Knowledge Diffusion and the Japanese Growth Miracle

Atsushi Oshima*†
University of Iowa

January 2009

JOB MARKET PAPER

Abstract

This paper develops a model in which measured total factor productivity (TFP) is endogenously determined, in order to quantitatively explain the following three facts of the Japanese growth miracle between 1956 and 1973. First, the growth rate of GDP per unit of labor in the second half of this period was 7.9 percent while in the first half it was only 5.5 percent. Second, the rate of return to capital increased from 24 percent to 37 percent in the first six years, and then gradually decreased. Third, the contribution of capital to the Japanese growth rate was 16.5 percent between 1956 and 1966, so TFP drove most of the growth. After the middle of the 1960s, however, capital contributed more than 50 percent to GDP growth. In our model, measured TFP growth is driven by endogenous shifts in labor input from low- to high-productivity workers. These shifts occur through a knowledge diffusion process. In this process, a low-productivity worker acquires knowledge from an existing high-productivity worker to become a high-productivity worker in the future. We calibrate the parameter values using Japanese data and our model quantitatively captures the three facts described above.

JEL Classification: E1, O4, O53

*Department of Economics, University of Iowa, atsushi-oshima@uiowa.edu

†I am grateful to B. Ravikumar and Raymond G. Riezman for invaluable advice and guidance. I thank Guillaume Vandenbroucke and Gustavo Ventura for their helpful comments. I also thank Mitsunari Ishida, Tae Okada and Keiji Saito for assistance in collecting the data.

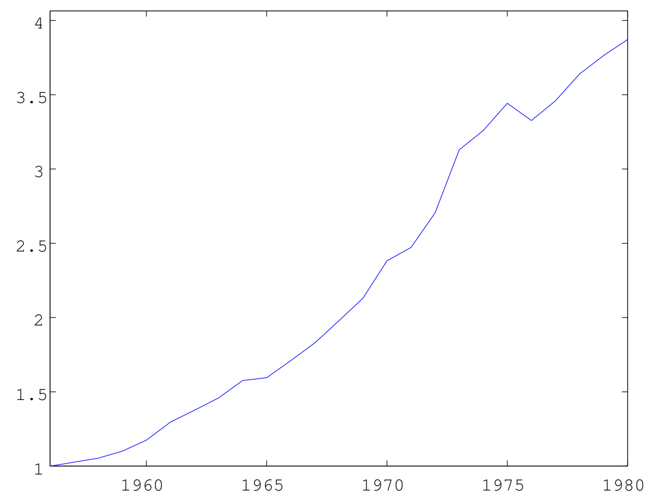

Figure 1: GDP per unit of labor

1 Introduction

Japan achieved remarkably high economic growth between 1956 and 1973, an accomplishment widely known as the Japanese growth miracle. The average annual growth rate of gross domestic product (here after GDP) per worker in Japan at this time was about three times higher than that in the United States. This paper documents three facts of the Japanese growth miracle between 1956 and 1973. First, we compute GDP per unit of labor by subtracting out the effects of changes in observed labor quality from real GDP.[1] Figure 1 shows normalized GDP (1956 = 1) per unit of labor from 1956 to 1980. The average annual growth rate of GDP per unit of labor between 1956 and 1973 was 6.9 percent. But while it had been 5.5 percent in the first half of this period, it accelerated to 7.9 percent during the second half.

Second, Figure 2 shows the rate of return in capital which is defined as the marginal product of capital.[2] The rate of return to capital increased from 24 to 37 percent in the first six years, then decreased gradually; that is, the rate of return has an inverted U-shape.[3]

[1]The growth rate of labor input here depends not only on changes in the number of workers but also changes in hours worked and observed quality. Since this observed labor quality is not part of our model, we take it out from GDP. To construct the series of the growth rate of labor input, we follow Christensen, Cummings and Jorgenson (1980). The details are in the Appendix.

[2]Our rate of return is defined as $\frac{\text{GDP - Labor Compensation}}{\text{Capital}} \times 100$. We construct the series of capital as follows: first, for capital in 1956, we use the value estimated by Hayashi and Prescott (2002). Second, we compute the series of the growth rate of capital using the same method as Christensen, Cummings and Jorgenson (1980). The details are in the Appendix.

[3]When we compute the series of capital using investment data and the perpetual inventory method, the pattern of rate of return does not change.

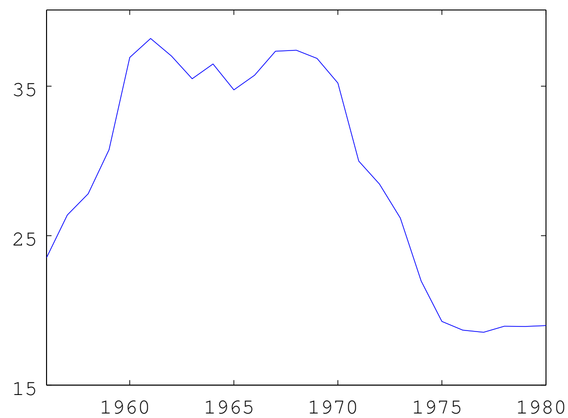

Figure 2: Rate of Return to Capital Input (%)

Third, Table 1 shows the results of our growth accounting in which we decompose the growth rate of Japanese GDP per unit of labor into the growth rate of capital per unit of labor and total factor productivity (TFP). The first column denotes the sample years, the second denotes the average annual growth rate of GDP per unit of labor, and the remaining columns denote the contribution of capital per unit of labor and TFP to GDP growth. According to Table 1, TFP was a more important contributor to the Japanese growth miracle than capital during the 1956 – 73 period. This result is consistent with the past growth accounting exercises, for example, in Christensen, Cummings, and Jorgenson (1980), in Dougherty (1991) and in Jorgenson and Yip (2001). The growth accounting exercises in the sub-sample periods show that the contribution of capital to GDP growth changed over time. Capital contributed only 16.5 percent to the Japanese growth rate between 1956 and 1966. But capital's contribution to GDP growth jumped to more than 50 percent after the middle of the 1960s. After the middle of the 1960s, however, capital contributed more than 50 percent to GDP growth. Christensen, Cummings and Jorgenson (1980) also pointed out this phenomenon and further report that similar phenomena were observed in other countries, such as Germany and Italy, that also recovered rapidly from World War II.

These three facts are interesting characteristics of the Japanese growth miracle. Suppose that the Japanese growth miracle is a recovery from the capital destruction due to World War II, so the level of Japanese capital was much smaller than its steady state level. In this case, the optimal growth model predicts that capital would be accumulated quickly because of its high marginal product, and as a result output would grow quickly. The theory also predicts, however, that output growth would decrease over time due to diminishing returns to capital. Furthermore, the rate of return to capital would decrease along with capital accumulation. These predictions contradict the data. The theory also faces difficulty in explaining the change in capital contribution

	Growth Rate (%)	Contribution (%)	
	GDP	Capital	TFP
1956 – 73	6.9	36.7	63.3
1956 – 66	5.5	16.5	83.5
1963 – 73	7.9	51.8	48.2

Table 1: GDP per unit of labor: Growth Accounting)

to output growth since the growth rate of capital would decrease as the economy converges toward the balanced growth path.

In order to understand the Japanese growth miracle, this paper develops a model in which the measured TFP is endogenously determined. In our model, TFP is defined as the weighted average of productivity across workers. Endogenous changes in the composition of low and high-productivity workers drive TFP growth. These changes occur through a diffusion process. In this process, a low-productivity worker acquires knowledge from an existing high-productivity worker in order to become a high-productivity worker in the future. Therefore, increases in the share of high-productivity workers not only lead to increases in TFP, but also help to turn low-productivity workers into high-productivity workers in the future.

Our model quantitatively explains the three facts of the Japanese growth miracle described above. The model works as follows: the economy starts with a small proportion of high-productivity workers. Initially, more output is used in order to increase the share of high-productivity workers than to invest in capital stock. As a result, both TFP and the rate of return to capital increase. Once the share of high-productivity workers is sufficiently large, more output is allocated to investment in capital. Capital grows rapidly due to the high marginal product. Furthermore, the rate of return decreases along with capital accumulation.

This paper is the first quantitative work that sheds light on the third fact: the contribution of capital to GDP growth increased over time. For example, Parente and Prescott (1994) developed a model of technology adoption and succeeded in replicating the growth of Japanese GDP quantitatively. However, their model cannot capture the other two facts: the inverted U-shape rate of return and the changes in the capital contribution. Our work, on the other hand, replicates all three facts at the same time.

The paper is organized as follows. Sections 2 through 4 show the model, the parameter calibration and the baseline results. The last section contains concluding remarks and some ideas for future research. We describe the details of the data manipulation and report the sensitivity analyses in the Appendix.

2 Model

In the model, the economy features a representative household, which is infinitely lived and maximizes its lifetime utility. Only one type of goods exists, and producing the goods requires capital and two types of workers: high and low type. A high-type worker has higher productivity than a low-type worker. In the production process, high- and low-type workers are perfect substitutes. Output is used for consumption, investment in physical capital, and payment for turning low-type workers into high-type workers. Households invest in physical capital, which depreciates at rate δ over time. The household can turn low-type workers into high-type workers by paying a cost. We call this process "training". Training occurs through a knowledge diffusion process. In this process, a low-type worker acquires knowledge from an existing high-type in order to become a high-type worker in the future. In the model, not all low-type workers will become a high-type workers because some low-type workers may not interact with existing high-type workers despite paying a training cost. We describe the details of the model below.

Household

A household, which we call a family, consists of a large number (continuum) of people distributed between $(0, N_t)$. Each family member corresponds to a worker. Furthermore, each worker is either high or low type, and inelastically supplies one unit of time regardless of types. The size of a family grows exogenously at rate n. All new born family members are counted as low-type workers. In this model, the decision-making unit is the family rather than an individual worker. The family maximizes the lifetime utility of all family members by choosing the level of consumption and capital stock, and the number of low-type workers who are trained. The objective function, the lifetime utility of all family members, is

$$\sum_{t=0}^{\infty} \beta^t N_t \frac{c_t^{1-\sigma}}{1-\sigma},$$

where c is consumption per family member, β is the subjective discount factor and σ is the inverse of the intertemporal elasticity of substitution.

The income of the family is a sum of the wage income of all family members and the return to capital. Therefore, the budget constraint is:

$$C_t + K_{t+1} - (1-\delta)K_t + p_t s_t (N_t - N_{Ht}) = r_t K_t + w_{Ht} N_{Ht} + w_{Lt}(N_t - N_{Ht})$$

Here, C is consumption, K is the family capital stock, and r is the rate of return to capital. The physical capital stock exogenously depreciates at rate δ, and w_H and w_L are the wage of high- and low-type workers, respectively. The number of high-type workers in the family is given by N_H and $(N - N_H)$ is the number of low type. The per-worker cost of training is given by p, and $s_t \in [0, 1]$ is the fraction of low-type workers who are being trained in period t to become high-type workers in the next period. Hence, $s(N - N_H)$ is the number of low-productivity workers being trained today. The second and third terms on the left-hand-side (hereafter LHS) of the previous equation express the net investment in physical capital. The fourth term is the expenditure for training. The LHS is the total expenditure while the expression on the rignt-hand-side (hereafter RHS) is total family income. Consumption is distributed to all family members equally, that is, $C_t = N_t c_t$.

The model makes three assumptions related to workers:

1. Low-type workers work and earn regardless of being trained.

2. The number of new high-type workers is limited by the current share of this type of worker in income. Not all low-type workers can become high-type workers, even after paying a training cost, because some low-type workers may not be able to interact with existing high-type workers.

3. The family takes into account both effects of increases in the share of high-type workers, increasing total income and decreasing training costs, when making decision both on consumption and investment and on how many low-type workers to train in the current period.

