KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE
OF
ECONOMIC RESEARCH

Discussion Paper No.1042

"Long Life-span and Optimal Recurrent Education"

Akira Momota

September 2020

KYOTO UNIVERSITY

KYOTO, JAPAN

Long Life-span and Optimal Recurrent Education

Akira Momota*

College of Economics, Ritsumeikan University

August, 2020

Abstract

This paper theoretically investigates the effect of increased longevity on the years of schooling and work. We consider a situation in which individuals have opportunities for recurrent education by assuming that the transition from schooling to work is reversible. We find that setting aside a period of time for recurrent education is optimal for individuals when the life-span is longer than a certain threshold number of years. As the life-span increases, the total schooling years and the retirement age increase. However, when the life-span becomes so long that recurrent education takes place, the effect of an increase in the active life by one year on the lifetime income is significantly smaller than in the situation where the life-span is less long.

JEL Classification No.: D15, I29, J24, J26

Keywords: Human capital, Life-cycle model, Longevity, Recurrent education.

*College of Economics, Ritsumeikan University, 1-1-1, Noji Higashi, Kusatsu, Shiga 525-8577, Japan (E-mail: momoakir123@yahoo.co.jp).

1 Introduction

Growth in average life expectancy is a common trend that has been observed in many developed countries. Ben-Porath's (1967) pioneering study analyzed the impact of increased life expectancy on investment in human capital and the lifetime labor supply. Using a life-cycle model, Ben-Porath (1967) showed that increased longevity increases the return on education investment and accordingly leads people to devote more time to education. This mechanism, known as the Ben-Porath mechanism, has been studied from theoretical and empirical perspectives.[1]

From a theoretical viewpoint, the effect of an increase in life expectancy on the length of an individual's schooling period for their human capital accumulation and working period has been explored by de la Croix and Licandro (1999), Kalemli-Ozcan et al. (2000), Boucekkine et al. (2003), Soares (2005), Zhang and Zhang (2005), Cervellati and Sunde (2005, 2013), and Cai and Lau (2017), among others. In these studies, an individual's retirement age is treated as exogenous (or ignored). Moreover, this issue is investigated in an endogenous retirement-age setting; some examples include Boucekkine et al. (2002), Echevarría (2004), Ferreira and Pessôa (2007), Hazan (2009), Sánchez-Romero et al. (2016), and Yasui (2016).[2]

In the preceding studies, it is assumed that schooling and working are indivisible[3] and that the transitions from the stage of schooling to work and from the stage of work to retirement are irreversible. As a result, it is commonly postulated that individuals follow an orderly progression through the three life stages: schooling, work, and retirement. In fact, this three-stage view of life is widely observed over the twentieth century, making it an appropriate assumption.

However, in societies experiencing increased life expectancy and a longer active life (in this paper, the active life is defined as the aggregate of the periods engaged in education and work), it is difficult for individuals equipped only with the education that they received during their youth to perform work over their lifetimes. It has been noted that recurrent education, namely, returning to a university or other educational institution after a certain period of work to relearn

[1] From the empirical perspective, the causal effect of life expectancy on investment in human capital has been examined. Although some studies do not support the causality (for example, Acemoglu and Johnson (2006), Lorentzen et al. (2008), and Hazan (2009)), it is supported by others, including Bils and Klenow (2000), Jayachandran and Lleras-Muney (2009), Cervellati and Sunde (2011, 2013), and Oster et al. (2013).

[2] Another related but slightly different line of study is undertaken by Ehrlich and Lui (1991), Zhang et al. (2003), and Zhang and Zhang (2009), who investigate the effect of a rise in the longevity of parents on the education investment for children. In addition, d'Albis et al. (2012) study the effect on the optimal retirement age of a change in mortality at an arbitrary age by abstracting the education investment problem. Nishimura et al. (2018) consider the effects of a rise in longevity on the optimal retirement and education expenditure, rather than years of education.

[3] Ben-Porath (1967) assumes that schooling and work are divisible in every period, and time is allocated to human capital investment and labor supply. In this setting, the life stage does not appear explicitly.

or to gain new knowledge and skills, will become increasingly important. Table 1 supports this view. This table describes the proportion of adults (25–64 years old) who participate in formal education[4] across 22 countries in the OECD.[5] We compare the rate of adult participation in 2007 (in 2005, 2006, or 2008 in some countries) with that in 2012 (or 2015 for some countries). We observe that this level rises in many countries, increasing, on average, by 3 percentage points in this short period. This trend is expected to continue in the future.

This paper theoretically investigates the effect of increased longevity on the years of schooling and work. The novel point is that we consider the situation in which individuals have opportunities for recurrent education.[6]

In line with the existing research, we assume that schooling and work are indivisible, and that individuals can engage only in education or work, but not both, at a given point in time. This indicates that individuals cannot update their skills to the latest versions available while they are engaged in work. In addition, we assume that, with the passage of time, the individuals' existing human capital gradually becomes outdated.[7][8] In effect, individuals consider whether to set aside periods of time for recurrent education during their active life and optimally decide on the timing and length of their recurrent education and work; that is, their life plans, during their active life to maximize their lifetime utility.

In the present model, we use a rectangular survival function, where individuals live with certainty until a certain age, at which time they all die. As pointed out by Wilmoth and

[4]According to the OECD (2011, 2017), formal education is defined as the planned education provided in the system of schools, colleges, universities, and other formal educational institutions. Recurrent education in the present model is closely related to formal education.

[5]Among these people, there are some students who are still completing tertiary education, rather than undertaking recurrent education, even though they are more than 25 years old. However, the proportion of adults (25–64 years old) who participate in formal education is considered to provide a reasonable approximation of the proportion of individuals undertaking recurrent education.

[6]In this regard, Tanaka (2017) is an exception. Tanaka (2017) incorporates recurrent education into a three-period overlapping-generations model, and investigates the effect of a decline in the mortality rate on the human capital. This paper differs from Tanaka (2017) in several points. Tanaka (2017) considers a situation in which people always undertake recurrent education. By contrast, this paper considers the conditions under which recurrent education is undertaken. Moreover, we derive our results in a more rigorous manner.

Tanaka (2017) focuses on the cases in which tertiary education and recurrent education are complements or substitutes. The present paper assumes that the relation between the two is neutral (they are neither complements nor substitutes).

[7]Berk and Weil (2015) observe that this occurs in the case of scientists and medical practitioners. It is also likely to apply to other white-collar workers, who then experience a need for recurrent education.

By contrast, Magnac et al. (2018) assume that the post-schooling human capital investment is divisible.

[8]We can also consider a situation in which work experience raises the individual's productivity, such as an on-the-job training effect or a learning-by-doing effect. If we incorporate such effects into a model, we consider that they attenuate the depreciation of the human capital during working periods (see footnote 14 on this point).

Horiuchi (1999) and Cervellati and Sunde (2013), the observed increase in life expectancy is the result of a process of rectangularization of the survival function. Moreover, Strulik and Vollmer (2013) maintain that recent improvements in life expectancy, since 1970 onward, are driven at least partly by an expanding human life-span (i.e., an increase in the possible maximum age). Taking this into account, it is presumably justifiable, to a certain extent, to model the life-span of individuals with a rectangular survival function when we focus on current and future economic situations.[9] Of course, this assumption assists in making the analysis tractable.

The findings of this analysis reveal that the question of whether recurrent education is an optimal choice for individuals depends crucially on the length of the individual's active life (or life-span). If the active life (or life-span) for individuals is below a certain threshold number of years, the traditional progression of life stages from education to work to retirement will be the optimal life plan for individuals. On the other hand, we find that setting aside a period for recurrent education will be part of the optimal life plan for individuals with an active life that surpasses this threshold number of years. Furthermore, we find outcomes concerning the properties of the optimal recurrent education and of the lifetime income.

First, it is desirable to acquire the latest skills available regardless of whether they are being acquired during one's initial education when young, or later, during a period of recurrent education. In effect, when striving to build human capital during youth, it is not optimal for individuals to lower their standards of effort on the assumption that they will engage in recurrent education at some future point in time. Moreover, if people undertake recurrent education, they should make an effort to acquire cutting-edge skills.

Second, an individual's lifetime income increases for each year added to one's active life. However, when the life-span becomes so long that recurrent education does take place, the effect of an increase in the active life by one year on the lifetime income is significantly smaller compared with the situation where the life-span is less long and the recurrent education does not take place. The implication is that, in an economy with a longer life-span, individuals cannot continue to expect income growth on a level commensurate with their past experience, even if they extend their periods of schooling based on recurrent education (and strive to acquire the latest skills, whether during youth or through recurrent education).

The remainder of this paper is structured as follows. Section 2 introduces the model. We treat the length of the active life (or equivalently, retirement age) as exogenous from Section 2 to Section 4. In Section 3, we analyze the optimal lifetime schedule of the individuals. In

[9] Oxborrow and Turnovsky (2017) build dynamical models by employing various formulations with respect to the survival function. They use the rectangular survival function in one model and the survival function formulated by Boucekkine et al. (2002) in another. They find that the properties of the equilibrium paths that are derived from these alternative models are similar to each other. De la Croix (2017) states that when current data are used, this result is not surprising, as the rectangularization process of the actual survival function is well advanced.

particular, we explore the relationship between the length of the active life and the optimal schedule concerning the recurrent education. Section 4 explores the effect of an increase in the length of the active life on the lifetime income. In Section 5, we extend the model by treating the retirement age as endogenous. We show that the main result obtained in Sections 3 and 4 holds under the extended model, and we examine the optimal retirement age. Finally, Section 6 concludes the paper.

2 Model

Time is continuous and indexed by real numbers.[10] Consider the individuals who just completed their primary education at time 0. For notational convenience, we label their age at time 0 as age 0. The individuals' active life occurs from time 0 to R (that is, from age 0 to R), where R denotes the retirement age. We treat R as an exogenous variable until Section 4. The individuals engage in either (higher) schooling or labor supply activities during their active life. As stated in the Introduction, we assume that schooling and the labor supply are indivisible. In addition, we assume that the transition from schooling to work is reversible, which is the novel point of this study. First, the individuals obtain an education at school and accumulate human capital, after which they leave school and begin work. The knowledge and skills that they learned at school become old-fashioned as time passes or, in other words, their human capital depreciates during the working period. They can update their human capital by entering school again after a certain period of working, which is referred to as recurrent education.

2.1 Life schedule

Let s_i be the length of the i-th period during which a person receives education at a higher learning institution during their lifetime. s_1 denotes tertiary education and s_i $(i \geq 2)$ represents recurrent education. Similarly, let w_i be the length of the i-th period during which a person engages in work during their lifetime. We call the sequence s_i and w_i the *life schedule* (during the active life). For simplicity, we assume that individuals have the opportunity to undertake recurrent education at most once, and the life schedule is expressed as the vector (s_1, w_1, s_2, w_2).[11] The following equation holds as the active life constraint:

[10] As a result of our continuous-time setting, the appearance of the model is somewhat complicated compared with a discrete-time model. However, we employ this setting because the main results (for example, the result concerning thresholds) can be presented in a simpler manner than those obtained with a discrete-time model.

[11] More generally, we could consider a situation where an individual undertakes recurrent education more than once and express the life schedule as the vector $(s_1, w_1, s_2, w_2, \cdots, s_n, w_n)$. From this perspective, we can interpret the present model as considering the situation where individuals choose $s_i = 0$ for $i \geq 3$. This situation

$$s_1 + w_1 + s_2 + w_2 = R. \quad (1)$$

In accordance with the life schedule, the active life $[0, R]$ is divided and the division points are represented as the set $\Delta \equiv \{0,\ t_{w_1},\ t_{s_2},\ t_{w_2},\ R\}$, where, for example, t_{w_1} is the date at which the first working period starts. (Note that $t_{s_1} = 0$ holds.) The set Δ has one-to-one correspondence to the life schedule (s_1, w_1, s_2, w_2); for example, $t_{s_2} - t_{w_1} = w_1$. Thus, we also refer to Δ as the life schedule throughout the analysis.

2.2 Human capital accumulation

Let A_t represent the newest knowledge or skills at time t, evaluated in terms of human capital. In other words, A_t represents the maximum level of human capital available at time t. We assume that A_t grows at an exogenous constant rate g.[12] We normalize A_0 as unity. That is, $A_t = e^{gt}$. Let B_t denote the basic human capital level obtained as a result of primary education. Put differently, B_t is possessed by the people aged 0 at time t. Taking into account that the content of the primary education is influenced by existing knowledge, we assume that it also grows at a rate g. (In this sense, we do not interpret B_t as the innate ability of individuals.) That is, $B_t = B_0 e^{gt}$, where $B_0 < A_0 = 1$.

We represent the human capital of the individuals who have the life schedule Δ by $h_\Delta(t)$, $t \in [0, R]$. The initial human capital level at age 0 is $h_\Delta(0) = B_0$. While they receive education, $h_\Delta(t)$ grows at a constant rate of β until they catch up to the frontier of knowledge, A_t. Once they reach A_t, $h_\Delta(t)$ grows at a pace of g. We assume that $\beta > g$, that is, that the speed of learning existing knowledge is greater than the speed of creating new knowledge.[13] The law of motion of human capital in the schooling periods (s_1 and s_2) is as follows:

$$\dot{h}_\Delta(t) = \begin{cases} \beta h_\Delta(t), & h_\Delta(t) < A_t \\ g h_\Delta(t), & h_\Delta(t) = A_t \end{cases}. \quad (2)$$

After the individuals finish schooling, they work, using their acquired human capital. Because new knowledge continues to emerge in the economy, the acquired knowledge becomes out of

is justified by taking into account the fact that entering school involves some fixed costs. Instead of introducing the fixed costs explicitly into the model, we assume that $n = 2$ to simplify the analysis.

[12] We can consider A_t as being created by researchers, and treating g as exogenous implies that we do not focus on the behavior of the researchers.

