数理解析研究所講究録 1337 巻 2003 年 129-165



# A Model of Growth with the Size Distribution of Firms and Economies of Scale

Hideyuki Adachi
Department of Economics, Kobe University, Kobe 657, Japan

November 27, 2002

**Abstract**

The size distribution of firms in each industry will usually be highly skew, and empirical evidence shows that it is approximated closely by the Pareto distribution. In this paper we make an attempt to explain why the Pareto law applies to the size distribution of firms based on their innovation and investment behavior, and then develop a model of economic growth that take into account this empirical law. First, we show that the Pareto distribution of firms is generated under the assumption that firms acquire the technology of operating efficiently on a larger scale through learning by doing, and expand their scale of operation through the accumulation of capital induced by profitability. Then, we construct a model of economic growth that is based on the Pareto distribution of firms and economies of scale. In our model the growth rate is determined endogenously, and it exhibits scale effects with respect to savings. Our model is different from the neoclassical growth model or the recently developed endogenous growth models in that it takes into account the size distribution of firms, and it yields quite realistic predictions.

## 1 Introduction

Empirical laws are rare in economics, and an example of them is given by the regular pattern of some statistical distributions, such as the distribution of persons according to the level of income or of business firms according to some measurement of size such as sales or the number of workers. Many of those distributions conform to the so-called the law of Pareto. Many economists attempted to explain the mechanisms that generate the Pareto distributions by constructing models with stochastic processes. Simon (1955), Champernowne (1953), Wold and Whittle (1957), Steindl (1965), etc., may be mentioned as pioneers of such models. The most ingenious model among them is the one developed by Simon (1955), which explains the Pareto distribution based on two simple and meaningful assumptions: one is 'the law of proportional effect', and the other is the constancy of new entry. When his model is applied to the size distribution of firms, however, it is not clear how those assumptions are related to the behavior of firms. Besides, there is no work, as far as I know, that makes use of this interesting empirical evidence on the size distribution of firms to analyze macroeconomic problems such as economic growth or income distribution

The first purpose of this paper is to explain why the size distribution of firms is approximated by the Pareto distribution, based on the innovation and investment behavior of firms. In our model we assume that new firms start their operation from the minimum size, because they lack not only the necessary know-how to operate efficiently at larger size, but also sufficient finance to start on a larger scale. They gradually acquire the technology of operating efficiently on a larger scale through learning by doing, and expand their scale of operation through accumulation of capital induced by profitability. We show that the size distribution of firms tends to be the Pareto distribution under such assumptions. To my knowledge, there is no attempt, until now, that relates the size distribution of firms to the learning by doing hypothesis. The learning by doing hypothesis was introduced into the growth theory by Allow (1962). Then, Romer (1986) used Arrow's set-up to develop an endogenous growth model with increasing returns. The learning by doing hypothesis is regarded as one of the analytical devices for explaining knowledge accumulation in endogenous growth models. However, there is scarcely any work, so far, that has applied this hypothesis to more microeconomically based models. In this paper, we will show that the learning doing hypothesis is quite appropriate for explaining how the size distribution of firms tends to be the Pareto distribution. Our model may be characterized as a kind of an evolutionary model instead of an equilibrium model, because it is based on stochastic processes.

Our second purpose is to present a model of growth and income distribution that is based on the assumption that the size distribution of firms is approximated by the Pareto distribution. That the size distribution of firms conforms to the Pareto law is empirically well established. Simon and Bonini (1958) and Ijiri and Simon (1964) (1971) (1974) apply this law to study industrial structures, and get many interesting results. However, this interesting empirical law is scarcely utilized in macroeconomic analyses. In this paper, we will construct a growth model that takes into account the size distribution of firms, and show that our growth model can explain some structural aspects of economic growth that previous models cannot explain. It should be noted that the size distribution of firms in our model is generated by learning by doing and investment behavior of individual firms under stochastic processes. Thus, our growth model assume a kind of evolutionary process of growth for individual firms. It will be shown that the growth rate of aggregate output is determined endogenously, and it exhibits scale effects with respect to the savings rate. In this respect, the properties of our model are different from the model of Solow (1955). It is rather similar to endogenous growth models by Romer (1986) (1990) or Aghion and Howitt (1992). However, it is pointed out by Solow (2000) that the scale effects in those endogenous growth models are due to some arbitrary assumptions about R&D processes. In our model, the scale effects will be shown to follow as a natural result of learning by doing and investment behavior of individual firms in the process of growth.

The paper is organized as follows. Section 2 reviews the Simon's model of explaining the Pareto law and its generalization by Sato (1970). Section 3 shows that the Pareto law for the size distribution of firms can be explained from learning by doing and investment behavior of individual firms. In Section 4, we construct a macroeconomic model based on the Pareto law and the learning by doing hypothesis, and analyze the determination of income distribution. In Section 5, we extend it to a growth model, and analyzes the steady-state properties of this model. It is shown that the steady growth equilibrium exhibits scale effects, and that it is unstable. In Section 6, we show that the steady growth equilibrium becomes stable if the substitutability between capital and labor is taken into account, and then examine the effects of changes in structural parameters. Conclusions are given is Section 7.

# 2 The Size Distribution of Firms and the Pareto Law

In the Appendix of his book Steindl (1965) gives many empirical data which illustrate the regularity and comparative stability of patterns found in the size distribution of firms in U.S. and Germany.[1] Those patterns are approximated by the Pareto distribution which is given by

$$N(x) = Ax^{-\rho}, \tag{1}$$

where $x$ represents the size of firms, $N(x)$ the number of firms with the size in excess of $x$, and $\rho$ is a parameter called the Pareto coefficient. The size of firms may be measured by sales, capital or employment depending on the availability of data. The above equation implies that the number of firms with the size in excess of $x$, plotted against $x$ on logarithmic paper, is a straight line. Figure 1 shows the size distribution of the Japanese manufacturing industry in 1998 when the size of firms is measured by the number of employees. It is almost entirely a straight line, illustrating the Pareto law quite well.[2]

The Pareto distribution is observed not only in the size distribution of firms but in many other fields in economics, such as distributions of persons according to income, distribution of cities according to inhabitants and so on. It is also found in non-economic phenomena, such as the distribution of scientists according to the number of papers published, the distribution words in a book and so on.[3] Why such a regular pattern is observed in many fields is a big puzzle. Attempts to reveal this puzzle have been made by many economists, including Champernowne (1953), Simon (1955), Wold and Whittle (1957), Mandelbrot (1961), and Steindl (1965). Among them, the solution given by Simon (1955) seems to me the simplest and the most ingenious. In this paper we use the

Simon's model to extend it to a growth model that incorporates the learning by doing hypothesis and investment behavior of firms.

The Simon's model was originally designed for a non-economic problem, namely the distribution of words in a book. We interpret it here as the model that explains the size distribution of firms. At any given moment, an economy (or an industry) consists of a large number of firms with given size distribution. We assume in the following that the size of each firm is measured by its productive capacity, namely its output at the normal utilization of capital, denoting it by $x$. For analytical convenience we express the size of firms by discrete numbers, putting the minimum size to be unity. Each firm expands its capacity over time, some firms growing more rapidly than others. The process of growth of firms is assumed to be stochastic. Summing up the capacity of all firms, we obtain the total capacity of the economy, which is denoted by $X$. Let us designate by $f(x,X)$ the number of firms whose capacity is $x$. Then, we must have

$$\sum_{x=1}^{X} xf(x,X) = X. \tag{2}$$

Simon makes two basic assumptions to prove the Pareto law. In the context of the size distribution of firms, those assumptions may be stated as follows.

**Assumption 1 (the law of proportionate effect):** *The probability that a unit increase in the total capacity of the economy from $X$ to $X+1$ is attributed to firms with the size class $x$ is proportional to $xf(x,X)$, i.e., the total capacity of that size class.*

**Assumption 2 (a constant birth rate for new firms):** *There is a constant probability, $\alpha$, that a unit increase in the total capacity from $X$ to $X+1$ is contributed by new firms which enter from the minimum size.*

The first assumption is called 'the law of proportionate effect', which was originally proposed by Gibrat (1930) to derive the log-normal distribution. It implies that the expected growth rate of a firm is independent of its size. The second assumption means that new firms are being born in the minimum-size class at a relatively constant rate. This assumption of a constant birth rate for new firms plays an crucial role in distinguishing the Pareto distribution from the log-normal.

Under the first assumption, the law of proportionate effects, the expected number of firms with size class $x$ when the total capacity of the economy is expanded to $X+1$ is determined by

$$E[f(x,X+1)] = f(x,X)+L(X)\{(x-1)f(x-1,X)-xf(x,X)\}, \quad x = 2,3,\cdots,X+1, \tag{3}$$

where $L(X)$ is the proportionality factor of the probabilities. The second assumption, a constant birth rate for new firms, together with the first one gives the following equation:

$$E[f(1,X)] = f(1,X) - L(X)f(1,X) + \alpha. \tag{4}$$

The proportionality factor $L(k)$ must satisfy

$$L(X)\sum_{x=1}^{X} xf(x,X) = 1-\alpha. \tag{5}$$

Substituting (2) into (5), we have

$$L(X) = \frac{1-\alpha}{X}. \tag{6}$$

Simon is concerned with steady-state distributions in which the expected values in (3) and (4) coincide with the actual frequencies. We thus have

$$f(x,X+1) = f(x,X)+L(X)\{(x-1)f(x-1,X)-xf(x,X)\} \quad x=2,3,\cdots,X+1, \tag{7}$$

$$f(1,X+1) = f(1,X) - L(x)f(1,X)+\alpha, \tag{8}$$

where the $f's$ may now be interpreted as either expected values or actual frequencies. From the definition of the steady-state distribution, we have

$$\frac{f(x,X+1)}{f(x,X)} = \frac{X+1}{X} \quad \mathit{for\ all\ x\ and\ X}. \tag{9}$$

This means that all the frequencies grow proportionately with $X$, and maintains the same relative size. The relative frequencies denoted by $f^*(x)$ may be defined as

$$f^*(x) = \frac{f(x,X)}{\alpha X}, \tag{10}$$

where $\alpha X$ is equal to the total number of firms, $N$, from the definition of $\alpha$. Thus, $f^*(x)$ represents the proportion of the number of firms with size $x$ in the total.