The backdrop of these assumptions is as follows: beginning in the middle of the 1950s, Japanese firms sent some of their workers to developed countries, mainly the United States, where they could improve their knowledge, technology, and skill. When these workers returned to Japan, they taught the other Japanese workers while they were working. In addition, the Japanese government and business organizations invited foreign engineers, managers and consultants to come to Japan, and they organized seminars for Japanese workers.[4] Firms had their employees visit the developed countries and participate in the seminars, not only because they wanted highly productive people but also because they expected trained workers to help other workers improve their knowledge. In other words, the firms took into account the effects of knowledge diffusion when they trained their workers.

The law of motion governing the share of high-type workers is:

$$(1 + n_t)\phi_{t+1} = s_t(1 - \phi_t)\phi_t + (1 - \delta_w)\phi_t,$$

where $\phi \equiv \frac{N_H}{N}$ is the share of high-type workers in the family, δ_w is the depreciation rate of high-type workers, and $s(1 - \phi)$ represents the share of low-type workers who are trained. Since a low-type worker has to acquire knowledge from an existing high-type worker, only the fraction ϕ out of $s(1 - \phi)$ can be high type. Therefore, the first term on the RHS shows the number of

[4] Although the data regarding these activities are very limited, we can see anecdotal evidence in government and business reports, for example, see reports issued by the Japan Productivity Center for Socio-Economic Development (1965 and 1985).

new high-type workers in the next period. The second term implies that some high-type workers become low type. The interpretation of δ_w is as follows: productivity levels grow exogenously, and, as a result, the embodied productivity of some high-type workers becomes obsolete. The term $(1+n_t)$ on the LHS comes from the assumption that the new born family members are counted as a low-type worker.

The law of motion of the share of high-type workers is the following.

$$(1+n_t)\phi_{t+1} = s_t(1-\phi_t)\phi_t + (1-\delta_w)\phi_t,$$

where $\phi \equiv \frac{N_H}{N}$ is the share of high-type workers in the family. δ_w is the depreciation rate of high-type workers, and $s(1-\phi)$ represents the share of low-type workers who are trained. Since a low-type worker has to acquire knowledge from an existing high- type worker, only the fraction ϕ out of $s(1-\phi)$ can be high type. Therefore, the first term on the RHS shows the number of new high-type workers in the next period. The second term implies that some high-type workers become low type. The interpretation of δ_w is as follows: productivity levels grow exogenously, and, as a result, the embodied productivity of some high-type workers becomes obsolete. The term $(1+n_t)$ on the LHS comes from the assumption that the new born family members are counted as a low-type worker.

To summarize, the optimization problem per worker is the following.

$$\begin{aligned}
& \max_{\{c_t, k_{t+1}, s_t, \phi_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \hat{\beta}_t \frac{c_t^{1-\sigma}}{1-\sigma} \\
s.t. \quad & c_t + k_{t+1}(1+n_t) - (1-\delta_t)k_t + p_t s_t(1-\phi_t) \\
& = r_t k_t + w_{Ht}\phi_t + w_{Lt}(1-\phi_t) \\
& (1+n_t)\phi_{t+1} = s_t(1-\phi_t)\phi_t + (1-\delta_w)\phi_t \\
& (1+n_t)\phi_{t+1} \leq (1-\phi_t)\phi_t + (1-\delta_w)\phi_t \\
& (1+n_t)\phi_{t+1} \geq (1-\delta_w)\phi_t \\
& \text{given } k_0 \text{ and } \phi_0 > 0.
\end{aligned}$$

where $\hat{\beta}_t \equiv \beta^t \left(\prod_{s=0}^{t-1}(1+n_s)\right)$. The first constraint is the budget constraint per worker. The depreciation rate of physical capital and the growth rate of the family size have time subscript because we allow these variables to exogenously change over time. The second is the law of motion of the share of high-type workers. The third and fourth constraints are the upper and lower bounds of ϕ_{t+1}.

Firm

The firm maximizes its profit taking prices as given. The production function is of the Cobb-Douglas type, and it depends on both the capital stock and labor input.

$$F(K_t, N_{Ht}, N_t) = K_t^{\alpha}[A_{Ht}N_{Ht} + A_{Lt}(N_t - N_{Ht})]^{1-\alpha}$$

Here A_H and A_L are the productivity level embodied, respectively, in a high- and low-type labor. As we mentioned above, high- and low-type workers are assumed to be perfect substitutes. A description of the firm's optimization behavior is given by:

$$\begin{aligned} r_t &= \alpha K_t^{\alpha-1}[A_{Ht}N_{Ht} + A_{Lt}(N_t - N_{Ht})]^{1-\alpha} \\ w_{Ht} &= (1-\alpha)A_{Ht}K_t^{\alpha}[A_{Ht}N_{Ht} + A_{Lt}(N_t - N_{Ht})]^{-\alpha} \\ w_{Lt} &= (1-\alpha)A_{Lt}K_t^{\alpha}[A_{Ht}N_{Ht} + A_{Lt}(N_t - N_{Ht})]^{-\alpha} \end{aligned}$$

where r is the rental rate (equal to the marginal product) of capital, and w_H and w_L are the wages of high- and low-type workers.

In the model, aggregate productivity is defined as

$$[A_H\phi_t + A_L(1-\phi_t)]^{1-\alpha}.$$

Exogenous Growth

We assume that the levels of productivity of both types of labor and the unit cost of training, p_t, exogenously grow over time. Their growth rates are common across labor types but can vary across time. That is

$$\begin{aligned} &A_{Ht} = Z_t A_H \\ &A_{Lt} = Z_t A_L \\ &p_t = Z_t p \\ &\text{where } Z_t \equiv \left(\prod_{s=0}^{t-1}(1+\gamma_s)\right). \end{aligned}$$

Here, $A_H$and A_L are the initial productivity levels, p is the initial unit training cost, and γ_t is the growth rate at time t. In balanced growth, both C_t and K_{t+1} grow at $(1+n)(1+\gamma)-1$, where n and γ are constants. The wage ratio between high- and low-type workers remains constant over time due to (1) the assumption of a common growth rate of the productivity across the two types of workers, (2) the assumption of perfect substitutability between the two types of workers. The next section shows the evidence supporting a constant wage ratio.

Equilibrium

First, we normalize C_t and K_t by productivity growth, Z_t, and the number of workers N_t, and denote the normalized variables by $\tilde{c}_t$ and $\tilde{k}_t$. In addition, normalized output, $\tilde{y}$, is defined as

$$\tilde{y}_t \equiv \tilde{k}_t^{\alpha} A_t^{1-\alpha},$$

where $A_t \equiv A_H \phi_t + A_L(1-\phi_t)$. The wage rates for high- and low-type workers is normalized by Z_t as well.

$$\tilde{w_H} = \frac{w_H}{Z_t}$$
$$\tilde{w_L} = \frac{w_L}{Z_t}$$

The competitive equilibrium is defined as follow.

Competitive Equilibrium

In this subsection, we define competitive equilibrium. A competitive equilibrium is an allocation $\{\tilde{c}_t,\ \tilde{k}_{t+1},\ \phi_{t+1}\}_{t=0}^{\infty}$ and prices $\{r_t,\ \tilde{w}_{Ht},\ \tilde{w}_{Lt}\}_{t=0}^{\infty}$ such that

- $\{\tilde{c}_t,\ \tilde{k}_{t+1},\ \phi_{t+1}\}_{t=0}^{\infty}$ is the solution to the household's following utility maximization problem, given the sequences of prices:

$$\max_{\{\tilde{c}_t, \tilde{k}_{t+1}, \phi_{t+1}\}} \sum_{t=0}^{\infty} \tilde{\beta}_t \frac{\tilde{c}_t^{1-\sigma}}{1-\sigma}$$
$$\begin{aligned} s.t.\quad & \tilde{c}_t + \tilde{k}_{t+1}(1+n_t)(1+\gamma) + p\left[\frac{(1+n_t)\phi_{t+1} - (1-\delta_w)\phi_t}{\phi_t}\right] \\ & = (r_t + 1 - \delta_t)\tilde{k}_t + \tilde{w}_{Ht}\phi_t + \tilde{w}_{Lt}(1-\phi_t) \\ & (1+n_t)\phi_{t+1} \le (1-\phi_t)\phi_t + (1-\delta_w)\phi_t \\ & (1+n_t)\phi_{t+1} \ge (1-\delta_w)\phi_t \\ & \text{given } \tilde{k}_0 \text{ and } \phi_0 > 0, \end{aligned}$$

  where $\tilde{\beta}_t = \hat{\beta}_t Z_t^{1-\sigma}$.

- Given the price sequences, a firm maximizes its profit:

$$\begin{aligned} r_t &= \alpha \tilde{k}_t^{\alpha-1} A_t^{1-\alpha} \\ \tilde{w}_{Ht} &= (1-\alpha) A_H \tilde{k}_t^{\alpha} A_t^{-\alpha} \\ \tilde{w}_{Lt} &= (1-\alpha) A_L \tilde{k}_t^{\alpha} A_t^{-\alpha} \end{aligned}$$

- E The market clearing condition is given by:

$$\tilde{c}_t + \tilde{k}_{t+1}(1+n_t)(1+\gamma) + p\left[\frac{(1+n_t)\phi_{t+1} - (1-\delta_w)\phi_t}{\phi_t}\right] = \tilde{y}_t + (1-\delta_t)\tilde{k}_t$$

The equilibrium system is given by:

- $(1+\gamma)\tilde{c}_t^{-\sigma} = \beta(1+\gamma)^{1-\sigma}\tilde{c}_{t+1}^{-\sigma}\left[\alpha\tilde{k}_{t+1}^{\alpha-1}A_{t+1}^{1-\alpha} + 1 - \delta_t\right]$
- $\tilde{c}_t^{-\sigma}\dfrac{p}{\phi_t} + \mu_{1t} - \mu_{2t}$
  $= \beta(1+\gamma)^{1-\sigma}\tilde{c}_{t+1}^{-\sigma}\left[(1-\alpha)\tilde{k}_{t+1}^{\alpha}A_{t+1}^{-\alpha}(A_H - A_L) + \dfrac{p(1+n_{t+1})\phi_{t+2}}{\phi_{t+1}^2}\right]$
  $+\beta(1+\gamma)^{1-\sigma}\left[\mu_{1t+1}(2 - 2\phi_{t+1} - \delta_w)\right] - \beta(1+\gamma)^{1-\sigma}\left[\mu_{2t+1}(1-\delta_w)\right]$
- $\tilde{k}_t^{\alpha}A_{t+1}^{1-\alpha} = \tilde{c}_t + \tilde{k}_{t+1}(1+\gamma)(1+n_t) - (1-\delta_t)\tilde{k}_t + p\left[\dfrac{(1+n_t)\phi_{t+1} - (1-\delta_w)\phi_t}{\phi_t}\right]$
- $\mu_{1t}\left[\dfrac{(1-\phi_t)\phi_t + (1-\delta_w)\phi_t}{1+n_t} - \phi_{t+1}\right] = 0$
- $\mu_{2t}\left[\phi_{t+1} - \dfrac{(1-\delta_w)\phi_t}{1+n_t}\right] = 0$
- $\mu_{1t}, \mu_{2t} \geq 0$ for all t,

Here, μ_{1t} and μ_{2t} are the Lagrangian multipliers for the upper- and lower-bound constraints of ϕ_{t+1}.

The first equation represents the trade-off between consumption and investment, and is quite standard in an optimum growth model. The second equation represents the trade-off between consumption and training of a low-type worker. The first term on the RHS expresses the marginal benefit of increasing ϕ_{t+1}. The large bracket in this term can be divided into two parts: $(1-\alpha)\tilde{k}_{t+1}^{\alpha}A_{t+1}^{-\alpha}(A_H - A_L)$ and $\frac{p(1+n_t)\phi_{t+2}}{\phi_{t+1}^2}$. The former corresponds to the benefit due to increases in labor income by increasing the share of the high-type workers. The latter, on the other hand, corresponds to the benefit due to the fact that increases in the share of high-type workers make it easier to turn low-type workers into high-type workers. The third equation is the resource constraint. The remaining equations are slackness conditions. Since we cannot avoid the case that either boundary condition is binding, the corner solution (when either μ_{1t} or μ_{2t} is strictly positive) is sometimes relevant.