[13] We make this assumption as simple as possible to focus on exploring the effect of a longer life-span on the optimal life schedule.

date. Thus, human capital depreciates at the rate δ and the law of motion of human capital in the working periods (w_1 and w_2) is:[14]

$$\dot{h}_{\Delta}(t) = -\delta h_{\Delta}(t). \tag{3}$$

2.3 Labor income

Individuals earn labor income in the working periods, with labor income at time t represented by $y_{\Delta}(t) = \phi h_{\Delta}(t)$. ϕ denotes the productivity of a unit of human capital. Let I_i denote the present value of the labor income earned during the i-th working period (from $t = t_{w_i}$ to $t = t_{w_i} + w_i$). I_i is expressed as:

$$I_i = \int_{t_{w_i}}^{t_{w_i}+w_i} y_{\Delta}(t) e^{-rt} dt = \phi \int_{t_{w_i}}^{t_{w_i}+w_i} h_{\Delta}(t) e^{-rt} dt,$$

where r is the interest rate. We assume that r is exogenous and constant. From (3), we obtain $h_{\Delta}(t) = h_{\Delta}(t_{w_i}) e^{-\delta(t-t_{w_i})}$ during the i-th working period. Thus, we can rewrite I_i as:

$$I_i = \phi h_{\Delta}(t_{w_i}) e^{-rt_{w_i}} \int_{t_{w_i}}^{t_{w_i}+w_i} e^{-(\delta+r)(t-t_{w_i})} dt = \phi h_{\Delta}(t_{w_i}) e^{-rt_{w_i}} \int_{0}^{w_i} e^{-(\delta+r)t} dt. \tag{4}$$

The lifetime income, denoted by I, is expressed as follows:

$$I = \sum_{i=1}^{2} I_i$$

Here, we impose the following assumption regarding the parameters.

Assumption 1 $\beta > r \geq g \geq 0$ *holds.*[15]

$\beta > r$ indicates that the rate of return of the human capital investment is higher than the rate of return of savings, and $r \geq g$ indicates that the potential economic growth rate is lower than the interest rate.

[14] It is natural to consider that work experience also increases human capital. Considering this aspect, the law of motion of $h_{\Delta}(t)$ in the working period will be represented as $\dot{h}_{\Delta}(t) = (\varepsilon - \delta) h_{\Delta}(t)$, where $\varepsilon \geq 0$ represents the work experience effect (for example, the on-the-job training effect). However, this modification does not change the qualitative result as long as $\varepsilon - \delta$ is less than g.

[15] When we explicitly consider ε (see footnote 14), we assume that $\beta > r \geq g \geq \varepsilon - \delta$.

2.4 Optimal problem of individuals

We analyze the optimal behavior of individuals aged 0 at time 0. Similar to Echevarría (2004), Ferreira and Pessôa (2007), Hazan (2009), and Yasui (2016), we use a rectangular survival function; that is, individuals live with certainty until age T, at which time they all die. There is no uncertainty about life expectancy, and the lifetime utility function is expressed as:

$$U = \int_0^T u\left(c_t\right) e^{-\rho t} dt, \tag{5}$$

where $u\left(c_t\right)$ is the instantaneous utility from consumption, and we specify it as $u\left(c_t\right) = \left(c_t^{1-\sigma} - 1\right) / \left(1 - \sigma\right)$ and $\sigma \geq 1$.[16] In Section 5, we extend the model by incorporating the disutility of work or study into the utility function, and we treat the retirement age as endogenous.

We consider the budget constraint of the individuals. They can access the perfect capital market and face no borrowing constraints. We assume that the foregone labor income consists only of the cost of schooling, and that the individuals have no initial assets.[17] In this situation, the lifetime budget constraint is represented as:

$$\int_0^T c_t e^{-rt} dt = I. \tag{6}$$

The individuals choose the consumption profile $\left\{c_t\right\}_{t=0}^T$ and the life schedule $\left(s_1, w_1, s_2, w_2\right)$ to maximize (5). The maximization problem can be decomposed into the following two steps. That is, we directly apply the separation theorem presented by Acemoglu (2009 Theorem 10.1) to the present model.

(Step 1) The individuals choose the life schedule, $\left(s_1, w_1, s_2, w_2\right)$ to maximize I. Let us denote the maximized lifetime income by I^*.

(Step 2) Based on I^*, the individuals maximize lifetime utility by selecting $\left\{c_t\right\}_{t=0}^T$.

The main contribution of the present study is in the Step 1 analysis.[18] We focus on Step 1

[16] The elasticity of intertemporal substitution in consumption is given by $1/\sigma$. This assumption means that the elasticity is less than unity. Havranek et al. (2015) summarize the estimated values of intertemporal substitution for 45 countries and show that in 41 of these countries, the mean elasticity is less than unity.

[17] In Section 5.2, we discuss the effect of the initial asset on the optimal behavior.

[18] Even if we explicitly consider lifetime uncertainty, the separation theorem can be applied to the optimal problem of an individual as long as the complete market is assumed. However, the functional form of I, the objective function of Step 1, varies depending on the assumption on the survival function. For example, it will be shown in Section 3.1 (the case of $s_2 = 0$) that I is hump-shaped. However, this property does not necessarily hold when the survival function is significantly different from the rectangular form, and the result obtained in Step 1 will vary. A similar discussion is applied to the property of I in Section 3.2.

in Section 3, and explore the property of I^* in Section 4.

Once I^* is obtained, we need only solve a standard utility maximization problem concerning the intertemporal choice of consumption. That is, in Step 2, we construct the Lagrangean as:

$$L = \int_0^T u(c_t) e^{-\rho t} dt + \lambda \left(I^* - \int_0^T c_t e^{-rt} dt \right).$$

We obtain the following optimal condition for c_t:

$$u'(c_t) e^{(r-\rho)t} = \lambda. \tag{7}$$

From (7), we obtain $u'(c_0) = u'(c_t) e^{(r-\rho)t}$.[19] Furthermore, noting that $u'(c_t) = c_t^{-\sigma}$, we obtain:

$$c_t = c_0 e^{\frac{r-\rho}{\sigma}t}. \tag{8}$$

From (6) and (8), we obtain:

$$c_0 = \frac{I^*}{\theta(T)}, \tag{9}$$

where $\theta(T)$ is defined as $\theta(T) \equiv \int_0^T e^{-\frac{(\sigma-1)r+\rho}{\sigma}t} dt$.

3 Optimal lifetime schedule

Let us explore the optimal life schedule $(s_1^*, w_1^*, s_2^*, w_2^*)$, which maximizes the lifetime income I. The individuals face the time constraint (1), the law of motion of human capital (2) and (3), and the nonnegative constraints, $s_1 \geq 0$, $w_1 \geq 0$, $s_2 \geq 0$, and $w_2 \geq 0$. In particular, we are interested in the situation where s_2^* is positive.

We solve this problem by dividing it into several steps. To begin, we explore the relationship between the individuals' human capital $h_\Delta(t)$ and A_t. In this regard, we obtain the following lemma under Assumption 1:

Lemma 1 *If $h_\Delta(t) = A_t$ is attained at time t, it is optimal for an individual to leave school at time t.*

---

[19] Taking the logarithm of (7) and differentiating it with respect to t, we obtain the Euler equation:

$$\frac{u''(c_t)}{u'(c_t)} \dot{c}_t + r - \rho = 0.$$

The proof is given in Appendix A. The intuition of this lemma is simple. Once $h_{\Delta}(t) = A_t$ is attained, the pace of the human capital accumulation slows down from β to g, as in (2). The situation of $g \leq r$ implies that an additional year of schooling reduces the lifetime income. (Although it raises the individual's income by $g \times 100\%$ because the receipt of the income is postponed by one year, the present value of the income is discounted by $r \times 100\%$.)[20]

Taking Lemma 1 into account, the law of motion of $h_{\Delta}(t)$, (2), and (3), is rearranged as:

$$\dot{h}_{\Delta}(t) = \begin{cases} \beta h_{\Delta}(t) & (s_1 \text{ and } s_2) \\ -\delta h_{\Delta}(t) & (w_1 \text{ and } w_2) \end{cases} . \tag{10}$$

Let us define $\bar{s}_1$ as $\bar{s}_1 \equiv \frac{b_0}{\beta - g}$, where $b_0 \equiv -\log B_0 > 0$, and note that $h_{\Delta}(t) = A_t$ is attained at $t = \bar{s}_1$. Lemma 1 argues that $s_1 \leq \bar{s}_1$ must hold at the optimum. In relation to this, we obtain the following lemma:

Lemma 2 *(i) $s_2 > 0$ can be the optimal solution only if $s_1 = \bar{s}_1$ holds (that is, if $h_{\Delta}(t) = A_t$ is attained in the first schooling period). (ii) If $s_1 < \bar{s}_1$, $s_2 = 0$ is optimal.*

We describe the proof in Appendix B. To understand the implication of this lemma, let us consider the case where $s_1 < \bar{s}_1$ and $s_2 > 0$. That is, individuals do not study as much as possible in their youth, and they go to school again later in their active life. Lemma 2 argues that such a plan does not maximize the total income; that is, it is not optimal. In Appendix B, we show that if the individuals increase s_1 by Δt and reduce s_2 by Δt (this is feasible under $s_1 < \bar{s}_1$ and $s_2 > 0$), the lifetime income rises. In other words, studying in their early life is more beneficial than studying in their later life if the individuals' productivity of learning, β, is the same across ages[21] because the former involves a longer period over which the individuals can receive returns on their human capital investment than does the latter.

It is useful to note that Lemma 2 suggests that either $s_2 = 0$ or $s_1 = \bar{s}_1$ (or both) holds at the optimum. Accordingly, we examine these situations in turn, and then unite the two cases to obtain the optimal solution. The case where $s_2 = 0$ means that the recurrent education does not take place. Thus, this situation corresponds closely with the situation considered in the preceding studies, where irreversibility of the transition from schooling to work is assumed. Conversely, when $s_1 = \bar{s}_1$, we will explicitly explore the optimal choice of the recurrent education, which is the novel point of this study.

[20] If we consider that A_t grows as a result of the activity of researchers, and if researchers are explicitly introduced into the model, we will have to consider their incentives; that is, the benefit that the researchers receive from successful research.

[21] This result is reinforced when the productivity falls as the age rises.

3.1 The case where $s_2 = 0$ holds

In this case, the active life $[0, R]$ is simply divided into the two subperiods, the schooling period $[0, s_1]$ and the working period $[s_1, R]$. When $s_2 = 0$ holds, the working period is not interrupted by the recurrent education. In other words, we do not have to distinguish between the first and second working periods, and only the total working years $w \equiv w_1 + w_2$ matter. The lifetime income I is represented as $I = \int_{s_1}^{R} y_\Delta(s) e^{-rs} ds$. Based on (4), I can be represented as follows (by noting that $t_{w_1} = s_1$ and $w = R - s_1$):

$$I = \phi h_\Delta(s_1) e^{-rs_1} \int_0^w e^{-(\delta+r)t} dt.$$

From (10), $h_\Delta(s_1) = B_0 e^{\beta s_1}$ holds, so that the above equation is calculated as:

$$I = \phi B_0 e^{(\beta - r)s_1} \int_0^w e^{-(\delta+r)t} dt = \frac{\phi B_0}{\delta + r}\left(e^{(\beta-r)s_1} - e^{-(\delta+r)R} e^{(\beta+\delta)s_1}\right). \tag{11}$$

Here, let us define $\bar{I}$ as:

$$\bar{I} = \frac{\phi B_0}{\delta + r}\left(e^{(\beta-r)\bar{s}_1} - e^{-(\delta+r)R} e^{(\beta+\delta)\bar{s}_1}\right). \tag{12}$$

$\bar{I}$ represents the lifetime income when both $s_1 = \bar{s}_1$ and $s_2 = 0$ hold. We will utilize $\bar{I}$ later.

Taking the logarithm of (11) (of the first equality) and calculating the total derivative, we obtain:

$$\frac{dI}{I} = (\beta - r) ds_1 + \frac{\delta + r}{e^{(\delta+r)w} - 1} dw. \tag{13}$$

The term $(\beta - r) ds_1$ indicates the effect of the accumulation of human capital on I. When s_1 rises by ds_1, human capital increases by $\beta ds_1 \times 100\%$. (Note that $dh_\Delta / h_\Delta = \beta dt$ from (10).) By contrast, the increase in s_1 by ds_1 delays the time when people start working, so that the present value of the income, I, is discounted by $rds_1 \times 100\%$. The term $(\beta - r) ds_1$, which is positive, represents the marginal benefit of human capital investment.

The last term of (13) indicates the effect of the length of the working period on I. It is useful to note that $\frac{\delta+r}{e^{(\delta+r)w}-1}$ is equal to $\frac{1}{I}\frac{\partial I}{\partial w}$. An increase in s_1 by ds_1 reduces w by the same amount ($dw = -ds_1$ from $w = R - s_1$). That is, this term represents the marginal cost of human capital investment.

From (13), we obtain:

$$\frac{1}{I}\frac{dI}{ds_1} = (\beta - r) - \frac{\delta + r}{e^{(\delta+r)(R-s_1)} - 1},$$

and, consequently, we obtain:

$$\frac{dI}{ds_1} \gtreqless 0 \Longleftrightarrow \beta - r \gtreqless \frac{\delta + r}{e^{(\delta+r)(R-s_1)} - 1}. \tag{14}$$

We consider the optimal s_1. Note that the marginal cost of the human capital investment, the right-hand side (RHS) of (14), decreases as R rises. This means that as the length of the active life increases, individuals have an incentive to receive more education. Let us define $\hat{R}_0$ as the level of R that satisfies $dI/ds_1 = 0$ at $s_1 = 0$. From (14), $\hat{R}_0$ is calculated as $\hat{R}_0 = \frac{1}{\delta+r}\log\frac{\beta+\delta}{\beta-r}$. Moreover, we define $\hat{R}_1$ as the level of R that satisfies $dI/ds_1 = 0$ at $s_1 = \bar{s}_1$, and it is obtained as $\hat{R}_1 \equiv \bar{s}_1 + \hat{R}_0 = \frac{b_0}{\beta-g} + \frac{1}{\delta+r}\log\frac{\beta+\delta}{\beta-r}$. As we will see later, $\hat{R}_0$ and, more importantly, $\hat{R}_1$ become the thresholds for the optimal choice of schooling.[22]

Let s_1^+ denote the level of s_1 satisfying $dI/ds_1 = 0$. s_1^+ is calculated as:

$$s_1^+ = R - \hat{R}_0, \tag{15}$$

and we obtain the following result in the case of $s_2 = 0$:

Lemma 3 *Conditional on $s_2 = 0$, I is maximized at (i) $s_1 = 0$ when $R \leq \hat{R}_0$; (ii) $s_1 = s_1^+$ (interior solution) when $\hat{R}_0 < R < \hat{R}_1$; and (iii) $s_1 = \bar{s}_1$ when $R \geq \hat{R}_1$.*

Proof. Because the RHS of (14) is a decreasing function of s_1, $dI/ds_1 \geq 0$ holds if and only if $s_1 \leq s_1^+$. That is, the graph of I is hump-shaped on the s_1-I plane, and its peak is at $s_1 = s_1^+$. Recall that the possible range of s_1 is $0 \leq s_1 \leq \bar{s}_1$. Moreover, we confirm the following from (15):

(i) When $R \leq \hat{R}_0$, $s_1^+ \leq 0$ holds, so that I is maximized at $s_1 = 0$.