Using (10) and (6), we rewrite (7) and (8) in terms of the relative frequencies as follows:

$$f^*(x) = f^*(x-1)\frac{(1-\alpha)(x-1)}{1+(1-\alpha)x}, \tag{11}$$

$$f^*(1) = \frac{1}{2-\alpha}. \tag{12}$$

Putting $\rho \equiv 1/(1-\alpha)$, we obtain from (11) the following solution:

$$f^*(x) = \frac{(x-1)(x-2)\cdots 2\cdot 1}{(x+\rho)(x+\rho-1)\cdots(2+\rho)}f^*(1) = \frac{\Gamma(x)\Gamma(\rho+2)}{\Gamma(x+\rho+1)}f^*(1), \tag{13}$$

where $\Gamma$ denotes the Gamma function. This is in fact the solution for the original equation (7), since it is confirmed to satisfy (7) by direct substitution. Simon called the expression (13) the *Yule distribution* after the name of a biologist who derived this function to explain the distribution of biological genera by numbers of species.

There is a well-known asymptotic property of the Gamma function that as $x \to \infty$,

$$\frac{\Gamma(x)}{\Gamma(x+k)} \sim x^{-k} \tag{14}$$

for any constant $k$. It follows from (13), therefore, that as $x \to \infty$,

$$f^*(x) \sim Ax^{-(\rho+1)}, \tag{15}$$

where $A \equiv \Gamma(\rho+2)f^*(1)$ is a constant, and $\rho$, the Pareto coefficient, must be greater than unity as long as $\alpha > 0$. This result shows that the size distribution function is approximated by the Pareto distribution for sufficiently large values of $x$, i.e., above a certain minimum size of firms. We can confirm that the expression (13) is a proper distribution function. For we have

$$\sum_{x=1}^{\infty} xf^*(x) \sim A\sum_{x=1}^{\infty} x^{-\rho}, \tag{16}$$

and this expression is convergent if $\rho > 1$. But this condition is satisfied as long as the birth rate of new firms, $\alpha$, is positive.

The result obtained above may be summarized by the following proposition:

**Proposition 1 (Simon):** *Under assumption 1 (the law of proportionate effect) and assumption 2 (a constant birth rate of new firms for new firms) stated above, the size distribution of firms is asymptotic to the Pareto distribution, namely, almost identical with the Pareto distribution for firms above a certain minimum size.*

Simon's model was extended by Sato (1970) to include the case where the law of proportionate effect in the usual sense does not apply in an exact sense. He replaces Assumption 1 by the following assumption:

**Assumption 1′ (the generalized law of proportionate effects):** *The probability that a unit increase in the total capacity of the economy from $X$ to $X+1$ is attributed to firms with size class $x$ is proportional to $(ax+b)f(x,X)$, where $a$ and $b$ are constant and subject to the following restrictions:*

$$\sum_{x=1}^{X}(ax+b)f(x,X) = \sum_{x=1}^{X} xf(x,X) = X. \tag{17}$$

This assumption generalizes the law of proportionate effect by including it as a special case where $a = 1$ and $b = 0$. Otherwise, the expected rate of a firm is not independent of its size, but increases or decreases with it. Sato (1970) shows, with this assumption together with Assumption 2, that the steady-state distribution becomes as

$$f^*(x) = \frac{\Gamma\left(x+\frac{b}{a}\right)\Gamma\left(\frac{\rho+b}{a}+2\right)}{\Gamma\left(x+\frac{\rho+b}{a}+1\right)}f^*(1), \tag{18}$$

where $a + b > 0$ is required for this value to be finite. From the asymptotic property of the Gamma function, we have, as $x \to \infty$,

$$f^*(x) \sim B\left(x + \frac{b}{a}\right)^{-\frac{\rho}{a}-1}, \tag{19}$$

where $B$ is a constant. This is called the Pareto distribution of the second kind, which is one of three forms of size distributions originally proposed by Pareto. The size distribution of the form (15) is called the Pareto distribution of the first kind.

The distribution function (19), when plotted on logarithmic paper, is not exactly a straight line. Notice, however, that this distribution function is rewritten as

$$B\left(x + \frac{b}{a}\right)^{-\frac{\rho}{a}-1} = Bx^{-\frac{\rho}{a}-1}\left(1 + \frac{b}{ax}\right)^{-\frac{\rho}{a}-1}, \tag{20}$$

and that for any given value of $a$ and $b$, we have $1 + (b/ax) \to 1$ as $x \to \infty$. It follows, therefore, that as $x \to \infty$,

$$f^*(x) \sim Bx^{-\frac{\rho}{a}-1}. \tag{21}$$

Thus, the Pareto distribution of the second kind is asymptotic to that of the first kind with the Pareto coefficient of $\rho/a$.

Substituting (10) into (17) and rewriting it, we have, as $X \to \infty$,

$$(1-a)\sum_{x=1}^{\infty} xf^*(x) = b\sum_{x=1}^{\infty} f^*(x), \tag{22}$$

so that

$$\frac{1-a}{b} = \frac{\sum_{x=1}^{\infty} f^*(x)}{\sum_{x=1}^{\infty} xf^*(x)} = \frac{N}{X}. \tag{23}$$

But, since $N/X = \alpha$, we obtain

$$\frac{1-a}{b} = \alpha, \tag{24}$$

which is the relation $a$ and $b$ are subject to. Considering that $\alpha > 0$, we must have

$$a \gtreqless 1 \textit{ according as } b \lesseqgtr 0. \tag{25}$$

Assumption 1′ implies that the expected growth rate of a firm with size $x$ is proportional to $a + (b/x)$. Therefore, the expected growth rate of a firm is independent of its size, if and only if $a = 1$ and $b = 0$. It increases with its size if $a > 1$ and $b < 0$, and it decreases with its size if $a < 1$ and $b > 0$. The Pareto coefficient $\rho/a$ is different from $\rho = 1/(1-\alpha)$ except for the case $a = 1$.These results may be summarized by the following proposition:

**Proposition 2 (Sato):** *Under Assumption 1′ and Assumption 2 stated above, the size distribution of firms is asymptotic to the Pareto distribution.*

*Depending on the values of a and b, we can distinguish the following three cases:*

(a) *If* $a = 1$ *and* $b = 0$, *the expected growth rate of a firm is independent of size. The Pareto coefficient is equal to* $\rho = 1/(1-\alpha)$ *as Simon demonstrated.*

(b) *If* $a > 1$ *and* $b < 0$, *the expected growth rate of a firm proportionately increases with size. The Pareto coefficient is less than* $\rho$.

(c) *If* $a < 1$ *and* $b > 0$, *the expected growth rate of a firm proportionately declines with size. The Pareto coefficient exceeds* $\rho$.

In order to see how these different cases occur, we have to relate the model to the innovation and investment behavior of firms.

# 3 Learning by Doing and the Pareto Law

In the last section we have reviewed Simon's model and its extension by Sato to see how the size distribution of firms becomes the Pareto distribution. The essential element common to the models is stochastic processes. In particular, the main assumptions governing such stochastic processes are the law of proportionate effect (or its variation) and a constant birth rate for new firms. Though these assumptions are simple and interesting, there are scarcely any discussion so far made to justify them by more basic hypotheses on the behavior of firms or the cost conditions of firms. In this section we attempt to give a microeconomic foundations, so to say, for the above model.

The standard theory of the firm we learn in the ordinary textbook is the neoclassical theory of the firm. In the neoclassical theory of the firm, it is assumed that the U-shaped curve, *LAC*, illustrated in Figure 2 is the long-run average cost curve of all firms in a particular industry, freely available to all including new entrants. It is not by empirical observation but by the assumption of perfect competition that the theory requires the long-run average cost to be U-shaped. If it is U-shaped, the size distribution of firms is expected to become a normal distribution around the optimum size at which the long-run average cost is minimum. But empirical data for many countries including Japan show that the size distributions of firms are highly skew, approximated closely by the Pareto distribution. This implies that the neoclassical theory of the firm is inconsistent with empirical observations on the size distribution of firms. The crucial assumptions leading to such unrealistic results are perfect knowledge of technology and perfectly competitive supply of finance.[4] We make changes in these two assumptions for the sake of realism, and develop a completely different theory of the firm. In our model developed below, we assume that knowledge of technology is imperfect and that outside finance is limited by the amount of retained profits.

Let us first explain the assumption that technological knowledge is imperfect. This assumption seems to be quite plausible. For much of technology—commercial, financial, organizational, and physical—at any moment of time is essentially private in the sense that it is acquired by efforts made by each firm and technology thus accumulated is different between firms. A firm has to master the technology of operating efficiently on a larger scale by trial and error, in the course of which a firm may acquire the additional technology. It is a process of learning by doing. Private technology gradually becomes more or less public through being copied by competitors. But all technological improvements are initially private, since they start by being the results of efforts made by a single enterprise. They do not fall like manna from heaven as the neoclassical theory assume.

To describe the process of knowledge accumulation, we will adopt Arrow's formulation of learning by doing. Arrow (1962) presented a model of economic growth based on the learning by doing hypothesis. In his model productivity increases through learning process. Learning is regarded as the product of experience, and as an index of experience Arrow chooses cumulative investment in the economy as a whole. Our model adopts his basic idea, but differs from his model in two respects. First, in Arrow's model, the learning is embodied in the new capital goods, while, in our model, it takes place in each firm with different size. Arrow's model is a vintage model in which technical change is completely embodied in new capital goods. At any moment of time, the new capital goods incorporate all knowledge then available, but once built their productive efficiency cannot be altered by subsequent learning. Contrastively, in our model, firms with different size and different efficiency coexist at any moment of time, the great majority of firms being still of small or medium size. Most firms have a tendency to expand their scales over time by mastering technology through learning by doing. Profitable firms tend to grow at higher rates than unprofitable firms, the proccess of which is stochastic.

Another difference between our model and Arrow's concerns an index of experience. For convenience of our analysis, we use capacity output of firms as an index of experience instead of cumulative investment as Arrow does. Many empirical studies have shown that cumulative output is a good index of experience. But, under the assumption that output is increasing exponentially, current output is proportional to cumulative output. Since current output of a firm is equal to its capacity output in our model, the latter may be an appropriate index of experience.

The learning curves may be different firm by firm, since some firms are more efficient in learning than others. Therefore, different firms will usually follow different paths in learning by doing. We assume, however, that a learning function of a typical firm with productive capacity $x$ is expressed as follows:[5]

$$\frac{l(x)}{x} = \gamma(x), \quad \gamma'(x) < 0, \quad x \in [1, X], \tag{26}$$

$$\frac{k(x)}{x} = \delta(x), \quad \delta'(x) \leq 0 \quad x \in [1, X]. \tag{27}$$

The notations are as follows: $l(x)$ is the amount of labor used in the production by a typical firm with size $x$, $k(x)$ is the amount of capital used in the production by a typical firm with size $x$, $\gamma(x)$ is a decreasing function, and $\delta(x)$ is a non-increasing function. A pair of equations (26) and (27) means that as a typical firm with size $x$ expands the scale, the amount of labor used per unit of output decreases while the amount of capital used per unit of output stays constant or decreases. Therefere, regardless of wages and rental value of capital, it always pay for the firm to expand its scale.