We would like to note that this system depends on the productivity ratio, $\frac{A_H}{A_L}$, rather than the individual productivity level. We can show it by normalizing $\tilde{c}_t$, $\tilde{k}_{t+1}$, μ_{1t}, μ_{2t}, A_H and p by A_L.

Denote these normalized variables and parameters using hat. Then, the new system will be

- $(1+\gamma)\hat{c}_t^{-\sigma} = \beta(1+\gamma)^{1-\sigma}\hat{c}_{t+1}^{-\sigma}\left[\alpha\hat{k}_{t+1}^{\alpha-1}\hat{A}_{t+1}^{1-\alpha} + 1 - \delta_t\right]$
- $$\begin{aligned}&\hat{c}_t^{-\sigma}\frac{p}{\phi_t} + \hat{\mu}_{1t} - \hat{\mu}_{2t}\\ &= \beta(1+\gamma)^{1-\sigma}\hat{c}_{t+1}^{-\sigma}\left[(1-\alpha)\hat{k}_{t+1}^{\alpha}\hat{A}_{t+1}^{-\alpha}(\hat{A_H}-1) + \frac{\hat{p}(1+n_{t+1})\phi_{t+2}}{\phi_{t+1}^2}\right]\\ &+\beta(1+\gamma)^{1-\sigma}\left[\hat{\mu}_{1t+1}(2-2\phi_{t+1}-\delta_w)\right] - \beta(1+\gamma)^{1-\sigma}\left[\hat{\mu}_{2t+1}(1-\delta_w)\right]\end{aligned}$$
- $\hat{k}_t^{\alpha}\hat{A}_{t+1}^{1-\alpha} = \hat{c}_t + \hat{k}_{t+1}(1+\gamma)(1+n_t) - (1-\delta_t)\hat{k}_t + \hat{p}\left[\frac{(1+n_t)\phi_{t+1} - (1-\delta_w)\phi_t}{\phi_t}\right]$
- $\hat{\mu}_{1t}\left[\frac{(1-\phi_t)\phi_t + (1-\delta_w)\phi_t}{1+n_t} - \phi_{t+1}\right] = 0$
- $\hat{\mu}_{2t}\left[\phi_{t+1} - \frac{(1-\delta_w)\phi_t}{1+n_t}\right] = 0$

where $\hat{A}_t \equiv \hat{A_H}\phi_t + (1-\phi_t)$. This characteristics allows us to set $A_L = 1$ without loss of generality.

Steady States

Given a parameter set (except for the initial conditions), there are two steady states where ϕ is positive. The first is a steady state in which both μ_1 and μ_2 are zero and the second is a steady state in which μ_1 is strictly positive. The former steady state, where both ϕ and $\tilde{k}$ are the interior solution, is knife edge.[5] Unless the economy starts with exactly the interior steady state allocation, the economy will converge either to the latter steady state ($\mu_1 > 0$) or the steady state with $\phi = 0$ ($\mu_2 > 0$). Our calibrated parameter values lead the economy toward the steady state with $\mu_1 > 0$.

The steady state conditions with $\mu_1 > 0$ are the following.

$$\begin{aligned}&(1+\gamma)^{\sigma} = \beta\left[\tilde{k}^{*\alpha-1}A^{*1-\alpha} + 1 - \delta\right]\\ &\tilde{k}^{*\alpha}A^{*1-\alpha} = \tilde{c}^* + \tilde{k}^*[(1+n)(1+\gamma) - (1-\delta)] + p(n+\delta_w)\\ &\phi^* = 1 - (n+\delta_w)\end{aligned}$$

where $A^* \equiv A_H\phi^* + A_L(1-\phi^*)$.

In the next section, we describe the strategy of parametrization.

3 Calibration

The parameter values are determined so that the model can capture the Japanese post-war economy. Table 2 summarizes our calibration strategy and the values used. The parameters in Panel A are

[5] Although we do not offer formal proof, the log-linearlized system of the equilibrium with calibrated parameter values is unstable around this interior steady state.

Panel A

	Description	Target	Value
α	Capital income share	National Account : $\frac{\text{Capital Income}}{\text{GDP}}$	0.361
n_t	Growth rate of the family size	Growth rate of the number of workers	0.01
δ_t	Depreciation rate of capital	Hayashi and Prescott (2002)	0.108
$\tilde{k}_0/\tilde{y}_0$	Initial capital-output ratio	Capital-output ratio in 1956	1.37
σ	Inverse of IMRS	-	1.5

Panel B

	Description	Target	Value
γ_t	Exogenous growth rate	Growth rate of GDP per worker in US	0.02
δ_w	Depreciation rate of high type	Growth rate of individual wage	0.03
β	Subjective discount factor	Steady state capital-output ratio =	0.97
		Average capital-output ratio in the 1990s	
A_H/A_L	Productivity ratio	$90-10$ ratio in wage	2.2
ϕ_0	Initial share of high type	Average growth rate of GDP between 1956 and 1980	0.07
p	Unit cost of training	Steady state training expenditure share =	1.12
		Training expenditure share in 2005	

Table 2: Calibration Summary

determined by direct observation; those in Panel B depend on the model specifications. We will now explain the parameters one by one. Since our data start in 1956, the initial period, $t = 0$ in our model, corresponds to 1956 in the real world. Capital's income share, α, is defined as $\frac{\text{Capital Income}}{\text{GDP}}$. To compute it, we calculate capital's income share every year from 1956 to 2000 in Japan, and then take the average of them. The value of α is 0.361. The source of these data is "Annual Report of National Accounts".

We use the growth rate of employment as the rate of growth in family size, n_t. The steady state level of n is the average of the growth rate of the number of workers between 1956 and 1998, which is 0.01. In the economy, $\{n_t\}_{t=0}^{43}$ is the actual growth rate of the number of workers and $\{n_t\}_{t=44}^{\infty} = 0.01$. The source is "Monthly Labor Survey" (*Maitsuki Kinro Tokei Chosa* in Japanese).

We also use the series of the depreciation rate, δ, to compute the equilibrium. Figure 3 shows the depreciation rate of capital between 1956 and 1980, which exhibited a decreasing trend. Thus, constant depreciation is not a reasonable assumption for analyzing the Japanese growth miracle. The steady state depreciation rate is assumed to be the average depreciation rate between 1956 and 2000, which is 0.108. When we compute the dynamics of the model, the depreciation rate for the first 45 years is the same as that shown by data from 1956 to 2000, and it remains constant after period 46. The data series of depreciation are taken from Hayashi and Prescott (2002).

For the initial capital-output ratio $\frac{\tilde{k}_0}{\tilde{y}_0}$, we use the capital-output ratio in 1956 estimated by

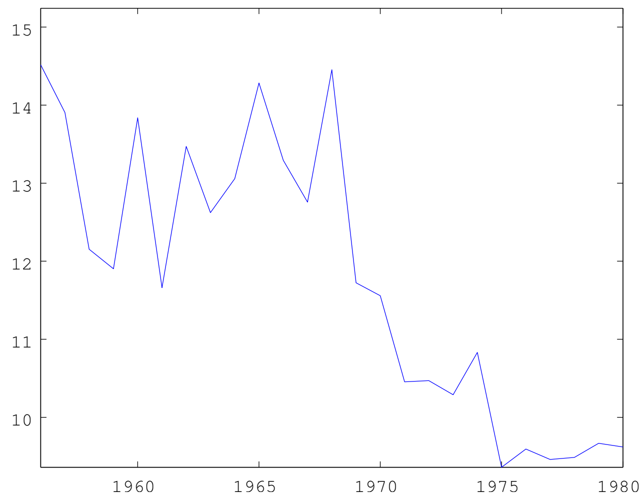

Figure 3: Depreciation Rate (%)

Hayashi and Prescott (2002)[6], which is 1.37.

The inverse of intertemporal elasticity of substitution, σ, is set to 1.5. This value is quite common. For example, it is used by He and Liu (2008), which is one of the most recently published studies of human capital.

We take four steps to compute the series of the growth rate of productivity (and training cost), γ_t, in order to make the model's variables consistent with the data described in the introduction. This is because our model ignores the observed quality changes mentioned, for instance schooling and hours worked. First, we define γ_t as

$$1 + \gamma_t = (1 + q_t)(1 + g)$$

where q_t is the growth rate of the observed labor quality. g is the growth rate of neutral technology, which is assumed to be constant over time. Second, we calculate $\{q_t\}_{t=0}^{\infty}$ as follows.

$$1 + q_t = \begin{cases} \frac{1+l_t}{1+n_t} \text{ if } t \leq 43 \\ \frac{1+l}{1+n} \text{ if } t \geq 44 \end{cases}$$

where l_t is the growth rate of the labor input computed from the growth accounting exercise. After $t = 44$, we assume l and n become the steady state values. The average values between 1956 and 1998 are used for the steady state values of l and n. This implies that q after $t = 43$ is -0.001. Since the observed quality here is affected by changes in hours worked (which exhibited a strong decreasing trend in Japan), the growth rate of labor quality in the steady state is negative. Third,

[6] Hayashi and Prescott (2002) estimate two capital-output ratios depending on accounting methods. The number we use is based on the capital excluding foreign factors.

we assume the productivity growth rate in the steady state, γ, is the same as the growth rate of US GDP per worker, which is 0.02.[7] The steady state relationship among γ, q and g,

$$1+g=\frac{1+\gamma}{1+q},$$

implies $g=0.021$. Finally, the series of γ_t is calculated using $\{q_t\}$ and g.

Our model's targets, GDP per unit of labor, is real GDP normalized by the observed labor input, $(1+n_t)(1+q_t)$, as described in the introduction. In the model, on the other hand, $\tilde{y}$ is defined as the output normalized by $(1+n_t)$ and $(1+q_t)(1+g)$ $(\equiv 1+\gamma_t)$. Therefore, the variables in the data and the model are conceptually consistent when we reconstruct the series of output by multiplying $\tilde{y}_t$ by neutral technology growth $(1+g)^t$.

In order to infer the depreciation rate of the share of high-type workers, δ_w, we assume that its value is the same as the depreciation rate of human capital in a life-cycle model. In such a model, the depreciation rate of human capital is calculated using wage profiles. First, we compute the average growth rate of real wage from workers aged 55 to 65 who retired during the period of 1973-2003.[8] We then calculate the average value over time. The average annual wage growth rate during the final 10 years before retirement is -1.1%. Since an older worker invests very little in his human capital, the growth rate of his wage can be expressed by $(1+\gamma)(1-\delta_w)$ in a stationary equilibrium.[9] The depreciation rate of high-type workers, δ_w, is 0.03 when $\gamma=0.02$.

Given γ, σ, δ and n, we can compute the subjective discount factor, β, from the first equation of the steady state system and capital-output ratio. We use 2.13 as the steady state capital-output ratio, which is the average capital-output ratio between 1988 and 1998.[10]

As we mentioned in the previous section, the levels of productivity is not relevant but the productivity ratio between high and low types is. Therefore, we have to consider only how to calibrate $\frac{A_H}{A_L}$, rather than each productivity level. In our model, the wage ratio $\frac{w_{Ht}}{w_{Lt}}=\frac{A_H}{A_L}$ is constant over time. Therefore, we use $90-10$ wage ratio as the target of the productivity ratio. We first calculate the average $90-10$ wage ratio of monthly Japanese male wage for each 5-year-age cohorts, and then take the weighted average of these values across age cohorts. Figure 4 shows the result. An individual's wage in Japan strongly depends on the length of his tenure at in the same company. It also strongly correlates with ages because of the Japanese life-time employment system. In that sense, the average $90-10$ wage ratio across the age cohorts expresses the difference in productivity without the effects of time-dependent experience. Furthermore, changes in age-dependent labor quality are captured as changes in labor input in the growth accounting exercises;

[7] Since GDP growth rate per worker in the United States is stable around two percent, we assume the US economy is on the balanced growth path.

