional schools and junior colleges have negative effects. In general, there is a trend the male wage gap has been diminishing from the 1960s birth bracket. These diminishing gaps would not be attributed to the effects of higher professional schools and junior colleges.

Result-4C indicates the ratio between the median and first quartile of female employees. Similarly, Result-4D indicates the ratio between the third quartile and the median of female employees. When higher quartiles proceed to higher education, the ratios decrease. When lower quartiles proceed to higher education, the ratios increase. According to the information recorded in the graphs, it is possible to consider that the human capital formation effect of higher education is inferior to that of the labor market.

Indeed, the quartile ratios approach according to published BSWS is imprecise to identify educational effects on wages. While the share of college graduates among male employees has exceeded 25% since the first baby boom (births from approximately 1947 to 1950), this share did not exceed 50% in 2003 (births in the 1980s). Published BSWS does not record the percentiles between these quartiles. In addition, the vertical axes on the birth bracket effects in Result-4 indicate that the change of ratios between quartiles are considerably smaller than those of Result-2A and Result-2B. I have no alternative but to focus on minor variations. Furthermore, there may be variations that are irrelevant to effects of higher education, such as the trend of the decreasing wage gap of male employees. It is possible that other factors are responsible for the changes that appear as net human capital losses. Meanwhile, if there are net human capital losses, the quartile ratio may change before the educational partitions pass over the quartiles. The time is dependent on the loss, distribution shape, and the position of the dividing line. If a bivariate distribution of labor productivity and academic ability is introduced, the measurement of the effects would become considerably complicated.

Therefore, further analysis of effects of higher education is extremely difficult because of insufficient data and difficulties in estimation. However, a majority of the quartile ratios imply the net human capital losses in higher education. These losses are, at most, a few percentage points. Although it would still be unwise to conclude that higher education has net negative effects on wages in Japan, it is safe to state that higher education does not have significant positive effects. The net effects range from slightly negative to insignificant. In either case, if the higher education effects are minor, the weighted average procedure is justified as an acceptable approximation.

5. The wage ratio on weighted averages of employees

This section provides analyses of the wage ratio on weighted averages of employees. The average wage of lower educational status is calculated from the weighted average of [1] graduates from junior high-schools and [2] graduates from high-schools. In the same manner, the average wage of higher educational status is calculated from the weighted average of [3] graduates from higher professional schools and junior colleges and [4] graduates from universities. The estimation method for the wage ratio is the same as (1).

Result-5A shows the estimation result of the wage ratio of male employees between the higher educational status and lower one. The transition of the birth brackets in Result-5A is smoother than that of Result-2A. The smooth transition is partly caused by time spreading of new graduates. The time spreading alleviates the initial effect of new graduates as shown in Appendix 2. In addition, Result-5A takes account of the entire distribution that was ignored in Result-2A. Taking account of the entire distribution stabilizes the estimates. Similarly, Result-5B shows the estimation result of the wage ratio of female employees between the higher educational status and lower one. The transition of birth bracket effects in Result-5B is smoother and more stable than that of Result-2B. As the share of graduates from junior colleges has been high in female employees, the transition of the birth bracket effects heavily depends on the manner in which graduates from junior colleges are handled.

Based on the same method as (2), Result-5C reports the result of the dummy variable regression for the share of the higher educational status among male employees. Result-5C indicates that the share of the higher educational status has risen monotonically over its cohorts. This monotonic increase of the share of the higher educational status is different from that of Result-2C. From the birth brackets from 1955 to 1970 in Result-5C, the stable share of higher educational status corresponds to the stable wage ratio in Result-5A. The wage ratio increased significantly with the rapid increase in the share of higher educational status from the 1970s birth brackets. Result-5D reports the share of the higher educational status among female employees. The share of the female higher educational status has risen monotonically, in a manner similar to that of the males. The wage ratio transition in Result-5B is caused by the monotonic increase in the share of the higher educational status in Result-5D.

Katz and Revenga (1992), Ohtake and Inoki (1997) and Genda (1997) analyze wage differential and the effects of birth brackets (generational effects) using BSWS, as the wage ratio between male college graduates and male high-school graduates. However, these transitions follow the entire distribution rule of the four statuses. Furthermore, the trick of wage ratio is the main cause for the transition.

Meanwhile, these results imply the answer to the question, "Is there substantial skill-biased technological progress in Japan?" It is unlikely that there is substantial skill-biased technological progress in Japan. The coefficient of determination in Result-5A is 0.9969. Only the birth bracket effects of columns and the age bracket effects of rows can explain almost all the wage ratio variations of male employees. If there was indeed a skill-biased technological progress in Japan, its influence would emerge on the slanting line. The column and row effects would not explain the influence on the slanting line. However, as there is little room for explanation by slanting lines, it appears that there is no substantial skill-biased technological progress in Japan.

6. Conclusion

This paper has presented empirical wage analyses in Japan using data arranged by cohorts. The effects of higher education on wages range from slightly negative to insignificant from the viewpoint of the entire society. Besides, it is unlikely that there has been any substantial skill-biased technological progress in Japan. The main reason for the transition in wage differentials in Japan is a statistical trick resulting from the distributional divide.

In the process of analyses, I delivered a judgement to the question, "Which is more practical in Japan, the human capital theory or the signaling theory?" Since the empirical results imply net human capital losses, it is safe to state that the signaling theory is more practical in Japan. In terms of wages, the effects of human capital formation due to higher education are either reduced or the same as those of

OJT. In general, at least in Japan, higher education is not suitable for fostering of employees.

Even if higher education does not contribute to human capital formation, a separating equilibrium in the signaling model may exist. From the viewpoint of household sectors, there may be an incentive to send out signals even if they pay tuition fee and the opportunity cost. There exists no problem in adding the net human capital loss to other costs. For example, in order to differentiate oneself from another whose capability is 1, a person with a capability of 2 might proceed to higher education even if his/her capability decreases to 1.9.

All analyses in this paper are based on labor productivity from the viewpoint of macroeconomic distribution. Net human capital loss or insignificant effect of higher education depends on the ratio of quartiles from the viewpoint of macroeconomic distribution. I quote from the *introduction to the second edition* of Becker (1964): "*Even if schooling also works in this way* [implying the signaling theory]*, the significance of private rates of return to education is not affected at all.*" Indeed, this statement is true even if there are net human capital losses. The concept of private returns remain unaffected. Therefore, the results mentioned in this paper would be different from those of causal effect analyses using instrumental variables.[14] However, although the causal effect would be profitable from the viewpoint of households, it remains an illusion, at least in Japan, from the viewpoint of the entire society.

The premise that schooling is a kind of investment for households is common to both the human capital theory and the signaling theory. In reality, there exist other investments that directly contribute to utility or welfare without affecting the economic or wage growth. Education and schools possess aspects of culture that entertains students and the public. However, I believe that if society turns away from accepting the facts, it is the role of the social scientists to dispassionately demonstrate the facts and solutions through research.

[14] It is conceivable that two people with the same productivity earn different wages because of a difference in their educational statuses. The reasons might arise from asymmetric information, bargaining power and career requirement in firms. Even though schools do not increase productivity, the causal effects of education may be positive.

Appendix 1: Dummy Variable Regression

1.A. The dummy variable regression leads to neither the customary standard error nor the t-value.

This explanation should be read along with the tables on the next page. This is an example that does not lead to a loss of generality.

I. Consider a situation where I estimate the effects of rows $(\alpha \cdots \gamma)$ and columns $(\delta \cdots \theta)$ from [$\langle 1 \rangle$ Original Table] The shaded cells are unobservable and missing.

II. Each independent variable is set as a dummy variable in [$\langle 2 \rangle$ Estimation Input].
I have to set reference groups, one for each row and column to avoid multicollinearity.
(In this example, the bases are α and δ.)