(ii) When $\hat{R}_0 < R < \hat{R}_1$, $s_1^+ \in (0, \bar{s}_1)$ holds, so that I is maximized at $s_1 = s_1^+$.

(iii) When $\hat{R}_1 \leq R$, $s_1^+ \geq \bar{s}_1$ holds, so that I is maximized at $s_1 = \bar{s}_1$. ■

Fig. 1 panels (i), (ii), and (iii) illustrate Lemma 3 (i), (ii), and (iii), respectively. From (12), $I = \bar{I}$ holds at $s_1 = \bar{s}_1$. (Note that the level of $\bar{I}$ differs across Fig. 1 panels (i), (ii), and (iii) because $\bar{I}$ changes as R changes.)

Lemma 3 indicates that the length of schooling increases as the length of the active life increases. $\hat{R}_0$ represents the threshold determining whether (further) education takes place. Lemma 3 (i) suggests that if the length of the active life R falls short of the threshold, then the cost of schooling always dominates the benefit and, thus, $s_1 = 0$ is optimal. Lemma 3 (ii) indicates that if R is longer than $\hat{R}_0$, the benefit of schooling outweighs its cost and, thus,

[22] As will be shown in Proposition 1, individuals choose recurrent education when $R > \hat{R}_1$. It can be observed that $\hat{R}_1$ is a decreasing function of δ. (In this regard, differentiating $\hat{R}_1$ with respect to δ yields $\partial\hat{R}_1/\partial\delta = [x + \log(1-x)] / (\delta + r)^2$, where $x \equiv (\delta + r) / (\beta + \delta) \in (0, 1)$. It is immediately confirmed that $\eta(x) \equiv x + \log(1-x) < 0$ for any $0 < x < 1$ because $\eta(0) = 0$ and $\eta'(x) < 0$ hold.) This indicates that the individuals have more incentive to undertake recurrent education as δ rises. We interpret this as indicating that recurrent education assists individuals to boost their depreciated income-earning ability. Incidentally, we confirm that even though $\delta = 0$ (that is, even though δ is absent), the recurrent education may take place.

$s_1 > 0$ is optimal. Furthermore, the optimal years of schooling $(s_1 = s_1^+)$ are an increasing function of R. This result is consistent with the preceding literature.[23]

3.2 The case where $s_1 = \bar{s}_1$ holds

Let us consider the case where R is larger than $\hat{R}_1$. In this case, $s_1 = \bar{s}_1$ holds, as shown in Lemma 3 (iii). By noting that $t_{w_1} = \bar{s}_1$ and $t_{w_2} = \bar{s}_1 + w_1 + s_2$, and that $h_\Delta(t_{w_i}) = B_0 e^{\beta \bar{s}_1}$ and $h_\Delta(t_{w_2}) = B_0 e^{\beta(\bar{s}_1 + s_2)} e^{-\delta w_1}$ hold from (10), we obtain I_1 and I_2 from (4):

$$I_1 = \phi B_0 e^{(\beta - r)\bar{s}_1} \int_0^{w_1} e^{-(\delta + r)t} dt = \frac{\phi B_0}{\delta + r} e^{(\beta - r)\bar{s}_1} \left(1 - e^{-(\delta + r)w_1}\right), \tag{16}$$

$$I_2 = \phi B_0 e^{(\beta - r)(\bar{s}_1 + s_2)} e^{-(\delta + r)w_1} \int_0^{w_2} e^{-(\delta + r)t} dt = \frac{\phi B_0}{\delta + r} e^{(\beta - r)(\bar{s}_1 + s_2) - (\delta + r)w_1} \left(1 - e^{-(\delta + r)w_2}\right). \tag{17}$$

Applying a procedure similar to the one used to derive (13) from (11), we derive the following equations from (16) and (17). Noting that $\bar{s}_1$ is constant (so far as the exogenous parameters do not change), we obtain:

$$\frac{dI_1}{I_1} = \frac{\delta + r}{e^{(\delta + r)w_1} - 1} dw_1, \tag{18}$$

$$\frac{dI_2}{I_2} = \left[(\beta - r)\, ds_2 - (\delta + r)\, dw_1\right] + \frac{\delta + r}{e^{(\delta + r)w_2} - 1} dw_2. \tag{19}$$

Let us interpret (18) and (19). As mentioned, when we interpret (13), the term $\frac{\delta + r}{e^{(\delta + r)w_i} - 1} dw_i$, which appears on the RHS of both (18) and (19), represents the effect of the length of the i-th working period w_i on I_i. The term in the square brackets in (19) represents the effect of a change in the human capital on I_2. An increase of ds_2 years in s_2 raises the human capital by βds_2 units, whereas a dw_1 increase in w_1 depreciates the human capital by δdw_1 units. Thus, the net change of the human capital at $t = t_{w_2}$ is $\beta ds_2 - \delta dw_1$. In addition, the time when w_2 starts, $t = t_{w_2}$, is delayed by $ds_2 + dw_1$, which discounts the present value of I_2 by $r\,(ds_2 + dw_1)$.

The lifetime income is expressed as:

$$I = I_1 + I_2 = \frac{\phi B_0}{\delta + r} \left\{ e^{(\beta - r)\bar{s}_1} \left(1 - e^{-(\delta + r)w_1}\right) + e^{(\beta - r)(\bar{s}_1 + s_2) - (\delta + r)w_1} \left(1 - e^{-(\delta + r)w_2}\right) \right\}.$$

The optimization problem is to maximize I, subject to the active life constraint:

$$\bar{s}_1 + w_1 + s_2 + w_2 = R, \tag{20}$$

[23] In particular, this result is very close to Hazan's (2009) Proposition 2.

the constraint concerning the upper bound of human capital, $h_{\Delta}(t_{w_2}) \leq A_{t_{w_2}}$, and the nonnegative constraints $w_1 \geq 0$, $s_2 \geq 0$, and $w_2 \geq 0$. Here, substituting (20) into the above equation to eliminate s_2, I is expressed as a function of w_1 and w_2, $I(w_1, w_2)$:

$$I(w_1, w_2) = \frac{\phi B_0}{\delta + r}\left\{e^{(\beta - r)\bar{s}_1}\left(1 - e^{-(\delta + r)w_1}\right) + e^{(\beta - r)(R - w_2) - (\beta + \delta)w_1}\left(1 - e^{-(\delta + r)w_2}\right)\right\}. \tag{21}$$

The constraint $h_{\Delta}(t_{w_2}) \leq A_{t_{w_2}}$ is equivalent to $B_0 e^{\beta(\bar{s}_1 + s_2)} e^{-\delta w_1} \leq e^{g(\bar{s}_1 + s_2 + w_1)}$. Taking the logarithm and using (20), we obtain:

$$\frac{\beta + \delta}{\beta - g} w_1 + w_2 \geq R - \bar{s}_1. \tag{22}$$

Furthermore, we rewrite $s_2 \geq 0$ using (20) as:

$$w_1 + w_2 \leq R - \bar{s}_1. \tag{23}$$

In sum, the optimization problem is to maximize (21) with respect to $(w_1, w_2) \in \Re_+^2$ subject to (22) and (23). Fig. 2 depicts the relationship between (22) and (23) on the (w_1, w_2) plane. The two lines intersect at $(w_1, w_2) = (0, R - \bar{s}_1)$ (Point A in the figure), and (22) is steeper than (23). The shaded triangle area represents the region in which both (22) and (23) are satisfied.

When (w_1, w_2) is on the border of (23), we obtain the following:

Lemma 4 *$I(w_1, w_2) = \bar{I}$ holds when (w_1, w_2) is on the border of (23).*

We can prove Lemma 4 simply by substituting $w_2 = R - \bar{s}_1 - w_1$ into (21) and making some arrangement. Noting that $s_2 = 0$ holds on the border of (23), the implication of Lemma 4 is straightforward. As stated in Section 3.1, when $s_2 = 0$, we do not have to distinguish between w_1 and w_2, and only w matters, which is equal to $R - \bar{s}_1$. As seen in (12), the lifetime income when $s_1 = \bar{s}_1$ and $s_2 = 0$ is $\bar{I}$.

Because $I(w_1, w_2)$ is not necessarily a concave function, we solve the maximization problem by the following steps.

(Step I) Taking w_2 as given, we seek a value of w_1 that maximizes $I(w_1, w_2)$. The solution is expressed as a function of w_2, $\Psi(w_2)$.

(Step II) We solve w_2, which maximizes $J(w_2) \equiv I(\Psi(w_2), w_2)$. We denote the solution by w_2^*. The optimal w_1 is obtained by $w_1^* = \Psi(w_2^*)$.

3.2.1 Property of $\Psi(w_2)$: Step I problem

First, treating w_2 as given, we maximize $I(w_1, w_2)$ with respect to w_1. To begin, for notational convenience, we introduce $\underline{\Psi}(w_2)$ and $\overline{\Psi}(w_2)$, and rewrite the constraints (22) and (23), respectively, as:

$$w_1 \geq \underline{\Psi}(w_2) \equiv \frac{\beta - g}{\beta + \delta}(R - \bar{s}_1 - w_2), \tag{24}$$

$$w_1 \leq \overline{\Psi}(w_2) \equiv R - \bar{s}_1 - w_2. \tag{25}$$

The solution $\Psi(w_2)$ is represented as:

$$\Psi(w_2) = \arg\max_{\underline{\Psi}(w_2) \leq w_1 \leq \overline{\Psi}(w_2)} I(w_1, w_2).$$

Differentiating (21) partially with respect to w_1 yields:

$$\frac{\partial I(w_1, w_2)}{\partial w_1} = \frac{\phi B_0}{\delta + r}\left\{(\delta + r)e^{(\beta - r)\bar{s}_1}e^{-(\delta + r)w_1} - (\beta + \delta)e^{-(\beta + \delta)w_1}e^{(\beta - r)(R - w_2)}\left(1 - e^{-(\delta + r)w_2}\right)\right\}. \tag{26}$$

Let us interpret (26). It is useful to note that:

$$dI = dI_1 + dI_2 = \frac{dI_1}{I_1} \cdot I_1 + \frac{dI_2}{I_2} \cdot I_2. \tag{27}$$

Given $dw_2 = 0$, and using (18) and (19), (27) is represented as:

$$\left.\frac{dI}{dw_1}\right|_{dw_2 = 0} = \frac{\delta + r}{e^{(\delta + r)w_1} - 1}I_1 + \left((\beta - r)\frac{ds_2}{dw_1} - (\delta + r)\right)I_2,$$

Furthermore, we obtain $ds_2 = -dw_1$ from the active life constraint (20). That is, an increase in the working period w_1 reduces the schooling period s_2 by the same amount, given that w_2 is constant. Thus, the above equation is rewritten as:

$$\left.\frac{dI}{dw_1}\right|_{dw_2 = 0} = \frac{\delta + r}{e^{(\delta + r)w_1} - 1}I_1 - (\beta + \delta)I_2. \tag{28}$$

It is immediately confirmed that the first and second terms of the RHS of (26) correspond, respectively, to the first and second terms of the RHS of (28). The first term represents the positive effect of increasing w_1 on I and the second term represents the negative effect of increasing w_1. The term $e^{-(\delta + r)w_1}$ in the first term of (26) indicates that the positive effect declines as w_1 rises, and the speed of the decline is $\delta + r$. Observe that the negative effect (the second term) also decreases, and its speed is $\beta + \delta$. Because $\beta > r$ holds under Assumption

1, the positive effect declines more slowly than the negative effect does as w_1 rises. In other words, the positive effect becomes (relatively) larger than the negative effect as w_1 rises. The following equation, derived from (26), clearly conveys this point:

$$\frac{\partial I(w_1, w_2)}{\partial w_1} \gtreqless 0 \Longleftrightarrow e^{(\beta-r)w_1} \gtreqless \frac{\beta+\delta}{\delta+r} e^{(\beta-r)(R-\bar{s}_1-w_2)} \left(1-e^{-(\delta+r)w_2}\right). \tag{29}$$

The following lemma summarizes the implication of (29):

Lemma 5 *Taking w_2 as given, the level of w_1 that maximizes $I(w_1, w_2)$ is obtained as a corner solution. That is, $\Psi(w_2)$ is either on $\underline{\Psi}(w_2)$ or on $\overline{\Psi}(w_2)$.*

Proof. The left-hand side (LHS) of (29) is an increasing function of w_1 under Assumption 1, and its value is one when $w_1 = 0$. Because w_2 is treated as given, the value of the RHS is constant. If it is less than one, the LHS is always greater than the RHS and, thus, $\frac{\partial I(w_1,w_2)}{\partial w_1} > 0$. In this case, the optimal w_1 is on $\overline{\Psi}(w_2)$.

On the other hand, if the RHS is greater than one, a value of w_1 that satisfies $\frac{\partial I(w_1,w_2)}{\partial w_1} = 0$ exists and is unique, and we denote it by w_1^-. It is immediately confirmed that when $w_1 < w_1^-$, $\frac{\partial I(w_1,w_2)}{\partial w_1} < 0$, and vice versa. Thus, $I(w_1, w_2)$ has a minimum value at $w_1 = w_1^-$. This indicates that the $\Psi(w_2)$ is characterized as the corner solution; that is, $\Psi(w_2)$ is either on $\underline{\Psi}(w_2)$ or on $\overline{\Psi}(w_2)$. ■

Lemma 5 indicates the features of the optimal recurrent education based on the assumption that $\beta > r$. $w_1 = \underline{\Psi}(w_2)$ means that $h_\Delta(t) = A_t$ is attained in the recurrent education period, and $w_1 = \overline{\Psi}(w_2)$ means that $s_2 = 0$. If individuals go to school to undertake recurrent education, they should acquire cutting-edge skills. Otherwise, they should not undertake recurrent education.

More details on the property of $\Psi(w_2)$ (in particular, on which border $\Psi(w_2)$ exists) will be provided in Section 3.2.2.

