A fallacy of wage differentials:

wage ratio in distribution

Keiji Saito*

Graduate School of Economics, The University of Tokyo

Abstract

This paper demonstrates a fallacy of the wage differential between educational statuses. A large number of studies have based their approaches upon the wage ratio between educational statuses. However, this wage ratio contains a simple but crucial fallacy. The wage ratio is tricky; it changes even if both the labor market and the school effects remain unchanged. Irrespective of the school effects, the wage ratio first decreases and then increases as the advancement rate increases.

JEL classification: D31; J24; J31

Keywords: Wage ratio; Wage distribution; Education; Human capital

* E-mail: master@keijisaito.info

1. Introduction

This paper demonstrates a fallacy of the wage differential between educational statuses. This fallacy has a vital and far-reaching influence on the empirical analysis of education and wages as a whole.

Education can be analyzed in economics from two major empirical viewpoints. The first is from the viewpoint of microeconomics, pioneered by Schultz (1963) and Becker (1964). Mincer (1974) stylizes the wage equation. The causal effects of education, self-selection and unobserved ability are keywords in this viewpoint. Educational status is a textbook example of an endogenous variable. Griliches (1977) summarizes the econometric problems. A number of papers analyze the causal effects of education using instrumental variables and fixed effect estimations, as summarized by Card (2001). Trostel et al. (2002) estimates the economic returns in 28 countries using data from the International Social Survey Programme.

The second viewpoint is that of macroeconomics, which deals with the endogenous growth theory and wage differentials. The focus on the macro wage differential between educational statuses has undergone a change in the last two decades. Katz and Murphy (1992) and Levy and Murnane (1992) analyze the widening disparity in the US using data from the Current Population Survey (CPS). This research subject has extended beyond the US. Freeman and Katz, eds (1995) compile cross-country and individual country studies, and Acemoglu (2003) reviews differential trends in wage inequality in the US/UK and Europe using data from the CPS and the Luxembourg Income Studies.

The widening disparity gives rise to the question of how it came about. The most common idea in the 1990s was that technological progress that occurred at that time was favorable to high-skilled labor. In general, labor statistics do not contain any direct indicators of labor-skill. This is partly because the definition of labor-skill is obscure. Leuven et al. (2004) explains international relative wage differences using literacy scores; however, most studies use educational status as a proxy for labor-skill. They have little choice but to view higher educated workers as higher skilled ones, and lower educated workers as lower skilled ones. They substantially consider the college wage premium as the skill wage premium.

While most of the studies with regard to the widening disparity lay emphasis on a change in qualitative labor demand, this paper demonstrates a labor supply trick. Elementary mathematics can demonstrate this simple but crucial trick that brings to light the fallacy of wage differentials. The fallacy is directly related to wage differentials from the macroeconomic viewpoint. Moreover, the idea presented in this paper has a vital influence on the microeconomic viewpoint. It is strongly related to the empirical analysis of education and wages as a whole.

This paper is organized as follows: Section 2 introduces the fallacy using elementary mathematics and examples. Section 3 demonstrates the relationship between the advancement rate to higher education and the wage differential by taking into account the shape of the distribution and human capital formation. Although univariate distributions require that the educational status must divide the wage distribution unrealistically, Section 4 demonstrates that bivariate distributions can duplicate a realistic wage distribution. Some conclusions are offered in Section 5.

2. Idea and examples

2.1. Idea

Average wages classified by educational status are basic information in analyses of wage differentials. A wage ratio can be calculated by dividing the average wage of higher educational status by that of lower educational status. The wage ratio is a primary indicator of the wage differential or the return from education; partly because of simplicity, but mainly because of the non-availability of data. In most cases, the higher educated workers imply college graduates and the lower educated workers imply high-school ones. The college/high-school wage differential is considered to be a primary indicator of educational effect and skill premium. Although in this field there is a tendency to take the logarithm, the essence remains unchanged as long as it is a monotonic transformation.

The wage ratio is the relative wage of higher educated workers based on that of lower educated workers. The ratio is invariant even if the numerator and the denominator change at the same rate, such as due to rising prices or changes in the entire real wages. While it may appear that the wage ratio is an easy and plausible indicator enabling the determination of educational effects, it is open to misconstruction. The wage ratio has a simple but crucial trick.

Let me present this with a simple example. I shall hereafter use the term *capability* to describe potential productivity. For simplicity, I tentatively assume that education has no effect on capability and wage.

Assumptions in a simple example

1. Three men whose capabilities are $A > B > C$, respectively come into the world.
2. Man proceeds to higher education in order of the capability.
3. Three men receive A, B, and C wages respectively in labor market according to the marginal productivity principle.

"How does the wage ratio of higher/lower educational status change if the educational advancement rate increases?" The results are as follows.

$$A \div \frac{B+C}{2} = \frac{A}{\frac{B+C}{2}} \qquad \langle \textbf{phase 1}: \text{when only } A \text{ proceeds to higher education} \rangle \tag{1}$$

$$\frac{A+B}{2} \div C = \frac{\frac{A+B}{2}}{C} \qquad \langle \textbf{phase 2}: \text{when both } A \text{ and } B \text{ proceed to higher education} \rangle \tag{2}$$

B transfers to the higher educational status as the phase advances.[1] From A's viewpoint, B has an inferior capability. B's participation in the higher educational status implies a decrease in the average wage of the higher educational status. From C's viewpoint, B has a superior capability. B's withdrawal from the lower educational status implies a decrease in the lower average wage. Therefore, both the numerator and the denominator decrease. The change in the higher/lower wage ratio is determined by these rates of decrease.

Next, I assign simple numerical values to A, B, and C, such that $A = 3$, $B = 2$, and $C = 1$. The ratios are as follows.

$$\frac{3}{\frac{2+1}{2}} = 2 \qquad \langle \textbf{phase 1}: \text{when only } A \text{ proceeds to higher education} \rangle \tag{3}$$

$$\frac{\frac{3+2}{2}}{1} = 2.5 \qquad \langle \textbf{phase 2}: \text{when both } A \text{ and } B \text{ proceed to higher education} \rangle \tag{4}$$

In these numerical values, due to the phase transition, the numerator decreases by approximately 16.7% from 3 to 2.5, and the denominator decreases by approximately 33.3% from 1.5 to 1. The wage ratio increases because the rate of decrease of the denominator exceeds that of the numerator. As referred to earlier, the wage ratio would be invariant if the numerator and the denominator change at the same rate. Conversely, advancement rates have an impact on the wage ratio, except in the special case of same-rate change. Each worker receives a constant wage in this example. It is obvious that the wage ratio changes as a trick even if both the labor market and the educational effects remain unchanged.[2]

2.2. A Case of uniform distribution

The aforementioned numerical values $A = 3$, $B = 2$, and $C = 1$ present the simplest example of uniform distribution. The case of uniform distribution can be easily solved with an analytical approach. Let me give you the general uniform distribution example. The frequency takes the same value over a continuous

[1] In this paper, I do not discuss the reason people pursue higher education. This paper focuses on the effects of changed advancement rates. It is conceivable that although firms can distinguish educational statuses, they cannot accurately perceive the difference in the same educational status, prior to employment. Nevertheless, in this paper, it is only necessary to assume that people prefer to study rather than work. As income levels rise and new schools open, an increasing number of people pursue higher education.