III. Including the intercept, the matrix of independent variables, is written in [$\langle 3 \rangle$ **X**].
The transpose of **X** is [$\langle 4 \rangle$ **X'**].

IV. The product of **X'** and **X** is [$\langle 5 \rangle$ **X'X**]

- Consider ϵ and ζ, these have the same pattern of observation and missing. Due to the similarity in the pattern, they have the same inner products with any independent variable vector ϕ; $\phi \cdot \epsilon = \phi \cdot \zeta$ $(\phi'\epsilon = \phi'\zeta)$ and $\epsilon \cdot \epsilon = \zeta \cdot \zeta$ $(\epsilon'\epsilon = \zeta'\zeta)$. It is clear that $\epsilon \cdot \phi = \zeta \cdot \phi$ $(\epsilon'\phi = \zeta'\phi)$ because of the symmetry of inner products. Then, the cofactors of (4,4) and (5,5) in [$\langle 5 \rangle$ **X'X**] are necessarily equal.

V. Rounding off to two decimal places, the inverse matrix of [$\langle 5 \rangle$ **X'X**] is written in $\left[\langle 6 \rangle\, (\mathbf{X'X})^{-1}\right]$.

- Since both the determinant and the cofactor are the same, the corresponding element in the inverse matrix is necessarily equal.

VI. The product of $\left[\langle 6 \rangle\, (\mathbf{X'X})^{-1}\right]$ multiplied by s^2 is the variance matrix of coefficients.

- The standard error of ϵ and ζ are equal, and fixed proportions of standard errors are predetermined. Although the predicted values are the same, the estimated coefficients depend on the reference groups. Therefore, the dummy variable regression leads to neither the customary standard error nor the t-value.

	δ	ϵ	ζ	η	θ
α	10	19.8	30	39.8	49.7
β	11.2	21	30.9	40.2	51
γ	12.1	22.2	32.1	42	52

[⟨1⟩ Original Table]

	β	γ	ϵ	ζ	η	θ
11.2	1	0	0	0	0	0
12.1	0	1	0	0	0	0
19.8	0	0	1	0	0	0
21	1	0	1	0	0	0
22.2	0	1	1	0	0	0
30	0	0	0	1	0	0
30.9	1	0	0	1	0	0
32.1	0	1	0	1	0	0
39.8	0	0	0	0	1	0
40.2	1	0	0	0	1	0
49.7	0	0	0	0	0	1
10	0	0	0	0	0	0
42	0	1	0	0	1	0
51	1	0	0	0	0	1
52	0	1	0	0	0	1

[⟨2⟩ Estimation Input]

C	β	γ	ϵ	ζ	η	θ
1	1	0	0	0	0	0
1	0	1	0	0	0	0
1	0	0	1	0	0	0
1	1	0	1	0	0	0
1	0	1	1	0	0	0
1	0	0	0	1	0	0
1	1	0	0	1	0	0
1	0	1	0	1	0	0
1	0	0	0	0	1	0
1	1	0	0	0	1	0
1	0	0	0	0	0	1

[⟨3⟩ $\mathbf{X}_{(11\times 7)}$]

C′	1	1	1	1	1	1	1	1	1	1	1
β'	1	0	0	1	0	0	1	0	0	1	0
γ'	0	1	0	0	1	0	0	1	0	0	0
ϵ'	0	0	1	1	1	0	0	0	0	0	0
ζ'	0	0	0	0	0	1	1	1	0	0	0
η'	0	0	0	0	0	0	0	0	1	1	0
θ'	0	0	0	0	0	0	0	0	0	0	1

[⟨4⟩ $\mathbf{X'}_{(7\times 11)}$]

11	4	3	3	3	2	1
4	4	0	1	1	1	0
3	0	3	1	1	0	0
3	1	1	3	0	0	0
3	1	1	0	3	0	0
2	1	0	0	0	2	0
1	0	0	0	0	0	1

[⟨5⟩ $\mathbf{X'X}_{(7\times 7)}$]

1.06	-0.51	-0.60	-0.69	-0.69	-0.80	-1.06
-0.51	0.63	0.40	0.17	0.17	0.20	0.51
-0.60	0.40	0.80	0.20	0.20	0.40	0.60
-0.69	0.17	0.20	0.90	0.56	0.60	0.69
-0.69	0.17	0.20	056	0.90	0.60	0.69
-0.80	0.20	0.40	0.60	0.60	1.20	0.80
-1.06	0.51	0.60	0.69	0.69	0.80	2.06

[⟨6⟩ $(\mathbf{X'X})^{-1}{}_{(7\times 7)}$]

Sum of Squared Residuals	0.2280
Unbiased Estimate of Variance (s^2)	0.0570
Standard Error of Regression (s)	0.2387

[⟨7⟩ Estimated Result of Regression]

	Estimated Coefficient	Standard Error	t-value
C	10.18	0.2455	41.47
β	0.88	0.1893	4.65
γ	2.06	0.2135	9.65
ϵ	9.84	0.2259	43.56
ζ	19.84	0.2259	87.83
η	29.38	0.2615	112.34
θ	39.52	0.3424	115.41

[⟨8⟩ Estimated Result of Coefficient]

1.B. The sum of share estimates always equals 1.

The sum of employees' share by educational status always equals 1 in Result-3. Consider a dummy variable regression in this situation:

$$\mathbf{Y_A} = \begin{pmatrix} Y_{A1} \\ Y_{A2} \\ \vdots \\ Y_{An} \end{pmatrix}, \quad \mathbf{Y_B} = \begin{pmatrix} Y_{B1} \\ Y_{B2} \\ \vdots \\ Y_{Bn} \end{pmatrix}, \quad \mathbf{Y_C} = \begin{pmatrix} Y_{C1} \\ Y_{C2} \\ \vdots \\ Y_{Cn} \end{pmatrix}, \quad \mathbf{Y_D} = \begin{pmatrix} Y_{D1} \\ Y_{D2} \\ \vdots \\ Y_{Dn} \end{pmatrix}$$

By setting $Y_{Ai} + Y_{Bi} + Y_{Ci} + Y_{Di} = 1$, the vector sum is

$$\mathbf{Y_A} + \mathbf{Y_B} + \mathbf{Y_C} + \mathbf{Y_D} = \mathbf{Y_T} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \mathbf{1}$$

Then, all the elements equal 1. Setting the independent variables matrix as $\mathbf{X}$, the predicted value $\mathbf{b_A}$ for $\mathbf{Y_A}$ is written in the following manner:

$$\mathbf{b_A} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'Y_A} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'}(\mathbf{1} - \mathbf{Y_B} - \mathbf{Y_C} - \mathbf{Y_D}) = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'1} - \mathbf{b_B} - \mathbf{b_C} - \mathbf{b_D}$$

$(\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'1}$ is the least squares estimate for $\mathbf{1}$. The estimates for a constant is 0 except for the intercept. Since all the dependent variables are 1, the estimated coefficient of the intercept is 1.

$$\mathbf{b_A} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'1} - \mathbf{b_B} - \mathbf{b_C} - \mathbf{b_D} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \mathbf{b_B} - \mathbf{b_C} - \mathbf{b_D}$$

Then, $\mathbf{b_A} + \mathbf{b_B} + \mathbf{b_C} + \mathbf{b_D} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

In this situation, the sum of the estimated coefficients of the intercept is necessarily 1. The sums of all other coefficients are necessarily 0. Therefore, Result-3 is obtained without any adjustment.