3.2.2 Property of $J(w_2)$: Step II problem

Substituting $\Psi(w_2)$ into $I(w_1, w_2)$, we represent I as a function of w_2, and we define $J(w_2)$ as $J(w_2) \equiv I(\Psi(w_2), w_2)$. From Lemma 5, $J(w_2)$ is expressed as:

$$J(w_2) = \max\left\{\underline{J}(w_2), \overline{J}(w_2)\right\},$$

where $\underline{J}(w_2)$ and $\overline{J}(w_2)$ are defined as, respectively:[24]

$$\begin{aligned}\underline{J}(w_2) &\equiv I\left(\underline{\Psi}(w_2), w_2\right)\\ &= \frac{B_0\phi}{\delta+r}\left\{e^{(\beta-r)\bar{s}_1}\left(1-e^{-(\delta+r)\underline{\Psi}(w_2)}\right)+e^{(\beta-g)\bar{s}_1}e^{-(r-g)(R-w_2)}\left(1-e^{-(\delta+r)w_2}\right)\right\}, \quad (30)\\ \overline{J}(w_2) &\equiv I\left(\overline{\Psi}(w_2), w_2\right) = \frac{B_0\phi}{\delta+r}\left(e^{(\beta-r)\bar{s}_1}-e^{-(\delta+r)R}e^{(\beta+\delta)\bar{s}_1}\right) = \overline{I}. \quad (31)\end{aligned}$$

The last equality of (31) is immediately confirmed from (12), and this is the restatement of Lemma 4.

Next, let us consider the property of $\underline{J}(w_2)$. In this case, (24) holds as an equality:

$$w_1 = \underline{\Psi}(w_2) = \frac{\beta-g}{\beta+\delta}\left(R-\bar{s}_1-w_2\right), \quad (32)$$

and the following equation is obtained from the active life constraint (20):

$$s_2 = \frac{g+\delta}{\beta+\delta}\left(R-\bar{s}_1-w_2\right) = \frac{g+\delta}{\beta-g}w_1. \quad (33)$$

Note that the first and the second terms of (30) correspond to I_1 and I_2, respectively.[25] Differentiating (30) with respect to w_2 yields:

$$\begin{aligned}\underline{J}'(w_2) = \frac{\phi B_0}{\delta+r}\{&e^{(\beta-r)\bar{s}_1}(\delta+r)e^{-(\delta+r)\underline{\Psi}(w_2)}\underline{\Psi}'(w_2)\\ &+e^{(\beta-g)\bar{s}_1}e^{-(r-g)(R-w_2)}\left[(r-g)\left(1-e^{-(\delta+r)w_2}\right)+(\delta+r)e^{-(\delta+r)w_2}\right]\}. \quad (34)\end{aligned}$$

Let us interpret this. Using (18) and (19), (27) is represented as:

$$\frac{dI}{dw_2} = \frac{\delta+r}{e^{(\delta+r)w_1}-1}\frac{dw_1}{dw_2}I_1+\left[(\beta-r)\frac{ds_2}{dw_2}-(\delta+r)\frac{dw_1}{dw_2}\right]I_2+\frac{\delta+r}{e^{(\delta+r)w_2}-1}I_2. \quad (35)$$

By applying (32) and (33) to the above equation (35), we confirm that the first, second, and third terms of (34) correspond to the first, second, and third terms of (35), respectively. (In

[24] When we explore the effect of a change in R on I in the next section, we explicitly express the parameter R as an argument, and represent the RHSs of (30) and (31) as $\underline{J}(w_2, R)$ and $\overline{J}(w_2, R)$, respectively.

[25] When (32) and (33) hold, it is confirmed that the following holds:

$$(\beta-g)\bar{s}_1-(r-g)(R-w_2) = (\beta-r)(\bar{s}_1+s_2)-(\delta+r)w_1.$$

Thus, $e^{(\beta-g)\bar{s}_1}e^{-(r-g)(R-w_2)}$ in (30) is equal to $e^{(\beta-r)(\bar{s}_1+s_2)-(\delta+r)w_1}$ in (17).

particular, note that $(\beta - r)\frac{ds_2}{dw_2} - (\delta + r)\frac{dw_1}{dw_2}$ is equal to $r - g$.) We can interpret (35) based on (18) and (19). The last term $\frac{\delta + r}{e^{(\delta + r)w_2} - 1} I_2$ indicates the direct effect of an increase in the length of the second working period w_2 on I_2. At the same time, as shown by (32) and (33), a rise in w_2 reduces w_1 and s_2, which also affects I (indirect effects). The first term $\frac{\delta + r}{e^{(\delta + r)w_1} - 1}\frac{dw_1}{dw_2} I_1$ indicates the effect of a decrease in w_1 on I_1. The second term $(\beta - r)\frac{ds_2}{dw_2} - (\delta + r)\frac{dw_1}{dw_2} (= r - g)$ represents the effect of a change in the human capital stock on I_2, as discussed earlier.

The first term of (35) is negative because $\frac{dw_1}{dw_2} = \underline{\Psi}'(w_2) < 0$, whereas the second term is positive on the assumption that $r > g$, and the third term is also positive. Let us investigate the (relative) strength of these negative and positive effects. We rearrange (34) as:

$$\begin{aligned}
\underline{J}'(w_2) &= \frac{\phi B_0}{\delta + r} e^{(\beta - g)\bar{s}_1} e^{-(r-g)(R-w_2)} \{ -\frac{\beta - g}{\beta + \delta}(\delta + r) e^{\left(r - g - \frac{\beta - g}{\beta + \delta}(\delta + r)\right)(R - w_2 - \bar{s}_1)} \\
&\quad + (r - g)\left(1 - e^{-(\delta + r)w_2}\right) + (\delta + r) e^{-(\delta + r)w_2} \} \\
&= \frac{\phi B_0}{\delta + r} e^{(\beta - g)\bar{s}_1} e^{-(r-g)(R-w_2)} \{ -\frac{\beta - g}{\beta + \delta}(\delta + r) e^{-\frac{(\beta - r)(g + \delta)}{\beta + \delta}(R - w_2 - \bar{s}_1)} + (r - g) + (g + \delta) e^{-(\delta + r)w_2} \}.
\end{aligned} \tag{36}$$

The last equality is obtained by noting that $r - g - \frac{\beta - g}{\beta + \delta}(\delta + r)$ is equal to $-\frac{(\beta - r)(g + \delta)}{\beta + \delta}$. Let us denote the terms in the curly brackets of (36) by $\kappa(w_2)$:

$$\kappa(w_2) \equiv -\frac{\beta - g}{\beta + \delta}(\delta + r) e^{-\frac{(\beta - r)(g + \delta)}{\beta + \delta}(R - w_2 - \bar{s}_1)} + (r - g) + (g + \delta) e^{-(\delta + r)w_2}.$$

From (36), it can be seen that the sign of $\underline{J}'(w_2)$ is equal to the sign of $\kappa(w_2)$. That is:

$$\underline{J}'(w_2) \gtreqless 0 \Longleftrightarrow \kappa(w_2) \gtreqless 0 \Longleftrightarrow \frac{\beta - g}{\beta + \delta}(\delta + r) e^{-\frac{(\beta - r)(g + \delta)}{\beta + \delta}(R - w_2 - \bar{s}_1)} \lesseqgtr (r - g) + (g + \delta) e^{-(\delta + r)w_2}. \tag{37}$$

The LHS of (37) corresponds to the negative effect (the first term of (35)) and the RHS of (37) corresponds to the positive effects (the sum of the second and the third terms). It is confirmed that as w_2 increases, the LHS increases, whereas the RHS decreases, which means that the negative effect dominates the positive effect when w_2 is large, and vice versa. In other words, $\kappa'(w_2) < 0$ holds. Let w_2^+ denote a value of w_2 that satisfies $\kappa(w_2) = 0$. The above property indicates that $\kappa(w_2) > 0$ ($\Longleftrightarrow$ $\underline{J}'(w_2) > 0$) holds when $w_2 < w_2^+$ and $\kappa(w_2) < 0$ ($\Longleftrightarrow$ $\underline{J}'(w_2) < 0$) holds when $w_2 > w_2^+$.

As seen in Fig. 2, the range of w_2 that satisfies the constraints is $[0, R - \bar{s}_1]$. We obtain the following property concerning $\underline{J}(w_2)$ and $\overline{J}(w_2)$ in the range of $w_2 \in [0, R - \bar{s}_1]$:

Lemma 6 *(i)* $\overline{J}(w_2) = \overline{I}$ *holds for any* $w_2 \in [0, R - \bar{s}_1]$.

(ii) $\underline{J}(0) < \overline{I}$ *and* $\underline{J}(R - \bar{s}_1) = \overline{I}$ *holds.*

(iii) When $R \leq \hat{R}_1$, $\underline{J}(w_2) \leq \overline{I}$ holds for any $w_2 \in [0, R - \bar{s}_1]$. That is, $\underline{J}(w_2)$ is maximized at $w_2 = R - \bar{s}_1$.

(iv) When $R > \hat{R}_1$, w_2^+ exists in $w_2^+ \in (0, R - \bar{s}_1)$ and $\underline{J}(w_2)$ is maximized at $w_2 = w_2^+$.

The proof is given in Appendix C. Based on Lemma 6, we depict $J(w_2)$ (the solid curve) as well as $\underline{J}(w_2)$ and $\overline{J}(w_2)$ (the dotted curves) in Fig. 3. Panel (i) illustrates the case where $R \leq \hat{R}_1$. From Lemma 6 (i) and (iii), we obtain $J(w_2) = \overline{I}$ for any $w_2 \in [0, R - \bar{s}_1]$ and, in regard to $\Psi(w_2)$, we obtain $w_1 = \Psi(w_2) = \overline{\Psi}(w_2) = R - \bar{s}_1 - w_2$.

Panel (ii) illustrates the case where $R > \hat{R}_1$. $\overline{J}(w_2)$ is continuous and, from Lemma 6 (ii) and (iv), it is guaranteed that there exists a $\hat{w}_2 \in \left(0, w_2^+\right)$ that satisfies $\overline{J}(\hat{w}_2) = \overline{I}$. Using it, $J(w_2)$ is expressed as:

$$J(w_2) = \begin{cases} \overline{I} & w_2 \in [0, \hat{w}_2) \\ \underline{J}(w_2) & w_2 \in [\hat{w}_2, R - \bar{s}_1] \end{cases} . \tag{38}$$

Furthermore, as regards $\Psi(w_2)$, we obtain:

$$w_1 = \Psi(w_2) = \begin{cases} \overline{\Psi}(w_2) = R - \bar{s}_1 - w_2 & w_2 \in [0, \hat{w}_2) \\ \underline{\Psi}(w_2) = \frac{\beta - g}{\beta + \delta}(R - \bar{s}_1 - w_2) & w_2 \in [\hat{w}_2, R - \bar{s}_1] \end{cases} . \tag{39}$$

$J(w_2)$ is a continuous function at $w_2 = \hat{w}_2$, whereas $\Psi(w_2)$ is discontinuous at this point.

Remember that the maximum of $J(w_2)$ corresponds to the maximum level of I conditional on $s_1 = \bar{s}_1$. We obtain the following result:

Lemma 7 *Conditional on $s_1 = \bar{s}_1$, the following hold.*

(i) When $R \leq \hat{R}_1$, the maximum level of I is $\overline{I}$. (ii) When $R > \hat{R}_1$, I is maximized at w_2^+, which satisfies $\underline{J}'\left(w_2^+\right) = 0$ (an interior solution).

3.3 Derivation of the optimal lifetime schedule

By combining Lemmas 3 and 7, we obtain the optimal lifetime schedule $(s_1^*, w_1^*, s_2^*, w_2^*)$. As mentioned earlier, when $s_2 = 0$, discriminating between w_1 and w_2 is not significant, and only the total working years $w \equiv w_1 + w_2$ matter. We obtain the following result:

Proposition 1 *(i) When $R \leq \hat{R}_0$, $s_1^* = 0$ and $s_2^* = 0$ hold and, in regard to the total working years, $w^* = R$ holds.*

(ii) When $\hat{R}_0 < R \leq \hat{R}_1$, $s_1^* = s_1^+ \equiv R - \hat{R}_0$, $s_2^* = 0$, and $w^* = \hat{R}_0$ hold.

(iii) When $R > \hat{R}_1$, $s_1^* = \bar{s}_1$ and $s_2^* > 0$ hold. In regard to the working years, $w_1^* > 0$ and $w_2^* = w_2^+ > 0$ hold.

Proof. (i) When $R \leq \hat{R}_0$, Fig. 1 (i) (that is, Lemma 3 (i)) represents I conditional on $s_2 = 0$, and Fig. 3 (i) (that is, Lemma 7 (i)) represents I conditional on $s_1 = \bar{s}_1$. By comparing these figures, we see that I is maximized when $s_1^* = 0$ and $s_2 = 0$. Furthermore, from the active life constraint (1), $w^* = R$ holds.

(ii) When $\hat{R}_0 < R < \hat{R}_1$, we compare Fig. 1 (ii) with Fig. 3 (i) and, when $R = \hat{R}_1$, we compare Fig. 1 (iii) with Fig. 3 (i). Consequently, we see that I is maximized when $s_1^* = s_1^+ \equiv R - \hat{R}_0$ and $s_2 = 0$. In this case, $w^* = \hat{R}_0$ holds from (1).

(iii) When $R > \hat{R}_1$, by comparing Fig. 1 (iii) with Fig. 3 (ii), we can see that I is maximized when $s_1 = \bar{s}_1$ and $w_2 = w_2^+$. Because $w_2^+ \in (\hat{w}_2, R - \bar{s}_1)$ holds, as shown in (39), (32) and (33) hold:

$$w_1 = \underline{\Psi}(w_2) = \frac{\beta - g}{\beta + \delta}(R - \bar{s}_1 - w_2),$$
$$s_2 = \frac{g + \delta}{\beta - g} w_1 = \frac{g + \delta}{\beta + \delta}(R - \bar{s}_1 - w_2).$$

When $w_2^* = w_2^+$, it is confirmed that $w_1^* > 0$ holds from (32), and $s_2^* > 0$ holds from (33). ■

This proposition indicates that the recurrent education takes place when $R > \hat{R}_1$. Let us discuss this case. Note that $w_2^* = w_2^+$ indicates that (37) holds as an equality. Combining it with (32) yields the following equation:

$$\frac{(\beta - g)(\delta + r)}{\beta + \delta} e^{-\frac{(\beta - r)(g + \delta)}{\beta - g} w_1} = r - g + (g + \delta) e^{-(\delta + r) w_2}. \tag{40}$$

s_2^*, w_1^*, and w_2^* are obtained from (32), (33), and (40), respectively, and we can characterize the optimal schedule by examining these three equations. Fig. 4 depicts (32) and (40) on the (w_1, w_2) plane. The graph of (32) is a downward-sloping line and the w_2-intercept is $w_2 = R - \bar{s}_1$. Conversely, the graph of (40) is an upward-sloping curve, and the w_2-intercept is $w_2 = \frac{1}{\delta + r} \log \frac{\beta + \delta}{\beta - r} = \hat{R}_0$. (Moreover, it is convex under Assumption 1, as we prove in Appendix D.) Thus, when $R > \hat{R}_0 + \bar{s}_1 = \hat{R}_1$, we observe that the intersection exists in the interior, that is, $(w_1^*, w_2^*) \in \Re^2_{++}$, and is unique.