[2] This idea has been referred to as a composition effect in Katz and Murphy (1992, P.67) and Levy and Murnane (1992, P.1360) in words. However, the following parts explicitly consider distribution shape, human capital formation and wage ratio transition. I believe that this is my contribution.

range from the highest capability represented by U to the lowest one represented by L, as shown in Figure 1. D indicates the dividing line. The right side of the line falls into the higher educational status such as the college graduates. The left side of the line falls into the lower educational status such as the high-school graduates. The focal point is the change in the ratio between these averages when the dividing line moves from the right towards the left.

An average can be derived by dividing an aggregated value by the total number. I calculate the average capability of the higher educational status.

$$\int_D^U Cxdx = \frac{1}{2}C(U^2 - D^2) \qquad \langle \text{the aggregated capability of the higher educational status} \rangle \quad (5)$$

$$\int_D^U Cdx = C(U - D) \qquad \langle \text{the total number of the higher educational status} \rangle \quad (6)$$

$$AU = \frac{\int_D^U Cxdx}{\int_D^U Cdx} = \frac{1}{2}(U + D) \qquad \langle AU: \text{ the average capability of the higher educational status} \rangle \quad (7)$$

AU is an increasing function of D. It applies to any other type of distribution [Proposition 1]. The implication of this proposition is obvious. When the advancement rate increases and the dividing line moves towards the left, it includes people whose capability is the most inferior from the viewpoint of the higher educational status. Hence, the average capability of the higher educational status decreases.

Similarly, I calculate the average capability of the lower educational status.

$$\int_L^D Cxdx = \frac{1}{2}C(D^2 - L^2) \qquad \langle \text{the aggregated capability of the lower educational status} \rangle \quad (8)$$

$$\int_L^D Cdx = C(D - L) \qquad \langle \text{the total number of the lower educational status} \rangle \quad (9)$$

$$AL = \frac{\int_L^D Cxdx}{\int_L^D Cdx} = \frac{1}{2}(D + L) \qquad \langle AL: \text{ the average capability of the lower educational status} \rangle \quad (10)$$

AL is an increasing function of D. It also applies to any other type of distribution [Proposition 2]. Similarly, the implication of this proposition is obvious. When the advancement rate increases and the dividing line moves towards the left, it excludes people whose capability is the most superior from the viewpoint of the lower educational status. Hence, the average capability of the lower educational status decreases.

As a consequence, the fractional number of the averages corresponding to the wage ratio or the college premium is

$$WR = \frac{AU}{AL} = \frac{U + D}{D + L}. \qquad \langle WR: \text{the wage ratio} \rangle \quad (11)$$

The ratio definitely depends on the dividing line D. Next, I differentiate the ratio with respect to D in order to examine the effect of D on the wage ratio.

$$\frac{\partial WR(D)}{\partial D} = \frac{(D+L)-(U+D)}{(D+L)^2} = \frac{L-U}{(D+L)^2} < 0 \qquad \langle \text{the first-order derivative of } WR \rangle \quad (12)$$

The first-order derivative is a negative value. Therefore, the wage ratio decreases when the dividing line moves towards the right. On the contrary, the shifting of the dividing line towards the left implies an increase in the advancement rate and causes an increase in the wage ratio. In addition,

$$\frac{\partial^2 WR(D)}{\partial D^2} = \frac{2(U-L)}{(D+L)^3} > 0. \qquad \langle \text{the second-order derivative of } WR \rangle \quad (13)$$

The second-order derivative is a positive value. In a uniform distribution, the wage ratio expands at an accelerated pace with an increase in the higher educational status rate. Table 1 shows a numerical example of a uniform distribution. The minimum wage ratio cell is represented in bold font. In this case, the minimum upper rate (20%) results in the minimum wage ratio between the two averages.

2.3. A Case of realistic distribution

It is natural to assume that the actual wage and capability follow a mountain-shaped distribution as typified by a normal or a log-normal distribution like Figure 2. The next section gives a detailed demonstration; however, here is an introduction of mountain-shaped distribution that can be examined by hand calculation. Table 2 indicates the relationship between the upper rate and the wage ratio on the mountain-shaped distribution. In this numerical value, the minimum upper rate (10%) brings the wage ratio to 2.143. Next, the second-minimum upper rate (20%) brings the minimum wage ratio to 2.118. Moreover, an upper rate of more than 50% increases the wage ratio. In summary, the wage ratio begins to increase reversely after a decrease if the upper rate increases.

The wage ratio expands monotonically in a uniform distribution with an increase in the upper rate. In a realistic distribution, the wage ratio increases after a decrease under the same conditions. In either case, the ratio between averages hinges on the dividing line.[3] In other words, the wage ratio strongly depends on the educational advancement rate. The wage ratio can change as a trick regardless of educational effects and individual wage variations.

3. Simulation of the ratio between averages

This section presents simulations taking into consideration human capital formation. Based on discrete distribution, Simulation-1 and Simulation-2 are organized as follows: The upper left table presents the

[3] As these tables show, the difference between averages is independent of the dividing line only in the case of uniform distribution.

given capability and frequency. For the sake of simplicity, the total number setting is considered to be 100. The upper right graph draws them visually. The table in the middle left is the result of division; the minimum wage ratio cell is represented in bold font. The figure in the middle right draws the relationship between the upper rate and the ratio between averages. The middle portion indicates a case where there is no human capital formation. It is nothing more than a detailed presentation of the numerical example given previously. The lower left figure illustrates a relationship based on human capital formation as a constant value addition. It suggests that when a person proceeds to higher education, his/her capability increases additively, such as (+2). The lower right figure presents a case of multiplicative human capital formation, such as (×1.2). In this case, the increment depends on the prior capability.

Simulation-2 is a parallel shift (+3) version of Simulation-1. Despite having the same distribution shape, the relationship between the the upper rate and the ratio between averages is different from that in Simulation-1. One question that arises here is regarding the measurement of capability. In Simulation-1, the person at the topmost position is ten times as capable as the one at the lowermost position. This gap reduces to 3.25 times in Simulation-2. An additional variation causes a change in the ratio between averages, but no change is caused by a multiplicative one.

It might be useful to recall the example on uniform distribution. The ratio in (11) is $WR(D) = \frac{U+D}{D+L}$. If I shift the distribution by t, the result is $WR^{+}(t, D) = \frac{U+D+2t}{D+L+2t}$ because of addition t to U, D and L. The derivative with respect to t is $\frac{\partial WR^{+}(t,D)}{\partial t} = \frac{2(L-U)}{(D+L+2t)^2} < 0$. On the other hand, multiplication of U, D, and L by t does not affect the ratio as they cancel out. In very simple terms, if I multiply the numerator and the denominator by the same number, the fractional number is invariant. On the other hand, if I add the same number to the numerator and the denominator, the fractional number decreases, converging with 1.[4] For this reason, the ratio between averages depends on their measurement. For example, the Graduate Record Examination (GRE) has a non-zero minimum score. It has a different dividing point for calculating the minimum ratio from a zero-minimum examination. If the framework can be empirically applied, the capability is measured in terms of the productivity or wage. Therefore, this paper treats the ratio between averages as the wage ratio.

Both in Simulation-1 and Simulation-2, little difference exists in the line shape between these lower figures and the middle figure (no increase). If it is assumed that capability or wage follows a mountain-shape, regardless of human capital formation, the ratio increases after it decreases once along with the increase in the upper rate.