Appendix 2: The Initial Effect of New Graduates in Japan

The wage ratio for the births in about 1973 has an irregular decrease regardless of gender as shown in Result-1, Result-2A, and Result-2B. These irregular shifts seem to reflect the business climate at the time of the graduation. The men who were born around 1973 got jobs during the bubble (extreme prosperity) period if they were high-school graduate employees. If they were college graduate employees, they got jobs during the post-bubble (depression) period. Result-6 draws a comparison between the birth bracket effects in Result-2A and job-offers-to-seekers ratio in the following manner:

1. Result-6A reports the job-offers-to-seekers ratios in order to describe the business climate existing at the time of the graduation.
2. In order to correspond to the overlapping birth brackets, Result-6B takes moving averages of five years' job-offers-to-seekers ratio.
3. In order to correspond to the time difference between high-school and college graduates, Result-6C displays differences from four years before.
4. Result-6D draws comparisons with the birth bracket effects in Result-2A.

Result-6D indicates that the business climate at the time of the graduation has an impact on the birth bracket effects. The college/high-school wage ratio of male employees also shows irregular shifts in the birth brackets of 1950s. These shifts appear to reflect the turbulent fluctuations of the 1970s, such as the two oil crises. Despite the passage of approximately 30 years, the initial effects of fluctuations of the 1970s strongly exist in the wages even to this day. The discovery that the business climate at the time of the graduation has a persistent effect on wages in Japan is consistent with Ohtake and Inoki [1997], Genda and Kurosawa [2001].

References

Becker, Gary S. (1964) *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education,* 1st ed. (Chicago: University of Chicago Press)

Brunello, Giorgio and Kenn Ariga (1997) 'Earnings and seniority in Japan: A re-appraisal of the existing evidence and a comparison with the UK.' *Labour Economics* 4(1), 47–69

Card, David (2001) 'Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems.' *Econometrica* 69(5), 1127–60

Card, David and John E. DiNardo (2002) 'Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles.' *Journal of Labor Economics* 20(4), 733–783

Deaton, Angus and Christina Paxson (1994) 'Intertemporal Choice and Inequality.' *Journal of Political Economy* 102(3), 437–67

Genda, Yuji (1997) 'Chansu ha ichido (in Japanese)[only one chance].' *The monthly journal of the Japan Institute of Labour* 449, 2–12

Genda, Yuji and Masako Kurosawa (2001) 'Transition from School to Work in Japan.' *Journal of the Japanese and International Economies* 15(4), 465–488

Griliches, Zvi (1977) 'Estimating the Returns to Schooling: Some Econometric Problems.' *Econometrica* 45(1), 1–22

Katz, Lawrence F. and Ana L. Revenga (1992) 'Changes in the Structure of Wages: The United States vs. Japan.' *Journal of the Japanese and International Economies* 3(4), 522–553

Ohtake, Fumio and Takenori Inoki (1997) 'Rodosijoni okeru sedaikoka (in Japanese) [generation effects in labor market].' In *Gendai Makurokeizai bunseki [contemporary macroeconomic analyses],* ed. Fukuda Shinichi Asako Kazumi and Yohino Naoyuki (University of Tokyo Press) chapter 10, pp. 297–320

Saito, Keiji (2005) 'A Fallacy of Wage Differentials: Wage Ratio in Distribution.' Unpublished manuscript (The university of Tokyo, Economics)

Spence, A Michael (1973) 'Job Market Signaling.' *The Quarterly Journal of Economics* 87(3), 355–74

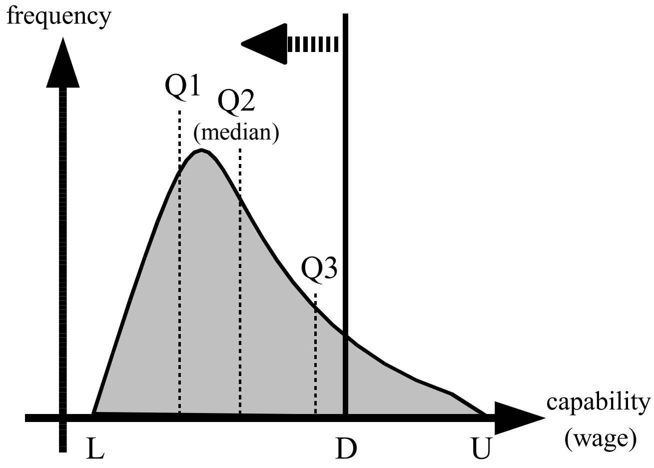


Figure 1 The Ratio of Quartiles

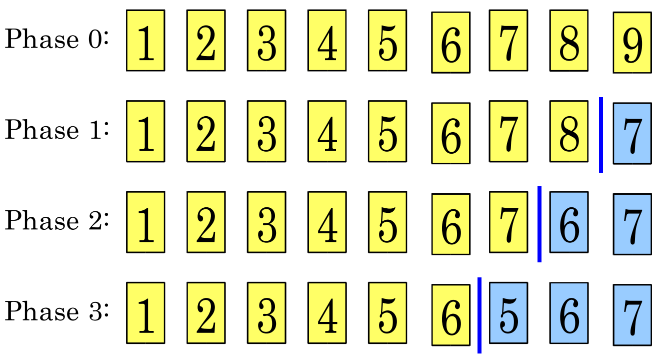


Figure 2 The Time of the Quartile Ratio Change on the Net Loss

average annual wage of male college graduates
———————————————
average annual wage of male high-school graduates

Compared with each immediate left birth bracket, two or more all comparable ratios from [20-24] to [55-59]

increase : ↗ decrease : ↘

are almost the same (the differences are less than 2%) : → fluctuate with the age brackets : ?

final numbers: 1 and 6			from the upper-left corner				1976	1981	1986	1991	1996	2001	survey		
birth from July 1	1906	1911	1916	1921	1926	1931	1936	1941	1946	1951	1956	1961	1966	1971	1976
to June 30	1911	1916	1921	1926	1931	1936	1941	1946	1951	1956	1961	1966	1971	1976	1981
20-24 years old								↘		0.920	0.922	0.967	0.989	0.963	0.998
25-29 years old							↘		1.064	1.025	1.103	1.110	1.125	1.111	
30-34 years old						↘		1.152	1.165	1.156	1.206	1.229	1.258		
35-39 years old					↘		1.299	1.217	1.267	1.227	1.304	1.334		↘	
40-44 years old						1.422	1.388	1.316	1.329	1.298	1.379		↗		
45-49 years old					1.590	1.514	1.493	1.359	1.399	1.374		↗			
50-54 years old				1.676	1.614	1.584	1.557	1.452	1.458		↗				
55-59 years old			1.631	1.707	1.704	1.610	1.566	1.488		↘					
60-64 years old			1.653	1.871	1.810	1.603	1.752		↗						
over 65 years old		1.722	1.988	2.031	2.016	2.156									

final numbers: 2 and 7			from the upper-left corner				1977	1982	1987	1992	1997	2002	survey		
birth from July 1	1907	1912	1917	1922	1927	1932	1937	1942	1947	1952	1957	1962	1967	1972	1977
to June 30	1912	1917	1922	1927	1932	1937	1942	1947	1952	1957	1962	1967	1972	1977	1982
20-24 years old								?		0.909	0.940	0.981	0.991	0.956	1.005
25-29 years old							↘		1.071	1.037	1.111	1.118	1.123	1.140	
30-34 years old						↘		1.138	1.165	1.162	1.207	1.239	1.267		
35-39 years old					↘		1.284	1.212	1.271	1.241	1.299	1.348		?	
40-44 years old						1.397	1.365	1.313	1.355	1.302	1.409		↗		
45-49 years old					1.586	1.511	1.467	1.377	1.393	1.380		↗			
50-54 years old				1.665	1.663	1.581	1.525	1.409	1.476		↗				
55-59 years old			1.707	1.751	1.726	1.568	1.508	1.522		↘					
60-64 years old			1.841	1.788	1.737	1.614	1.849		↗						
over 65 years old		1.813	2.254	2.061	2.041	2.220									