Moreover, by drawing the minus 45 degree line that passes Point A (the dotted line in Fig. 4), we can represent the optimal life schedule visually on the w_2 axis. First, the length of the line segment OB is $w_1^* + w_2^*$. Next, the length BC is equal to s_2^*, which is seen from (20). Finally, the length OR is equal to R and depicting Point R on the w_2 axis, we obtain s_1^* as the segment CR. That is, OR is divided into four regions, s_1^*, s_2^*, w_1^*, and w_2^*.

3.4 Effects of an increase in R on the optimal schedule

When the value of the parameter R changes, the optimal life schedule changes. Let us explore the effect of a rise in R on the optimal life schedule. In particular, we are interested in the case where $R > \hat{R}_1$. Using (32) and (40), we obtain the following result.[26]

Proposition 2 *Suppose that $R > \hat{R}_1$. When R rises marginally, s_1^* remains at $s_1^* = \bar{s}_1$, and w_1^*, $w_2^* \left(= w_2^+\right)$, and s_2^* increase.*

Proof. First, $s_1^* = \bar{s}_1$ is directly obtained from Proposition 1 (iii). We observe that the graph of (32) shifts rightward as R increases, and that (40) remains unchanged. As seen in Fig. 5, the intersection moves up and to the right, from Point A to Point A'. Thus, both w_1^* and w_2^* rise. Moreover, taking (33) into account, we obtain that s_2^* increases. ■

This proposition indicates the optimal timing for the recurrent education and its optimal length. We observe that when R is close to $\hat{R}_1$, w_1^* is zero,[27] and, from (33), s_2^* also becomes zero. As R rises, individuals work longer in the first working period and, thus, delay the timing of recurrent education. At the same time, the duration of the recurrent education, s_2, becomes longer. We note that both the total schooling period, s^*, and the total working period, w^*, increase.

4 Relationship between R and lifetime income

Based on the optimal schedule, we derive the maximized lifetime income I^*. When R varies, the optimal lifetime schedule changes, as examined in Section 3.4, and, thus, I^* changes. That is, I^* is a function of R. Thus, let us represent I^* as $I^* = I(R)$.

Let us explore the property of $I(R)$. When $R \leq \hat{R}_1$, $I(R)$ is obtained from (11) and Lemma 3 (i) and (ii), and when $\hat{R}_1 < R$, it is obtained from Lemma 7. That is, $I(R)$ is expressed as:

$$I(R) = \begin{cases} \phi B_0 \int_0^R e^{-(\delta+r)t} dt, & 0 \leq R \leq \hat{R}_0 \\ \phi B_0 e^{(\beta - r)\left(R - \hat{R}_0\right)} \int_0^{\hat{R}_0} e^{-(\delta+r)t} dt, & \hat{R}_0 < R \leq \hat{R}_1 \\ \underline{J}\left(w_2^+, R\right), & \hat{R}_1 < R \end{cases}, \tag{41}$$

[26]When $R < \hat{R}_1$, the following results are immediately confirmed from Proposition 1 (i) and (ii), respectively: (i) when $R < \hat{R}_0$, only the total working years $w^* \equiv w_1^* + w_2^*$ increase, and the total schooling years $s^* \equiv s_1^* + s_2^*$ remain zero as R rises; and (ii) when $\hat{R}_0 < R < \hat{R}_1$, s^* increases, whereas w^* does not change as R rises.

[27]As can be observed from Fig. 5, when R approaches $\hat{R}_1$ (from above), (32) shifts to the left and the intersection of (32) and (40) approaches $(w_1, w_2) = (0, \hat{R}_0)$.

where:[28]

$$\underline{J}\left(w_2^+, R\right) = \frac{B_0\phi}{\delta + r}\left\{e^{(\beta-r)\bar{s}_1}\left(1 - e^{-\frac{\beta-g}{\beta+\delta}(\delta+r)\left(R-\bar{s}_1-w_2^+\right)}\right) + e^{(\beta-g)\bar{s}_1}e^{-(r-g)\left(R-w_2^+\right)}\left(1 - e^{-(\delta+r)w_2^+}\right)\right\}.$$

It is immediately confirmed that $\lim_{R\to\hat{R}_0-0} I(R) = \lim_{R\to\hat{R}_0+0} I(R)$ and $\lim_{R\to\hat{R}_1-0} I(R) = \lim_{R\to\hat{R}_1+0} I(R)$ hold and, thus, that $I(R)$ is a continuous function for $R \geq 0$. Furthermore, we obtain the following proposition:

Proposition 3 *(i) $I(R)$ is continuously differentiable (that is, $I(R)$ is of class C^1).*

(ii) $I(R)$ is an increasing function of R, that is, $I'(R) > 0$.

(iii) We obtain the second derivative of $I(R)$ as:

$$I''(R) = \begin{cases} -(\delta + r) I'(R), & 0 < R < \hat{R}_0 \\ (\beta - r) I'(R), & \hat{R}_0 < R < \hat{R}_1 \\ -\left[(r-g) + (g+\delta)\frac{dw_2^+}{dR}\right] I'(R), & \hat{R}_1 < R \end{cases}. \tag{42}$$

That is, $I(R)$ is strictly concave when $0 \leq R \leq \hat{R}_0$ and $\hat{R}_1 \leq R$, whereas it is strictly convex when $\hat{R}_0 \leq R \leq \hat{R}_1$.

The proof of Proposition 3 is given in Appendix E. Proposition 3 (ii) argues that the lifetime income increases as the active life becomes longer, which is a natural result. We obtain an interesting finding in Proposition 3 (iii). We depict $I(R)$ in Fig. 6. When $0 \leq R \leq \hat{R}_0$, individuals choose no education, that is, $s_1 = s_2 = 0$ (refer to Proposition 1 (i)). In this case, an increase in R raises $I(R)$, but $I'(R)$ declines. Conversely, when $\hat{R}_0 \leq R \leq \hat{R}_1$ and $s_1 > 0$ and $s_2 = 0$ is chosen (refer to Proposition 1 (ii)), $I'(R)$ increases. This result indicates that the human capital investment significantly contributes to an increase in the lifetime income when R is smaller than $\hat{R}_1$.

When R increases further and $R > \hat{R}_1$ holds, individuals choose $s_2 > 0$ (refer to Proposition 1 (iii)). Proposition 3 (iii) argues that $I'(R)$ declines again as R increases, although the human capital investment increases.

Eq. (42) suggests the reason why this happens. The main reason is $r \geq g$. As discussed in Lemmas 2 and 5, individuals update their skills twice, first at the end of the first schooling period and then again at the end of their recurrent education. In this case, although the individuals accumulate human capital at a speed of β during the schooling periods, the effect of the human capital accumulation on $I(R)$ is eventually determined by the growth rate of the cutting-edge knowledge, g, which is lower than r.

[28] $\underline{J}\left(w_2^+, R\right)$ is expressed as $\underline{J}\left(w_2^+\right)$ in the previous section. Refer to footnote 24 on this point. Note also that w_2^+ is a function of R.

The implication is that, in an economy with a longer life-span, individuals cannot expect income growth to continue on a level commensurate with past experience, even if they extend their periods of schooling based on recurrent education (and strive to acquire the latest skills, whether during youth or through recurrent education) and work longer.

5 An extension: Endogenous retirement age

We extend the model by incorporating the disutility of work or study into the utility function, and we treat the retirement age R as an endogenous variable. We show that the main result obtained in Sections 3 and 4 still hold under the extended model. The utility function is given as:

$$U = \int_0^T u(c_t) e^{-\rho t} dt - \int_0^R v(t, T) dt, \tag{43}$$

where $v(t, T)$ is the instantaneous disutility of work or study at age t, evaluated at time 0.[29][30] We impose the following assumption on $v(t, T)$.

Assumption 2 *(i)* $v_1(t, T) \equiv \partial v(t, T)/\partial t > 0$ *and* $v_2(t, T) \equiv \partial v(t, T)/\partial T \leq 0$. *(ii)* $v(0, T) = 0$ *and* $\lim_{t \to T} v(t, T) = +\infty$ *hold.*

$v_1(t, T) > 0$ indicates that the disutility increases with age, and $v_2(t, T) \leq 0$ means that the disutility at each age decreases as the longevity T increases. This is interpreted as indicating that when the longevity increases, the health status at each age tends to improve (or at least, it does not worsen), which leads to a decrease in the disutility of work. This idea is consistent with the relative compression of morbidity assumption presented in Bloom et al. (2007). (A similar idea is seen in Nishimura et al. (2018).) Assumption 2 (ii) ensures that the optimal R is determined as an interior solution, as we will prove in Section 5.1.

The budget constraint is the same as (6):

$$\int_0^T c_t e^{-rt} dt = I.$$

In this setting, individuals choose the consumption profile $\{c_t\}_{t=0}^T$, the retirement age R, and the life schedule (s_1, w_1, s_2, w_2) to maximize (43). It is of interest to note that the separation

[29] Of course, we can express the disutility as $\tilde{v}(t, T) e^{-\rho t}$, where $\tilde{v}(t, T)$ is the disutility evaluated at time t.

[30] Many empirical studies suggest that individuals have different preferences between study and work; the examples include Heckman et al. (1998), Bils and Klenow (2000), Card (2001), Oreopoulos (2007), and Restuccia and Vandenbroucke (2013). However, to make the analysis tractable, we assume that study and work induce the same disutility, in line with many of the studies in the existing literature. (A notable exception is Sánchez-Romero et al. (2016). They consider the case where the agents may have different preferences between schooling time and working time.)

theorem can also be applied to the model where the retirement age R is endogenous. The maximization problem in Section 2.4 is modified as follows:

(Step 1) Given the retirement age R, the individuals choose the life schedule (s_1, w_1, s_2, w_2) to maximize I. The maximized lifetime income is denoted by $I(R)$.

(Step 2) Based on $I(R)$, the individuals maximize lifetime utility by choosing $\{c_t\}_{t=0}^{T}$ and R.

Note that Step 1 is essentially equal to Step 1 in Section 2.4. That is, by considering R as given, we obtain the same optimal life schedule as presented in Section 3 and the same $I(R)$ as derived in Section 4. In other words, because of the separation theorem, the main result is independent of whether R is an exogenous or an endogenous variable.

5.1 Longevity and optimal retirement age

Contrary to the exogenous retirement age model, by exploring Step 2, we obtain the relationship between the longevity T and the optimal retirement age. Substituting $I(R)$ into I in (6), we construct the Lagrangean as:

$$L = \int_0^T u(c_t) e^{-\rho t} dt - \int_0^R v(t, T) dt + \lambda \left(I(R) - \int_0^T c_t e^{-rt} dt \right).$$

As will be seen below, the optimal solution is obtained as an interior solution. In this case, the first-order conditions are:

$$\lambda I'(R) = v(R, T), \tag{44}$$

$$u'(c_t) e^{(r-\rho)t} = \lambda. \tag{45}$$

Equation (44) is the optimal condition with respect to R. Equation (45) expresses the optimal condition for c_t, which is identical to (7). Using a similar procedure to derive (9), we obtain:

$$c_0 = \frac{I(R)}{\theta(T)}, \tag{46}$$

where $\theta(T) \equiv \int_0^T e^{-\frac{(\sigma-1)r+\rho}{\sigma}t} dt$.

Combining (44), (45) as of time 0, and (46) yields the following key equation:

$$u'\left(\frac{I(R)}{\theta(T)}\right) I'(R) = v(R, T). \tag{47}$$

The LHS of (47) expresses the marginal benefit of working at age R (the increase in utility from increasing the lifetime income) and the RHS expresses the marginal disutility of delaying the retirement age. Using this equation, we derive the optimal retirement age R^*.

Given T, the marginal benefit is a decreasing function of R. To see this, differentiating it with respect to R, and noting that $-u''(c)\,c/u'(c) \equiv \sigma$, we obtain:

$$\frac{\partial}{\partial R}\left(u'\left(\frac{I(R)}{\theta(T)}\right)I'(R)\right) = u'\left(\frac{I(R)}{\theta(T)}\right)I'(R)\left(\frac{I''(R)}{I'(R)} - \sigma\frac{I'(R)}{I(R)}\right) = v(R,T)\left(\frac{I''(R)}{I'(R)} - \sigma\frac{I'(R)}{I(R)}\right). \tag{48}$$

It can be seen that the term $\frac{I''(R)}{I'(R)} - \sigma\frac{I'(R)}{I(R)}$ determines the sign of $\frac{\partial}{\partial R}\left(u'\left(\frac{I(R)}{\theta(T)}\right)I'(R)\right)$. Noting that $I(R)$ and $I'(R)$ are positive, we immediately confirm that (48) is negative when $I''(R) < 0$. That is, when $I'(R)$ decreases with age, the marginal benefit of continuing to work necessarily decreases with age. Conversely, if $I''(R) > 0$, it is possible that the marginal benefit may increase with age. Using (53) and (42), we examine the case of $\hat{R}_0 < R < \hat{R}_1$, and we obtain:

$$\frac{I''(R)}{I'(R)} - \sigma\frac{I'(R)}{I(R)} = (1-\sigma)(\beta - r) \leq 0.$$

The last inequality holds on the assumption of $\sigma \geq 1$. Consequently, the LHS of (47) is a decreasing function of R.

On the other hand, the RHS of (47), which represents the marginal disutility of postponing the retirement age, is an increasing function of R under Assumption 2 (i), and the range of $v(R,T)$ is from 0 to $+\infty$ under Assumption 2 (ii). Fig. 7 depicts the LHS and RHS of (47), which are drawn as downward- and upward-sloping curves, respectively. Thus, the solution of (47) exists in $R \in (0,T)$ and is unique. (In other words, the optimal retirement age R^* is derived as an interior solution.) Moreover, we obtain the relationship between the longevity T and the optimal retirement age R^*.

Proposition 4 *When T increases, R^* rises.*

Proof. Regarding the marginal benefit, we obtain:

$$\frac{\partial}{\partial T}\left(u'\left(\frac{I(R)}{\theta(T)}\right)I'(R)\right) = -u''\left(\frac{I(R)}{\theta(T)}\right)\frac{I(R)}{\theta(T)^2}\phi'(T)\,I'(R) > 0.$$

The last inequality comes from $\phi'(T) = e^{-\frac{(\sigma-1)r+\rho}{\sigma}T} > 0$. This indicates that the marginal benefit rises as T rises. The downward-sloping curve in Fig. 8, the marginal benefit, shifts upward. This effect increases the length of the optimal active life.