Simulation-3 is based on continuous log-normal distribution with mean 5 and standard deviation 1. The integral of the frequency function from 0 to ∞ equals 1. In this case, I can consider the density

[4] L'Hôpital's rule is used to compute the limit.

function of capability. Simulation-3A represents the case in which there is no human capital formation. Simulation-3B assumes that the capability increases by 10% on proceeding to higher education. The essence is already mentioned in the discrete simulations. Additionally, the upper rate that brings about minimum value is invariant if human capital is formed in multiplicative forms [Proposition 3].

The previous section illustrates that the wage ratio changes as a trick even if there is neither individual wage variations nor the human capital formation. The result is independent of human capital formation. Even though education has real effects, the line shape of the wage ratio is almost a carbon copy of no effect. Conversely, the wage ratio is inappropriate to assess educational effects. Furthermore, in realistic mountain-shaped distributions, the wage ratio increases after decrease as the upper rate increases, regardless of educational effects.

Previous studies have proposed a number of hypotheses to explain the cause of the wage ratio change.[5] The share of higher educational status has been increasing in developed countries. Under these circumstances, the change in the wage ratio is a natural result regardless of educational effects or the labor market transition.

4. Simulation of bivariate distributions

This paper has hitherto assumed that people proceed to higher education in strict order of potential capability. It is a strong assumption. The decision to pursue higher education depends not only on capability or test scores but also on household income, location, character and so forth. Moreover, if this framework can be directly applied, the educational status would divide the wage distribution sharply. However, it is observed that real wage distributions by educational status overlap. Some people earn poor scores but make a fortune and vice versa. This section demonstrates that although a one-to-one correspondence between educational status and wage fails to satisfy, this theory remains suitable for application using bivariate distributions.

First, supposing that there exists a joint distribution of academic ability ($a > 0$) and labor productivity ($w > 0$), I define the density function of a bivariate distribution.

$$f(a, w) \geq 0 \qquad \langle\text{joint density function of academic ability and labor productivity}\rangle \quad (14)$$

$$f_a(a) = \int_0^\infty f(a, w)dw \qquad \langle\text{marginal density function of academic ability}\rangle \quad (15)$$

$$f_w(w) = \int_0^\infty f(a, w)da \qquad \langle\text{marginal density function of labor productivity}\rangle \quad (16)$$

The labor productivity distribution in an educational status can be expressed as the integral taken

[5] Levy and Murnane (1992) use the ratio of medians instead of averages. Although the ratio of medians seems to have a higher level than the ratio of averages, the change is similar to that of average in mountain-shaped distributions.

between the academic ability of D_l and D_u.

$$g(w \mid D_l \leq a < D_u) = \frac{\int_{D_l}^{D_u} f(a,w)da}{\int_0^\infty \int_{D_l}^{D_u} f(a,w)da\ dw} \tag{17}$$

I then build an intermediate function.

$$h(a,w \mid D_l, D_u) = \begin{cases} f(a,w) & \text{if } D_l \leq a < D_u \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

An academic ability distribution in an educational status is as follows.

$$i(a \mid D_l \leq a < D_u) = \frac{\int_0^\infty h(a,w \mid D_l, D_u)dw}{\int_0^\infty \int_{D_l}^{D_u} f(a,w)da\ dw} = \frac{\int_0^\infty h(a,w \mid D_l, D_u)dw}{\int_0^\infty \int_0^\infty h(a,w \mid D_l, D_u)da\ dw} \tag{19}$$

Second, I specify the bivariate distribution. The marginal density function of academic ability follows a log-normal distribution in analogy with Simulation-3.

$$f_a(a \mid \hat{\mu}, \hat{\sigma}) \ = \frac{1}{\sqrt{2\pi}\ a\hat{\sigma}} \exp\left\{-\frac{(\log a - \hat{\mu})^2}{2\hat{\sigma}^2}\right\} \tag{20}$$

It is highly unlikely to attain a one-to-one correspondence between academic ability and labor productivity. However, it is safe to assume that there exists a positive correlation between the two. labor productivity scatters on following a normal distribution with the mean of the academic ability and the standard deviation as σ_b. I specify the conditional density function of w given that $a = \bar{a}$.

$$f(w \mid \bar{a}, \sigma_b) = \frac{1}{\sqrt{2\pi}\sigma_b} \exp\left\{-\frac{(w-\bar{a})^2}{2\sigma_b^2}\right\} \ \langle\text{the conditional density function of } w \text{ given that } a = \bar{a}\rangle \tag{21}$$

Given an academic ability, the mean of labor productivity corresponds to the academic ability by the assumption.

$$E[f(w \mid \bar{a})] = \bar{a} \qquad \langle\text{the conditional expected value of } \ w \text{ given that } a = \bar{a}\rangle \tag{22}$$

The joint density function of academic ability and labor productivity is

$$\begin{aligned} f(a,w) = f_a(a \mid \hat{\mu}, \hat{\sigma})\ f(w \mid a, \sigma_b) &= \left[\frac{1}{\sqrt{2\pi}\ a\hat{\sigma}} \exp\left\{-\frac{(\log a - \hat{\mu})^2}{2\hat{\sigma}^2}\right\}\right]\left[\frac{1}{\sqrt{2\pi}\sigma_b} \exp\left\{-\frac{(w-a)^2}{2\sigma_b^2}\right\}\right] \\ &= \frac{1}{2\pi\ a\hat{\sigma}\sigma_b} \exp\left\{-\frac{(\log a - \hat{\mu})^2}{2\hat{\sigma}^2} - \frac{(w-a)^2}{2\sigma_b^2}\right\}. \end{aligned} \tag{23}$$

Third, I specify the parameters of the distribution. In analogy with Simulation-3, the marginal density function of academic ability follows a log-normal distribution with mean 5 and standard deviation 1. $E[a] = \mu_a = \exp(\hat{\mu} + \frac{1}{2}\hat{\sigma}^2) = 5$ and $\sqrt{Var[a]} = \sigma_a = \sqrt{\exp(2\hat{\mu} + \hat{\sigma}^2)[\exp(\hat{\sigma}^2) - 1]} = 1$. Then, $\hat{\mu} \approx 1.5898$ and $\hat{\sigma} \approx 0.1980$.

Simulation-4 and Simulation-5 are based on the parameters that σ_b are 0.2 and 0.7, respectively. In these simulations, the upper portion indicates that a dividing line visually partitions the bivariate distribution in a manner similar to that of a knife cutting a mountain. The six figures in the middle portion show the transitions of wage distribution divided by one line. The lower figure illustrates each labor productivity distribution partitioned by three dividing lines.

σ_b is the only difference between Simulation-4 and Simulation-5. It is the standard deviation between the academic ability and the labor productivity. Since σ_b in Simulation-4 is small, the educational status divides the wage distribution bluntly. Since σ_b in Simulation-5 is relatively large, the dividing point becomes so indistinct that it is difficult to believe that the other dimension divides the distribution. The approach of σ_b to zero implies the approach to a univariate distribution. The identification of a with w is a special case of the bivariate distribution. In the special case, the educational status divides the wage distribution sharply.