final numbers: 3 and 8			from the upper-left corner				1978	1983	1988	1993	1998	2003	survey		
birth from July 1	1908	1913	1918	1923	1928	1933	1938	1943	1948	1953	1958	1963	1968	1973	1978
to June 30	1913	1918	1923	1928	1933	1938	1943	1948	1953	1958	1963	1968	1973	1978	1983
20-24 years old								↘		0.907	0.962	0.977	0.996	0.979	1.025
25-29 years old							↘		1.051	1.043	1.112	1.121	1.126	1.154	
30-34 years old						↘		1.150	1.162	1.165	1.215	1.251	1.247		
35-39 years old					↘		1.270	1.226	1.253	1.249	1.325	1.368		?	
40-44 years old						1.410	1.346	1.302	1.346	1.351	1.393		→		
45-49 years old					1.529	1.492	1.425	1.395	1.403	1.391		↗			
50-54 years old				1.686	1.627	1.561	1.501	1.466	1.429		↗				
55-59 years old			1.742	1.722	1.664	1.633	1.563	1.479		→					
60-64 years old		1.662	1.799	1.817	1.711	1.609	1.755		?						
over 65 years old	1.731	1.990	2.151	2.178	1.918	1.968									

final numbers: 4 and 9			from the upper-left corner				1979	1984	1989	1994	1999	survey			
birth from July 1	1909	1914	1919	1924	1929	1934	1939	1944	1949	1954	1959	1964	1969	1974	1979
to June 30	1914	1919	1924	1929	1934	1939	1944	1949	1954	1959	1964	1969	1974	1979	1984
20-24 years old								?		0.920	0.960	0.999	0.985	0.986	
25-29 years old							↘		1.038	1.052	1.118	1.115	1.121		
30-34 years old						↘		1.151	1.145	1.184	1.211	1.252			
35-39 years old					↘		1.243	1.245	1.256	1.271	1.320		→		
40-44 years old						1.390	1.319	1.329	1.335	1.340		?			
45-49 years old					1.510	1.477	1.417	1.405	1.396		↗				
50-54 years old				1.673	1.584	1.579	1.485	1.464		↗					
55-59 years old			1.699	1.749	1.708	1.625	1.541		→						
60-64 years old		1.623	1.761	1.989	1.686	1.720									
over 65 years old	1.776	1.883	2.112	1.984	2.038										

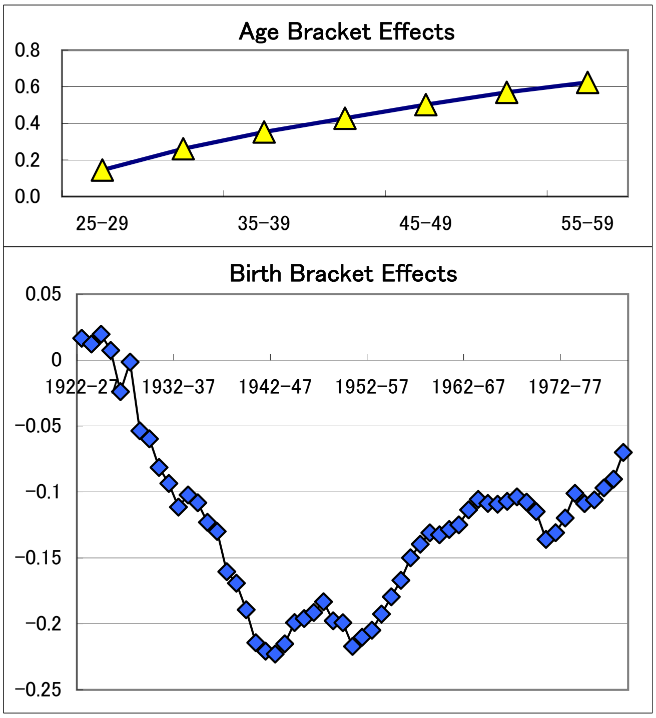
final numbers: 5 and 0			from the upper-left corner				1980	1985	1990	1995	2000	survey			
birth from July 1	1910	1915	1920	1925	1930	1935	1940	1945	1950	1955	1960	1965	1970	1975	1980
to June 30	1915	1920	1925	1930	1935	1940	1945	1950	1955	1960	1965	1970	1975	1980	1985
20-24 years old								↗		0.921	0.971	0.990	0.970	0.989	
25-29 years old							↘		1.024	1.075	1.114	1.123	1.094		
30-34 years old						↘		1.148	1.141	1.190	1.219	1.252			
35-39 years old					↘		1.227	1.255	1.237	1.282	1.322		↘		
40-44 years old						1.390	1.300	1.336	1.322	1.360		↗			
45-49 years old					1.515	1.483	1.408	1.408	1.357		↗				
50-54 years old				1.649	1.565	1.561	1.459	1.463		↗					
55-59 years old			1.667	1.749	1.656	1.578	1.487		↘						
60-64 years old		1.659	1.783	1.837	1.731	1.688									
over 65 years old	1.811	1.992	2.003	2.116	2.097										

* The shaded cells [over 65 years old] do not correspond to the column birth brackets.

Result-1

average annual wage of male college graduates / average annual wage of male high-school graduates

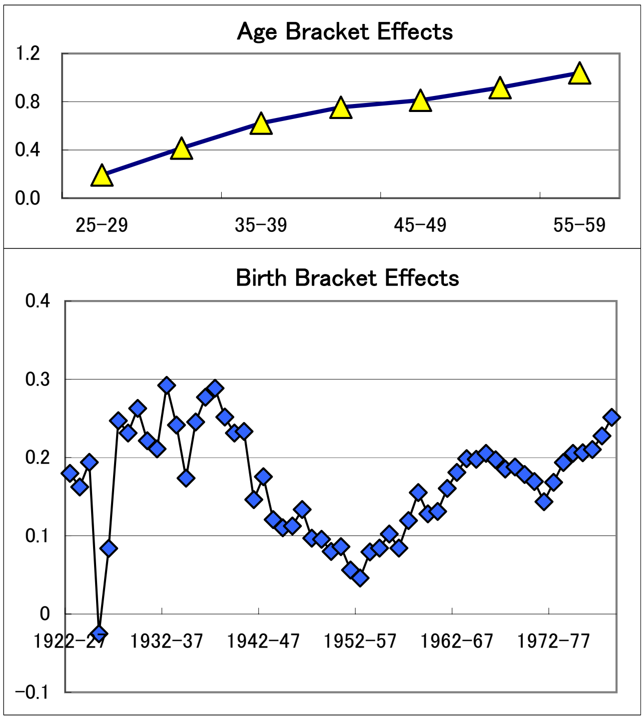
R-squared	0.9958		
Adjusted R-squared	0.9940	Constant	1.0952



Result-2A

average annual wage of female college graduates / average annual wage of female high-school graduates

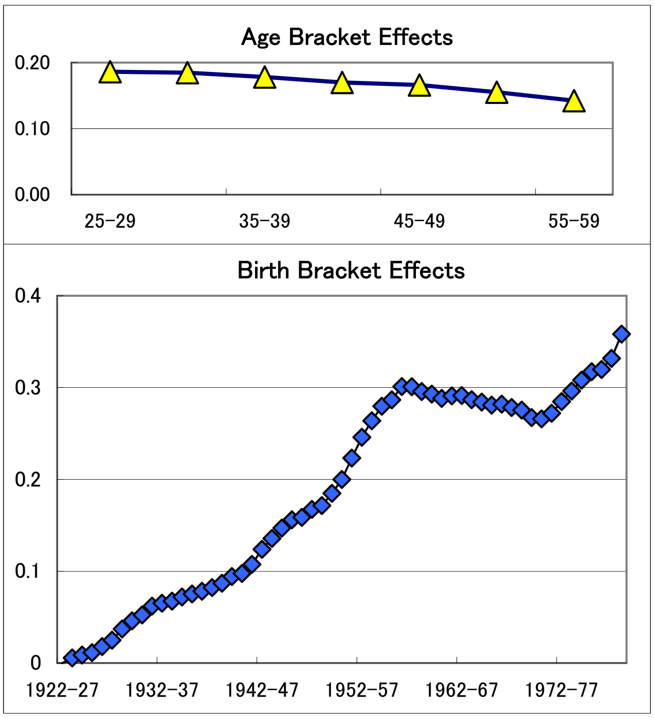
R-squared	0.9810		
Adjusted R-squared	0.9731	Constant	0.9191