Furthermore, Assumption 2 (ii) indicates that the marginal disutility $v(R,T)$ decreases as T rises. The upward-sloping curve in Fig. 8, the marginal disutility, shifts downward. This effect also increases R^*. ■

The positive relationship between longevity and the length of the active life is obtained in this model. The result is consistent with the preceding studies, which treat the retirement age as endogenous.[31]

Moreover, by combining Proposition 4 with Proposition 2, we obtain the effect of a rise in the longevity on the optimal lifetime schedule. Let $\hat{T}_j$ $(j = 0, 1)$ denote T, which satisfies:

$$u'\left(\frac{I\left(\hat{R}_j\right)}{\theta\left(T\right)}\right) I'\left(\hat{R}_j\right) = v\left(\hat{R}_j, T\right).$$

From Proposition 4, it is confirmed that $\hat{T}_0 < \hat{T}_1$ holds. We obtain the following result as a corollary of Proposition 2:

Corollary 1 *Suppose that $T > \hat{T}_1$. When T rises marginally, s_1^* remains at $s_1^* = \bar{s}_1$, and w_1^*, w_2^*, and s_2^* increase.*[32]

5.2 Effect of initial asset

So far, we assume that individuals have no initial assets. Here, let us consider a situation where an individual has some initial asset, denoted by k_0, and examine the effect of k_0 on the optimal choice. Suppose that k_0 is given exogenously. Instead of (6), the budget constraint is given as:

$$\int_0^T c_t e^{-rt} dt = I + k_0. \tag{49}$$

Although the initial asset is incorporated, the procedure to solve the optimization problem remains unchanged. That is,

(Step 1) Given the retirement age R, the individuals choose the life schedule (s_1, w_1, s_2, w_2) to maximize I. The maximized lifetime income is denoted by $I(R)$.

(Step 2) Based on $I(R)$, the individuals maximize lifetime utility by choosing $\{c_t\}_{t=0}^T$ and R.

[31] For example, refer to Boucekkine et al. (2002), Echevarría (2004), Ferreira and Pessôa (2007), Hazan (2009), Sánchez-Romero et al. (2016), and Yasui (2016).

Sánchez-Romero et al. (2016) employ a general survival function and show that the retirement age increases (falls) when there is a decline in the mortality rate during the retirement (working) period. In the present model, we use a rectangular survival function. An increase in the life-span T can be considered as a decrease in mortality during the retirement period.

[32] Related to footnote 26, we also obtain the following: (i) when $T < \hat{T}_0$, only total working years $w^* \equiv w_1^* + w_2^*$ increase, and total schooling years $s^* \equiv s_1^* + s_2^*$ remain at zero; and (ii) when $\hat{T}_0 < T < \hat{T}_1$, s^* increases, whereas w^* does not change.

Observe that the Step 1 problem is not affected by the constant term k_0. Put differently, when R is given as exogenous, the initial asset has no impact on the optimal schedule of the individuals.

Conversely, when R is endogenous, the individual's optimal behavior is affected by the initial asset. Note that (47) is modified as:

$$u'\left(\frac{I(R)+k_0}{\theta(T)}\right)I'(R)=v(R,T). \tag{50}$$

We obtain the following result:

Proposition 5 *(i) The optimal retirement age R^* falls as the initial asset k_0 increases.*

(ii) When $R>\hat{R}_1$, the optimal duration of recurrent education s_2^* decreases when k_0 rises. Moreover, w_1^* and w_2^* fall, and s_1^* remains at $s_1^*=\overline{s}_1$.

Proof. Noting that the LHS of (50) is a decreasing function of k_0 because $u''<0$, the graph of the LHS of (50) shifts downward as k_0 rises, as seen in Fig. 9. Consequently, we confirm that R^* falls when k_0 rises and, thus, (i) is proved. Furthermore, recalling Proposition 2, we obtain statement (ii) as the effect of a decrease in R^* on the optimal schedule. ■

Proposition 5 maintains that as the amount of the initial asset increases, the marginal utility of income decreases and, thus, the duration of the active life R^* falls. Moreover, the durations of both total working time $w_1^*+w_2^*$ and the recurring education s_2^* decrease.

6 Concluding remarks

We have investigated the effect of a longer life-span on the optimal life schedule during an individual's active life. We focused on the situation where individuals choose to undertake recurrent education. In the present analysis, we made Assumption 1 in regard to the relationship between the return of education, the interest rate, and the economic growth rate, and we have explored the property of the optimal schooling schedule. We found that the optimal schooling years are characterized as a corner solution; that is, people should obtain cutting-edge skills in all schooling periods if they choose to undertake recurrent education. In addition, we found that the total years of schooling and working increase and, thus, the retirement age is delayed as the life-span increases.

Although the lifetime income increases as R rises, the important point is that the marginal effect of a rise in R on the lifetime income decreases significantly when recurrent education takes place. This indicates that we cannot expect vigorous income growth in an economy in which individuals have a longer life-span, even though these individuals work longer and study longer and harder.

These findings suggest that the development of new technologies will be more important than ever as a determinant of income growth in an economy where recurrent education is a common practice. Regarding this point, it will be interesting to explicitly incorporate R&D activities that produce new technologies into a model involving recurrent education. In the present study, we have not considered this point. To undertake such an analysis, it is necessary to build a model where a researcher's incentives for the new inventions are considered. We defer this issue to future research.

Acknowledgments

The author would like to thank the two anonymous referees and the editor, Charles Z. Zheng, for their helpful comments and suggestions. In addition, the author would like to thank Yuichi Furukawa, Tetsugen Haruyama, Ken-ichi Hashimoto, Kouki Sugawara, Ken Tabata, Takashi Unayama, Daishin Yasui, and seminar participants at the Romacs Workshop 2020 at Chukyo University for their valuable comments and suggestions. This study was financially supported by the joint research program of KIER and Grants-in-Aid for Scientific Research from the JSPS [grant number 18K01524].

Appendix A Proof of Lemma 1

Suppose that $h_{\Delta}(t) = A_t$ is attained at time t, and consider the case where individuals continue to study until time $t+a$ (where $a \geq 0$) and work from time $t+a$ to R.[33] The human capital stock at time $t+a$ is $A_t e^{ga}$ and, from (4), the present value of the income earned from time $t+a$ to time R is expressed as:

$$I = \phi A_t e^{ga} e^{-r(t+a)} \int_0^{R-t-a} e^{-(\delta+r)s} ds = \frac{\phi A_t}{\delta + r} e^{-rt} e^{-(r-g)a} \left(1 - e^{-(\delta+r)(R-t-a)}\right).$$

Here, note that both $e^{-(r-g)a}$ and $1 - e^{-(\delta+r)(R-t-a)}$ are decreasing functions of a under Assumption 1, so that I is a decreasing function of a. Thus, I is maximized when $a = 0$ and, therefore, Lemma 1 is proved.

[33] Instead of R, we can consider that the individuals work until $\tilde{t} < R$, and study from time $\tilde{t}$ (recurrent education). We confirm that although the description is more complicated, the main conclusion does not change.

Appendix B Proof of Lemma 2

Consider a life schedule $\left(s_1^O, w_1^O, s_2^O, w_2^O\right)$ (let us call it the original life schedule), and suppose that $s_1^O < \bar{s}_1$ and $s_2^O > 0$ hold. We show that this schedule does not maximize the lifetime income I; that is, it is not the optimal life schedule. We can prove this statement by showing that there is another feasible life schedule under which I is larger than the original life schedule. Let us consider an alternative feasible life schedule $\left(s_1^A, w_1^A, s_2^A, w_2^A\right) = \left(s_1^O + \Delta t, w_1^O, s_2^O - \Delta t, w_2^O\right)$, where $0 < \Delta t < s_2^O$, and show that I under $\left(s_1^A, w_1^A, s_2^A, w_2^A\right)$ is larger than under $\left(s_1^O, w_1^O, s_2^O, w_2^O\right)$.

First, I_2 is the same between the original life schedule and the alternative life schedule because (a) and (b) below hold.

(a) $s_1^A + w_1^A + s_2^A = s_1^O + w_1^O + s_2^O$ holds, so that t_{w_2}, the time when the recurrent education ends and when the second working period starts, is the same between the two life schedules. The length of the second working period is also the same ($w_2^A = w_2^O$).

(b) $\beta\left(s_1^A + s_2^A\right) - \delta w_1^A = \beta\left(s_1^O + s_2^O\right) - \delta w_1^O$ holds and, thus, the human capital at time t_{w_2} is the same between the two life schedules.

Second, let us focus on I_1. It is expressed as $\phi e^{\beta s_1^O} \int_{s_1^O}^{s_1^O + w_1^O} e^{-\delta\left(t - s_1^O\right)} e^{-rt} dt = \phi e^{(\beta - r) s_1^O} \int_0^{w_1^O} e^{-(\delta + r)t} dt$ under the original life schedule, and $\phi e^{(\beta - r) s_1^A} \int_0^{w_1^A} e^{-(\delta + r)t} dt$ under the alternative life schedule. We calculate the difference between the two, denoted by ΔI_1, as:

$$\begin{aligned}
\Delta I_1 &= \phi e^{(\beta - r) s_1^A} \int_0^{w_1^A} e^{-(\delta + r)t} dt - \phi e^{(\beta - r) s_1^O} \int_0^{w_1^O} e^{-(\delta + r)t} dt \\
&= \phi e^{(\beta - r)\left(s_1^O + \Delta t\right)} \int_0^{w_1^O} e^{-(\delta + r)t} dt - \phi e^{(\beta - r) s_1^O} \int_0^{w_1^O} e^{-(\delta + r)t} dt \\
&= \left(e^{(\beta - r)\Delta t} - 1\right) \phi e^{(\beta - r) s_1^O} \int_0^{w_1^O} e^{-(\delta + r)t} dt > 0.
\end{aligned}$$

Because $e^{(\beta - r)\Delta t} - 1 > 0$, on the assumption that $\beta > r$, $\Delta I_1 > 0$ holds. This indicates that I_1 on the alternative life schedule is higher than I_1 on the original life schedule.

Consequently, the lifetime income $I = I_1 + I_2$ under $\left(s_1^A, w_1^A, s_2^A, w_2^A\right)$ is larger than under $\left(s_1^O, w_1^O, s_2^O, w_2^O\right)$. Thus, the original life schedule cannot be the optimal life schedule. In other words, if $s_2 > 0$ is optimal, $s_1 = \bar{s}_1$ must hold. Therefore, Lemma 2 (i) is proved.

Furthermore, Lemma 2 (ii) is the contraposition of Lemma 2 (i). Hence, Lemma 2 (ii) is true because Lemma 2 (i) is true.

Appendix C Proof of Lemma 6

As Lemma 6 (i) has been already proved in the text, we begin by proving (ii). Substituting $\underline{\Psi}(w_2) = -\frac{\beta-g}{\beta+\delta}(R-\bar{s}_1-w_2)$ into (30), we obtain:

$$\underline{J}(w_2) = \frac{B_0\phi}{\delta+r}\left\{e^{(\beta-r)\bar{s}_1}\left(1-e^{-\frac{\beta-g}{\beta+\delta}(\delta+r)(R-\bar{s}_1-w_2)}\right)+e^{(\beta-g)\bar{s}_1}e^{-(r-g)(R-w_2)}\left(1-e^{-(\delta+r)w_2}\right)\right\}.$$

First, we show that $\underline{J}(0) < \overline{I}$.

$$\underline{J}(0) = \frac{B_0\phi}{\delta+r}\left\{e^{(\beta-r)\bar{s}_1} - e^{-\frac{\beta-g}{\beta+\delta}(\delta+r)R}e^{\left[\beta-r+\frac{\beta-g}{\beta+\delta}(\delta+r)\right]\bar{s}_1}\right\}.$$

Recall that $\overline{I} = \frac{\phi B_0}{\delta+r}\left\{e^{(\beta-r)\bar{s}_1} - e^{-(\delta+r)R}e^{(\beta+\delta)\bar{s}_1}\right\}$. On comparing the second term in the curly brackets of $\underline{J}(0)$ with $\overline{I}$, it is confirmed that $\underline{J}(0) < \overline{I}$ if and only if the following inequality holds:

$$-\frac{\beta-g}{\beta+\delta}(\delta+r)R + \left[\beta-r+\frac{\beta-g}{\beta+\delta}(\delta+r)\right]\bar{s}_1 > -(\delta+r)R + (\beta+\delta)\bar{s}_1.$$

On arranging it, we confirm that this inequality is equivalent to $R > \bar{s}_1$ and, thus, $\underline{J}(0) < \overline{I}$ holds true.

Second, we show that $\underline{J}(R-\bar{s}_1) = \overline{I}$. Recall that $\underline{\Psi}(w_2) = \overline{\Psi}(w_2)$ holds at $w_2 = R-\bar{s}_1$ (see Fig. 3) and, thus, $\overline{J}(R-\bar{s}_1) = \underline{J}(R-\bar{s}_1)$ holds. From (i), $\overline{J}(R-\bar{s}_1) = \overline{I}$ and, therefore, $\underline{J}(R-\bar{s}_1) = \overline{I}$ holds.

Next, we examine (iii) and (iv). We have argued that $\kappa(w_2)$, the terms in the curly brackets of (36), is a decreasing function of w_2, and that $\underline{J}'(w_2) > 0$ (respectively, $\underline{J}'(w_2) < 0$) holds if and only if $\kappa(w_2) > 0$ (respectively, $\kappa(w_2) < 0$). Thus, we can prove (iii) and (iv) by showing that (a) $\kappa(0) > 0$ and (b) $\kappa(R-\bar{s}_1) < 0$ if and only if $R > \hat{R}_1$. Let us recall (37):

$$\underline{J}'(w_2) \gtreqless 0 \iff \kappa(w_2) \gtreqless 0 \iff \frac{\beta-g}{\beta+\delta}(\delta+r)e^{-\frac{(\beta-r)(g+\delta)}{\beta+\delta}(R-w_2-\bar{s}_1)} \lesseqgtr (r-g)+(g+\delta)e^{-(\delta+r)w_2}.$$

When $w_2 = 0$, the following holds:

$$LHS = \frac{\beta-g}{\beta+\delta}(\delta+r)e^{-\frac{(\beta-r)(g+\delta)}{\beta+\delta}(R-\bar{s}_1)} < \delta+r = RHS,$$

where the inequality holds because $\frac{\beta-g}{\beta+\delta} < 1$ and $e^{-\frac{(\beta-r)(g+\delta)}{\beta+\delta}(R-\bar{s}_1)} < 1$ hold on the assumption of $\beta > r$. Thus, (a) $\kappa(0) > 0$ holds.