[8] The data source is "Basic Survey of Wage Structure (*Chingin Kouzou Kihon Chosa* in Japanese) issued by Ministry of Health, Labour and Welfare.

[9] We refer to Hugget, Ventura and Yaron (2006) for this method.

[10] Capital-output ratio here means capital input over output.

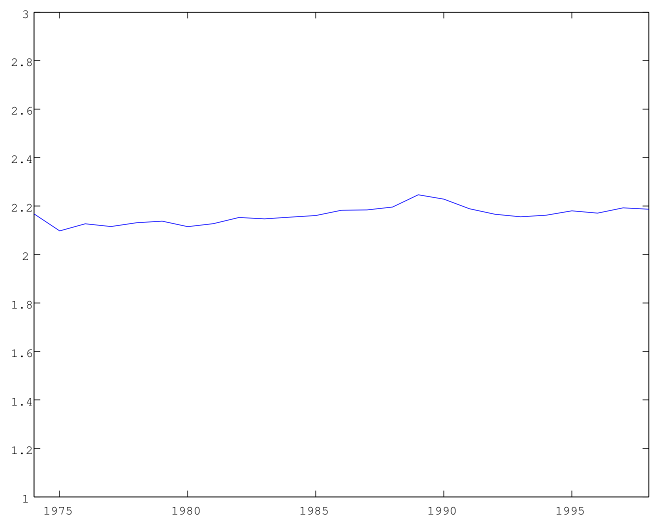

Figure 4: Wage Skill-Premium: 1974-1998

that is, the measured TFP does not include improvement of age-dependent labor quality. Since changes in TFP in our model must be free from the measured qualities of labor, the average wage ratio across age cohorts is more reasonable than the wage ratio of all workers. This ratio has been quite stable since 1974.[11] The average between 1974 and 1998 is 2.2. The data before 1973 are limited. However, the average ratio from 1958 to 1960 is also 2.2. These facts support the implication of the model, i.e., that the wage ratio is approximately constant. This conclusion results from two assumptions: that the two types of workers are perfectly substitutable, and that both types show the same growth in productivity. We set $\frac{A_H}{A_L} = 2.2$.

Finally, we simultaneously calibrate the initial share of high-type workers (ϕ_0) and unit cost of training (p), using following targets: First, the average growth rate of GDP in the first 25 periods matches the average annual growth rate of GDP per unit of labor between 1956 and 1980. Second, the steady state training expenditure share of output matches the training expenditure share of total sales in Japan in 2005.

The average annual growth rate of GDP per unit of labor between 1956 and 1980 was 5.8 percent.

The unit cost of training is set such that the steady state expenditure share of training is 1.4%. This number is based on "Basic Survey of Skill Development" (*Nouryoku Kaihatsu Kihon Chosa* in Japanese) in 2006 and "Survey of Incorporated Business Statistics" (*Hojin Kigyo Tokei Chosa* in Japanese) in 2006. According the former survey, the average cost *Off-the-Job Training* per firm was approximately 3669000 yen per company. The cost relative to total sales taken from the latter

[11]The data source is "Basic Survey of Wage Structure" (*Chingin Kouzou Kihon Chosa* in Japanese) issued by Ministry of Health, Labour and Welfare.

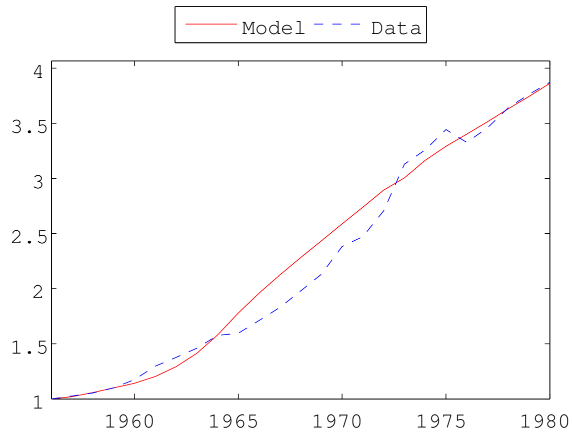

Figure 5: GDP per unit of labor: Model vs Data

survey is 0.67% in 2005. Since the costs of training while working should also include costs of On-the-Job Training[12] which are unobservable, we simply doubled the number as a total training cost share.

The values of the initial share of high-type workers, the levels of productivity of high- and low-type workers and the unit cost of training are $\phi_0 = 0.07$ and $p = 1.12$.

Now we use these parameter values for our baseline results.

4 Results

Figure 5 shows the output per unit of labor. The solid line shows the unit output predicted by the model and the dashed line shows the actual unit output as given in the data. The GDP path in the model is similar to that in the data. The striking result of our model is that it predicts just the GDP growth acceleration that we observe in the data. In the model, the average annual growth rate is 6.9 percent (in the $1956-66$ period), while it is 7.9 percent (in the $1963-73$ period). The actual data give these two period percentages as 5.5 percent and 7.9 percent, respectively.

Figure 6 shows the rate of return to capital. The rate from the data (dashed line) increased from 25 percent to 37 percent, and then decreased gradually. Our model also predicts the inverted U-shape of the rate of return during the miracle period. The rate of return first increases from 26 percent to 45 percent, and then decreases.

[12]"training while working" includes both off-the-job and on-the-job training. "Off-the Job Training" means training that a firm asks the other firms to provide. "On-the-Job Training", on the other hand, is training provided inside of a firm.

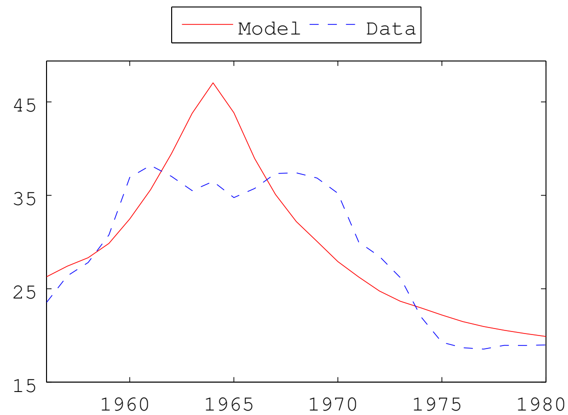

Figure 6: Rate of Return to Capital (%): Model vs Data

Table 3 shows the growth accounting in the model and the data between 1956 and 1973 (Panel A), 1956 and 1966, and 1963 and 1973 (Panel B). The row labeled "Growth Rate" shows the average annual growth rate of GDP per unit of labor, capital per unit of labor, and TFP respectively during each sample period. The row labeled "Contribution" shows the percentage contribution of capital and TFP to GDP growth. For example, from 1956 to 1973, the average annual growth rate of capital in the model is 7.4 percent while it is 7.0 percent in the data. The contribution of capital to GDP growth is 40.0 percent in the model and 36.7 percent in the data during this time period.

Our model captures the important characteristics of the factor contributions: the contribution of capital was small at the beginning while capital became more important than TFP later. Capital stock contribution is only 14.8 percent in the first decade. Indeed, the growth rate of capital is close to that observed in the data. However, since the growth rate of TFP in the model is higher than that in the data, the contribution of capital is slightly smaller. The contribution of capital is, on the other hand, 66.8 percent between 1963 and 1973; that is, capital contributes more than 50 percent to GDP growth, as we can observe in the data.

These results are based on the following mechanics: The economy starts with a small proportion of high type workers. Initially, more output is used in order to increase the share of high-type workers than to invest in capital stock. As a result, both TFP and the rate of return to capital increase. Once the share of high-type workers becomes sufficiently large, more output is allocated to invest in physical capital. Capital grows rapidly due to its high marginal product. In other words, our model has two stages: a stage of TFP growth, followed by a stage of capital growth.

Table 4 shows the comparison of our results with those of Parente and Prescott (1994). We reverse-engineer the growth rate of output, capital, and technology from Table 2 in their paper. Since the initial period in their model is 1960, and since we can compute the levels of output,

Panel A

1956-73

Data	Output	Capital	TFP
Growth Rate(%)	6.9	7.0	4.4
Contribution(%)	–	36.7	63.3

Model	Output	Capital	TFP
Growth Rate(%)	6.6	7.4	4.0
Contribution(%)	–	40.0	60.0

Panel B

	1956-66			**1963-73**		
Data	Output	Capital	TFP	Output	Capital	TFP
Growth Rate(%)	5.5	2.5	4.6	7.9	14.7	2.6
Contribution(%)	–	16.5	83.5	–	51.8	48.2

	1956-66			**1963-73**		
Model	Output	Capital	TFP	Output	Capital	TFP
Growth Rate(%)	6.9	2.8	5.9	7.9	12.3	1.3
Contribution(%)	–	14.8	85.2	–	66.8	33.2

Table 3: Growth Accounting: Model vs Data

capital and TFP every five years in their model, the period for the growth accounting is slightly different from that in Table 3. Panel A shows the annual growth rate and the factor contributions between 1960 and 1975. Panel B shows the results between 1960 and 1970, and between 1965 and 1975. Although their model replicates the Japanese GDP growth, most of their model's GDP growth is driven by TFP growth, which is inconsistent with the data. Furthermore, technology growth remains the main driver in the first 20 years, and its contribution does not change over time (see Panel B in Table4). In their model, GDP growth accelerates due to decreases in the cost of intangible capital accumulation over time. (Parente and Prescott (1994) call intangible capital "technology capital.") Furthermore, in their parameterizations, the share of capital income is approximately one third of the share of technology capital income. As a result, TFP growth drives the GDP growth rate for the entire period. In our model, on the other hand, TFP growth comes first and capital growth follows, and these changes in the factor contributions are consistent with the data.

To summarize, our model is consistent with the three characteristics of the Japanese post-war economy described in the Introduction.

Finally, our results depend on our law of motion describing the share of high-type worker and on some of our parameter values, such as the unit cost of training. In order to facilitate the reader's deeper understanding of the model's mechanics , we report many sensitivity analyses in the Appendix.

5 Conclusion and Future Research

This paper develops a growth model in which aggregate productivity is driven by endogenous shifts in labor input from low to high-productivity workers. Our model captures the following three facts of the Japanese growth miracle: the accelerated GDP growth, the inverted U-shape of the rate of return to capital, and the changes in factor contributions.

Our model has two key features. First, we define the measured TFP as the weighted average of productivity across workers so that endogenous changes in the composition of low- and high-productivity workers drive TFP growth. This feature implies that increases in the share of high-productivity workers lead to increases in TFP. Second, we introduce a knowledge diffusion process in which a low-productivity worker must acquire knowledge from an existing high-productivity worker in order to become a high-productivity worker in the future. In the model, changes in the composition of labor are limited by the current share of high-productivity workers. The implication of this idea is that increases in the share of high-productivity workers also help to turn low-productivity workers into high-productivity workers in the future.

The model works as follows: at the beginning, the share of high-productivity workers is small in the economy. Initially, people allocate more output in order to train low productivity workers

Panel A

1960-75

Data	Output	Capital	TFP
Growth Rate (%)	7.4	10.3	3.7
Contribution (%)	–	50.0	50.0

Model	Output	Capital	TFP
Growth Rate (%)	7.3	10.1	3.7
Contribution (%)	–	49.9	50.1

Parente and Prescott	Output	Capital	TFP
Growth Rate (%)	6.7	6.2	10.5
Contribution (%)	–	14.7	85.3

Panel B

	1960-70			**1965-75**		
Data	Output	Capital	TFP	Output	Capital	TFP
Growth Rate (%)	7.3	8.9	4.1	8.0	12.2	3.6
Contribution (%)	–	44.2	55.8	–	54.9	45.1

Model	Output	Capital	TFP			
Growth Rate (%)	8.5	10.2	4.8	6.3	13.8	1.5
Contribution (%)	–	43.4	56.6	–	77.1	22.9

Parente and Prescott	Output	Capital	TFP			
Growth Rate (%)	7.1	6.7	6.1	5.3	9.5	
Contribution (%)	–	15.1	84.9	–	14.0	86.0

Table 4: Growth Accounting: Comparison with Parente and Prescott (1994)

than to invest in capital stock. As a result, both TFP and the rate of return to capital increase. Once the share of high-productivity workers is sufficiently large, people use more output to invest in capital. Capital grows rapidly due to its initial high marginal product, but the rate of return decreases along with capital accumulation.