To simplify the analysis without loss of reality, we will specify the above learning function as follows:

$$\gamma(x) = cx^{\lambda-1}, \quad 0 < \lambda < 1, \; x \in [1, X], \tag{28}$$

$$\delta(x) = dx^{\mu-1}, \quad 0 < \mu \leq 1, \; x \in [1, X]. \tag{29}$$

Then, inputs of labor and capital as functions of capacity output becomes as follows:

$$l(x) = cx^{\lambda}, \quad 0 < \lambda < 1, \; x \in [1, X], \tag{30}$$

$$k(x) = dx^{\mu}, \quad 0 < \mu < 1 \; x \in [1, X]. \tag{31}$$

It is verified that these relations fit quite well to the data of Japanese manufacturing industries.[6]

Under the learning function expressed by a pair of equations (28) and (29), there will be an inherent tendency for firms to grow in size over time, because entrepreneurs expect to enjoy economies of scale by expansion. However, it should be noted that the above learning function is not a given schedule perfectly known to firms. Since knowledge of technology is imperfect, the additional technology for operating efficiently on a larger scale can be acquired only through a process of learning by doing. Imperfect knowledge of technology acts as a brake on expansion of firms.

Another brake is that such a firm must raise the necessary finance for expansion. The sources of finance for investment of firms consist of retained profits, issues of equity or bond, and borrowing from banks. The retained profits of firms are bounded by their expected profits. If the amount of investment a firm desires to carry out exceeds retained profits, it has to raise outside finance in the form of equity, bond or borrowing from banks. Normally, there will be limits to the amounts that can be raised in this way, or penalty rates to be paid beyond certain remits. The total amount of extra finance raised externally will tend to be limited by the amount of retained profits. Most firms, especially those of small or medium size, the extent of future expansion depends on the volume of previously accumulated profits. Therefore, the expected rate of profit is a key variable as the determinants of the expected rate of growth of a firm. Assuming that the learning function of a typical firm with size $x$ is give by a pair of equations (30) and (31), we can express its rate of profit as follows:

$$e(x) = \frac{x - wl(x)}{k} = \frac{x - wcx^{\lambda}}{dx^{\mu}} = \frac{1}{d}x^{1-\mu} - \frac{c}{d}wx^{\lambda-\mu}, \quad x \in [1, X], \tag{32}$$

where $w$ denotes the wage rate.

The average wage per worker tends to be an increasing function with respect to the size of firms, although not to the same degree as decreases of labor input.

One reason for this is that larger firms will usually have a more detailed division of labor, with a larger proportion of higher-paid skilled or managerial workers. Another reason is that trade unions are usually more powerful in larger firms, and may succeed in extracting part of extra profit created by economies of scale. Because of these reasons, we assume that the average wage rate increases with the size of firms according to the following equation:[7]

$$w(x) = w(1)x^{\omega}, \ \textit{where} \ \omega > 0, \ x \in [1, X]. \tag{33}$$

To simplify the following analysis, we assume $\mu = 1$ in equation (31). In other words, the capital-output ratio is assumed to be constant independently of the size of firms. This assumption is approximately supported by the actual data of Japanese manufacturing industry.[8] With this assumption and (33), the rate of profit of a typical firm with size $x$ becomes as follows:

$$e(x) = \frac{1}{d}\{1 - w(1)cx^{\lambda+\omega-1}), \ \ x \in [1, X]. \tag{34}$$

As is obvious from this function, the rate of profit is constant irrespective of the firm size if $\lambda + \omega = 1$. When $\lambda + \omega \neq 1$, the rate of profit increases or decreases with firm size $x$ depending on whether $\lambda + \omega < 1$ or $\lambda + \omega > 1$.[9]

The incentive of firms to expand arises from the prospect of improving their profitability by increasing their scale of operation. The accumulated profits can be used for further expansion either directly or as security for raising external finance. We assume, therefore, that the expected growth rate of capital, $E(\Delta k/k)$ of a typical firm with size $x$ depends on the rate of profit earned by the firm, $e(x)$. Notice that $E(\Delta k/k) = E(\Delta x/x)$ is obtained from (31), since we assume $\mu = 1$. Hence, the expected growth rate of capacity output, $E(\Delta x/x)$, depends on $e(x)$. Specifying this relation to be linear, we have

$$E\left(\frac{\Delta x}{x}\right) = \tau + \xi e(x), \ \textit{where} \ \tau > 0, \ \xi > 0, \ x \in [1, X]. \tag{35}$$

Substituting (34) into (35), we can express (35) as follows:

$$E\left(\frac{\Delta x}{x}\right) = \tau + \frac{\xi}{d}(1 - wcx^{\lambda+\omega-1}) = p - qx^{\lambda+\omega-1}, \tag{36}$$

where $p \equiv \tau + \xi/d$ and $q \equiv \xi wc/d$, both being constant. In order that the expected rate of growth to be positive for a typical minimum firm whose capacity output $x$ is equal to 1, we must have $p > q$. We assume this condition to hold in the following.

Let us first consider the case where $\lambda + \omega = 1$. In this case, the expected rate of capacity output becomes as

$$E\left(\frac{\Delta x}{x}\right) = p - q, \tag{37}$$

where $p - q$ is constant. In other words, the expected rate of growth of a firm is independent of its size $x$. In this case, Assumption 1 in Simon's model, that is 'the law of proportionate effect', applies to our model. We may also assume that all new firms start their operations from the minimum size. This assumption seems plausible in our model, because the two brakes mentioned above will affect not only existing firms but also new firms. First of all, since new entrants do not have the necessary know-how to operate efficiently at large size from the start, they have to enter the industry from the minimum size. Secondly, new entrants usually cannot raise sufficient finance to start on a large scale. These reasons justify the assumption that all new firms make entries from the minimum size. In addition, we assume that the birth rate of new firms is constant. Then, Assumption 2 in Simon's model, a constant rate of entry of new firms from the minimum size, also applies to our model. Therefore, due to Proposition 1, the size distribution of firms in our model becomes the Pareto distribution of the form (15).

This result may be summarized by the following proposition:

**Proposition 3:** *Suppose that new firms are being born in the smallest-size class, and that they account for a constant rate $\alpha$ of an increase in the total capacity output of the economy. Suppose also that a typical firm of each size class masters technology of operating more efficiently on a larger scale by learning by doing as represented by (30) and (31), and that its rate of expansion depends on the rate of profit as is expressed by (36). Then, if $\lambda + \omega = 1$, the size distribution of firms converges to the Pareto distribution of the form* (15), *and the Pareto coefficient $\rho$ is equal to* $1/(1-\alpha)$.

In this proposition, the condition $\lambda + \omega = 1$ implies that the unit wage cost for the average firm at each size class as well as its gross profit rate is constant irrespective of size. Since $\lambda < 1$, entrepreneurs expect to enjoy economies of scale by expansion. But, as $\omega = 1 - \lambda$, the wage rate increases with size just to offset the scale economies. The law of proportionate effect applies as the result.

Let us next consider the case where $\lambda + \omega \neq 1$. In this case, as is obvious from (36), the expected rate of growth of a firm increases or decreases with size $x$ depending on whether $\lambda + \omega < 1$ or $\lambda + \omega > 1$. The problem is how to relate these cases to Assumption 1′ in Sato's model. Equation (36) can be rewritten as

$$E(\Delta x) = px - qx^{\lambda+\omega}. \tag{38}$$

Linearizing the righthand side of this equation around a certain value of $x$ denoted by $x^*$, we obtain

$$E(\Delta x) = p(x - x^*) - (\lambda+\omega)q(x - x^*) = (p-q)\frac{p-(\lambda+\omega)q}{p-q}(x - x^*), \tag{39}$$

where we assume $p > (\lambda+\omega)q$ in order to focus on economically meaningful case. The above equation can be rewritten as

$$E(\Delta x) = (p-q)(ax+b) \ \ or \ \ E\left(\frac{\Delta x}{x}\right) = (p-q)\left(a + \frac{b}{x}\right), \tag{40}$$

where

$$a \equiv \frac{p-(\lambda+\omega)q}{p-q}, \quad b \equiv \frac{p-(\lambda+\omega)q}{p-q}x^*. \tag{41}$$

In addtion, we assume that $a$ and $b$ defined by (41) satisfy (24). Then, the value of $x^*$ is determined as

$$x^* = \frac{a-1}{a\alpha}. \tag{42}$$

Therefore, $a$ and $b$ are determined by the parameters that are given in our model. From (41) and (42), we can confirm that if $\lambda+\omega<1$, then $a>1$ and $b<0$, while if $\lambda+\omega>1$, then $a<1$ and $b>0$.

Equation (40) implies that the expected growth rate of a firm with size $x$ is proportional to $a+b/x$. In this case, Assumption 1′ in Sato's model, that is the generalized law of proportionate effect, exactly applies to our model. Therefore, due to Proposition 2, the size distribution of firms in our model becomes the Pareto distribution of the form (19).

This result may be summarized by the following proposition:

**Proposition 4:** *Let us make the same assumptions as in Proposition 3 except $\lambda+\omega=1$. Then, the size distribution of firms converges to the Pareto distribution of the form (19), the Pareto coefficient being equal to $\rho/a$. Depending on whether $\lambda+\omega<1$ or $\lambda+\omega>1$, we can distinguish the following two cases:*

*(a) If $\lambda+\omega<1$, then $a>1$ and $b<0$, in which case, the expected growth rate increases with size, and the Pareto coefficient is less than $\rho=1/(1-\alpha)$.*

*(b) If $\lambda+\omega>1$, then $a<1$ and $b>0$, in which case, the expected growth rate decreases with size, and the Pareto coefficient exceeds $\rho=1/(1-\alpha)$.*

In this proposition, the condition $\lambda+\omega<1$ implies that the unit wage cost for an average firm at each size class decreases and its gross profit rate increases with size. Thus, the expected growth rate increases with size. The exact opposite holds for the case $\lambda+\omega>1$. It should be noted, however, that, even if the gross profit rate is decreasing with size, the level of gross profits will increase as long as $xe(x)$ increases with $x$. Therefore, even in the case where $\lambda+\omega>1$, firms may have some incentives to expand, though the growth rate is decreasing with size. Even if it is so, however, equation (38) implies that expansion will stop at the capacity output $x^{\#}$ such that $x^{\#}=(p/q)^{1/(\lambda+\omega-1)}$. Moreover, it is quite unrealistic to assume that $\lambda+\omega>1$ holds for the whole range of size classes, because the rate of profit becomes highest for a minimum size firm in this case. Therefore, we assume that $\lambda+\omega\leq 1$ holds in the following discussion.