As shown in (21) and (22), both Simulation-4 and Simulation-5 are built on the basis of the assumption that academic ability generates labor productivity while preserving the mean. Therefore, these simulations satisfy the condition that

$$\begin{aligned}
&\int_0^\infty w \cdot g(w \mid \sigma_b,\ D_l \leq a < D_u)dw \\
=&\int_0^\infty a \cdot i(a \mid D_l \leq a < D_u)]da \\
=&\int_0^\infty w \cdot g(w \mid \sigma_b',\ D_l \leq a < D_u)dw \\
=&E[w \mid \sigma_b,\ D_l \leq a < D_u)] = E[a \mid D_l \leq a < D_u)] = E[w \mid \sigma_b',\ D_l \leq a < D_u)]. \qquad (\sigma_b \neq \sigma_b') \qquad (24)
\end{aligned}$$

If one dividing line corresponds to the other simulation, the average labor productivity in an educational status also corresponds. These average wage ratios are equivalent to the special case of univariate distribution. As far as the mean is concerned, it is independent of σ_b.

Though the real wage differential between the educational statuses tends to widen with age, the wage distributions by educational status overlap within each age bracket. The bivariate simulation can duplicate the overlapping wage distributions. The bivariate simulations are based on the assumption that capable persons do not always pursue higher education despite their inclination to do so. Therefore, the overlapping wage distribution is anything but a disproof of the dividing theory. It is a natural result of reasonable assumptions. Furthermore, with regard to indicators of averages, a univariate distribution can substitute for a bivariate one.

5. Conclusion

This paper began with an introduction to the fallacy of wage ratios. A simple numerical example reveals this fallacy. Univariate distribution simulations show that the wage ratio changes regardless of educational effects. Bivariate distribution simulations demonstrate that although a one-to-one correspondence between educational status and wage is not satisfied, this theory remains suitable for application.

This fallacy is directly related to wage differentials from the macroeconomic viewpoint. Even if there has truly been a change in the wage ratio between educational statuses, there is a possibility that individual wages have not changed. On the other hand, this fallacy has a vital influence on the microeconomic viewpoint. The wage ratios of averages are inappropriate to measure the educational effects. If the wage equation cannot control potential productivity perfectly, the estimated return to education must be biased. Further, the magnitude of the bias depends on the advancement rate to higher education. In a realistic mountain-shaped distribution, the biased return increases after a decrease as the advancement rate increases, regardless of educational effects.

Appendix: Proofs of Propositions

Proposition 1:

Suppose that capability is $x > 0$, the density function is $f(x) \geq 0$ and $h(x) = f(x) \cdot x$

$$\int_D^U f(x) \cdot x \, dx = H(U) - H(D) \qquad \langle \text{the aggregated capability of the higher educational status} \rangle \quad (25)$$

$$\int_D^U f(x) \, dx = F(U) - F(D) \qquad \langle \text{the total number of the higher educational status} \rangle \quad (26)$$

$$AU(D) = \frac{H(U) - H(D)}{F(U) - F(D)} \qquad \langle AU: \text{ the average capability of the higher educational status} \rangle \quad (27)$$

$$\begin{aligned} \frac{\partial AU(D)}{\partial D} &= \frac{\{H(U) - H(D)\} f(D) - \{F(U) - F(D)\} f(D) \cdot D}{\{F(U) - F(D)\}^2} \\ &= \frac{AU(D) \{F(U) - F(D)\} f(D) - \{F(U) - F(D)\} f(D) \cdot D}{\{F(U) - F(D)\}^2} \\ &= \frac{\{AU(D) - D\} f(D)}{\{F(U) - F(D)\}} \geq 0 \qquad \langle \text{the derivative of } AU \text{ with respect to } D \rangle \quad (28) \end{aligned}$$

The differential equation (28) is positive because each component is positive. Then, if D increases, the average capability of higher educational status increases.

Proposition 2:

$$\int_L^D f(x) \cdot x \, dx = H(D) - H(L) \qquad \langle \text{the aggregated capability of the lower educational status} \rangle \tag{29}$$

$$\int_L^D f(x) \, dx = F(D) - F(L) \qquad \langle \text{the total number of the lower educational status} \rangle \tag{30}$$

$$AL(D) = \frac{H(D) - H(L)}{F(D) - F(L)} \qquad \langle AL: \text{ the average capability of the lower educational status} \rangle \tag{31}$$

$$\begin{aligned}
\frac{\partial AL(D)}{\partial D} &= \frac{\{F(D) - F(L)\} f(D) \cdot D - \{H(D) - H(L)\} f(D)}{\{F(D) - F(L)\}^2} \\
&= \frac{\{F(D) - F(L)\} f(D) \cdot D - AL(D) \{F(D) - F(L)\} f(D)}{\{F(D) - F(L)\}^2} \\
&= \frac{\{D - AL(D)\} f(D)}{\{F(D) - F(L)\}} \geq 0 \qquad \langle \text{the derivative of } AL \text{ with respect to } D \rangle
\end{aligned} \tag{32}$$

The differential equation (32) is positive because each component is positive. Then, if D increases, the average of capability lower educational status increases.

Proposition 3:

<u>I. Remains the same</u>

If education has no effect on the capability, the average capability of the higher educational status (AU) and the wage ratio (WR) are as follows.

$$AU(D) = \frac{\int_D^U f(x) \cdot x \, dx}{\int_D^U f(x) \, dx} = \frac{H(U) - H(D)}{F(U) - F(D)} \tag{33}$$

$$WR(D) = \frac{AU(D)}{AL(D)} \tag{34}$$

The derivative of WR with respect to D is

$$\frac{\partial WR(D)}{\partial D} = \frac{AU'(D) \cdot AL(D) - AU(D) \cdot AL'(D)}{\{AL(D)\}^2}. \tag{35}$$

<u>II. Constant rate increase $[a \geq 0]$</u>

If education has a multiplicative effect on the capability, the increased average capability of the higher educational status (AU_a) and the increased wage ratio (WR_a) are as follows.

$$AU_a(D) = \frac{\int_D^U f(x) \cdot ax \, dx}{\int_D^U f(x) \, dx} = \frac{a \{H(U) - H(D)\}}{F(U) - F(D)} = a \, AU \tag{36}$$

$$WR_a(D) = \frac{AU_a(D)}{AL(D)} = \frac{a \cdot AU(D)}{AL(D)} \tag{37}$$

The derivative of WR_a with respect to D is

$$\frac{\partial WR_a(D)}{\partial D} = \frac{a\{AU'(D) \cdot AL(D) - AU(D) \cdot AL'(D)\}}{\{AL(D)\}^2}. \tag{38}$$

(35) and (38) are the same except for the coefficient a. These differential equations always have the same sign.