Indeed, this story is consistent with what occurred in Japan between 1950 and 1970. Beginning of 1950, Japanese firms sent their workers to the United States and other developed countries to improve their skills and knowledge. In addition, they had their workers attend seminars and meetings for workers and managers, inviting foreign engineers, managers and consultants to come to Japan. In the 1950s, the speakers and teachers in such seminars were foreigners. By the 1960s, the seminars were still held, but the teachers were now Japanese. Although hard data to directly support these facts are difficult to come by, anecdotal evidence is abundant.

For future research, our model may provide a basis for studying other countries' growth miracles. Germany would be a good example. Germany also achieved rapid recovery after World War II, and we can observe similar patterns in the rate of return and the factor contributions. Is the German post-war economic growth similar to the Japanese growth miracle? Our work provides a useful tool for answering this question.

References

[1] Chen, Kaiji, Ayse Imrohoroglu, and Selahattin Imrohoroglu, 2006, "The Japanese Saving Rate.", *American Economic Review*, Vol. 96(5), pp1850-1858

[2] Christensen, Laurits R., Dianne Cummings and Dale W. Jorgenson, 1980, "Economic Growth, 1947-1973 : An International Comparison.", *New Developments in Productivity Measurement and Analysis*, John W. Kendrick and Beatrice N. Vaccara, eds. National bureau of Economic Research, Studies in Income and Wealth, Vol. 44. Chicago: University of Chicago Press.

[3] Dougherty, John Chrysostom, 1991, "A Comparison of Productivity and Economic Growth." mimeo.

[4] Gilchrist, Simon and John C. Williams, 2004, "Transition Dynamics in Vintage Capital Models: Explaining the Postwar Catch-Up of Germany and Japan", NBER Working Paper, No. W10732

[5] Hayashi, Fumio and Edward C. Prescott, 2002, "Japan in the 1990s: A Lost Decade.", *Review of Economic Dynamics, January*, Vol. 5, pp206-235.

[6] He, Hui and Zheng Liu, 2008, 'Investment-Specific Technological Change, Skill Accumulation and Wage Inequality." *Review of Economic Dynamics*, Vol. 11(2), pp314-334

[7] Huggett, Mark, Gustavo Ventura and Amir Yaron, 2006, "Human capital and earnings distribution dynamics.", *Journal of Monetary Economics*, Vol. 53(2), pp265-290

[8] Japan Productivity Center for Socio-Economic Development, 1965, "The Decade History of Productivity Improvement Campaign." (*Seisansei Undo 10 Nen no Ayumi* in Japanese), Japan Productivity Center for Socio-Economic Development

[9] Japan Productivity Center for Socio-Economic Development, 1985, "The Three-Decade History of Productivity Improvement Campaign." (*Seisansei Undo 30 Nen Shi* in Japanese), Japan Productivity Center for Socio-Economic Development

[10] Jorgenson, Dale W. and Eric Yip, 2001, "Whatever Happened to Productivity Growth?", *New Developments in Productivity Analysis*, E.R. Dean, M.J. Harper, and Hulten, eds. Chicago: University of Chicago Press.

[11] Mincer, J., 1974, "Schooling, Experience, and Learning.", New York: Columbia University Press

[12] Parente, Stephen L. and Edward C. Prescott, 1994, "Barriers to Technology Adoption Development.", *Journal of Political Economy*, Vol. 102(2), pp298-321

[13] Stokey, Nancy L., 1996, "Free Trade, Factor Returns, and Factor Accumulation.", *Journal of Economics Growth*, Vol. 1(4), pp421-447

[14] Young, Alwyn, 1995, "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience.", *Journal of Political Economy*, Vol. 106(1), pp41-63

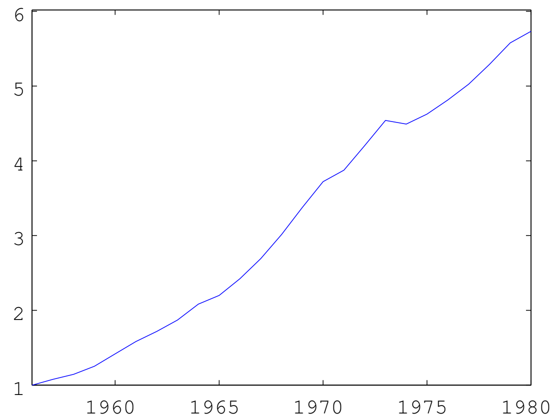

Figure 7: Japanese Real GDP

A Japanese Miracle

This section explains the details of the three facts that characterized the Japanese growth miracle: the accelerated GDP growth, the inverted U-shape rate of return, and the factor contribution changes. Figure 7 depicts the Japanese real GDP from 1956 to 1980. In Figure 7, the series are normalized by their levels in 1956. The average annual growth rate between 1956 and 1973 was 9.3 percent, the highest in the world during this time. For example, the growth rate of GDP during the same period was 4.0 percent in the United States, 6.6 percent in Germany, and 5.2 percent in Italy. This rapid economic growth is widely known as the Japanese growth miracle.

Figure 8 depicts the rate of return to capital from 1956 to 1980, which is defined as capital income per capital stock. Capital income is defined as

$$\text{capital income} \equiv \text{GDP - labor compensation}$$

and calculated using the data from the "Annual Report of National Accounts" in Japan. Capital stock data are taken from Hayashi and Prescott (2002). The dynamics of the rate of return during the miracle period have an inverted U-shape; that is, the rate of return increased at first and then gradually decreased.

Growth accounting exercises in which GDP growth is decomposed into input and productivity growth reveal changes in the factor contributions to GDP growth. In order to measure the growth rates of labor and capital input, this paper uses the method of Christensen, Cummings, and Jorgenson (1980). The details of method are described in the next section.

Table 5 shows the results of the growth accounting. The first column denotes the sample years, the second denotes the average annual percentage growth rate of GDP, and the remaining columns

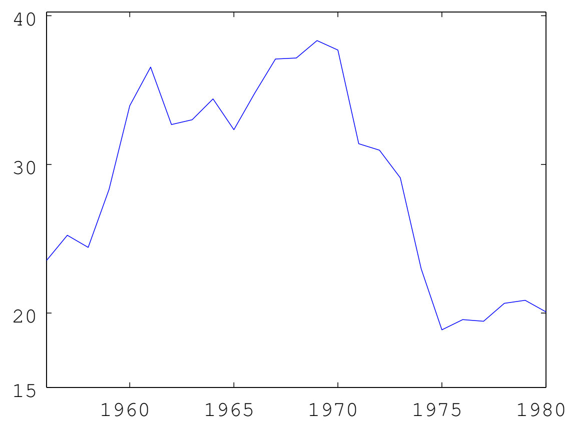

Figure 8: Rate of Return to Capital Stock (%)

	Growth Rate (%)	Contribution (%)		
	GDP	Labor	Capital	TFP
1956 – 73	9.2	15.4	37.5	47.1
1956 – 66	9.1	25.0	25.7	49.3
1963 – 73	9.2	8.6	49.6	41.8

Table 5: GDP Growth and Factor Contribution

denote, respectively, the factor contributions of labor, capital and TFP. The contribution of capital to the Japanese GDP growth between 1956 and 1973 was 38 percent, and the contribution of TFP was 47 percent. On the other hand, the contribution of labor input was 15 percent.[13]

The growth accounting exercises in the sub-sample periods show that the main driver of GDP growth changed over time. TFP made the highest contribution (49 percent) in the first decade, when it was twice as high as that of capital and labor. Between 1963 and 1973, however, capital made the greatest contribution. This pattern is observed even when the length of the sub-sample period changes.

[13]These accounting results imply that the Japanese growth miracle may have occurred by a mechanism different from that of the other Asian growth miracles. Young (1995) finds that most of the Asian miracles were driven mainly by physical and human capital accumulation rather than TFP growth, even though the GDP growth rates in East Asian countries from the end of the 1960s to the beginning of the 1990s were as high as that in Japan from the middle of the 1950s to the beginning of the 1970s.

	Growth Rate (%)	Contribution (%)	
	GDP	Capital	TFP
1956 – 73	6.9	36.7	63.3
1956 – 66	5.5	16.5	83.5
1963 – 73	7.9	51.8	48.2

Table 6: GDP Growth and Factor Contribution (per unit of labor)

GDP per unit of labor

According to the growth accounting exercise, the growth of labor input made a small contribution to the growth rate of Japanese GDP compared to the other two factors. We construct the series of GDP and capital both of which are normalized by growth of labor input. We call them GDP and capital per unit of labor. These manipulations remove the effects of the measured labor inputs from the original series of GDP and capital. These series will be used (directly or indirectly) as the targets of our model.

Figure 9 and Figure 10 are the same as Figure 1 and 2, respectively. Figure 9 shows GDP per unit of labor. Again, the series are normalized by their levels in 1956. The figure shows that the growth of GDP per unit of labor accelerated until the end of the 1960s. Figure 10 shows the rate of return to capital input. The series of capital input values are calculated using the series of the growth rate of capital input in the growth accounting and the level of the capital stock in 1956. The level of capital stock in 1956 has been estimated by Hayashi and Prescott (2002). In order to compute the series of quality-adjusted capital input, we assume that its level in 1956 is equal to the level of the capital stock. The inverted-U pattern of the rate of return does not change after controlling for changes in the quality of capital.

Table 6 is the same as Table 1. GDP per unit of labor grew at 6.9 percent annually between 1956 and 1973. Furthermore, the average annual growth rate increased from 5.5 percent in the first decade to 7.9 percent in the following decade. This observation contradicts the prediction of a typical optimum growth model, in which the output growth rate decreases as the economy converges toward the balanced growth path, when the economy starts with a small amount of capital.

The contribution of TFP was higher than that of capital during the growth miracle. From 1956 to 1966, the capital contribution was only 16.5 percent. Most of the Japanese GDP growth was driven by TFP growth during this period. Capital had, however, a higher contribution to GDP growth than TFP between 1963 and 1973. The contribution of capital was approximately 52 percent in the final 10 years. Again, the contribution of capital increased over time so that the pattern did not change even if the length of the sub-sample periods changed.