Result-2B

total number of male employees in college graduate / total number of male employees

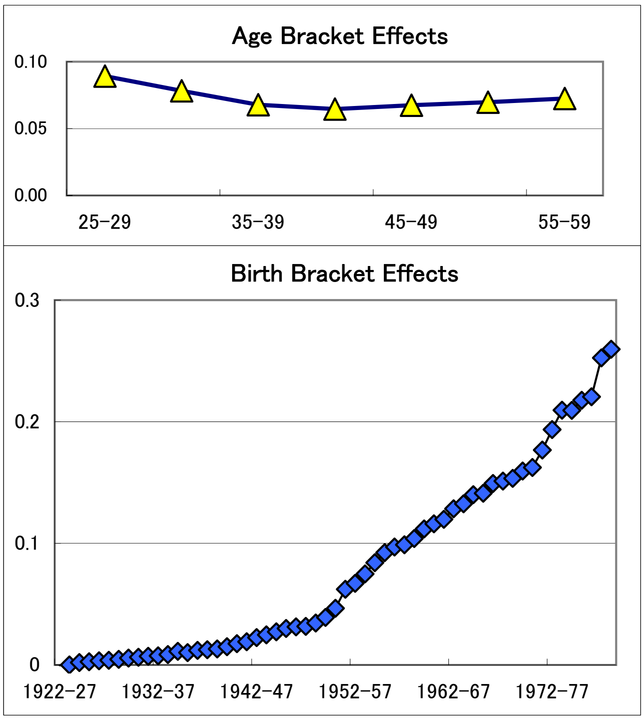
R-squared	0.9982		
Adjusted R-squared	0.9975	Constant	-0.0765



Result-2C

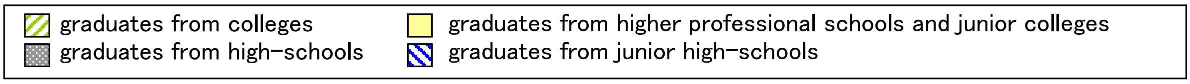
total number of female employees in college graduate / total number of female employees

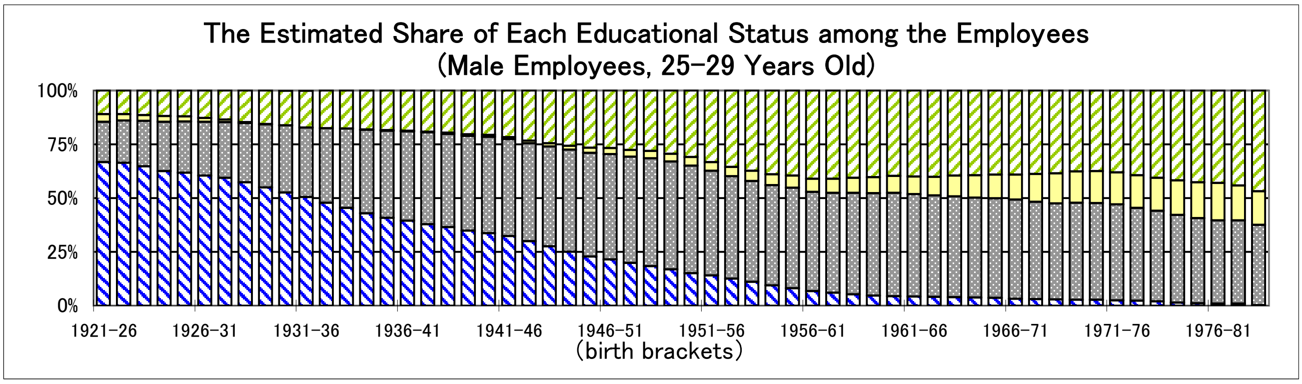
R-squared	0.9806		
Adjusted R-squared	0.9725	Constant	-0.0643