References

Acemoglu, Daron (2003) 'Cross-Country Inequality Trends.' *Economic Journal* 113(127), F121–F149

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

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

Freeman, Richard B. and Lawrence F. Katz, eds (1995) *Differences and Changes in Wage Structures* (University of Chicago Press)

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

Katz, Lawrence F. and Kevin M. Murphy (1992) 'Changes in Relative Wages, 1963-1987: Supply and Demand Factors.' *The Quarterly Journal of Economics* 107(1), 35–78

Leuven, Edwin, Hessel Oosterbeek and Hans van Ophem (2004) 'Explaining international differences in male skill wage differentials by differences in demand and supply of skill.' *Economic Journal* 114(127), 466–486

Levy, Frank and Richard J Murnane (1992) 'U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations.' *Journal of Economic Literature* 30(3), 1333–81

Mincer, Jacob (1974) *Schooling, Experience and Earnings* (Proquest Info & Learning)

Schultz, Theodore W. (1963) *The economic value of education* (Columbia University Press)

Trostel, Philip, Ian Walker and Paul Woolley (2002) 'Estimates of the economic return to schooling for 28 countries.' *Labour Economics* 9(1), 1–16

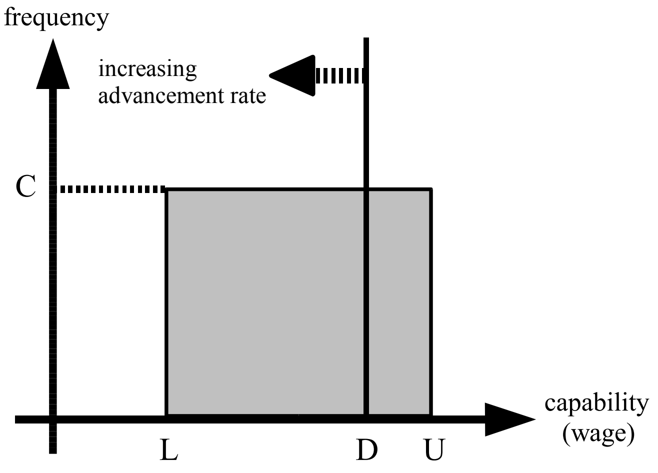


Figure 1 Uniform Distribution

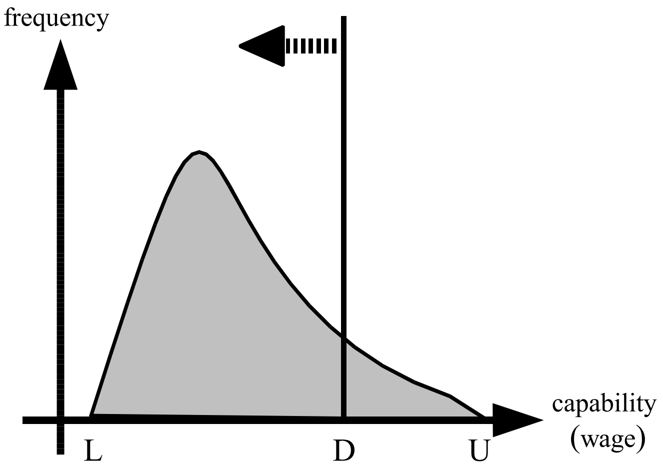


Figure 2 Generally Assumed Distribution

Table 1

An Example of Uniform Distribution

<table>
<tr><th>capability (number) / dividing line</th><th>1 (2)</th><th>2 (2)</th><th>3 (2)</th><th>4 (2)</th><th>5 (2)</th><th>upper rate</th><th>WR</th><th>the logarithm of WR</th></tr>
<tr><td>between 4 and 5</td><td colspan="4">2.5</td><td>5</td><td>20%</td><td>2.000</td><td>0.693</td></tr>
<tr><td>between 3 and 4</td><td colspan="3">2</td><td colspan="2">4.5</td><td>40%</td><td>2.250</td><td>0.811</td></tr>
<tr><td>between 2 and 3</td><td colspan="2">1.5</td><td colspan="3">4</td><td>60%</td><td>2.667</td><td>0.981</td></tr>
<tr><td>between 1 and 2</td><td>1</td><td colspan="4">3.5</td><td>80%</td><td>3.500</td><td>1.253</td></tr>
</table>

Table 2

An Example of Realistic Distribution

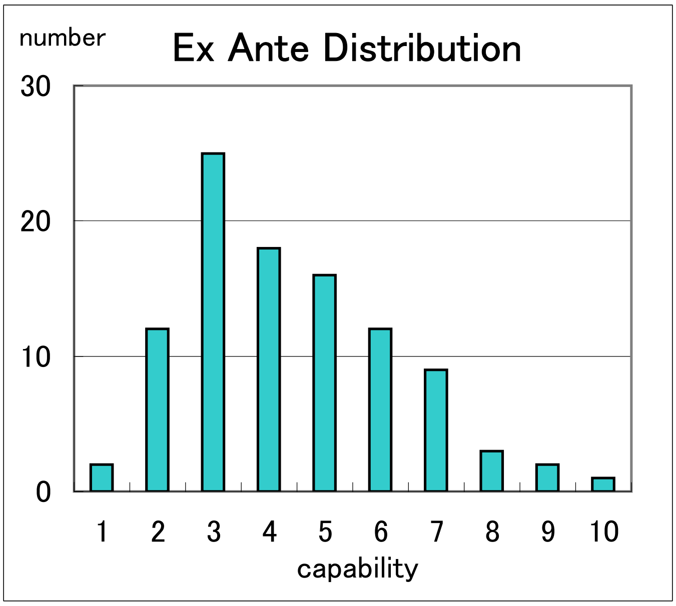
<table>
<tr><th>capability (number) / dividing line</th><th>1 (2)</th><th>2 (3)</th><th>3 (3)</th><th>4 (1)</th><th>5 (1)</th><th>upper rate</th><th>WR</th><th>the logarithm of WR</th></tr>
<tr><td>between 4 and 5</td><td colspan="4">2.333</td><td>5</td><td>10%</td><td>2.143</td><td>0.762</td></tr>
<tr><td>between 3 and 4</td><td colspan="3">2.125</td><td colspan="2">4.5</td><td>20%</td><td>2.118</td><td>0.750</td></tr>
<tr><td>between 2 and 3</td><td colspan="2">1.6</td><td colspan="3">3.6</td><td>50%</td><td>2.250</td><td>0.811</td></tr>
<tr><td>between 1 and 2</td><td>1</td><td colspan="4">3</td><td>80%</td><td>3.000</td><td>1.099</td></tr>
</table>

Ratio between Averages in a Discrete Distribution

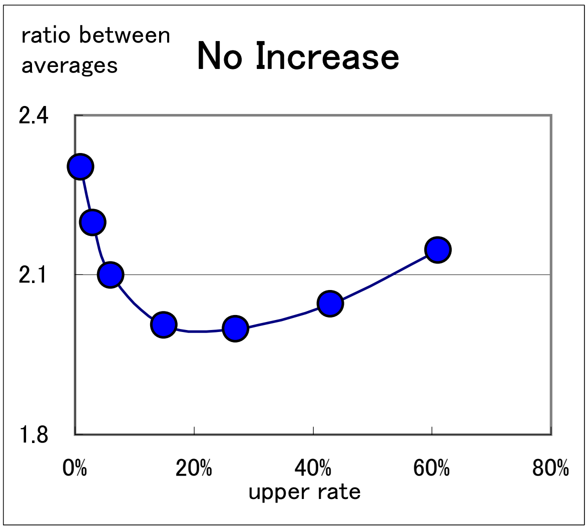
(A Right-skewed Distribution: Log-normal Distribution Type)