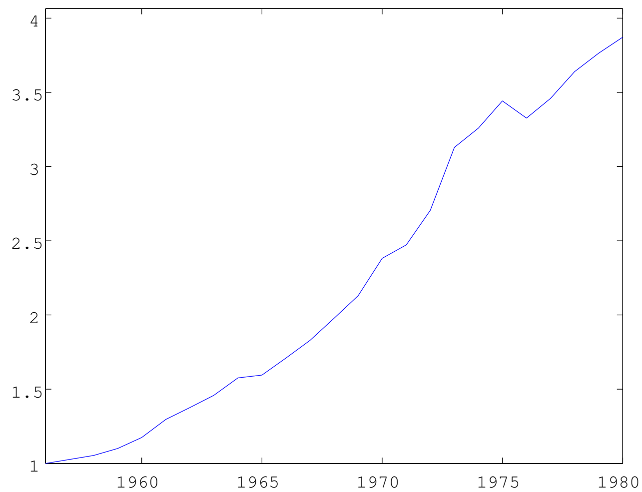

Figure 9: GDP per unit of labor

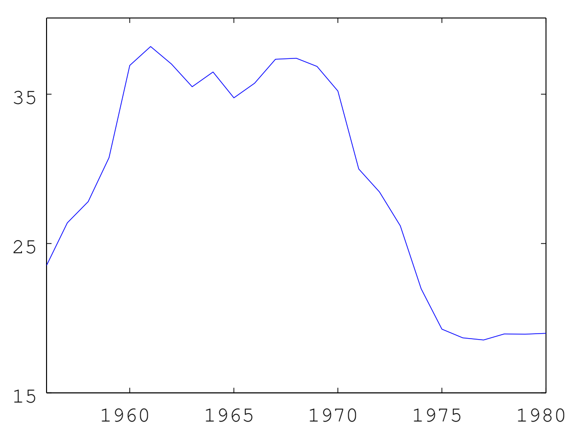

Figure 10: Rate of Return to Capital Input (%)

B Growth Accounting Exercises

In this section, we explain how we construct Table 5. Our growth accounting exercise is a simplified version of that described by Christensen, Cummings and Jorgenson (1980) in the sense that we assume a Cobb-Douglas production function and take capital's income share from National Account data in order to compute TFP growth. We use the method of Christensen et al. (1980) to compute the growth rate of labor input and capital input. The equations we used are the following.

$$
\begin{aligned}
&\ln A_t - \ln A_{t-1} \equiv \ln Y_t - \ln Y_{t-1} - \alpha\left[\ln K_t - \ln K_{t-1}\right] - (1-\alpha)\left[\ln L_t - \ln L_{t-1}\right] \\
&\ln K_t - \ln K_{t-1} \equiv \sum_j w_j^K \left[\ln K_{j,t} - \ln K_{j,t-1}\right] \\
&\ln L_t - \ln L_{t-1} \equiv \sum_i w_i^L \left[\ln L_{i,t} - \ln L_{i,t-1}\right], \\
&w_j^K \equiv (w_{j,t}^K + w_{j,t-1}^K)/2 \\
&w_i^L \equiv (w_{i,t}^L + w_{i,t-1}^L)/2
\end{aligned}
$$

where A, Y, K, and L are TFP, real GDP, real capital input and labor input. α is the capital income share, which is assumed to be constant over time. $w_{j,t}^K$ ($w_{i,t}^L$) is the weight of type j (i) capital (labor). We describe the details of how we construct capital and labor input in the next two subsections.

Capital

We categorize capital into four types: residential dwelling, non-residential dwelling, machinery, and inventory. Real investment data for each category is available in "Annual Report on National Accounts" (*Kokumin Keizai Keisan Nenpo* in Japanese). Since we do not know the real return rate and the depreciation rate in each category, we assume both of them to be the same across all categories. Therefore, the weight will be defined as

$$w_{j,t}^K = \frac{p_{j,t}^K K_{j,t}}{\sum_j p_{j,t}^K K_{j,t}}.$$

$p_{j,t}^K$ is the price of capital j at t,

$$p_{j,t}^K = p_{j,t-1}^I \left[r_t - \frac{p_{j,t}^I - p_{j,t-1}^I}{p_{j,t-1}^I} \right] + \delta_t p_{j,t}^I.$$

$p_{j,t}^I$ is the price of investment in capital j at t which is also available in Annual Report on National Accounts. r is the real rate of return which is the same for all types of capital and we compute the series of rate of return from National Account, that is,

$$r \equiv \frac{\text{GDP-Labor Compensation}}{\text{Capital Stock}}.$$

δ is the depreciation rate which is taken from Hayashi and Prescott (2002).

From the investment data and the depreciation rate, we construct the real capital stock of each category by the perpetual inventory method. We use the average depreciation rate between 1956 and 2000 as the steady state level of the depreciation rate. We set 1956 as the initial period.

The overall trend in our capital contribution is smaller than that in Christensen, Cummings and Jorgenson (1980) although the patterns of the changes in the capital contribution are the same. Two reasons for this difference exist. First, our capital income share, α, is taken from the National Account data rather than from estimation. In order to keep the production function used in the model and in the growth accounting consistent, we assume a Cobb-Douglas function in the growth accounting exercise rather than a general translog production function. Our α is 0.361 while it is 0.415 in their work. Second, we allow the depreciation rate to change over time when we construct the series of capital, while they use a fixed depreciation rate. As we pointed out in Section 4, the depreciation rate decreased over time. As a result, the growth rate of the capital stock in our calculation is lower than in theirs, particularly until the middle of the 1960s.

Labor

We categorize worker by age, sex and education. The subcategories of age are the following: $15-19$, $20-24$, $25-29$, $30-34$, $35-39$, $40-44$, $45-49$, $50-54$, $55-59$, $60-64$, and 65 or over. The subcategories of sex are male and female. The subcategories of education are the following: junior high school (9th grade), high school (12th grade), junior or technical college, and university and post-graduate programs. Labor input in each category $L_{ase,t}$ is defined as

$$L_{ase,t} \equiv H_t N_{ase,t}$$

H is hours worked. Since the data for H in each category are not available, we assume that the changes in H are the same for all categories. A change in labor input is expressed as,

$$\begin{aligned} &\ln L_t - \ln L_{t-1} = [\ln H_t - \ln H_{t-1}] + \sum_i w_i^L [\ln N_{i,t} - \ln N_{i,t-1}] \\ &w_i^L \equiv (w_{i,t}^L + w_{i,t-1}^L)/2 \\ &w_{i,t}^L = \frac{W_{i,t} N_{i,t}}{\sum W_{i,t} N_{i,t}} \end{aligned}$$

The data on the number of workers are available from 1968 in "Basic Survey of Wage Structure" (*Chingin Kouzou Kihon Chosa* in Japanese), and the monthly wage data in each category are available from 1954 in the same source. Since we cannot calculate the series before 1968, we use the indices constructed by Christensen, Cummings and Jorgenson (1980) for the labor input data before 1968. The data of hours worked are available in "Monthly Labor Survey" (*Maitsuki Kinro Tokei Chosa* in Japanese).

C Sensitivity Analysis

This section shows sensitivity analyses in order to understand the baseline results and the model's mechanics more deeply. Since high and low-type workers in the model are unobservable in the data, we examine the effects of the law of motion on the share of high-type workers and the parameter values related to the workers: the initial share of high-type workers, the depreciation rate of the high type, the skill premium and the cost of training low-type workers to become high type.

Specification of The Law of Motion Regulating the Share of high-type workers

One of the key ideas in this paper is the knowledge diffusion from high type to low type. In our model, the law of motion regulating the share of high-type workers reflects this idea by assuming that changes in the composition of labor are limited by the current share of high-type workers. We will show the four other specifications as follow: fixed fraction, flexible but exogenous fraction, fixed fraction with a different cost specification, and more general form of the law of motion.

Fixed Fraction

In this subsection, we change the law of motion of the share of high-type workers from the original one as follows: only a fraction x (which remains fixed over time) out of $s_t(1-\phi_t)$ will become new high-type workers. The transition equation is given by:

$$(1+n_t)\phi_{t+1} = s_t(1-\phi_t)x + (1-\delta_w)\phi_t.$$

This assumption implies that the effects of increases in the share of high-type workers on the training costs are ignored. Therefore, the household increases the share of high-type workers in the household only because labor income increases. Figure 11 shows the results of GDP path with several values of x. The solid line is the baseline GDP. The dashed line is the data. The doted lines are GDP with certain levels of x. To compute the dynamics of the economy, we recalibrate the other parameter values with the same strategy described in Section 4. x affects the share of high-type workers in the steady state as well as the optimal path toward the steady state. As a result, the initial level of the unit cost of training p will change. Therefore, increases in the fraction x leads to increases in the share of high-type workers in the steady state. Furthermore, a higher x implies a higher unit cost of training, p.

Due to this specification, the GDP path becomes more concave than the baseline, regardless of the value of x. The higher is the value of x, the more concave is the GDP path.

Figure 12 shows the rate of return to capital with the different values of x. Again, the dashed line is the one from the baseline model. The peak of the rate of return arrives earlier than the baseline. Furthermore, the higher the value of x is, the higher the peak level of the rate of return.

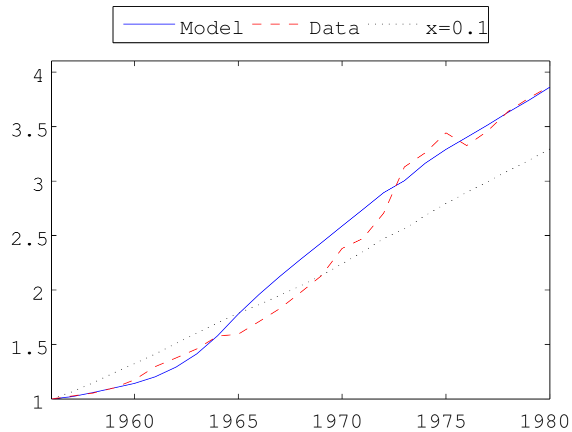

Figure 11: GDP: Fixed Fraction of Being High Type

	Data	$x = 0.7$	$x = 0.1$	Baseline
1956 – 1966	16.5	36.7	34.2	14.8
1963 – 1973	51.8	67.2	55.8	66.8

Table 7: Capital Contribution:1956 – 1966

With the fixed fraction greater than 0.07, the share of high-type workers grows more quickly than the baseline specification at the beginning of the period. Therefore, the family makes the decision earlier to switch some of its resources from the training of low-type workers to investing in physical capital, thus generating a more concave GDP path and an earlier peak timing.

Table 7 shows the contributions of capital in the fist and the second decades for the different values of x. With the fixed fraction, the contribution of capital is much larger than that observed in the data, although a contribution ranking change still occurs.

Exogenous Changes in x

This subsection generalizes the previous exercises. The fraction is still exogenous but it increases over time. That is, the household maximizes its utility given the prices and the sequence $\{x_t\}_{t=0}^{\infty}$. We try the following three sequences.

- Sequence 1: x_t grows linearly from 0.1 to 0.85. It takes 30 years to reach 0.85 and $x_t = 0.85$ for all $t \geq 30$.
- Sequence 2: x_t grows linearly from 0.1 to 0.85. It takes 20 years to reach 0.85 and $x_t = 0.85$ for all $t \geq 20$.

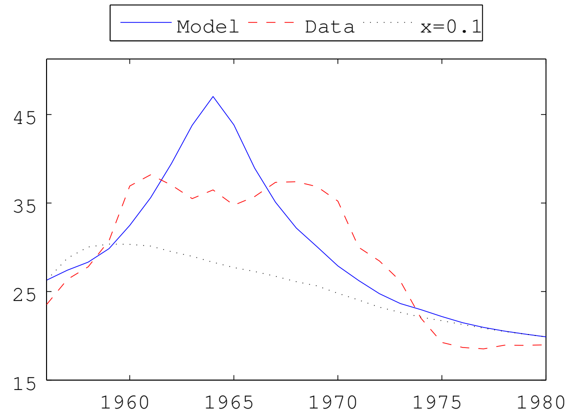

Figure 12: Rate of Return (%): Fixed Fraction of Being High Type

	Data	Seq. 1	Seq. 2	Seq. 3	Baseline
1956 – 1966	16.5	15.2	16.5	26.3	14.8
1963 – 1973	51.8	47.6	56.9	76.7	66.8

Table 8: Capital Contribution (%): Exogenous Changes in Fraction

- Sequence 3: x_t grows linearly from 0.1 to 0.85. It takes 10 years to reach 0.85 and $x_t = 0.85$ for all $t \geq 20$.

Again, the other parameter values are determined by the same strategy in Section 4. Figure 13 shows the GDP path. The GDP path produced with Sequence 1 more closely matches that in the actual data than the one produced with the baseline specification. Furthermore, Table 8 shows that the capital contributions with Sequence 1 are closer to the data than those with the baseline. However, Figure 13 implies that the model with this specification does not explain the rate of return in the 1950s. GDP growth at the very beginning is driven by capital accumulation rather by than productivity growth. As a result, the rate of return decreases at first. Now, this specification generates three stages. First, capital accumulation is faster than TFP growth. Second, TFP drives GDP growth so that the rate of return increases. Third, capital becomes more important than TFP for economic growth (see also Figure 15.)