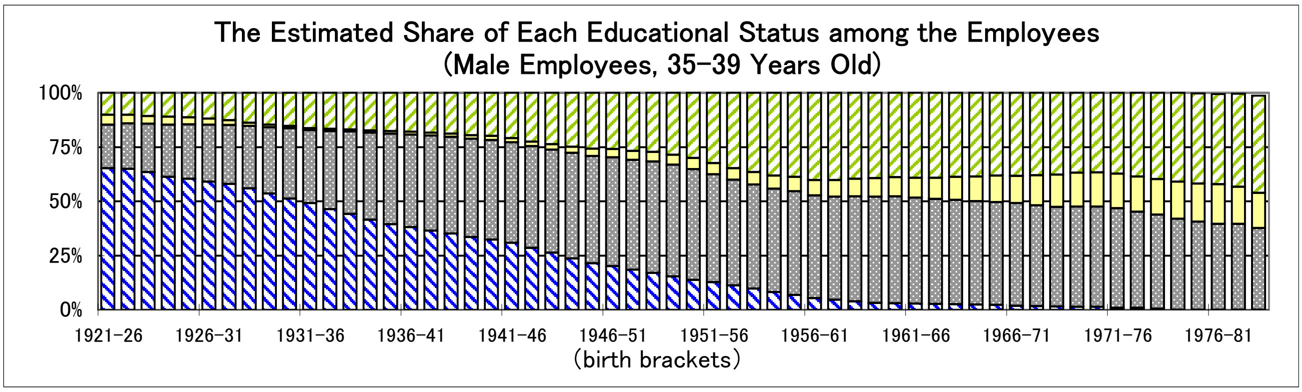
Result-2D

Result-2

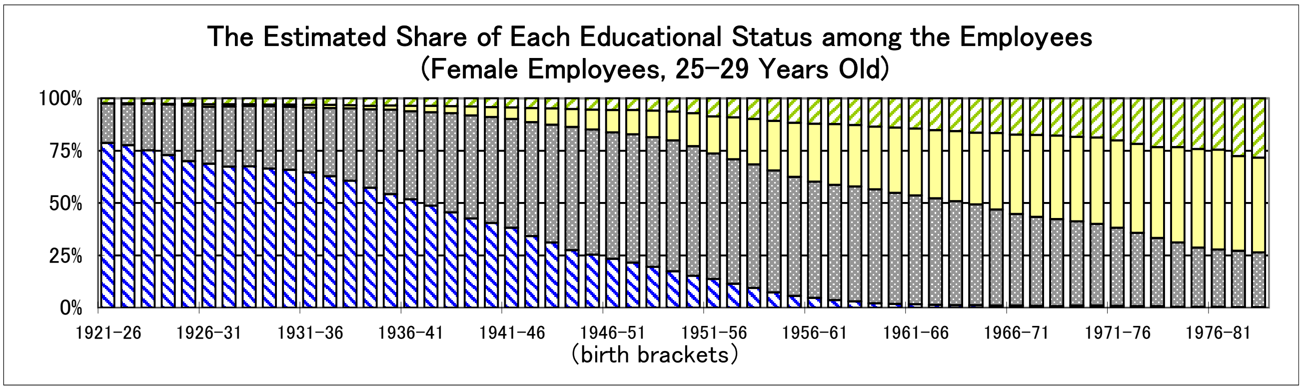




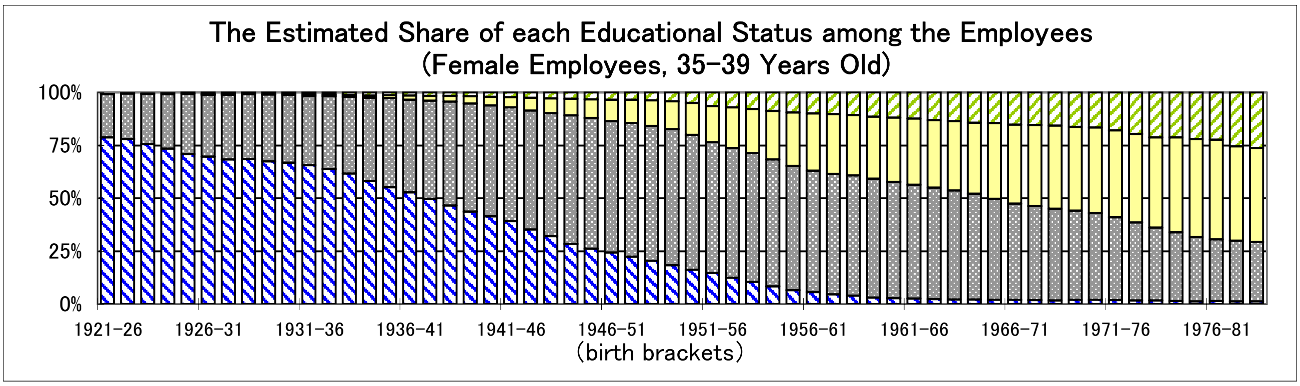
Result-3A



Result-3B



Result-3C



Result-3D

Result-3

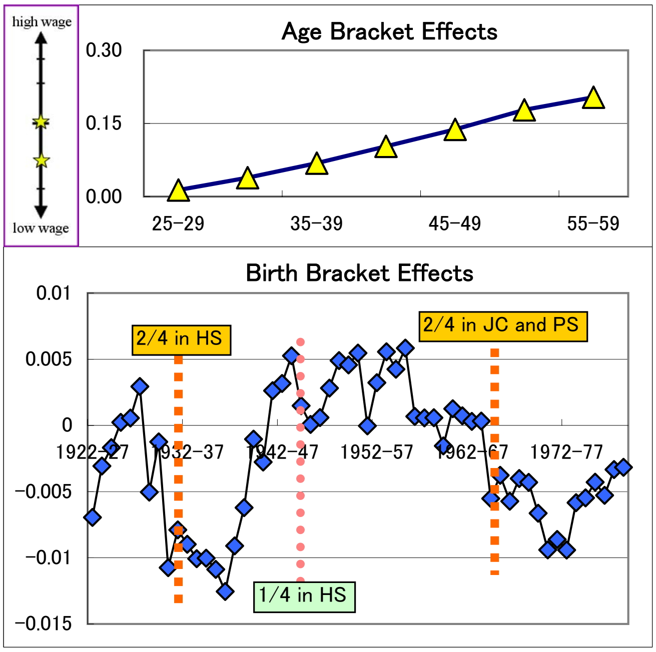
the median scheduled cash earnings of males

the first quartile scheduled cash earnings of males

R-squared 0.9818

Adjusted R-squared 0.9742

Constant 1.1206



Result-4A

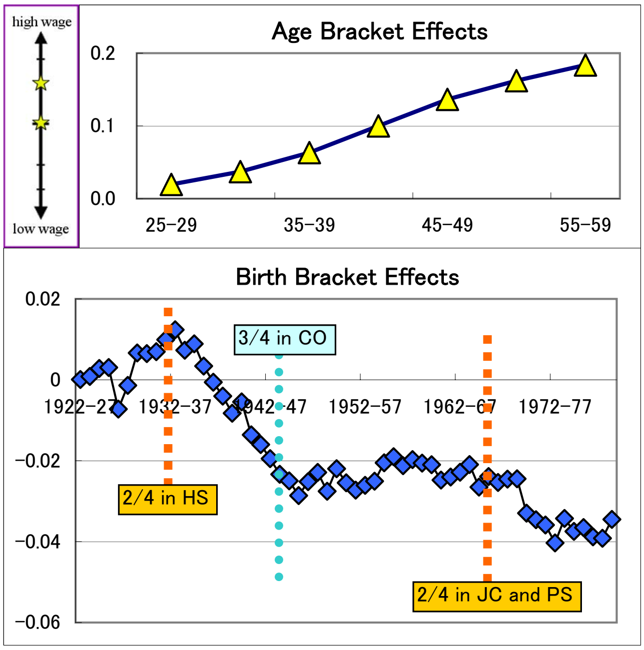
the third quartile scheduled cash earnings of males

the median scheduled cash earnings of males

R-squared 0.9922

Adjusted R-squared 0.9889

Constant 1.1528



Result-4B

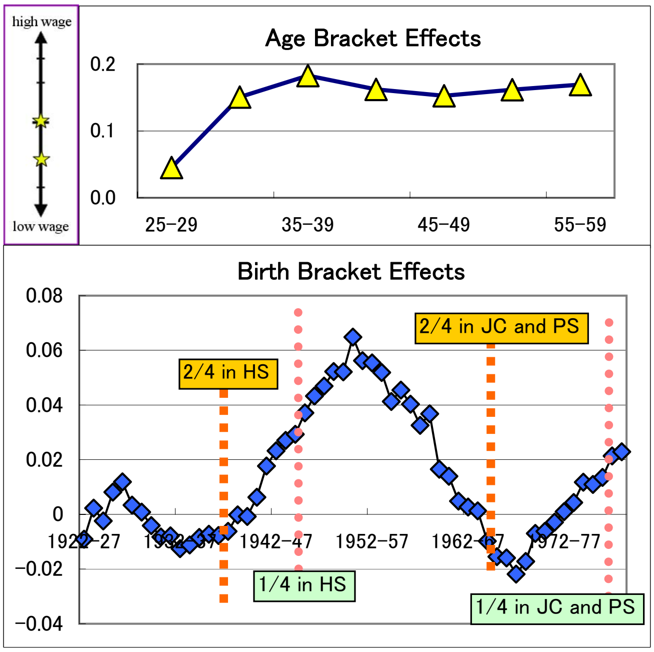
the median scheduled cash earnings of females

the first quartile scheduled cash earnings of females

R-squared 0.9515

Adjusted R-squared 0.9314

Constant 1.0977



Result-4C

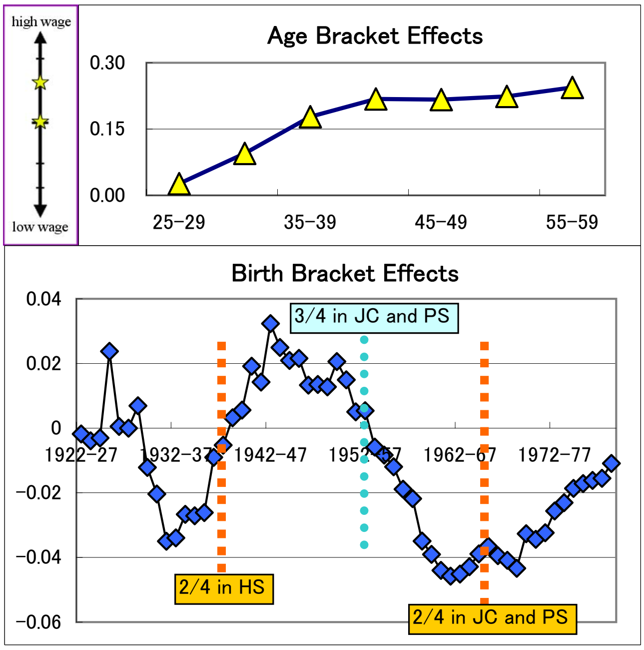
the third quartile scheduled cash earnings of females