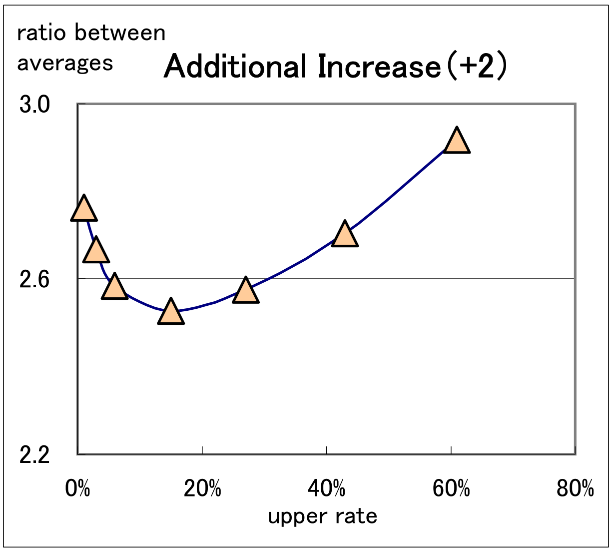
capability	number	product
1	2	2
2	12	24
3	25	75
4	18	72
5	16	80
6	12	72
7	9	63
8	3	24
9	2	18
10	1	10

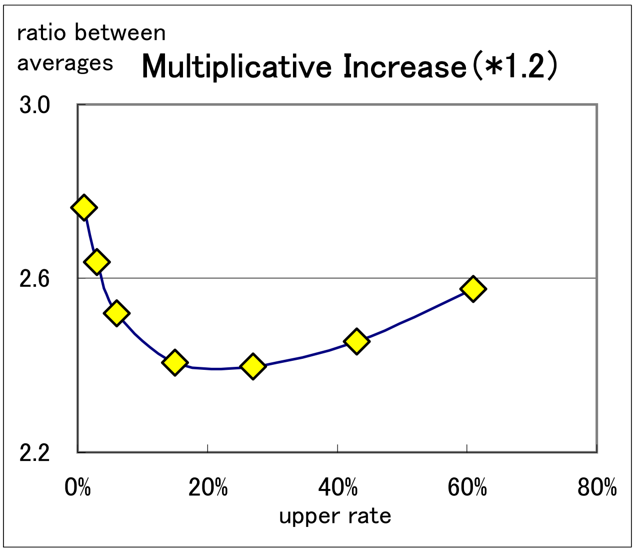
Total Number	100
Total Product	440
Ex Ante Aggregate Average	4.4



division between	upper numbers	upper rate	higher average	lower average	ratio between averages
9 and 10	1	1%	10.00	4.34	2.302
8 and 9	3	3%	9.33	4.25	2.197
7 and 8	6	6%	8.67	4.13	2.100
6 and 7	15	15%	7.67	3.82	2.005
5 and 6	27	27%	6.93	3.47	1.998
4 and 5	43	43%	6.21	3.04	2.046
3 and 4	61	61%	5.56	2.59	2.146
2 and 3	86	86%	4.81	1.86	2.592
1 and 2	98	98%	4.47	1.00	4.469





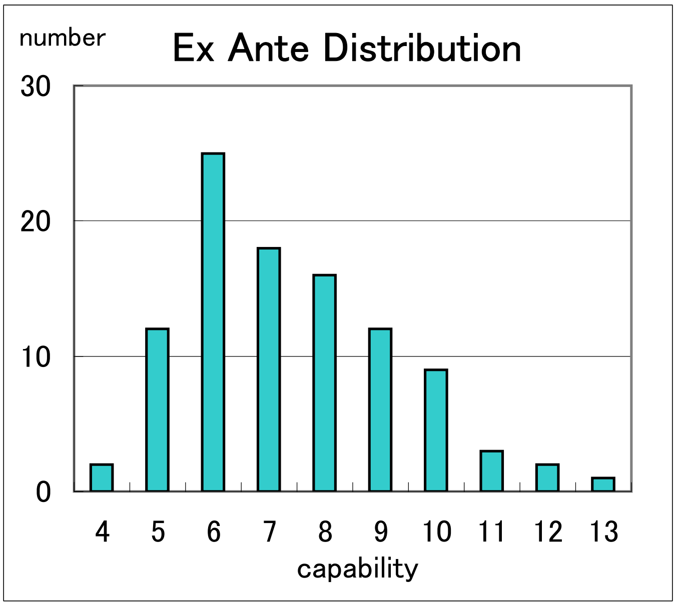


Ratio between Averages in a Discrete Distribution

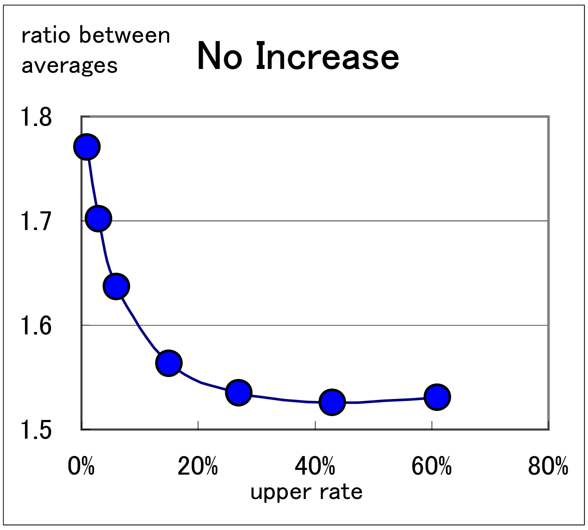
(A Right-skewed Distribution <+3>: Log-normal Distribution Type)

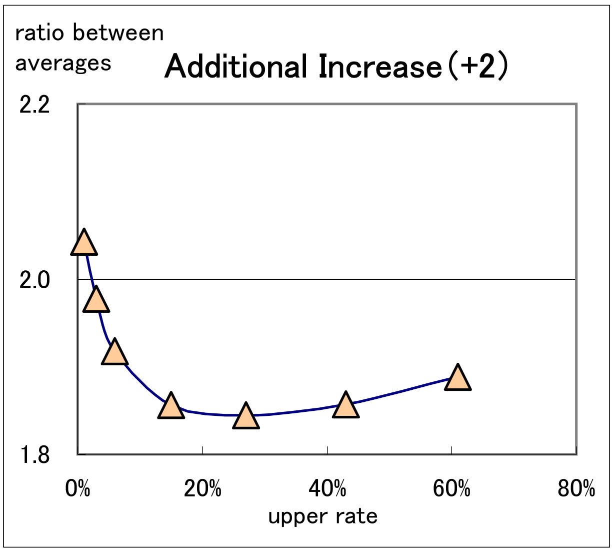
capability	number	product
4	2	8
5	12	60
6	25	150
7	18	126
8	16	128
9	12	108
10	9	90
11	3	33
12	2	24
13	1	13

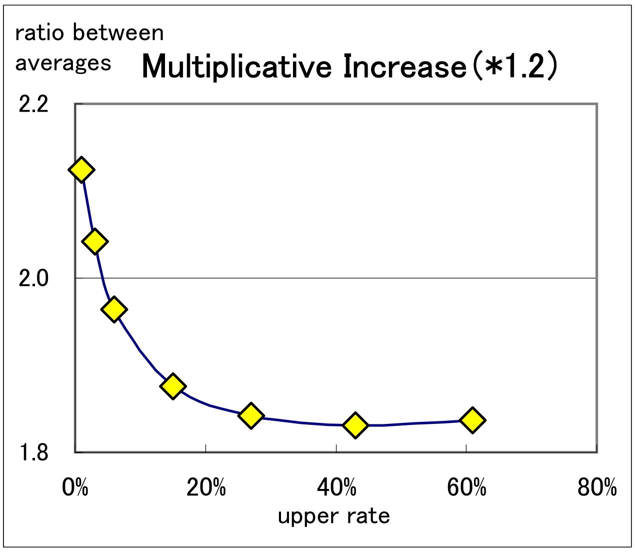
Total Number	100
Total Product	740
Ex Ante Aggregate Average	7.4



division between	upper numbers	upper rate	higher average	lower average	ratio between averages
12 and 13	1	1%	13.00	7.34	1.770
11 and 12	3	3%	12.33	7.25	1.702
10 and 11	6	6%	11.67	7.13	1.637
9 and 10	15	15%	10.67	6.82	1.563
8 and 9	27	27%	9.93	6.47	1.535
7 and 8	43	43%	9.21	6.04	1.526
6 and 7	61	61%	8.56	5.59	1.531
5 and 6	86	86%	7.81	4.86	1.609
4 and 5	98	98%	7.47	4.00	1.867