# 4 Determinants of Income Shares

It has been shown in the previous section that the size distribution of firms converges to the Pareto distribution under the quite plausible assumptions on the behavior of firms about technology and investment. In this section, we will turn to the analysis of the economy as a whole. We assume that the size distribution of firms is approximated by the Pareto distribution in the economy as a whole. We also assume that learning function described by (30) and (31) is applicable to the whole economy.

When the size distribution of firms is approximated by the Pareto distribution over the entire range, the number of firms with capacity output $x$, denoted by $n(x)$, can be expressed by the following frequency function:

$$n(x) = \rho A x^{-(\rho+1)}, \quad (\rho > 1,\ A > 0). \tag{43}$$

The Pareto coefficient $\rho$ is determined by the birth rate of new firms $\alpha$ and the scale factor, $a$, as $\rho = 1/a(1-\alpha)$.

Denoting capacity output of the minimum size firm by $x_0$ and that of maximum size firm by $x_T$, we can express the total number of firms, $N$, as follows:

$$N = \int_{x_0}^{x_T} n(x)dx = A(x_0^{-\rho} - x_T^{-\rho}). \tag{44}$$

Let us denote the ratio of $x_T$ to $x_0$ by $m$, namely,

$$x_T = m x_0, \tag{45}$$

where $m > 1$. We call $m$ the 'size differential ratio' in the following. Using this variable, we can rewrite (45) as follows:

$$N = A(1 - m^{-\rho})x_0^{-\rho}. \tag{46}$$

Similarly, total output, $X$, is given by

$$X = \int_{x_0}^{x_T} x n(x)dx = \frac{\rho A}{\rho - 1}(1 - m^{1-\rho})x_0^{1-\rho}. \tag{47}$$

Total output here is measured by the value added. We can also calculate total labor employment, $L$, and total capital stock, $K$, by taking into account (30) and (31), to get

$$L = \int_{x_0}^{x_T} l(x)n(x)dx = \frac{\rho A c}{\rho - \lambda}(1 - m^{\lambda-\rho})x_0^{\lambda-\rho}, \tag{48}$$

$$K = \int_{x_0}^{x_T} k(x)n(x)dx = \frac{\rho A d}{\rho - \mu}(1 - m^{\mu-\rho})x_0^{\mu-\rho}. \tag{49}$$

In the last section, $\mu$ is assumed to be unity when we show that the size distribution of firms becomes the Pareto distiribution under the learning by doing hypothesis. We will keep this assumption in the following analysis. In this case, (49) becomes as

$$K = \frac{\rho A d}{\rho - 1}(1 - m^{1-\rho})x_0^{1-\rho} = dX. \tag{50}$$

Let us next examine the determination of wages and profits. As is expressed by equation (33), we assume that the wage rate increases with size at a rate $\omega$. When the wage rate of a minimum size firm is $w(x_0)$, that of a firm with size $x$, namely $w(x)$, is given by

$$w(x) = w(x_0)\left(\frac{x}{x_0}\right)^{\omega}. \tag{51}$$

As for the determination of $w(x_0)$, we assume that a minimum size firm (or we may call them a 'marginal firm') set pruduct price $P$ with mark-up factor $\beta$ on wage costs.

$$P = \beta\frac{W(x_0)l(x_0)}{x_0}, \tag{52}$$

where $W(x_0)$ denotes the nominal wage rate paid by a maginal firm. We assume that marginal firms are under perfect competition, so that mark-up factor $\beta$ is determined at the level just sufficient to cover capital costs. Then, the real wage rate of a typical marginal firm, $w(x_0) \equiv W(x_0)/P$, is determined as

$$w(x_0) = \frac{1}{\beta}\frac{x_0}{l(x_0)} = \frac{1}{\beta c}x_0^{1-\lambda}. \tag{53}$$

Thus, the real wage rate of a marginal firm depends on its output, $x_0$, and technological coefficient, $c$.

In our model, all the new entrants into the economy are small enterprises of minimum size. They will grow by improving their technology through experience. As successful firms expand their scale, they will, on the average, be able to reduce their costs through learning by doing processes. Especially, when $\lambda+\omega < 1$ and $\mu = 1$, the larger firms attain more favorable profit margins than smaller firms.

From (47), (48), (51) and (53), the aggregate share of wages in value added, $S_w$, becomes as

$$S_w = \frac{\rho-1}{\beta(\rho-\lambda-\omega)}\frac{1-m^{\lambda+\omega-\rho}}{1-m^{1-\rho}}. \tag{54}$$

Thus, the aggregate share of wages in value added depends on the Pareto coefficient, $\rho$, the degree of economies of scale, $\lambda$, the wage differentials, $\omega$, the size differential ratio, $m$, and the mark-up factor, $\beta$. If $\lambda+\omega = 1$, it is obvious from (54) that $S_w = 1/\beta$. If $\lambda+\omega < 1$, however, the aggregate income shares depend not only on the mark-up factor of marginal firms, $\beta$, but also on the structural parameters of the economy, such as $\rho$, $\lambda$, $\omega$, and $m$. This theory of income shares is quite different from the neoclassical marginal productivity theory.[10]

Assuming that $\rho > 1 > \lambda+\omega$, we can prove by comparative static analysis of equation (54) that the aggregate wage share, $S_w$, depends on its determinants as follows:[11]

$$\frac{\partial S_w}{\partial \rho} > 0,\ \frac{\partial S_w}{\partial \lambda} > 0,\ \frac{\partial S_w}{\partial \omega} > 0,\ \frac{\partial S_w}{\partial \beta} < 0,\ \frac{\partial S_w}{\partial m} < 0. \tag{55}$$

In other words, the aggregate wage share increases with $\rho$, $\lambda$ and $\omega$, and diminishes with $\beta$ and $m$. An increase in $\rho$ represents a reduction in the proportion of large firms, and hence in the average profit margin in the whole population of firms. An increase in $\lambda$ represents a decrease in the degree of economies of scale, and hence a reduction in the profit margins of larger firms. An increase in $\omega$ represents an increase in the wage differential paid by large firms, and hence a reduction in their profit margins. An increase in $\beta$ represents an increase in the mark-up factor of a marginal firm, and hence an increase in the profit margin of all firms. An increase in $m$ represents an increase in size differential, and hence an increase in the profit margin of larger firms.

## 5 A Model of Growth with Economies of Scale

In this section, we will construct a growth model that takes into account the size distribution of firms, and examine the dynamics of aggregate variables obtained in the previous section. Taking the time derivatives of equations (46), (47), (48) and (49), we can express the growth rates of $N$, $X$, $N$, and $K$, as follows:

$$\frac{\dot{N}}{N} = \frac{\dot{A}}{A} + \frac{\rho}{m^{\rho} - 1}\frac{\dot{m}}{m} - \rho\frac{\dot{x}_0}{x_0}, \tag{56}$$

$$\frac{\dot{X}}{X} = \frac{\dot{A}}{A} + \frac{\rho - 1}{m^{\rho-1} - 1}\frac{\dot{m}}{m} - (\rho - 1)\frac{\dot{x}_0}{x_0}, \tag{57}$$

$$\frac{\dot{L}}{L} = \frac{\dot{A}}{A} + \frac{\dot{c}}{c} + \frac{\rho - \lambda}{m^{\rho-\lambda} - \lambda}\frac{\dot{m}}{m} - (\rho - \lambda)\frac{\dot{x}_0}{x_0}, \tag{58}$$

$$\frac{\dot{K}}{K} = \frac{\dot{A}}{A} + \frac{\dot{d}}{d} + \frac{\rho - 1}{m^{\rho-1} - 1}\frac{\dot{m}}{m} - (\rho - 1)\frac{\dot{x}_0}{x_0}. \tag{59}$$

Thus, all these growth rates depend on the rate of shift in the Pareto curve, $\dot{A}/A$, the rate of increase in the size differential ratio, $\dot{m}/m$, and the rate of growth in output of a marginal firm, $\dot{x}_0/x_0$. In addition to them, the growth rate of $L$ depends on the rate of change in the labor input coefficient $c$, and the growth rate in $K$ depends on the rate of change in the capital input coefficient, $d$. As is obvious from (30), a decrease in $c$ leads to a reduction in labor input per unit of output for every size class of firms. If there is exogenous technological progress that is common to all firms, $\dot{c}/c$ takes negative value. Thus, $(-\dot{c}/c)$ represents labor-augmenting technological progress. Similarly, as is obvious from (31), a decrease in $d$ leads to a reduction in capital input per unit of output for every size class of firms. Therefore, $(-\dot{d}/d)$ represents capital-augmenting technological progress. In the following, we assume that there is no capital-augmenting technological progress, so that $d$ is constant.

It is assumed in our model that new entrants start their operation at the minimum size $x_0$, and that the proportion $\alpha$ of the increment in total output, $\dot{X}$, is apportioned to new firms, $\dot{N}$. In other words, we have

$$\alpha\dot{X} = x_0\dot{N}, \tag{60}$$

which can be rewritten as

$$\frac{\dot{N}}{N} = \alpha\frac{X}{x_0 N}\frac{\dot{X}}{X}. \tag{61}$$

Substituting (46) and (47) into this equation, and taking into account the relation $\rho = 1/a(1-\alpha)$, we obtain the following relationship between the growth rate of output and the growth rate of the number of firms:

$$\frac{\dot{N}}{N} = \frac{\rho-(1/a)}{\rho-1}\frac{m^\rho - m}{m^\rho - 1}\frac{\dot{X}}{X}. \tag{62}$$

We assume here that $\rho > 1 \geq 1/a$ holds. The condition $\rho > 1$ is required for the size distribution of firms to converge to the Pareto distribution. The condition $a \geq 1$ holds true since we assume $\lambda + \omega \leq 1$. Now, substituting (56) into (62), we can solve it with respect to the rate of shift in the Pareto curve to obtain:

$$\frac{\dot{A}}{A} = \frac{\rho-(1/a)}{\rho-1}\frac{\dot{X}}{X} - \frac{\rho}{m^\rho - 1}\frac{\dot{m}}{m} + \rho\frac{\dot{x}_0}{x_0}. \tag{63}$$

As for the supply of labor, we assume that the population of labor grows at a constant rate, $n$. We also assume that the labor market is always in equilibrium. Then, we have

$$\frac{\dot{L}}{L} = n. \tag{64}$$

To complete the model, we have to specify the equation for the capital accumulation. In our model, we assume that a fraction of $s_p$ of total profits and a fraction of $s_w$ of total wages are saved and devoted to investment, and that $0 \leq s_w < s_p < 1$. For simplicity, we assume that there is no depreciation of capital. Then the growth rate of capital is expressed by the following equation:[12]

$$\frac{\dot{K}}{K} = \frac{X}{K}[s_p(1-S_w) + s_w S_w], \tag{65}$$

where $S_w$ is the share of wages in value added defined by (54). It is a decreasing function with respect to $m$ as is shown by (55). In other words, $S_w$ may be denoted as $S_w(m)$, whose derivative becomes as $S_w'(m) < 0$.