the median scheduled cash earnings of females

R-squared 0.9622

Adjusted R-squared 0.9465

Constant 1.1421



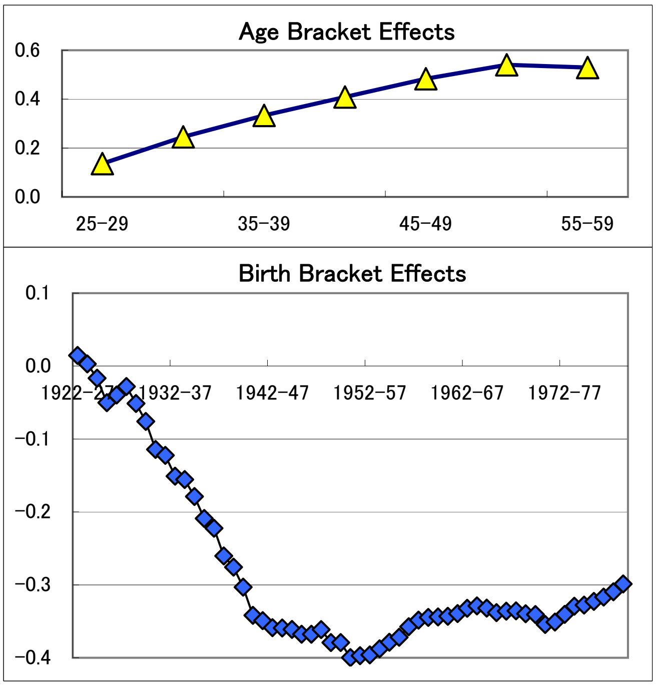
Result-4D

acronyms: HS stands for high school. JC and PS stands for junior college and professional school, respectively. CO stands for college.
If the (upper) numerator quartile gets into the next educational status, a label is attached at the head of the dotted line.
If the (lower) denominator quartile gets into the next educational status, a label is attached at the bottom of the dotted line.

Result-4

weighted average annual wage of male employees in the higher educational status / weighted average annual wage of male employees in the lower educational status

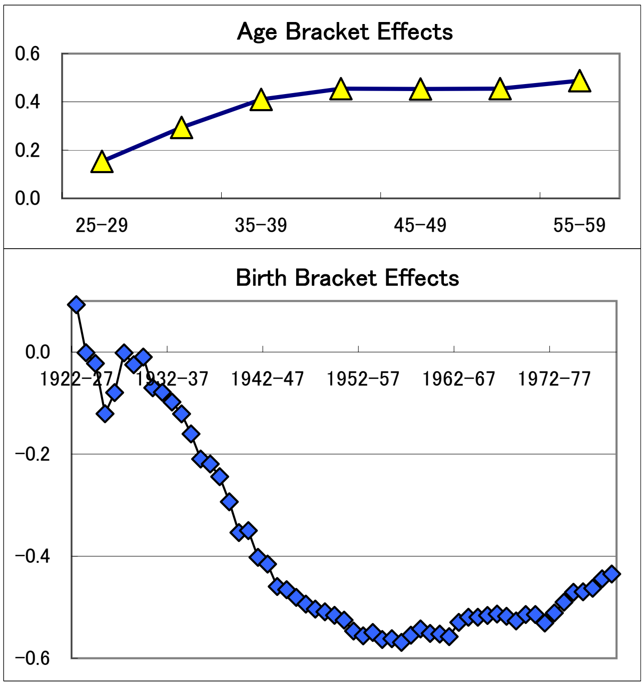
R-squared	0.9969		
Adjusted R-squared	0.9956	Constant	1.3024



Result-5A

weighted average annual wage of female employees in the higher educational status / weighted average annual wage of female employees in the lower educational status

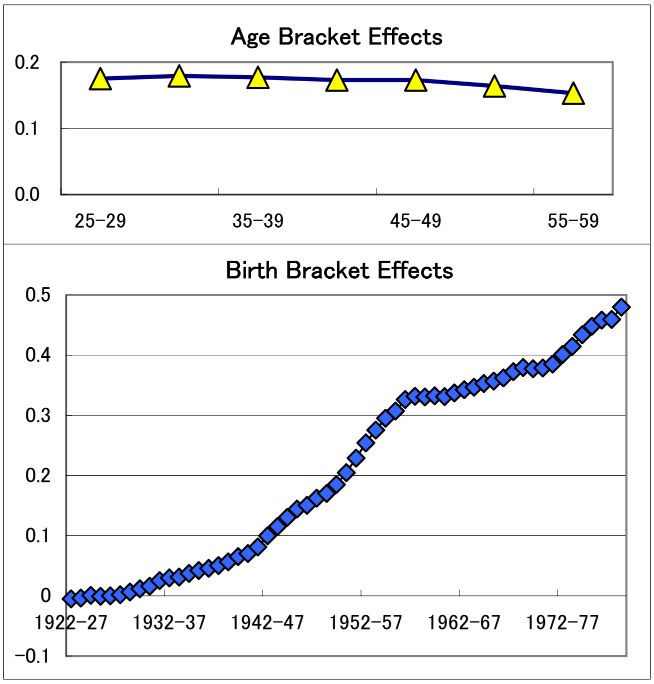
R-squared	0.9913		
Adjusted R-squared	0.9877	Constant	1.5739



Result-5B

total number of male employees in the higher educational status / total number of male employees

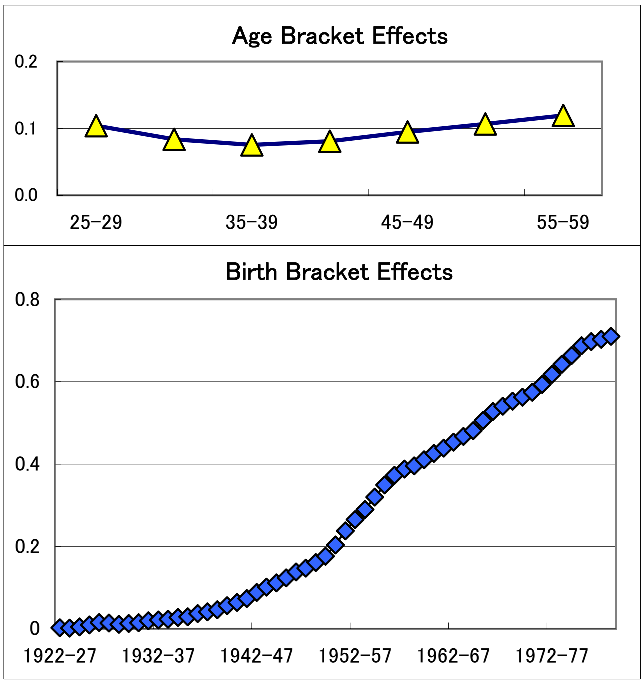
R-squared	0.9981		
Adjusted R-squared	0.9973	Constant	-0.0299



Result-5C

total number of female employees in the higher educational status / total number of female employees