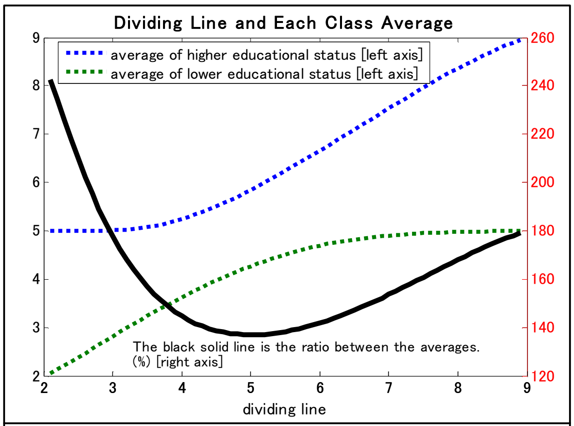


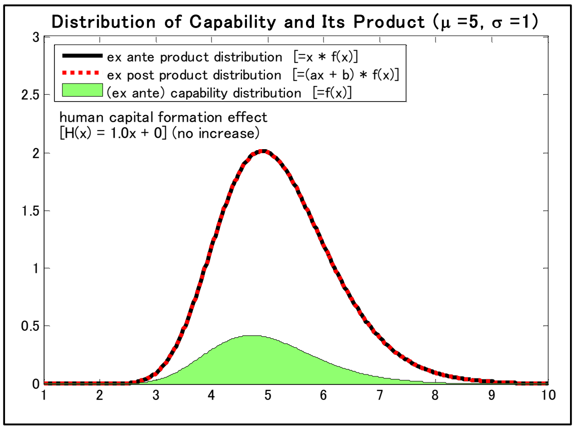


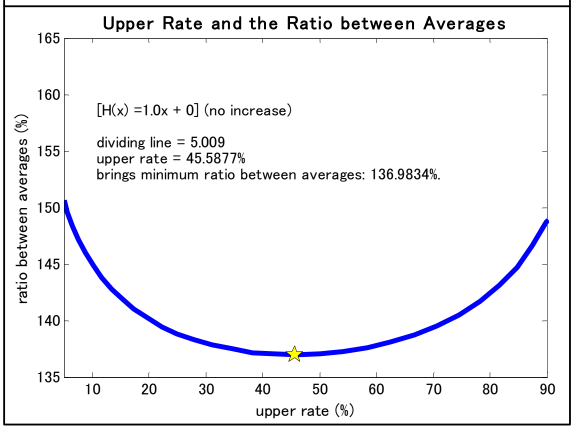


Simulation of a Log−normal Distribution

Distribution Setting	
log−normal distribution	
ex ante aggregate average	5
standard deviation	1
Human Capital Formation [H(x) = ix + j]	
i = 1.0, j = 0	(no increase)
The Minimum Ratio between Averages	
dividing line	5.009
upper rate	45.59%
average of the higher educational status	5.854
average of the lower educational status	4.274
the ratio between averages	136.98%

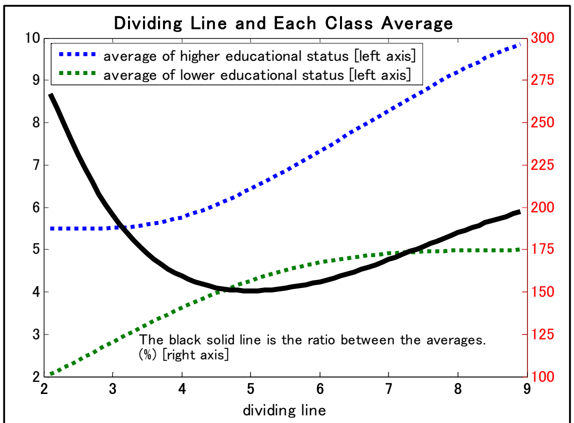


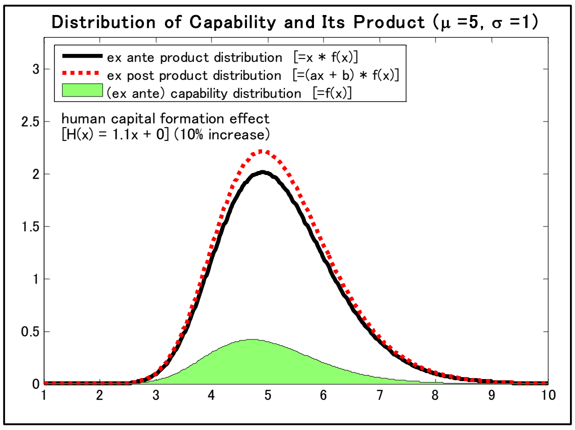


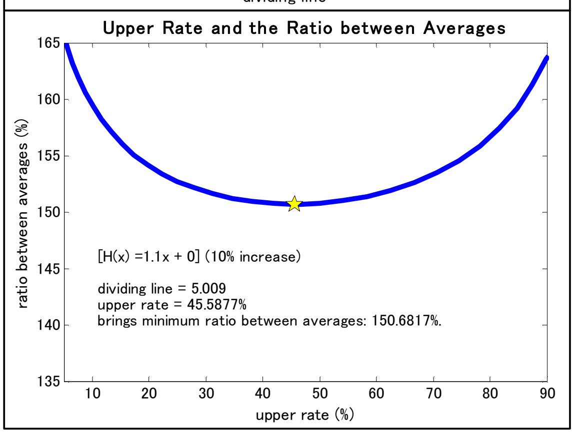


Simulation-3A

Distribution Setting	
log−normal distribution	
ex ante aggregate average	5
standard deviation	1
Human Capital Formation [H(x) = ix + j]	
i = 1.1, j = 0	(10% increase)
The Minimum Ratio between Averages	
dividing line	5.009
upper rate	45.59%
average of the higher educational status	6.440
average of the lower educational status	4.274
the ratio between averages	150.68%