In view of (50), we have $K = dX$. As is mentioned above, $d$ is assumed to be constant, and hence we have $\dot{K}/K = \dot{X}/X$. Then, equation (65) is rewritten as

$$\frac{\dot{X}}{X} = \frac{1}{d}[(s_p - s_w)\{1 - S_w(m)\} + s_w]. \tag{66}$$

Since $S_w(m)$ is a decreasing function of $m$, the growth rate of $X$ is an increasing function of $m$. Thus, if we denote the growth rate of $X$ as $g_X$, equation (66) may be rewritten as follows:

$$\frac{\dot{X}}{X} = g_X(m) = \frac{1}{d}[(s_p - s_w)\{1 - S_w(m)\} + s_w], \quad g_X(m) > 0. \tag{67}$$

To summarize, independent equations in the above model are (56) through (59), (62), (63), (64) and (66), and the variables to be determined are $N, X, L, K, A, m$ and $x_0$. Thus, our model consists of 7 equations which include 7 variables, so that it is complete. This model can be reduced to a system consisting of two equations as follows. Substituting (63) into (57), and using the notation $g_X(m)$ for the growth rate of $X$, we have

$$\left(\frac{\rho-1}{m^{\rho-1}-1} - \frac{\rho}{m^\rho - 1}\right)\frac{\dot{m}}{m} + \frac{\dot{x}_0}{x_0} = \left(1 - \frac{\rho-(1/a)}{\rho-1}\frac{m^\rho - m}{m^\rho - 1}\right) g_X(m). \tag{68}$$

This condition represents the equilibrium condition for the capital goods market.

Similarly, substituting (63) and (64) into (58), we have

$$\left(\frac{\rho-\lambda}{m^{\rho-\lambda}-1} - \frac{\rho}{m^\rho - 1}\right)\frac{\dot{m}}{m} + \lambda\frac{\dot{x}_0}{x_0} = (n+g) - \frac{\rho-(1/a)}{\rho-1}\frac{m^\rho - m}{m^\rho - 1} g_X(m), \tag{69}$$

where $g \equiv -\dot{c}/c$ represents the rate of exogenous technological progress. This equation represents the equilibrium condition for the labor market. The system consisting of (68) and (69) includes two variables, $m$ and $x_0$, so that it is a complete system.

Solving this system of equations with respect to $\dot{m}/m$, we obtain

$$\frac{\dot{m}}{m} = \frac{1}{D(m)}[\Phi(m)g_X(m) - (n+g)], \tag{70}$$

where

$$D(m) \equiv \lambda\left(\frac{\rho-1}{m^{\rho-1}-1} - \frac{\rho}{m^\rho - 1}\right) - \left(\frac{\rho-\lambda}{m^{\rho-\lambda}-1} - \frac{\rho}{m^\rho - 1}\right). \tag{71}$$

$$\Phi(m) \equiv \lambda + (1-\lambda)\frac{\rho-(1/a)}{\rho-1}\frac{m^\rho - m}{m^\rho - 1}. \tag{72}$$

The expression of $D(m)$ seems quite complicated, but it can be shown that it takes the positive value for the relevant range of $m$. Actually, it is proved that there exists a certain value of $m$ denoted by $\bar{m}$ such that both of them are positive for $m > \bar{m}$:[13]

$$D(m) > 0 \quad for \quad m > \bar{m}. \tag{73}$$

The magnitude of $\bar{m}$ is small enough compared to the practically relevant range of $m$, so that we may safely asuume that $D(m) > 0$ always holds in our model. As for (72), on the other hand, it is easily seen that

$$\Phi(m) > 0 \; and \; \Phi'(m) > 0 \quad for \;\; m > 1. \tag{74}$$

Next, if we solve the above system of equations (68) and (69) with respect to $\dot{x}_0/x_0$, we have

$$\frac{\dot{x}_0}{x_0} = \frac{1}{D(m)}[\Psi(m)(n+g) - \Omega(m)g_X(m)], \tag{75}$$

where

$$\Psi(m) \equiv \frac{\rho-1}{m^{\rho-1}-1} - \frac{\rho}{m^{\rho}-1}, \tag{76}$$

$$\Omega(m) \equiv \left(\frac{\rho-\lambda}{m^{\rho-\lambda}-1} - \frac{\rho}{m^{\rho}-1}\right) + \frac{\rho-(1/a)}{\rho-1}\frac{m^{\rho}-m}{m^{\rho}-1}\left(\frac{\rho-1}{m^{\rho-1}-1} - \frac{\rho-\lambda}{m^{\rho-\lambda}-1}\right). \tag{77}$$

Both of these functions are shown to be positive for $m > 1$:[14]

$$\Psi(m) > 0 \; and \; \Omega(m) > 0 \quad for \;\; m > 1. \tag{78}$$

Thus, the dynamic determination of our system proceeds as follows. Equation (70) determines the dynamic path of $m$ starting from its initial value. Corresponding to the path of $m$, the growth rate of output, $g_X(m)$, is determined by (67), and the minimum size of firms, $x_0$ is determined by (75). The growth rate of capital, $g_K(m)$, is always equal to the growth rate of output, $g_X(m)$, under the assumption of fixes coefficients. This assumption will be relaxed later.

In view of the dynamic equation (70), the steady growth equilibrium is attained at $m^*$ that satisfies the following equation:

$$g_X(m^*) = \frac{1}{d}[(s_p - s_w)\{1 - S_w(m^*)\} + s_w] = \frac{n+g}{\Phi(m^*)}. \tag{79}$$

This equation exhibits quite interesting implications of our model. First of all, our model is different the Solow model in that the steady growth rate depends not only on the sum of population growth and technological progress, $n+g$, but also on economies of scale arising from learning by doing of firms. As we will show below, $0 < \Phi(m^*) < 1$, and hence the steady growth rate of output, $g_X(m^*)$, is greater than $n+g$. The function $\Phi(m^*)$ reflects economies of scale, and the smaller its value, the larger the effects of economies of scale. The contribution of economies of scale to the steady growth rate may be measured by

$$g_X(m^*) - (n+g) = (n+g)\frac{1-\Phi(m^*)}{\Phi(m^*)}. \tag{80}$$

Secondly, in contrast to the Solow model in which the steady growth rate is independent of the saving rate, in our model, it is increasing with the saving rates, $s_p$ or $s_w$. This can be shown as follows. In view of (79), the growth rate of output, $g_X$ depends not only on $m$ but also on $s_p$ and $s_w$. To make it explicit, we may write the steady growth rate of output as $g_X(m^*; s_p, s_w)$. This function has the properties that $\partial g_X/\partial s_p > 0$ and $\partial g_X/\partial s_w > 0$. Calculating the effects of changes in $s_p$ or $s_w$ on $m^*$ from (79), we have

$$\frac{dm^*}{ds_p} = -\frac{\frac{\partial g_X}{\partial s_p}/g_X}{(g_X'/g_X)+(\Phi'/\Phi)} < 0, \quad \frac{dm^*}{ds_w} = -\frac{\frac{\partial g_X}{\partial s_w}/g_X}{(g_X'/g_X)+(\Phi'/\Phi)} < 0, \quad (81)$$

where $g_X'$ and $\Phi'$ denote derivatives of $g_X$ and $\Phi$ with respect to $m^*$. These results show that increases in $s_p$ or $s_w$ decreases $m^*$. But a decrease in $m^*$ decreases $\Phi(m^*)$ in view of (74), and hence increases $g_X(m^*)$ in view of (79). Thus, increases in the saving rates leads to a rise in the steady growth rate in our model.

Let us next examine how the growth rate of output of a minimum firm, $\dot{x}_0/x_0$, is determined at the steady state equilibrium. Substituting (79) into (75), we get, after simplifying,

$$\left(\frac{\dot{x}_0}{x_0}\right)^* = \frac{n+g}{1-\lambda}\frac{1-\Phi(m^*)}{\Phi(m^*)}. \quad (82)$$

Comparing this expression with (80), we find that the larger the degree of economies of scale, the higher the steady growth rate of output of a minimum firm.

We have to notice that there exists a lower limit to this growth rate. In view of (30), the labor input function for a minimum firm is written as $l_0 = cx_0^\lambda$. If we assume that the minimum firm is a one-man firm, then we must have $1 = cx_0^\lambda$. Differentiating this relationship with respect to time and solving it for the growth rate of $x_0$, we obtain the following expression for the growth rate of output of a one-man firm.

$$\left(\frac{\dot{x}_0}{x_0}\right)_1 = -\frac{1}{\lambda}\frac{\dot{c}}{c} = \frac{g}{\lambda}. \quad (83)$$

As long as the exogenous rate of technological progress, $g$, is positive, this expression must be positive. Since, minimum firms must be greater than one-man firms, we must have, from (82) and (83),

$$\frac{n+g}{1-\lambda}\frac{1-\Phi(m^*)}{\Phi(m^*)} \geq \frac{g}{\lambda}. \quad (84)$$

From this inequality, it follows that $0 < \Phi(m^*) < 1$.

The steady growth rate of output of a minimum firm represented by (82) decreases with $m^*$, since $\Phi(m^*)$ is an increasing function. In other words, the smaller the value of $m^*$, the larger the degree of economies of scale, and hence the larger the value of $(\dot{x}_0/x_0)^*$. When $(\dot{x}_0/x_0)^* > g/\lambda$ holds, marginal firms including one-man firms are exiting, because their productivity growth is too low. This case happens when larger firms are growing faster than one-man firms, namely, $(\dot{x}_T/x_T) > g/\lambda$. When larger firms are growing fast, workers tend to move to those those firms since they can get higher wages there. Marginal firms are forced to exit as a result.

In our model, firms with different size coexist, and they grow over time by taking advantage of potential economies of scale through learning by doing. This feature of our model may seem somewhat similar to the endogenous growth model of the Arrow type. However, our model differs from the existing endogenous growth models in that it takes into account of the size distribution of firms.