Conversely, when $w_2 = R-\bar{s}_1$, $\kappa(R-\bar{s}_1) < 0$ if and only if the following holds:

$$\frac{\beta-g}{\beta+\delta}(\delta+r) > (r-g)+(g+\delta)e^{-(\delta+r)(R-\bar{s}_1)}.$$

Noting that $\frac{\beta-g}{\beta+\delta}(\delta+r)-(r-g)$ is equal to $\frac{(\beta-r)(g+\delta)}{\beta+\delta}$, the above inequality can be rewritten as:

$$\frac{\beta-r}{\beta+\delta}>e^{-(\delta+r)(R-\bar{s}_1)}.$$

From the definition of $\hat{R}_1$, we immediately confirm that the above equality is equivalent to $R>\hat{R}_1$. Therefore, (b) is proved.

Taking (a), (b), and $\kappa'(w_2)<0$ into account, when $R\le\hat{R}_1$, we confirm that $\kappa(w_2)>0$ ($\iff \underline{J}'(w_2)>0$) holds for any $w_2\in(0,R-\bar{s}_1)$ and, together with Lemma 6 (ii), $\underline{J}(w_2)$ is maximized at $w_2=R-\bar{s}_1$. Thus, Lemma 6 (iii) is proved. On the other hand, when $R>\hat{R}_1$, w_2^+, a value of w_2 that satisfies $\kappa(w_2)=0$ exists in $w_2^+\in(0,R-\bar{s}_1)$, and $\underline{J}(w_2)$ is maximized at $w_2=w_2^+$. Thus, Lemma 6 (iv) is proved.

Appendix D The concavity of (40) in the case of $r\ge g$

Taking the logarithm, we rewrite (40) as:

$$w_2=\Phi(w_1)\equiv\frac{1}{\delta+r}\left\{-\log\varphi(w_1)+\log(g+\delta)\right\}. \tag{51}$$

where $\varphi(w_1)$ is defined as:

$$\varphi(w_1)=\frac{(\beta-g)(\delta+r)}{\beta+\delta}e^{-\frac{(\beta-r)(g+\delta)}{\beta-g}w_1}-(r-g).$$

Because $\frac{dw_2}{dw_1}=\frac{-1}{\delta+r}\frac{\varphi'(w_1)}{\varphi(w_1)}$, we calculate $\frac{\varphi'(w_1)}{\varphi(w_1)}$. It is obtained as:

$$\begin{aligned}\frac{\varphi'(w_1)}{\varphi(w_1)}&=\frac{-(\delta+r)(g+\delta)(\beta-r)e^{-\frac{(\beta-r)(g+\delta)}{\beta-g}w_1}}{(\beta-g)(\delta+r)e^{-\frac{(\beta-r)(g+\delta)}{\beta-g}w_1}-(r-g)(\beta+\delta)}\\&=\frac{-(\delta+r)(g+\delta)(\beta-r)}{(\beta-g)(\delta+r)-(r-g)(\beta+\delta)e^{\frac{(\beta-r)(g+\delta)}{\beta-g}w_1}}.\end{aligned}$$

Thus, we obtain:

$$\frac{dw_2}{dw_1}=\Phi'(w_1)=\frac{(g+\delta)(\beta-r)}{(\beta-g)(\delta+r)-(r-g)(\beta+\delta)e^{\frac{(\beta-r)(g+\delta)}{\beta-g}w_1}}. \tag{52}$$

From the above, we observe that the denominator of dw_2/dw_1 is a decreasing function of w_1 under Assumption 1. That is, dw_2/dw_1 increases as w_1 rises.

Appendix E Proof of Proposition 3

(i) We investigate the first derivative of $I(R)$. Because it holds that $\frac{\partial \underline{J}\left(w_2^+, R\right)}{\partial w_2} = 0$ (expressed as $\underline{J}'\left(w_2^+\right) = 0$ in the Section 3), $\frac{d\underline{J}\left(w_2^+, R\right)}{dR} = \frac{\partial \underline{J}\left(w_2^+, R\right)}{\partial w_2}\frac{dw_2^+}{dR} + \frac{\partial \underline{J}\left(w_2^+, R\right)}{\partial R} = \frac{\partial \underline{J}\left(w_2^+, R\right)}{\partial R}$ holds. Thus, we obtain the following from (41):

$$I'(R) = \begin{cases} \phi B_0 e^{-(\delta+r)R}, & 0 < R < \hat{R}_0 \\ (\beta - r) I(R), & \hat{R}_0 < R < \hat{R}_1 \\ \phi B_0 e^{(\beta-g)\bar{s}_1} e^{-\left[(r-g)R + (g+\delta)w_2^+\right]}, & \hat{R}_1 < R \end{cases}. \tag{53}$$

We examine whether $I'(R)$ is continuous at $R = \hat{R}_0$ and $R = \hat{R}_1$. From (53), we obtain $\lim_{R \to \hat{R}_0 - 0} I'(R) = \phi B_0 e^{-(\delta+r)\hat{R}_0}$, and from (41) and (53), we obtain:

$$\lim_{R \to \hat{R}_0 + 0} I'(R) = (\beta - r) \lim_{R \to \hat{R}_0 + 0} I(R) = (\beta - r)\phi B_0 \int_0^{\hat{R}_0} e^{-(\delta+r)t} dt.$$

Thus, $\lim_{R \to \hat{R}_0 - 0} I'(R) = \lim_{R \to \hat{R}_0 + 0} I'(R) \left(\equiv I'(\hat{R}_0)\right)$ holds if and only if the following equation holds:

$$\frac{\beta - r}{\delta + r}\left(1 - e^{-(\delta+r)\hat{R}_0}\right) = e^{-(\delta+r)\hat{R}_0}. \tag{54}$$

Recalling that $\hat{R}_0 \equiv \frac{1}{\delta+r}\log\frac{\beta+\delta}{\beta-r}$, we confirm that (54) is true. Thus, $I'(R)$ is continuous at $R = \hat{R}_0$.

Next, we consider $\lim_{R \to \hat{R}_1 - 0} I'(R)$ and $\lim_{R \to \hat{R}_1 + 0} I'(R)$. Recalling that $R - \hat{R}_0 = \bar{s}_1$ at $R = \hat{R}_1$, we obtain $\lim_{R \to \hat{R}_1 - 0} I'(R) = (\beta - r)\phi B_0 e^{(\beta - r)\bar{s}_1} \int_0^{\hat{R}_0} e^{-(\delta+r)t} dt$. Furthermore, recalling that $\lim_{R \to \hat{R}_1 + 0} w_2^+ = \hat{R}_0$ (refer to footnote 27), we obtain:

$$\lim_{R \to \hat{R}_1 + 0} I'(R) = \phi B_0 e^{(\beta-g)\bar{s}_1} e^{-\left[(r-g)\hat{R}_1 + (g+\delta)\hat{R}_0\right]} = \phi B_0 e^{(\beta-r)\bar{s}_1} e^{-(\delta+r)\hat{R}_0}.$$

From (54), we confirm that $\lim_{R \to \hat{R}_1 - 0} I'(R) = \lim_{R \to \hat{R}_1 + 0} I'(R) \left(\equiv I'(\hat{R}_1)\right)$ holds. That is, $I'(R)$ is continuous at $R = \hat{R}_1$. Hence, $I'(R)$ is a continuous function. In other words, $I(R)$ is a continuously differentiable function.

(ii) From (53), it can be seen that $I'(R) > 0$ holds for $R \in \Re_{++} \backslash \{\hat{R}_0, \hat{R}_1\}$, and we have shown that $I'(R)$ is continuous at $R = \hat{R}_0$ and $R = \hat{R}_1$ above. That is, $I'(R) > 0$ holds at $R = \hat{R}_0$ and at $R = \hat{R}_1$, which completes the proof.

(iii) We obtain the second derivative of $I(R)$ from (53):

$$I''(R) = \begin{cases} -(\delta + r) I'(R), & 0 < R < \hat{R}_0 \\ (\beta - r) I'(R), & \hat{R}_0 < R < \hat{R}_1 \\ -\left[(r - g) + (g + \delta)\frac{dw_2^+}{dR}\right] I'(R), & \hat{R}_1 < R \end{cases}.$$

It can be seen immediately from (42) that $I''(R) < 0$ when $0 < R < \hat{R}_0$ and that $I''(R) > 0$ when $\hat{R}_0 < R < \hat{R}_1$. Moreover, because $r \geq g$ (Assumption 1) and $\frac{dw_2^+}{dR} > 0$ (Proposition 2), $I''(R) < 0$ when $\hat{R}_1 < R$. (This indicates that $I''(R)$ is not continuous at $R = \hat{R}_0$ and $R = \hat{R}_1$. That is, $I(R)$ is not of class C^2.) $I(R)$ is a continuous function on $R > 0$, so we can conclude that $I''(R)$ is strictly concave for $0 \leq R \leq \hat{R}_0$, strictly convex for $\hat{R}_0 \leq R \leq \hat{R}_1$, and strictly concave for $\hat{R}_1 \leq R$.

References

[1] Acemoglu, D., 2009, "Introduction to modern economic growth," Princeton University Press, Princeton, NJ.

[2] Acemoglu, D. and S. Johnson, 2006, "Disease and development: The effect of life expectancy on economic growth," Journal of Political Economy, 115(6), 925–985.

[3] Ben-Porath, Y., 1967, "The production of human capital and the life cycle of earnings," Journal of Political Economy, 75, 352–365.

[4] Berk, J. and D. Weil, 2015, "Old teachers, old ideas, and the effect of population aging on economic growth," Research in Economics, 69, 661–670.

[5] Bils, M. and P.J. Klenow, 2000, "Does schooling cause growth?" American Economic Review, 90, 1160–1183.

[6] Bloom, D.E., D. Canning, R.K. Mansfield, and M. Moore, 2007, "Demographic change, social security systems, and savings," Journal of Monetary Economics, 54(1), 92–114.

[7] Boucekkine, R., D. de la Croix, and O. Licandro, 2002, "Vintage human capital, demographic trends, and endogenous growth," Journal of Economic Theory, 104(2), 340–375.

[8] Boucekkine, R., D. de la Croix, and O. Licandro, 2003, "Early mortality declines at the dawn of modern growth," Scandinavian Journal of Economics. 105, 401–418.

[9] Cai, Z. and S.-H.P. Lau, 2017, "Impact of mortality reductions on years of schooling and expected lifetime labor supply," Journal of Mathematical Economics, 72, 134–144.

[10] Card, D., 2001, "Estimating the return to education: progress on some persistent economic problems," Econometrica, 69(5),1127–1160.

[11] Cervellati, M. and U. Sunde, 2005, "Human capital, life expectancy, and the process of development," American Economic Review, 95(5), 1653–1672.

[12] Cervellati, M. and U. Sunde, 2011, "Life expectancy and economic growth: The role of the demographic transition," Journal of Economic Growth, 16(2), 99–133.

[13] Cervellati, M. and U. Sunde, 2013, "Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited," Econometrica, 81(5), 2055–2086.

[14] d'Albis, H., S.-H.P. Lau, and M. Sánchez-Romero, 2012, "Mortality transition and differential incentives for early retirement," Journal of Economic Theory, 147(1), 261–283.

[15] de la Croix, D. and O. Licandro, 1999, "Life expectancy and endogenous growth," Economics Letters, 65, 255–263.

[16] de la Croix, D., 2017, "Did longer lives buy economic growth? From Malthus to Lucas and Ben-Porath," In M. Cervellati and U. Sunde (eds), Demographic Change and Long-Run Development, 69–89, Cambridge MA: MIT Press.

[17] Echevarría, C.A., 2004, "Life expectancy, schooling time, retirement, and growth," Economic Inquiry, 42(4), 602–617.

[18] Ehrlich, I. and F. Lui, 1991, "Intergenerational trade, longevity, and economic growth," Journal of Political Economy, 99(5), 1029–1059.

[19] Ferreira, P.C. and S.A. Pessôa, 2007, "The effects of longevity and distortions on education and retirement," Review of Economic Dynamics, 10(3), 472–493.

[20] Havranek, T., R. Horvath, Z. Iršová, and M. Rusnak, 2015, "Cross-country heterogeneity in intertemporal substitution," Journal of International Economics, 96(1), 100–118.

[21] Hazan, M., 2009, "Longevity and lifetime labor supply: Evidence and implications," Econometrica, 77(6), 1829–1863.

[22] Heckman, J.J., L. Lochner, and C. Taber, 1998, "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents," Review of Economic Dynamics, 1(1),1–58.

[23] Jayachandran, S. and A. Lleras-Muney, 2009, "Life expectancy and human capital investments: Evidence from maternal mortality declines," Quarterly Journal of Economics, 124(1), 349–397.

[24] Kalemli-Ozcan, S., H.E. Ryder and D.N. Weil, 2000, "Mortality decline, human capital investment, and economic growth," Journal of Development Economics, 62, 1–23.

[25] Lorentzen, P., J. McMillan and R. Wacziarg, 2008, "Death and development," Journal of Economic Growth, 13, 81–124.

[26] Magnac, T., N. Pistolesi and S. Roux, 2018, "Post-schooling human capital investments and the life cycle of earnings," Journal of Political Economy, 126(3), 1219–1249.

[27] Nishimura, Y., P. Pestieau, and G. Ponthiere, 2018, "Education choices, longevity and optimal policy in a Ben-Porath economy," Mathematical Social Sciences, 94, 65–81.

[28] OECD, 2011, "Indicator C5 How many adults participate in education and learning?", in Education at a Glance 2011: OECD Indicators, OECD Publishing, Paris, https://doi.org/10.1787/eag-2011-27-en.

[29] OECD, 2017, "Indicator C6 How many adults participate in education and learning?" in Education at a Glance 2017: OECD Indicators, OECD Publishing, Paris, https://doi.org/10.1787/eag-2017-28-en.

[30] Oreopoulos, P., 2007, Do dropouts drop out too soon? Wealth, health, and happiness from compulsory schooling," Journal of Public Economics, 91(11), 2213–2229.

[31] Oster, E., I. Shoulson, and E. Dorsey, 2013, "Limited life expectancy, human capital and health investments," American Economic Review, 103(5), 1977–2002.

[32] Oxborrow, D. and S.J. Turnovsky, 2017, "Closing the small open economy model: A demographic approach," Review of International Economics, 25(1), 44–75.

[33] Restuccia, D. and G. Vandenbroucke, 2013, "A century of human capital and hours," Economic Inquiry, 51(3),1849–1866.

[34] Sánchez-Romero, M., H. d'Albis and A. Prskawetz, 2016, "Education, lifetime labor supply, and longevity improvements," Journal of Economic Dynamics and Control, 73, 118–141.