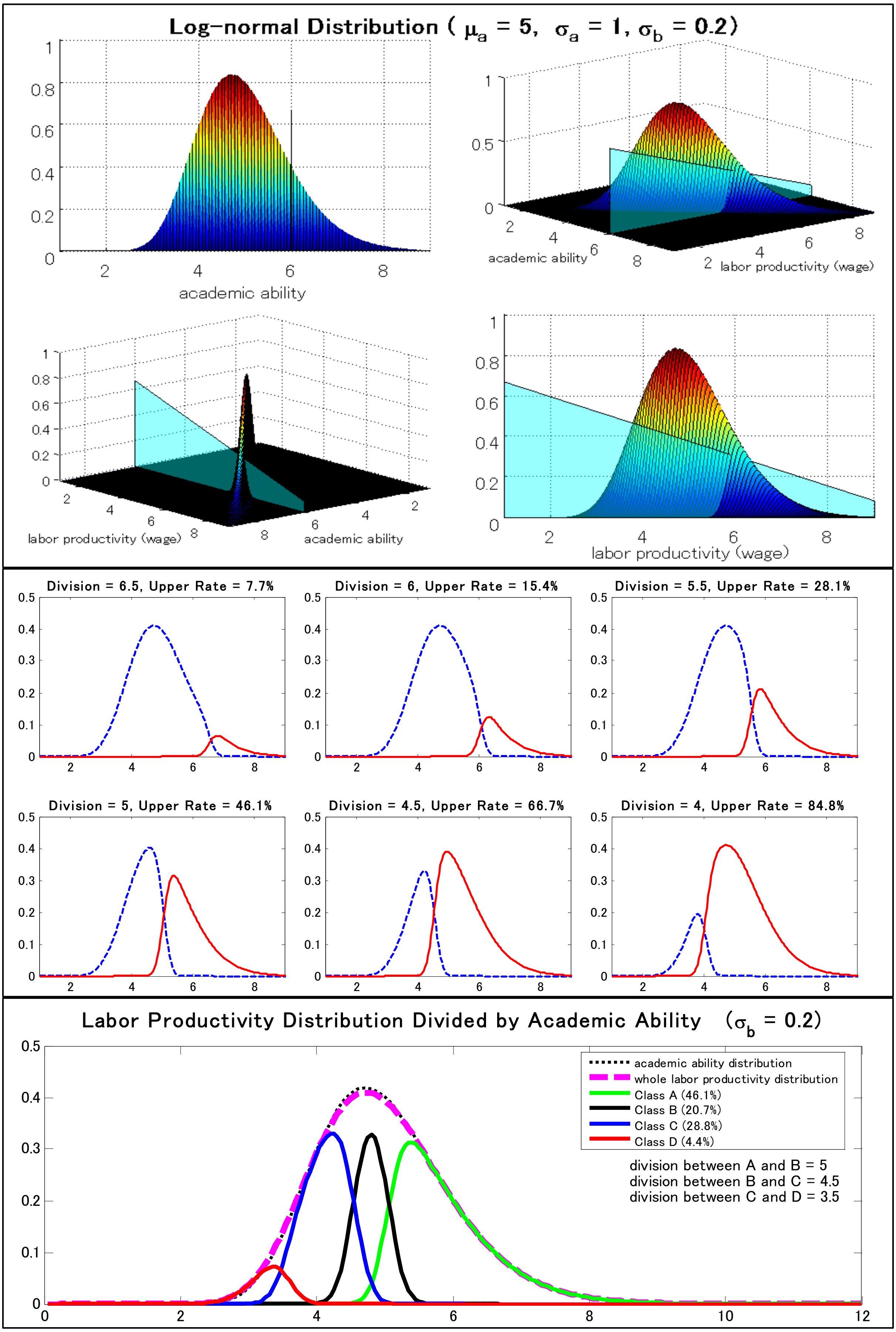


Simulation-3B

Simulation-3

Simulation of Bivariate Distribution

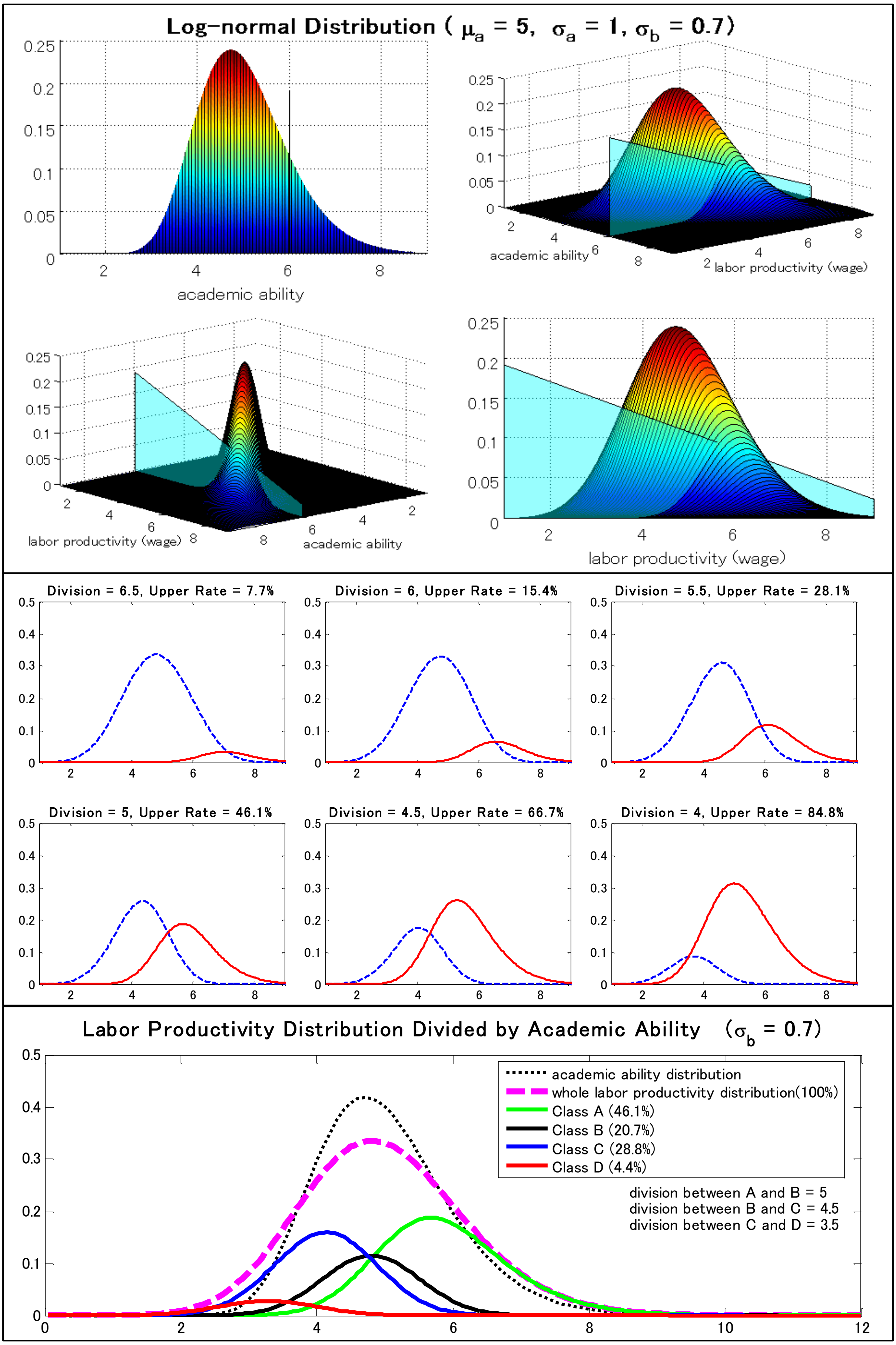
(Standard Deviation between Abilities: 0.2)



Simulation-4

Simulation of Bivariate Distribution

(Standard Deviation between Abilities: 0.7)



Simulation-5