Let us next examine the stability of the steady growth equilibrium derived above. For this purpose we focus on the dynamic equation (70). It is a differential equation that determines the path of $m$ over time. Since $D(m)$, $\Phi(m)$, and $g_X(m)$ are all increasing functions, $\dot{m}/m$ increases with $m$ in the neighborhood of the steady state equilibrium, $m = m^*$. Hence, the steady state equilibrium is unstable. Figure 3 provides a graphical representation of this instability property. Suppose that $m > m^*$ holds initially. Then, $m$ and $\dot{m}/m$ will increase over time, and so will $g_X(m)$. In this case, the equilibrium condition for the labor market (69) will be violated sooner or later. Conversely, suppose that $m < m^*$ holds initially. Then, $m$ and $\dot{m}/m$ decreases over time, and so will $g_X(m)$. In this case, the equilibrium condition for the capital goods market (68) will be violated sooner or later. Thus, the steady growth equilibrium will not be maintained, unless $m = m^*$ is satisfied initially.

# 6 Factor Substitution and the Stability of the Steady Growth Equilibrium

So far we have assumed that the production process for each size of firms is characterized by fixed coefficients, so that a fixed amounts of labor and capital are used corresponding to a given amount of output. In this section, we take into account the substitutability between labor and capital, and show that it stabilizes the system.

When there is substitutability between labor and capital, the production function of a typical firm with size $x$ may be expressed as

$$x = F\left(\frac{1}{\gamma(x)}l, \frac{1}{\delta(x)}k\right), \tag{85}$$

where $\gamma(x)$ and $\delta(x)$ are the learning functions defined by (28) and (29). We assume that this production function exhibits constant returns to scale and other neoclassical properties. We also keep the assumption for the learning function $\delta(x) = dx^{\mu-1}$, that $\mu = 1$ and $d$ is constant. So, we may put $\delta(x) = 1$ for convenience. For simplicity, we specify the production function as the Cobb-Douglas type. With these assumptions, we can rewrite (85) as follows:

$$x = \left(\frac{l}{\gamma(x)}\right)^{1-\eta} k^{\eta} = \frac{l}{\gamma(x)}\left(\gamma(x)\frac{k}{l}\right)^{\eta}, \text{ where } 0 < \eta < 1. \tag{86}$$

A typical firm with size $x$, chooses an optimum capital-labor ratio that minimize the total cost, given the level of output and technological knowledge. Thus, the problem for the typical firm is formulated as follows:

$$\min\ wl + rk, \quad s.t.\ \bar{x} = \frac{l}{\bar{\gamma}}\left(\bar{\gamma}\frac{k}{l}\right)^{\eta}. \tag{87}$$

The first order condition for this minimization problem is

$$\frac{w}{r} = \frac{1-\eta}{\eta}\frac{k}{l}. \tag{88}$$

Solving this equation with respect to $k/l$, and substituting it into (86), we can express the production function (86)

$$x = \frac{l}{\gamma(x)}\left(\frac{\eta}{1-\eta}\frac{\gamma(x)w}{r}\right)^{\eta}, \tag{89}$$

where $\gamma(x)w$ represents the wage rate per efficiency unit of labor for a typical firm with size $x$, which will be denoted by $w^e(x)$ in the following. Then, this production function may be transformed into the labor-input function of a firm with size $x$ as follows:

$$\frac{l(x)}{x} = \theta\left(\frac{w^e(x)}{r}\right)^{-\eta}\gamma(x), \tag{90}$$

where $\theta \equiv \{\eta/(1-\eta)\}^{-\eta}$. Compare this function with the learning function (26). Then, this function may be interpreted as the generalized learning function that takes into account the substitutability between labor and capital.

In our model, the wage rate for each size of firms is determined endogenously by (51) and (53), but the rate of interest is given exogenously. So, we assume $r$ to be constant, and put it equal to unity for convenience. We use (28) for the learning function $\gamma(x)$. Then, (90) is rewritten as

$$l(x) = c\{w^e(x)\}^{-\eta}x^{\lambda}, \tag{91}$$

where constant parameter $\theta$ is ignored without loss of generarity. From the definition of $w^e(x)$, we can express it in terms of the wage rate per man-hour labor, $w$, as

$$w^e(x) = \gamma(x)w = wcx^{\lambda-1}. \tag{92}$$

But, $w$ is also a function of $x$, as is shown by (51). Hence, we have

$$w^e(x) = cx_0^{-\omega}w(x_0)x^{\lambda+\omega-1}. \tag{93}$$

Substituting (93) into (92) yields

$$l(x) = c^{1-\eta} x_0^{\omega\eta} \{w(x_0)\}^{-\eta} x^{\lambda'}, \tag{94}$$

where $\lambda' = \lambda + \eta(1 - \lambda - \omega)$.Since we assume $\lambda + \omega \leq 1$, we must have $\lambda' \geq \lambda$. It is also assumed that $\lambda' < 1$.

Substituting (94) into (48), we can calculate the total labor employment as follows:

$$L = \int_{x_0}^{x_T} l(x) n(x) dx = \frac{\rho A c^{1-\eta} x_0^{\omega\eta} \{w(x_0)\}^{-\eta}}{\rho - \lambda'} (1 - m^{\lambda' - \rho}) x_0^{\lambda' - \rho}. \tag{95}$$

The wage rate for a marginal firm,$w(x_0)$, is determined by $w(x_0) = (1/\beta)\{x_0/l(x_0)\}$, as is shown by (53). Substituting (94) into this relation, and solving with respect to $w(x_0)$, we have

$$w(x_0) = \frac{1}{c} \left(\frac{1}{\beta}\right)^{\frac{1}{1-\eta}} x_0^{1-\lambda}. \tag{96}$$

Substituting this equation into (95) and arranging it, we obtain the total employment as follows:

$$L = \frac{\rho A c \beta^{\eta/(1-\eta)}}{\rho - \lambda'} (1 - m^{\lambda' - \rho}) x_0^{\lambda' - \rho - \varepsilon}, \tag{97}$$

where $\varepsilon = \eta(1 - \lambda - \omega)$ and $\lambda' = \lambda + \epsilon$. Differentiating (97) with respect to time, we obtain the following expression for the growth rate of total employment:

$$\frac{\dot{L}}{L} = \frac{\dot{A}}{A} + \frac{\dot{c}}{c} + \frac{\rho - \lambda'}{m^{\rho - \lambda'} - 1} \frac{\dot{m}}{m} + (\rho - \lambda' + \epsilon) \frac{\dot{x}_0}{x_0}. \tag{98}$$

Thus, when we take into consideration the substitutability between capital and labor, the equation for the level of employment (48) is replaced by (97), and the equation for the growth rate of employment (58) is replaced by (98). The share of wages in the value added (54) is also replaced by

$$S_w = \frac{\rho - 1}{\beta(\rho - \lambda' - \omega)} \frac{1 - m^{\lambda' + \omega - \rho}}{1 - m^{1-\rho}}. \tag{99}$$

Similarly, we can caluculate the capital-input function and total capital stock for this case. The capital-input function of a firm with size $x$ becomes as

$$k(x) = d\{w^e(x)\}^{1-\eta} x, \tag{100}$$

and the total capital stock becomes as

$$K = \frac{\rho A d \beta^{-1}}{\rho + \kappa - 1} (1 - m^{1-\rho-\kappa}) x_0^{1-\rho}, \tag{101}$$

where $\kappa = (1-\eta)(1-\lambda-\omega)$. Then, from (47) and (100), the total output-capital ratio becomes as

$$\frac{X}{K} = \frac{d(\rho+\kappa-1)}{\beta(\rho-1)} \frac{1-m^{1-\rho}}{1-m^{1-\rho-\kappa}}. \tag{102}$$

This equation shows that, unless $\kappa = 0$, the output-capital ratio is not constant as in the fixed coefficient case, but depends on $m$. Differentiating this equation with respect to time, we obtain the following relationship between the growth rate of output and the growth rate of capital:

$$\frac{\dot{X}}{X} = \frac{\dot{K}}{K} + \left(\frac{\rho-1}{m^{\rho}-m} - \frac{\rho+\kappa-1}{m^{\rho+\kappa-1}-m}\right)\frac{\dot{m}}{m}. \tag{103}$$

Thus, the growth rate of output deviates from the growth rate of capital if $\kappa \neq 0$ and $\dot{m} \neq 0$.

Let us assume as before that the growth rate of capital is determined by (65). Then, substituting (102) into (65) yields

$$\frac{\dot{K}}{K} = g_K(m) = \frac{d(\rho+\kappa-1)}{\beta(\rho-1)} \frac{1-m^{1-\rho}}{1-m^{1-\rho-\kappa}}[(s_p - s_w)(1-S_w) + s_w]. \tag{104}$$

where $S_w$ is defined by (99). It is easily shown that equation (99) is increasing and equation (102) is decreasing with respect to $m$, and hence $g_K(m)$ is an increasing function. From (103), the growth rate of output is expressed as follows:

$$g_X(m) = g_K(m) + \left(\frac{\rho-1}{m^{\rho}-m} - \frac{\rho+\kappa-1}{m^{\rho+\kappa-1}-m}\right)\frac{\dot{m}}{m}. \tag{105}$$

Now, substituting this equation into (68) yields

$$\left(\frac{\rho-1}{m^{\rho-1}-1} - \frac{\rho}{m^{\rho}-1} - \Theta(m)\right)\frac{\dot{m}}{m} + \frac{\dot{x}_0}{x_0} = \left(1 - \frac{\rho-(1/a)}{\rho-1}\frac{m^{\rho}-m}{m^{\rho}-1}\right) g_K(m), \tag{106}$$

where

$$\Theta(m) \equiv \left(1 - \frac{\rho-(1/a)}{\rho-1}\frac{m^{\rho}-m}{m^{\rho}-1}\right)\left(\frac{\rho-1}{m^{\rho}-m} - \frac{\rho+\kappa-1}{m^{\rho+\kappa-1}-m}\right) > 0 \;\; for\; m > 1. \tag{107}$$

Equation (106) represents the equilibrium condition for the capital goods market in the case where capital and labor are substitutable. Thus, equation (68) in the fixed coefficient model is changed to (106) when the factor substitutability is taken into account.