Fixed Fraction with a Different Cost Scheme

Now, we keep the law of motion as follow.

$$(1 + n_t)\phi_{t+1} = s_t(1 - \phi_t)x + (1 - \delta_w)\phi_t.$$

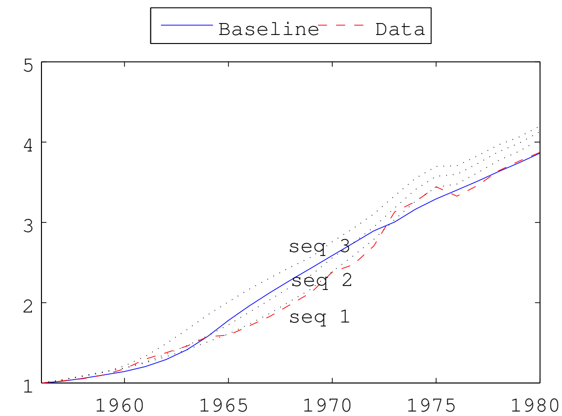

Figure 13: GDP: Exogenous Changes in Fraction

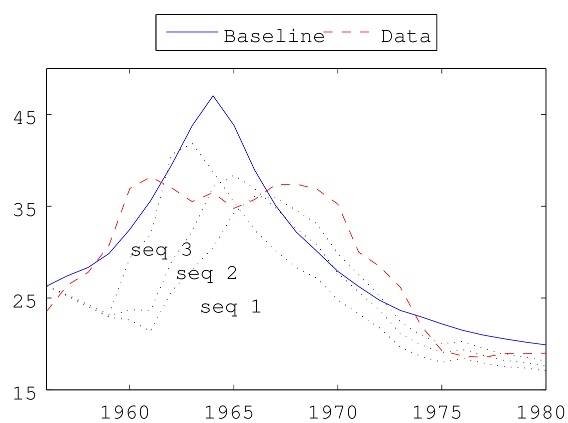

Figure 14: Rate of Return: Exogenous Changes in Fraction

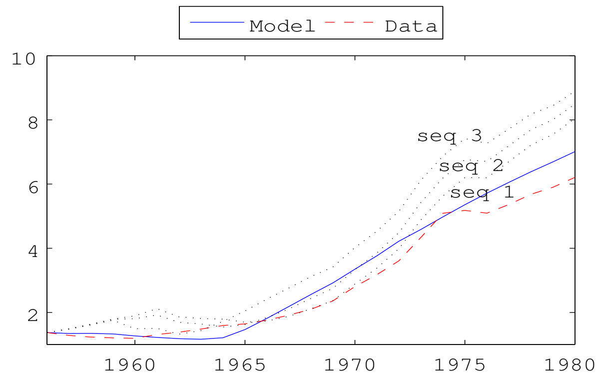

Figure 15: Capital Input: Exogenous Changes in Fraction

At the same time, we change the assumption of cost. Instead of a fixed unit cost of training, we assume that the total cost is

$$[s_t(1-\phi_t)]^\eta \, ,$$

where $\eta > 1$. $\eta > 1$ implies that the marginal cost of turning low-type workers into high-type workers rises as the number of low-type workers being trained increases. Figure 16 and Figure 17 are the results of the GDP path and the rate of return. x is fixed at 1. The value of training expenditure share in the steady state given by $\eta = 1.02$ is the same as that given in the baseline specification. $\eta = 2$ is chosen arbitrarily as a large number. The other parameter values are the same as the baseline.

Although the training cost is a convex function of the number of low-type workers, the share of high-type workers rises quickly. Under the calibrated parameter values, the cost itself is negligible at the beginning of the economy so that ϕ quickly increases to its upper bound. At smaller x, the GDP path becomes closer to the data.

Table 9 shows the contributions of capital in the first decade and the final decade. In either values of η, the contribution of capital is higher than the baseline result. Since the share of high-type workers grows quickly, the timing when capital accumulation occurs sooner than in the baseline specification. Furthermore, all of the GDP growth in the final decade is driven by capital because ϕ reaches the upper bound before 1963.[14]

[14]The contribution of capital is more than 100 percent because the share of high-type workers can decrease due to depreciation and the flexible growth rate in the size of the family.

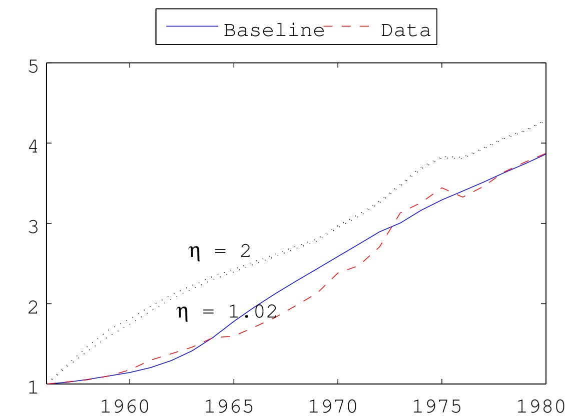

Figure 16: GDP: Exogenous Changes in Fraction

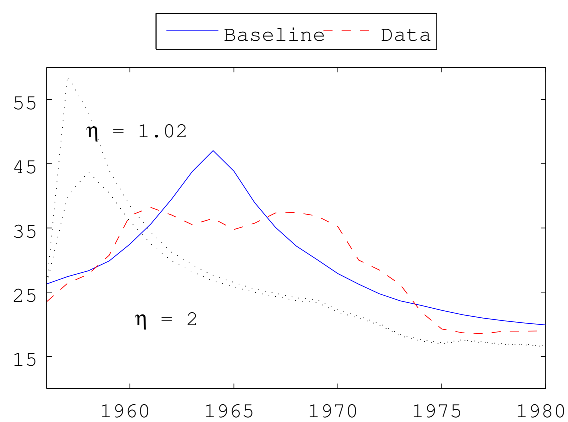

Figure 17: Rate of Return: Exogenous Changes in Fraction

	Data	$\eta = 2$	$\eta = 1.02$	Baseline
1956 – 1966	16.5	38.6	37.8	14.8
1963 – 1973	51.8	72.3	73.2	66.8

Table 9: Capital Contribution: Different Cost Scheme

.

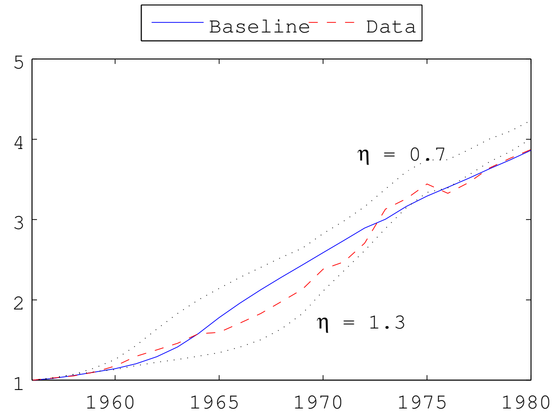

Figure 18: GDP: General Form of Fraction of Being High Type

General Form of the Law of Motion

In this subsection, we make the law of motion of the share of high-type workers slightly more general. Now, the law of motion of the share of high-type workers is

$$(1+n_t)\phi_{t+1} = s_t(1-\phi_t)\phi_t^{\eta} + (1-\delta_w)\phi_t, \quad \eta \geq 0.$$

$\eta = 1$ corresponds to the baseline case while $\eta = 0$ corresponds to the fraction case with $x = 1$. The lower the value of η, the smaller are the effects of the current share of high-type workers. The fraction is a concave (convex) function of the high-type share with $\eta < 1 (\eta > 1)$. We try both cases with $\eta = 0.7$ and $\eta = 1.3$. Again, we recalibrate the other parameter values with changes in η. η affects the share of high-type workers in the steady state, and as a result the unit cost of training, p, changes.

Figure 18 shows the results on GDP. The dashed line is the baseline. If the success probability is a concave function of the share of high-type workers, the transition path of GDP also becomes concave. This means that our model predicts the rapid growth of GDP to occur in the early stage of the miracle with low η. For $\eta > 1$, we can say the opposite. The intuition of these characteristics is as follows: The fraction of workers becoming new high-type workers at the beginning of the economy with $\eta < 1$ ($\eta > 1$) is larger (smaller) than the one in the baseline. In other words, the diffusion effect is larger (smaller) with $\eta < 1$ ($\eta > 1$) than the baseline economy. As a result, the expenditure for training is larger (smaller), and the economy grows more quickly (slowly) than the baseline case.

Figure 19 shows the results of the rate of return to capital. η has two effects: the timing of the peak, on one hand, and the height of the peak, on the other. The higher the value of η, the later the arrival of the peak and the lower is the peak's height. The contributions of capital are shown

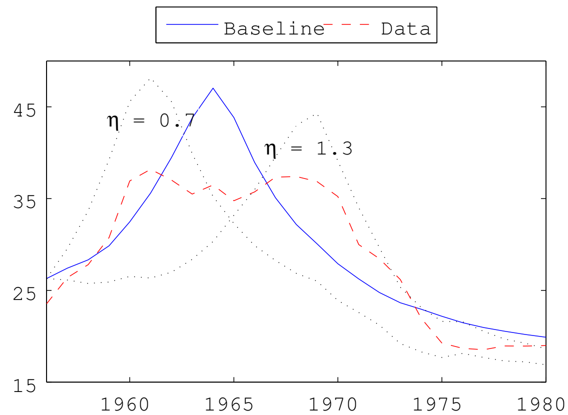

Figure 19: Rate of Return (%): General Form of Fraction of Being High Type

	Data	$\eta = 1.3$	$\eta = 1$	$\eta = 0.7$
1956 – 1966	16.5	3.2	14.8	30.6
1963 – 1973	51.8	41.6	66.8	80.1

Table 10: Capital Contribution: General Form of Fraction of Being High Type

.

in Table 10. $\eta = 1$ corresponds to the baseline case. With low η, we can still observe the change in relative contributions. On the other hand, with high η, TFP is the leader in both periods. This is because more output is allocated to train low-type workers but the speed of increases in the share of high-type workers is slow at the beginning.

Initial Share of high-type workers

In the baseline calibration strategy, we assume the initial share of high-type workers is the same as the share of college degree holders among total workers because it is difficult to directly observe our model's "high-type workers" in the data. This subsection examines the sensitivity of our model to the initial share of high-type workers, ϕ_0, by computing the equilibrium paths with $\phi_0 = 0.1$ and $\phi_0 = 0.05$. Changes in ϕ_0 affect the levels of productivity because we choose the levels of A_H and A_L to match the initial TFP level given the observed skill premium and the initial share of high-type workers. The lower the value of ϕ_0, the higher the value of A_H. As a result, the steady state level of output changes such that the unit cost of training p also changes. Changes in the initial share of high-type workers directly affect the optimal decision-making as well. This is because the lower the value of ϕ_0, the more difficult it is to be a high-type worker in the early

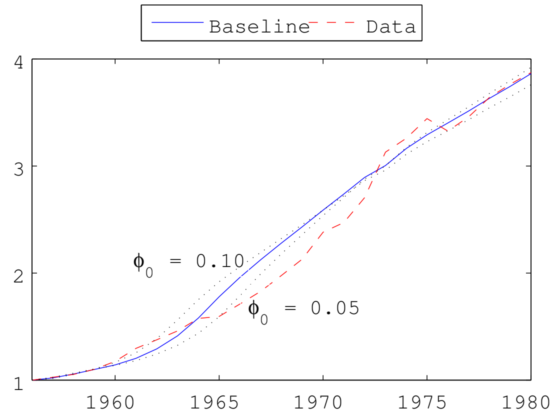

Figure 20: GDP: Initial Share of High-Type Workers

	Data	$\phi_0 = 0.10$	$\phi_0 = 0.07^*$	$\phi_0 = 0.05$
1956 – 1966	16.5	23.6	14.8	1.8
1963 – 1973	51.8	76.9	66.8	56.3

Table 11: Capital Contribution: Initial Share of High-Type Worker

.

stage of the period.