[35] Soares, R., 2005, "Mortality reductions, educational attainment, and fertility choice," American Economic Review, 95(3), 580–601.

[36] Strulik, H. and S. Vollmer, 2013, "Long-run trends of human aging and longevity," Journal of Population Economics, 26(4), 1303–1323.

[37] Tanaka, M., 2017, "Human capital accumulation through recurrent education (in Japanese)," Keizai Bunseki (Economic Analysis), 196, 49–81.

[38] Wilmoth, J.R. and S. Horiuchi, 1999, "Rectangularization revisited: Variability of age at death within human populations," Demography, 36(4), 475–495.

[39] Yasui, D., 2016, "Adult longevity and growth takeoff," Macroeconomic Dynamics 20(1), 165–188.

[40] Zhang, J., J. Zhang, and R. Lee, 2003, "Rising longevity, education, savings, and growth," Journal of Development Economics, 70(1), 83–101.

[41] Zhang, J. and J. Zhang, 2005, "The effect of life expectancy on fertility, saving, schooling, and economic growth: Theory and evidence," Scandinavian Journal of Economics, 107(1), 45–66.

[42] Zhang, J. and J. Zhang, 2009, "Longevity, retirement, and capital accumulation in a recursive model with an application to mandatory retirement," Macroeconomic Dynamics, 13(3), 327–348.

Table 1: Participation rate of adults (25–64 years old) in formal education

Country	2007	2012	Difference
Australia	11.7	16.7	5.0
Austria	4.2	6.4	2.2
Canada	9.9***	14.1	4.2
Czech Republic	3.9***	5.6	1.7
Denmark	10.1***	14.0	3.9
Estonia	5.0	9.1	4.1
Finland	10.2**	15.5	5.3
France	5.1	4.6	–0.5
Germany	5.2	6.7	1.5
Greece	2.3	5.5****	3.2
Ireland	6.2***	15.2	9.0
Italy	4.4**	5.6	1.2
Korea	5.7	4.8	-0.9
Netherlands	6.8***	14.0	7.2
New Zealand	20.3	17.6****	-2.7
Norway	9.9	15.7	5.8
Poland	5.5**	7.5	2.0
Slovak Republic	6.1	5.7	-0.4
Slovenia	8.7	10.6****	1.9
Spain	5.9	12.6	6.7
Sweden	12.7*	13.1	0.4
United States	8.6*	14.3	5.7
Average	7.7	10.7	3.0

Note: The symbols *, **, ***, and *** indicate that the surveys were implemented in 2005, 2006, 2008, and 2015, respectively.

Source: OECD Education at a Glance 2011 and 2017.

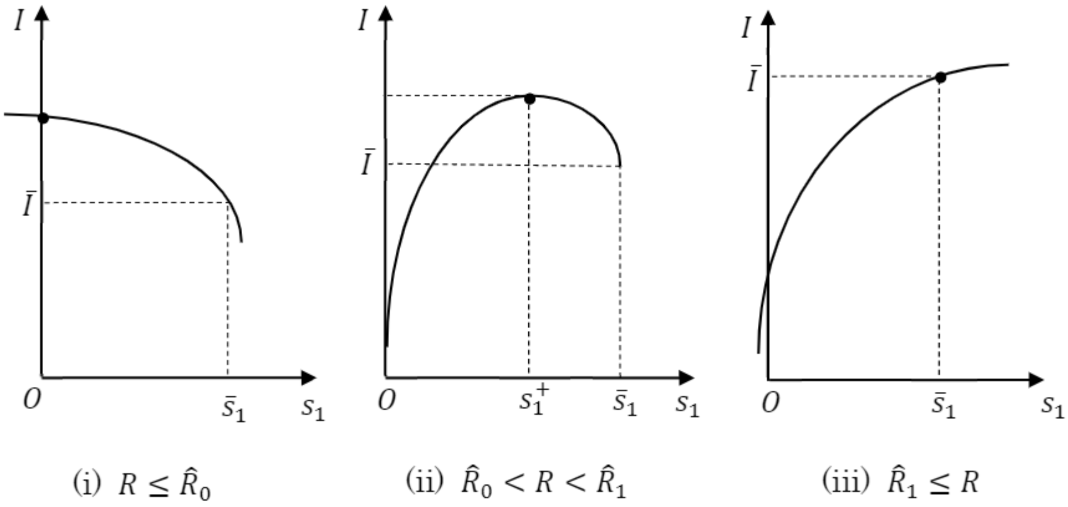

Fig. 1: Lifetime income when $s_2 = 0$

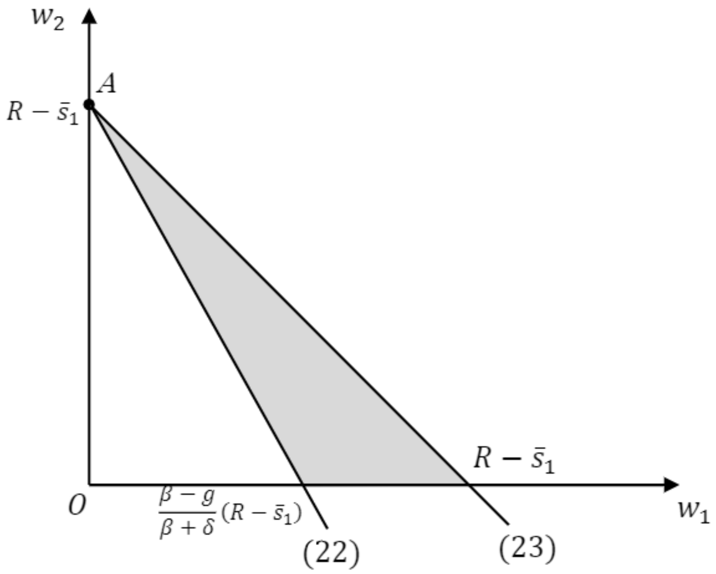

Fig. 2: Relationship between (22) and (23)

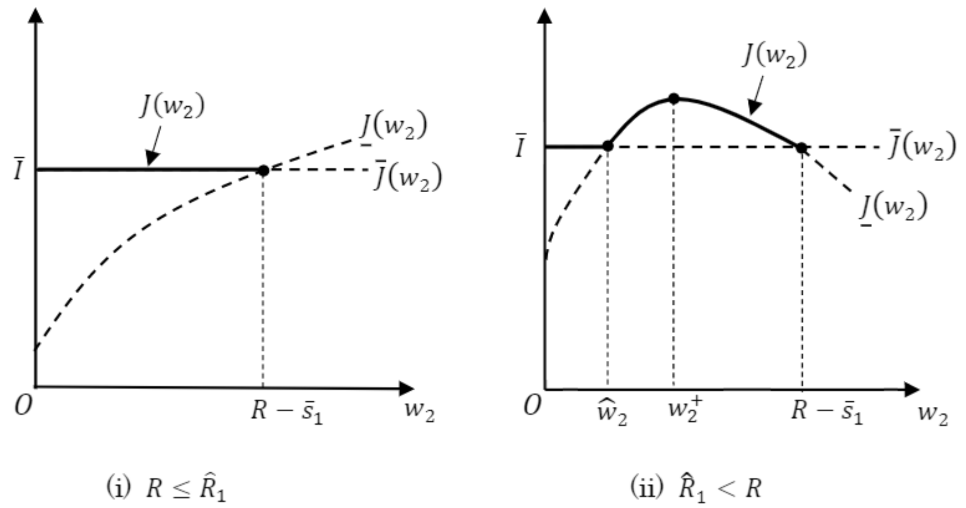

Fig. 3: Graphs of $\underline{J}(w_2), \bar{J}(w_2)$*, and* $J(w_2)$

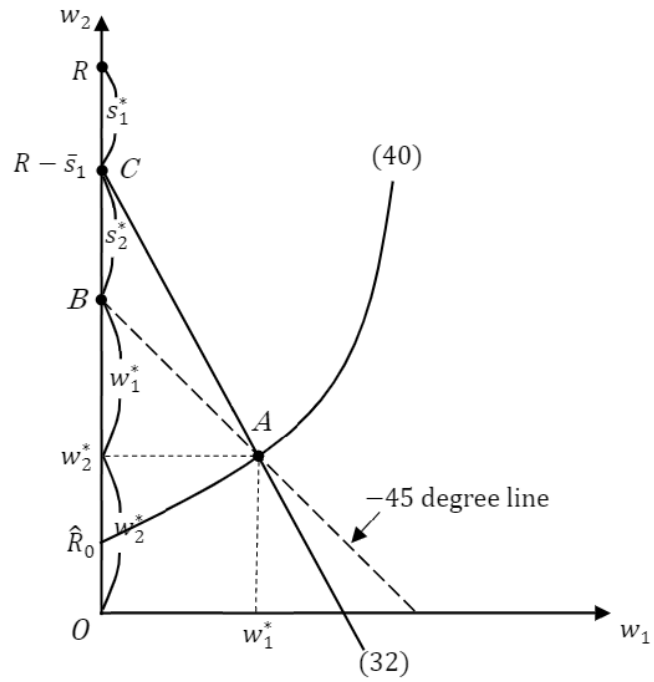

Fig. 4: Graphs of (32) and (40), and the optimal schedule

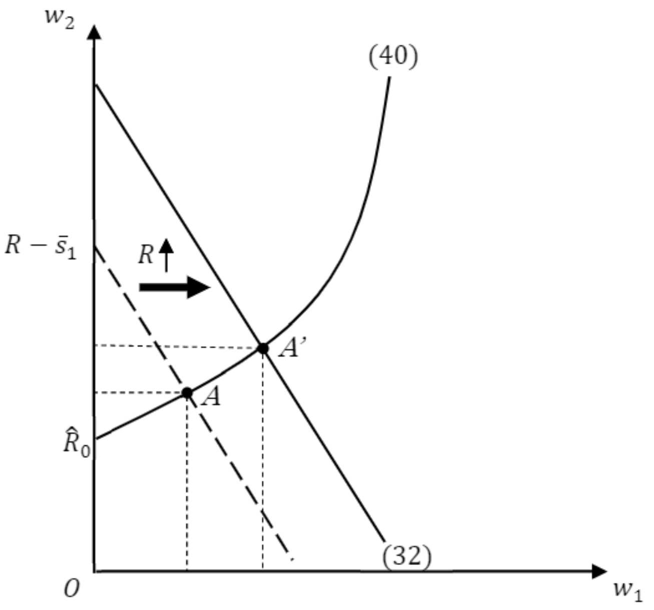

Fig. 5: Effect of an increase in R

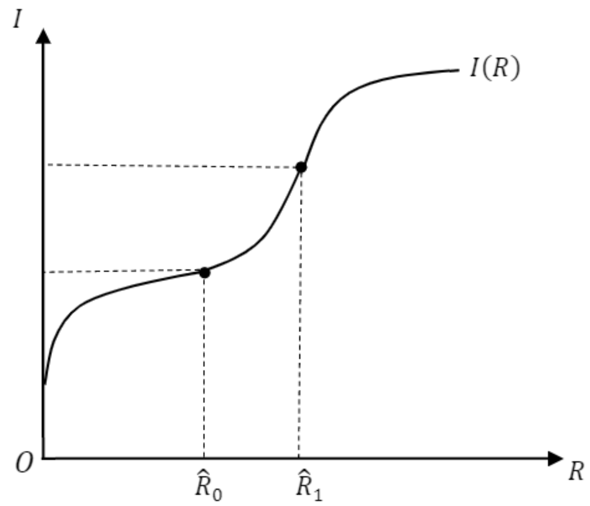

Fig. 6: Graph of $I(R)$

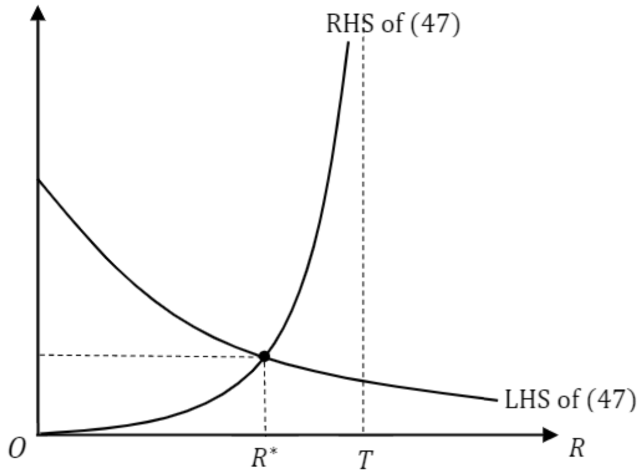

Fig. 7: Optimal retirement age

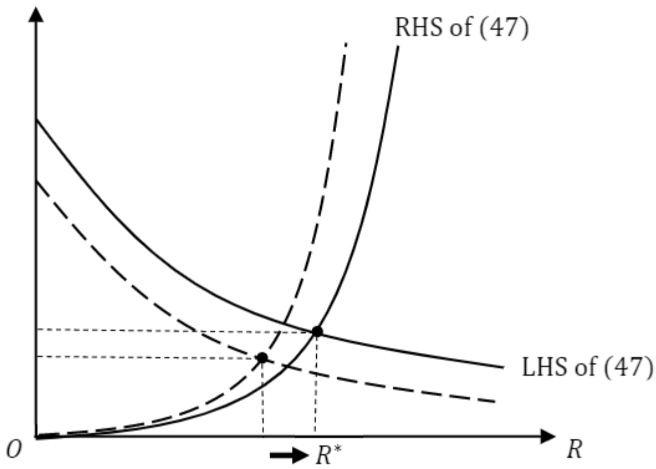

Fig. 8: Effect of a rise in T

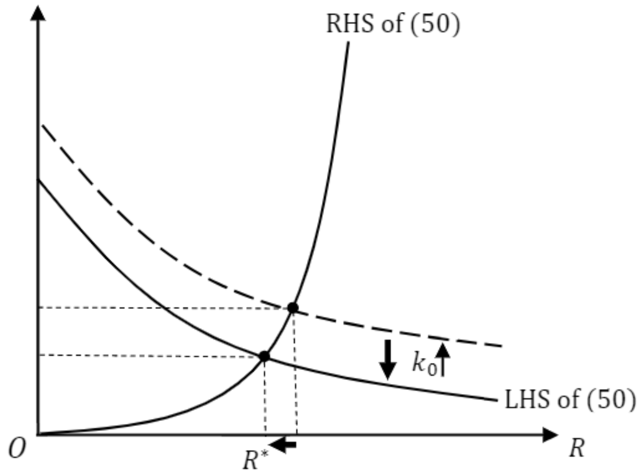

Fig. 9: Effect of a rise in k_0