Similarly, the equilibrium condition for the labor market is obtained by substituting (63), (64) and (105) into (98) as follows:

$$\left(\frac{\rho-\lambda'}{m^{\rho-\lambda'}-1} - \frac{\rho}{m^{\rho}-1} + \Lambda(m)\right)\frac{\dot{m}}{m} + (\lambda'-\varepsilon)\frac{\dot{x}_0}{x_0} = (n+g) - \frac{\rho-(1/a)}{\rho-1}\frac{m^{\rho}-m}{m^{\rho}-1} g_K(m), \tag{108}$$

$$\Lambda(m) \equiv \frac{\rho-(1/a)}{\rho-1}\frac{m^{\rho}-m}{m^{\rho}-1}\left(\frac{\rho-1}{m^{\rho}-m}-\frac{\rho+\kappa-1}{m^{\rho+\kappa-1}-m}\right) > 0 \ \ for\ m > 1. \quad (109)$$

Thus, equation (69) in the fixed coefficient model is changed to (100) when the factor substitutability is taken into account.

Solving (106) and (108) with respect to $\dot{m}/m$, we obtain the following dynamic equation for the firm size ratio:

$$\frac{\dot{m}}{m} = \frac{1}{\Delta(m)}[\Phi(m)g_K(m) - (n+g)], \quad (110)$$

where

$$\Delta(m) \equiv D(m;\lambda') - \Pi(m). \quad (111)$$

Here, $D(m;\lambda')$ represents the same expression as (71) except that $\lambda$ is replaced by $\lambda' = \lambda + \varepsilon$, and hence $D(m;\lambda') > 0$. It is assumed here that $\lambda' < 1$. On the other hand, $\Pi(m)$ represents the following expression, which is proved to be positive:[15]

$$\Pi(m) \equiv \varepsilon\left(\frac{\rho-1}{m^{\rho-1}-1}-\frac{\rho}{m^{\rho}-1}\right)+(\lambda'-\varepsilon)\Theta(m)+\Lambda(m) > 0 \ \ for\ m > 1. \quad (112)$$

The steady state equilibrium of (110) is attained at $m^*$ that satisfies

$$g_K(m^*) = \frac{n+g}{\Phi(m^*)}. \quad (113)$$

This equation is the same as (79) except that $g_X(m)$ is replaced by $g_K(m)$, and this function is defined by (104). Since $g_K'(m) > 0$, the steady state solution has qualitatively similar properties as in the fixed coefficient model.

How about the stability of this model. From (111), we have $\Delta(m) < 0$ if $D(m;\lambda') < \Pi(m)$. This condition is satisfied if the magnitudes of $\varepsilon$ and $\kappa$ are sufficiently large. From the definition of these parameters, $\varepsilon$ and $\kappa$ are the larger, the larger $1-\lambda-\omega$. In this case, the steady state equilibrium for (110) is locally stable, since $\dot{m}/m$ is decreasing with $m$ in the neighborhood of the equilibrium value of $m$ denoted by $m^*$. Figure 4 shows this stability property. Suppose that $m > m^*$ holds initially. Then, since $\dot{m}/m < 0$, $m$ will decrease over time, until it gets equal to $m^*$. Conversely, if $m < m^*$, then $\dot{m}/m > 0$, so that $m$ will increase over time until it reaches $m^*$. Thus, we find that the substitutability between capital and labor serves as a stabilizing factor.

The behavior of aggregate variables on the steady growth path in our model is similar to that in the Solow model. The steady growth of our model is consistent with Kaldor's 'stylized facts' as is the case in the Solow model. However, our model differs from the Solow model greatly in that it includes many structural parameters or variables of the economy, such as the Pareto coefficient, $\rho$, the learning coefficient, $\lambda$, the wage differential rate, $\omega$, or the size differential ratio, $m$. So our model gives more abundant implications about growth than the Solow model.

Let us examine how those parameters affect the steady state equilibrium. We first examine the effects of a change in $\rho$. In the steady growth equilibrium condition (113), it should be noted that both $g_K(m^*)$ and $\Phi(m^*)$ are functions of $\rho$. To take it into consideration explicitly, we may express those functions as $g_K(m^*;\rho)$ and $\Phi(m^*;\rho)$. It can be shown that $\partial g_K/\partial\rho < 0$ and $\partial\Phi/\partial\rho < 0$, though we omit the proof for saving space. Then, the effects of a change in $\rho$ on $m^*$ is derived from (113) as follows:

$$\frac{dm^*}{d\rho} = -\frac{(\frac{\partial g_K}{\partial\rho}/g_K) + (\frac{\partial\Phi_\rho}{\Phi}/\Phi)}{(g_K'/g_K) + (\Phi'/\Phi)} > 0, \tag{114}$$

where $g_K'$ and $\Phi'$ denote derivatives of $g_K$ and $\Phi$ with respect to $m^*$, both of them being positive. Thus, an increase in the Pareto coefficient, $\rho$, leads to an increase in the size differential ratio, $m$. This in turn leads to a decrease in the growth rate of a marginal firm, $\dot{x}_0/x_0$, in view of (82). Notice that $\rho$ is determined by $\rho = 1/a(1-\alpha)$. An increase in $\rho$ is caused by an increase in $\alpha$ or a decrease in $a$, where $\alpha$ is the rate of entry of new firms and $a$ represents the degree of size effects on the expected growth of firms. Thus, the higher the rate of entry of new firms, the greater will be the size differential ratio and the lower will be the growth rate of marginal firms. In this case, chances of smaller firms to survive will be greater. As for the parameter $a$, we assume that $a \geq 1$. The expected rate of growth of a firm is independent of size when $a = 1$, and it increases with size when $a > 1$. Thus, the smaller the value of $a$, the lower the degree of size effects on the expected growth of firms, and it results in the higher size differential ratio and the lower growth rate of marginal firms.

The effects of changes in $\lambda$ and $\omega$ are similarly examined. Notice that, from (104) and (72), $\partial g_K/\partial\lambda < 0$, $\partial\Phi/\partial\lambda > 0$, $\partial g_K/\partial\omega < 0$, and $\partial\Phi/\partial\omega = 0$. Then, it follows from (113) that

$$\frac{dm^*}{d\lambda} = -\frac{(\frac{\partial g_K}{\partial\lambda}/g_K) + (\frac{\partial\Phi}{\partial\lambda}/\Phi)}{(g_K'/g_K) + (\Phi'/\Phi)} \lesseqqgtr 0, \quad \frac{dm^*}{d\omega} = -\frac{\frac{\partial g_K}{\partial\omega}/g_K}{(g_K'/g_K) + (\Phi'/\Phi)} > 0. \tag{115}$$

In other words, the effect of an increase in $\lambda$ (*i.e.*, a decrease in learning by doing effects) on $m^*$ is indefinite, because it affects $g_K$ and $\Phi$ in the opposite direction. On the other hands, an increase in $\omega$ (*i.e.*, an increase in the wage differential with respect to size) leads to an increase in the size differential ratio, $m$, which will in turn lead to a decrease in the growth rate of a marginal firm, $\dot{x}_0/x_0$. Thus, an increase in the wage differential tends to benefit the survival of marginal firms.

It should be noted that changes in $\rho$ may not be independent of changes in $\lambda$ or $\omega$. In Proposition 4, we have shown that $a \gtreqqless 1$ depending on $\lambda + \omega \lesseqqgtr 1$. Thus, an increase in $\lambda$ or $\omega$ leads to a decrease in $a$, and hence to an increase in $\rho$. Therefore, though the effect of an increase in $\lambda$ on $m$ is indefinite by itself, it will most likely increase $m$ when its effect through $\rho$ is also taken into account. An increase in $\omega$ will lead to an increase in $m$ by itself and by its the effect through $\rho$ as well.

The effects of changes in these parameters on the growth rate of total output is indeterminate, because their direct effects and indirect effects through $m$ work in the opposite directions. Take a change in $\rho$ for example. An increase in $\rho$ decreases $\Phi(m^*)$ in (113) for given $m^*$, and hence the steady rate of growth. However, as is shown by (114), it increases $m^*$ at the same time, and hence increases $\Phi(m^*)$. As the result, the total effects of a change in $\rho$ on the steady rate of growth is indeterminate. To summarize, an increase in the Pareto coefficient, $\rho$, leads to a more equal size distribution of firms by benefiting smaller firms. But its effect on the steady rate of growth is uncertain.

Finally, it should be noted that a change in $\rho$, if it may happen, is a very slow process. An increase in the rate of entry of new firms, $\alpha$, does not change $\rho$ immediately, because most firms are incumbent firms. It will take so many years for new firms to grow to become larger firms.

# 7 Conclusion

The size distribution of firms in Japanese manufacturing, the data of which are available from the Census of Manufactures, exhibits a beautiful illustration of the Pareto law, not just for some particular years but over many years. This interesting empirical fact has motivated me to write this paper.

This paper makes two new attempts. First, I have explained why the Pareto law applies to the size distribution of firms based on the assumption that firms acquire the technology of operating efficiently on a larger scale through learning by doing and expand their scale of operation through accumulation of capital induced by profitability. Second, I set up a model of economic growth that is based on the Pareto distribution of firms and economies of scale.

It is found that the learning by doing hypothesis is quite suitable for explaining how the size distribution of firms tends to be the Pareto distribution. Especially, the coefficient for learning by doing effects, in addition to the birth rate of new firms, is found to be related to the Pareto coefficient. The derivation of the Pareto distribution, which owes to Simon (1955), is carried out by a stochastic process model. Thus, the Pareto distribution is attained through the evolutionary process of changes in position of individual firms. In this sense, our model is not an equilibrium model, but a kind of evolutionary model.

Based on this microeconomic foundation, we construct a model of economic growth which takes into consideration the size distribution of firms and economies of scale. In our model, firms with different size and different efficiency coexist, forming a regular pattern of distribution, namely, the Pareto distribution. The steady growth equilibrium of our model can consistently explain Kaldor's stylized facts, as the Solow model does. But, in contrast to the Solow model, the steady growth rate in our model depends not only on population growth and technological progress, but also on economies of scale generated by learning by doing. Our model has some similarities to endogenous growth models, but differs from them in that it take into account the size distribution of firms, and hence allows to analyze the effects of changes in the structural parameters.

# Notes

1. See also Simon and Bonini (1958) for some U.S. data.

2. Taking logarithm of equation (1) and estimating the regression for the data from Japanese manufacturing industry, we get the following result:

$$\log N = \underset{(0.06)}{6.38} - \underset{(0.027}{1.17} \log x \quad (R^2 = 0.995)$$

where the numerical value below each coefficient represents its standard error.

3. See Simon (1955) for such examples of the Pareto distribution.

4. Lydall (1998) criticizes the neo-classical theory of the firm form this point of view, and proposes an alternative theory. His ideas presented in his book is quite interesting, and this paper owes to his ideas. However, he does not present a formal model.