R-squared	0.9977		
Adjusted R-squared	0.9968	Constant	-0.0778

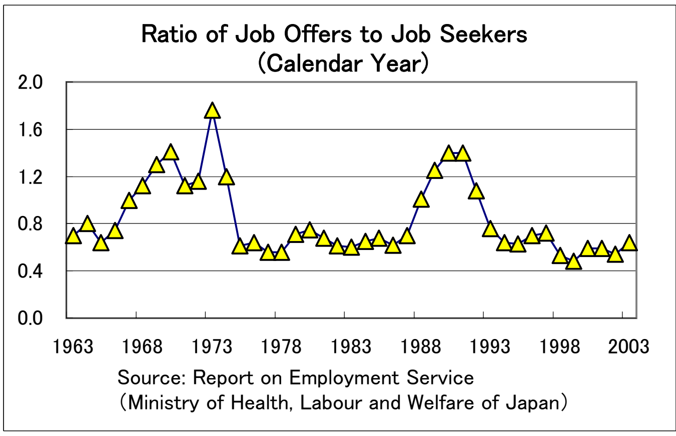


Result-5D

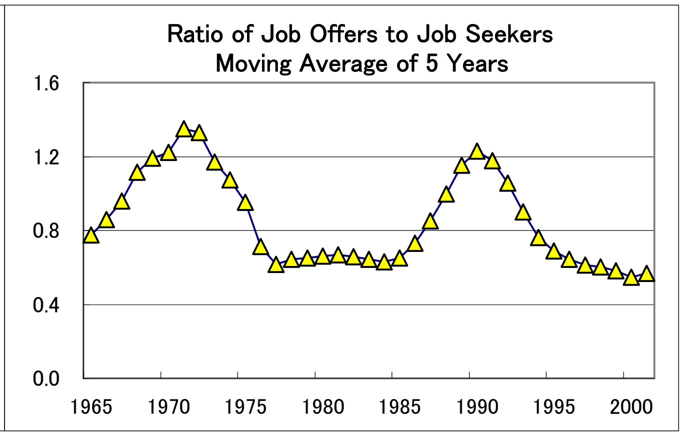
Result-5

Comparison of the birth bracket effects in Result-2A with ratios of JTS (job offers to job seekers) at the time of the graduation

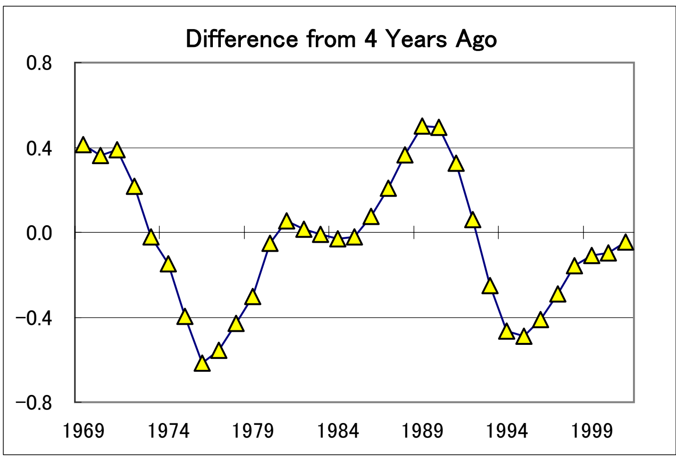
birth brackets (overlapping)	the estimated birth bracket effects on college/high-school wage ratio of males	recruiting time of high-school graduates	average JTS ratio	recruiting time of college graduates	average JTS ratio	JTS ratio difference
from July 1946 to June 1950	−0.1960	1963 – 1967	0.776	1967 – 1971	1.190	0.414
from July 1947 to June 1951	−0.1914	1964 – 1968	0.860	1968 – 1972	1.222	0.362
from July 1948 to June 1952	−0.1833	1965 – 1969	0.960	1969 – 1973	1.350	0.390
from July 1949 to June 1953	−0.1976	1966 – 1970	1.114	1970 – 1974	1.330	0.216
from July 1950 to June 1954	−0.1991	1967 – 1971	1.190	1971 – 1975	1.170	−0.020
from July 1951 to June 1955	−0.2170	1968 – 1972	1.222	1972 – 1976	1.074	−0.148
from July 1952 to June 1956	−0.2101	1969 – 1973	1.350	1973 – 1977	0.954	−0.396
from July 1953 to June 1957	−0.2048	1970 – 1974	1.330	1974 – 1978	0.714	−0.616
from July 1954 to June 1958	−0.1926	1971 – 1975	1.170	1975 – 1979	0.616	−0.554
from July 1955 to June 1959	−0.1795	1972 – 1976	1.074	1976 – 1980	0.644	−0.430
from July 1956 to June 1960	−0.1670	1973 – 1977	0.954	1977 – 1981	0.652	−0.302
from July 1957 to June 1961	−0.1498	1974 – 1978	0.714	1978 – 1982	0.662	−0.052
from July 1958 to June 1962	−0.1395	1975 – 1979	0.616	1979 – 1983	0.670	0.054
from July 1959 to June 1963	−0.1310	1976 – 1980	0.644	1980 – 1984	0.658	0.014
from July 1960 to June 1964	−0.1325	1977 – 1981	0.652	1981 – 1985	0.644	−0.008
from July 1961 to June 1965	−0.1284	1978 – 1982	0.662	1982 – 1986	0.632	−0.030
from July 1962 to June 1966	−0.1250	1979 – 1983	0.670	1983 – 1987	0.650	−0.020
from July 1963 to June 1967	−0.1135	1980 – 1984	0.658	1984 – 1988	0.732	0.074
from July 1964 to June 1968	−0.1057	1981 – 1985	0.644	1985 – 1989	0.852	0.208
from July 1965 to June 1969	−0.1086	1982 – 1986	0.632	1986 – 1990	0.996	0.364
from July 1966 to June 1970	−0.1092	1983 – 1987	0.650	1987 – 1991	1.152	0.502
from July 1967 to June 1971	−0.1071	1984 – 1988	0.732	1988 – 1992	1.228	0.496
from July 1968 to June 1972	−0.1037	1985 – 1989	0.852	1989 – 1993	1.178	0.326
from July 1969 to June 1973	−0.1077	1986 – 1990	0.996	1990 – 1994	1.056	0.060
from July 1970 to June 1974	−0.1149	1987 – 1991	1.152	1991 – 1995	0.902	−0.250
from July 1971 to June 1975	−0.1359	1988 – 1992	1.228	1992 – 1996	0.762	−0.466
from July 1972 to June 1976	−0.1308	1989 – 1993	1.178	1993 – 1997	0.690	−0.488
from July 1973 to June 1977	−0.1197	1990 – 1994	1.056	1994 – 1998	0.644	−0.412
from July 1974 to June 1978	−0.1011	1991 – 1995	0.902	1995 – 1999	0.612	−0.290
from July 1975 to June 1979	−0.1088	1992 – 1996	0.762	1996 – 2000	0.604	−0.158
from July 1976 to June 1980	−0.1060	1993 – 1997	0.690	1997 – 2001	0.582	−0.108
from July 1977 to June 1981	−0.0969	1994 – 1998	0.644	1998 – 2002	0.546	−0.098
from July 1978 to June 1982	−0.0903	1995 – 1999	0.612	1999 – 2003	0.568	−0.044



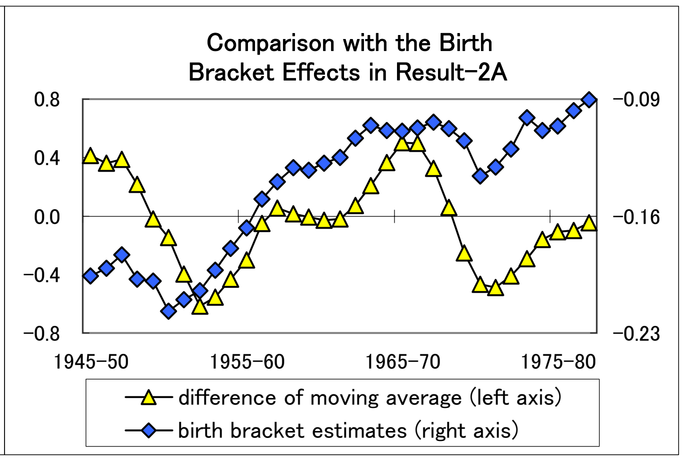
Result-6A



Result-6B



Result-6C



Result-6D