Figure 20 shows the results on GDP with different ϕ_0 (doted lines). The solid line is the baseline, and the dashed line is the data. GDP grows more quickly with $\phi_0 = 0.1$ than with $\phi_0 = 0.05$. However, the higher initial share implies a lower steady state level of output due to lower level of high-type productivity A_H. Therefore, GDP in 1973 with $\phi_0 = 0.1$ is lower than that with $\phi_0 = 0.07$ (baseline value). On the other hand, when the economy starts with the low share of high-type workers, GDP growth increases more slowly. But in 1973, GDP is nearly the same as the baseline.

If the economy starts with a lower share of high-type workers, the speed of TFP growth is slower. As a result, the peak of the rate of return to capital arrives later than in the baseline economy. (See Figure 21.) Changes in the initial share of high-type workers change the levels of productivity such that the height of the peak in the rate of return changes.

Table 11[15] shows the contribution of capital with different initial values of ϕ_0. The same pattern of changes in the factor contributions as the data is observed in both cases.

[15] $\phi_0 = 0.07$ is the baseline.

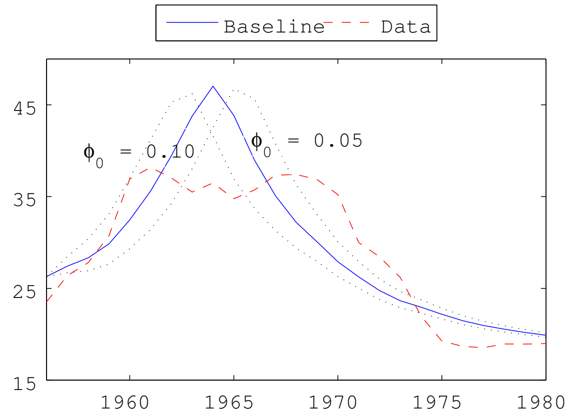

Figure 21: Rate of Return (%): Initial Share of High-Type Workers

Depreciation Rate of High-Type Workers

The following sensitivity analysis concerns the effects of changes in δ_w (the depreciation rate of high-type workers) on the economy. We assume the depreciation rate of the share of high-type workers is the same as that of human capital in a life-cycle model. This is simply because we cannot observe our model's "high-type workers" in the data. Thus, this subsection explores the effects of changes in δ_w by computing the path of GDP with $\delta_w = 0.08$ and $\delta_w = 0.012$. The former number is taken from Sorkey (1996) and the latter is taken from Mincer (1974). Changes in the depreciation rate directly affect the steady state level of the share of high-type workers. As a result, the unit cost of training, p, should also change. On one hand, decreases in δ_w help the household to increase the share of high-type workers. On the other hand, however, decreases in δ_w increases p such that the training becomes more costly.

Figure 22 and 23 show the results of GDP and the rate of return. Under our parameter values, the impact of the latter effect dominates the that of the former. That is, the higher the depreciation of the share of high-type workers, the more quickly the share of high-type workers increases. As a result, GDP grows more quickly with $\delta_w = 0.08$ than with $\delta_w = 0.012$.

Table 12 shows the contribution of capital. The column labeled $\delta_w = 0.03$ corresponds to the baseline results. With $\delta_w = 0.012$, the high unit cost of training slows growth of GDP, capital, and aggregate productivity. Although more output is allocated to turn low-type workers into high-type ones (as we observe with the baseline parameter values), the share of high-type workers does not increase as quickly as in the baseline case. As a result, the beginning of capital growth is delayed. TFP always drives GDP growth during the miracle period.

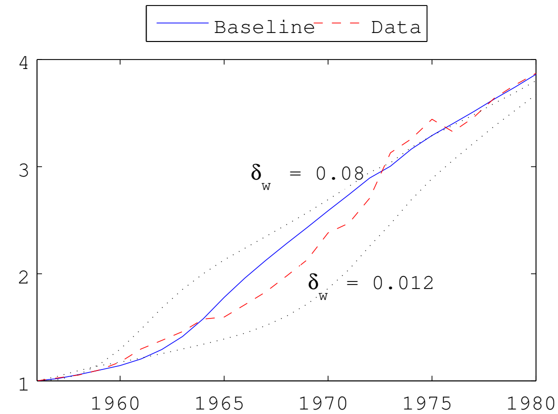

Figure 22: GDP: Depreciation Rate of High Type

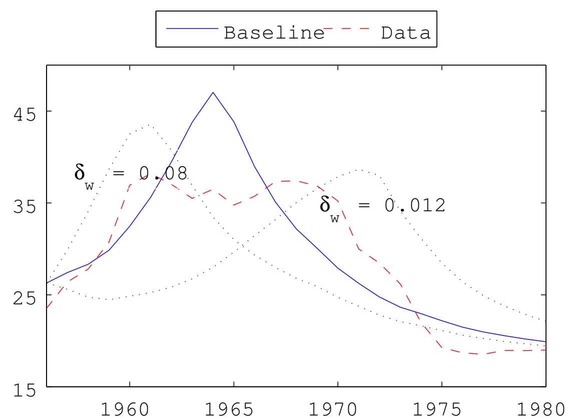

Figure 23: Rate of Return (%): Depreciation Rate of High Type

	Data	$\delta_w = 0.08$	$\delta_w = 0.03$	$\delta_w = 0.01$
1956 – 1966	16.5	31.3	14.8	18.3
1963 – 1973	51.8	74.5	66.8	22.3

Table 12: Capital Contribution: Depreciation Rate of High Type

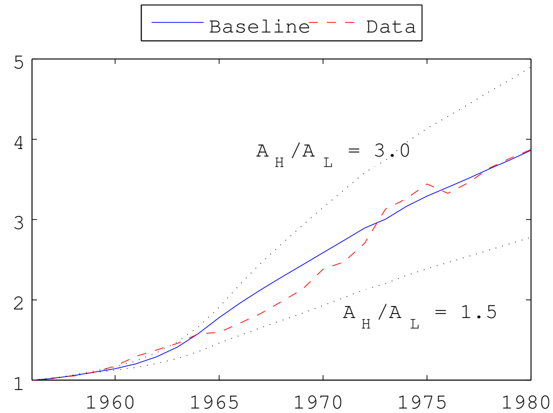

Figure 24: GDP: Skill Premium

	Data	$\frac{A_H}{A_L} = 3.0$	$\frac{A_H}{A_L} = 2.2^*$	$\frac{A_H}{A_L} = 1.5$
$1956 - 1966$	16.5	7.6	14.8	19.0
$1963 - 1973$	51.8	61.6	66.8	62.5

Table 13: Capital Contribution: Skill Premium

.

Skill Premium

A $90 - 10$ ratio of monthly wage is used to infer the skill premium in this paper. In Japan, this number is relatively low and very stable compared to that in other countries, especially developing countries. This subsection examines the effect of changes in skill premium on our model. Changes in the skill premium change the level of GDP in the steady state. A high skill premium implies a high productivity level of high-type workers so that the steady state level of GDP also becomes high.

Figure 24 and Figure 25 show GDP and the rate of return with $\frac{A_H}{A_L} = 3.0$, $\frac{A_H}{A_L} = 1.5$ and $\frac{A_H}{A_L} = 2.2$ (baseline, dashed line). The growth rate of GDP is higher with $\frac{A_H}{A_L} = 3$ than that with the baseline value due to the higher productivity level of high-type worker. However, the model does not change qualitatively. That is, GDP growth accelerates over time. The dynamic patter of the rate of return has the shape of an inverted U. Furthermore, the contribution of capital during the first decade is very small (range from -2% to 4%) while it is approximately 70% in the following decade. (See Table 12.)

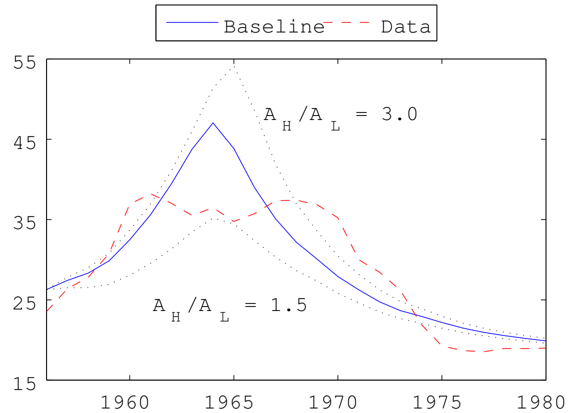

Figure 25: Rate of Return (%): Skill Premium

Cost of Training

In the baseline calibration, the training cost within a firm is assumed to be the same as the off-job training cost.[16] However, in practice it is difficult to observe how much training costs in the overall economy. Thus, this section analyzes the effects of the steady state level of training expenditure share (TES). In our model, it has a one-to-one relationship with the unit cost of training. The higher the steady state TES, the higher the unit training cost. For the analysis, 3.5% and 0.07% are used. The former is slightly higher than the average government's education expenditure share of GDP in 2005(3.5%). The latter is almost the same level as the average off-job training cost in a firm from the data in the "Basic Survey of Skill Development".

GDP growth with $TES = 0.035$ is smaller than that with the baseline setting. Since training workers is costly, the more output is allocated to accumulating capital in the beginning of the period. As a result, GDP growth is driven by capital accumulation in the first decade. (See Table 14.) Because of the high cost of training and the high depreciation rate of physical capital, the household cannot invest in physical capital sufficiently to support a growth miracle (see Figure 26.) Furthermore, the rate of return to capital monotonically decreases along with capital accumulation (Figure 27.) In the following decade, however, the household begins training workers such that GDP growth is driven by TFP and the rate of return gradually increases. On the other hand, with low expenditure share, although the GDP path becomes more concave, the level of GDP in 1973 predicted by the model is close to that in the data. Furthermore, we still observe an inverted U-shape in the rate of return and a pattern of changes in the factor contribution with lower expenditure share.

[16] "Off-job training" here refers to training provided by firms other than the trainees' employer.

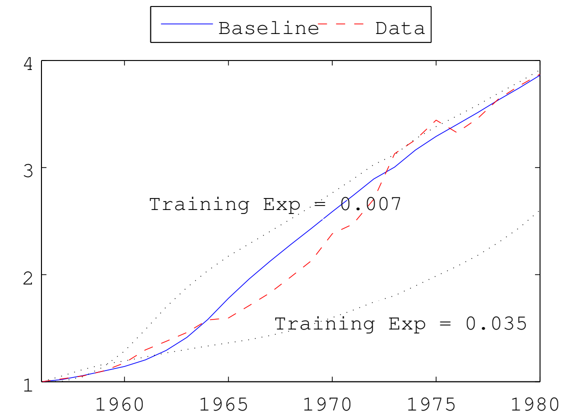

Figure 26: GDP: Training Expenditure Share

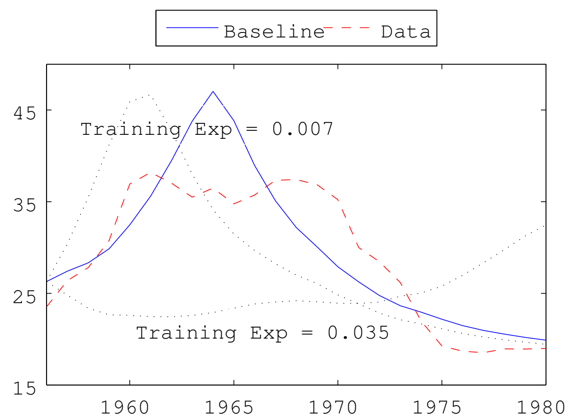

Figure 27: Rate of Return (%): Training Expenditure Share

	Data	$TSS = 3.5\%$	$TSS = 1.34\%$	$TSS = 0.7\%$
1956 – 1966	16.5	47.1	14.8	31.0
1963 – 1973	51.8	26.2	66.8	75.5

Table 14: Capital Contribution: Training Expenditure Share

.