5. The form of the learning function assumed here is the same as Arrow's except that we take capacity output of an individual firm as an index of experience of the firm. It may be better to adopt cumulative output rather than current capacity output as an index of experience. However, as will be discussed later, we assume that every firm starts their operation from the minimum size and expands its scale of operation through learning by doing. Under this assumption together with the assumption that output of a typical firm is increasing exponentially, current capacity output of the firm is proportional to cumulative output. Therefore, we can use the former variable as an index of experience.

6. Taking logarithm of these equations and estimating for the data from Japanese manufacturing industry, we get the following results:

$$\log l = \underset{(0.02)}{-0.51} + \underset{(0.006)}{0.83} \log x \quad (R^2 = 0.999),$$

$$\log k = \underset{(0.03)}{-0.13} + \underset{(0.009)}{1.005} \log x \quad (R^2 = 0.999),$$

where the new merical value below each coefficient represents its standard error. These results show that $\lambda = 0.83$ and $\mu = 1.005$ in the case of Japanse manufacturing industry, and hence our assumption that $0 < \lambda < 1$ and $\mu \geq 1$ may be justified. Moreover, the second regeression equation shows that the value of $\mu$ is approximately equal to 1. The assumption $\mu = 1$, which will be made later in our model, may also be justified in the case of Japanese manufacturing industry by this result.

7. Regression of this equation to the data of Japanese manufacturing industry gives the following result:

$$\log w = \underset{(0.01)}{0.36} + \underset{(0.04)}{0.09} \log x \quad (R^2 = 0.978)$$

where numerical value below each coefficient represdents its standard error. This result show that the positve relation between the wage rate and the size of firms is statistically significant in the case of Japanese manufacturing industry.

8. See note 6.

9. In the case of Japanese manufacturing industry, $\lambda+\omega = 0.83+0.09 = 0.91$, in view of notes 6 and 7. Hence, to assume $\lambda + \omega \leq 1$ may be realistic.

10. This theory of income distribution was first presented by Lydall (1971). The comparison of this theory with other theories of distribution is made in detail by Lydall (1979).

11. See Appendix (A) for the proof.

12. In recent macoroeconomics, more orthodox approach to the determination of saving is to assume that the households maximize lifetime utility. As Solow (2000) argues, however, we lose very little, from the steady-state point of view, by adopting the assumption of constant savings rates for profits and wages. Moreover, this assumption is more consistent with the model of firms in Section 3 in which we assumed the existence of retained profits.

13. See Appendix (B) for the proof.

14. These are easily proved from (B.3) in Appendix (B).

15. It is easlily proved from (B.3) in Appedix (B) that functions $\Theta(m)$, $\Lambda(m)$ and $\Pi(m)$ are all positive.

# Appendix

## (A) Mathematical Notes to Section 4

We give here mathematical proofs to (55) in the text. The share of wages in the value added is determined by (54), that is,

$$S_W = \frac{\rho - 1}{\beta(\rho - \lambda - \omega)} \frac{1 - m^{\lambda+\omega-\rho}}{1 - m^{1-\rho}}, \qquad ((A.1))$$

where it is assumed that $\rho > 1 > \lambda + \omega$. Under this assumption, we prove that (55) holds for $m > 1$.

**Proposition A1.** $\partial S_w/\partial\rho > 0$, *if* $\rho > 1 > \lambda + \omega$ *and* $m > 1$.

*proof* Taking logarithm of (A.1) and differentiating it with respect to $\rho$, we have

$$\begin{aligned}\frac{1}{S_w}\frac{\partial S_w}{\partial\rho} &= \frac{1}{\rho-1} - \frac{1}{\rho-\lambda-\omega} + \frac{\log m}{m^{\rho-\lambda-\omega}-1} - \frac{\log m}{m^{\rho-1}-1} \\ &= \frac{m^{\rho-1}[(1-\lambda-\omega)(m^{\rho-\lambda-\omega}-1)(1-m^{1-\rho}) + (\rho-1)(\rho-\lambda-\omega)(1-m^{1-\lambda-\omega})\log m]}{(\rho-1)(\rho-\lambda-\omega)(m^{\rho-\lambda-\omega}-1)(m^{\rho-1}-1)}\end{aligned}$$

The denominator of this last expression is positive since $\rho > 1 > \lambda + \omega$ and $m > 1$. To examine the numerator, we put the expression in [ ] of the numerator as $f(m)$:

$$f(m) = (1-\lambda-\omega)(m^{\rho-\lambda-\omega}-1)(1-m^{1-\rho}) + (\rho-1)(\rho-\lambda-\omega)(1-m^{1-\lambda-\omega})\log m.$$

In order to prove $f(m) > 0$ for $m > 1$, it is necessary and sufficient to prove $f(1) \geq 0$ and $f'(m) > 0$ for $m > 1$. Straightforwardly, we have $f(1) = 0$. The derivative $f(m)$ becomes, after some arrangement , as follows:

$$\begin{aligned}f'(m) &= m^{-\lambda-\omega}[(1-\lambda-\omega)\{(\rho-\lambda-\omega)(m^{\rho-1}-1) + (\rho-1)(1-m^{\lambda+\omega-\rho})\} \\ &\quad -(\rho-1)(\rho-\lambda-\omega)\{(1-\lambda-\omega)\log m - (1-m^{1-\lambda-\omega})\}].\end{aligned}$$

To prove the expression in [ ] of the right hand side of this equation to be positive, let us put it as $g(m)$. Then, we have only to prove $g(1) \geq 0$ and $g'(m) > 0$ for $m > 1$. It is trivial that $g(1) = 0$. Taking the derivative of $g(m)$ and arranging it, we have

$$g'(m) = (\rho-1)(\rho-\lambda-\omega)(1-\lambda-\omega)m^{\rho-2}(1+m^{1-\lambda-\omega})(1-m^{\lambda+\omega-\rho}).$$

Under the assumption that $\rho > 1 > \lambda + \omega$, we have $g'(m) > 0$ for $m > 1$. This concludes the proof.

**Proposition A2.** $\partial S_w/\partial\lambda > 0$ *if* $\rho > 1 > \lambda + \omega$ *and* $m > 1$.

*Proof* Taking the logarithmic derivative of (A.1) with respect to $\lambda$, we obtain, after some arrangement, the following:

$$\begin{aligned}\frac{1}{S_w}\frac{\partial S_w}{\partial\lambda} &= \frac{1}{\rho-\lambda-\omega} - \frac{\log m}{m^{\rho-\lambda-\omega}-1} \\ &= \frac{m^{\rho-\lambda-\omega} - 1 - (\rho-\lambda-\omega)\log m}{(\rho-\lambda-\omega)(m^{\rho-\lambda-\omega}-1)}.\end{aligned}$$

The denominator of this last expression is positive, since $\rho > \lambda+\omega$. To examine the numerator, let us put it as $f(m)$:

$$f(m) = m^{\rho-\lambda-\omega} - 1 - (\rho-\lambda-\omega)\log m.$$

Then, we have $f(1) = 0$. Taking the derivative of this function, we obtain

$$f'(m) = m^{-1}(\rho-\lambda-\omega)(m^{\rho-\lambda-\omega} - 1).$$

This expression is positive for $m > 1$, since $\rho > \lambda+\omega$. Hence, $f(m) > 0$, and the proposition is proved.

**Proposition A3.** $\partial S_w/\partial\omega > 0$ *if* $\rho > 1 > \lambda+\omega$ *and* $m > 1$.

*Proof* The same procedure as the proof of Proposition 2 can be applied.

**Proposition A4.** $\partial S_w/\partial\beta < 0$ *if* $\rho > 1 \geq \lambda+\omega$ *and* $m > 1$.

*Proof* Trivial.

**Proposition A5.** $\partial S_w/\partial m < 0$ *if* $\rho > 1 > \lambda+\omega$ *and* $m > 1$.

*Proof* Taking the logarithmic derivative of (A.1) with respect to $m$, we obtain, after some arrangement, the following:

$$\begin{aligned}\frac{1}{S_w}\frac{\partial S_w}{\partial m} &= \frac{\rho-\lambda-\omega}{m(m^{\rho-\lambda-\omega}-1)} - \frac{\rho-1}{m(m^{\rho-1}-1)} \\ &= \frac{(\rho-\lambda-\omega)(m^{\rho-1}-1) - (\rho-1)(m^{\rho-\lambda-\omega}-1)}{m(m^{\rho-\lambda-\omega}-1)(m^{\rho-1}-1)}\end{aligned}$$

The denominator of this last expression is positive, since $\rho > 1 > \lambda+\omega$ and $m > 1$. To examine the numerator, we put it as $f(m)$:

$$f(m) = (\rho-\lambda-\omega)(m^{\rho-1}-1) - (\rho-1)(m^{\rho-\lambda-\omega}-1).$$

Then, we have $f(1) = 0$. Taking the derivative of this function we obtain

$$f'(m) = (\rho-\lambda-\omega)(\rho-1)m^{\rho-2}(1 - m^{1-\lambda-\omega}).$$

This expression is negative for $m > 1$, since $\rho > 1 > \lambda+\omega$. Hence, $f(m) < 0$, and the proposition is proved.

### (B) Mathematical Notes to Section 5

(1) Properties of function $D(m)$ defined by (71) in the text.

Function $D(m)$ may be rewritten as follows defined as follows:

$$D(m) \equiv \lambda\left[\left(\frac{\rho-1}{m^{\rho-1}-1} - \frac{\rho}{m^{\rho}-1}\right) - \frac{1}{\lambda}\left(\frac{\rho-\lambda}{m^{\rho-\lambda}-1} - \frac{\rho}{m^{\rho}-1}\right)\right]. \quad ((B.1))$$

To examine the properties of this function, let us define function $f(m,x)$ as follows:

$$f(m,x) = \frac{1}{x}\left(\frac{\rho-x}{m^{\rho-x}-1} - \frac{\rho}{m^{\rho}-1}\right), \quad \textit{where } \rho > 1 \geq x > 0 \textit{ and } m > 1.$$

((B.2))

It is easy to prove that

$$\frac{\rho - x}{m^{\rho-x} - 1} - \frac{\rho}{m^{\rho} - 1} > 0, \quad if\ \rho > 1 \geq x > 0\ and\ m > 1. \tag{(B.3)}$$

Hence, we have $f(m,x) > 0$. Then, $D(m)$ is positve or negative depending on whether $f_x(m,x)$ is positive or negative. We first prove the following lemma.

**Lemma B1.** *There exists some value of $m$ denoted by $\bar{m}$, which is larger than 1, such that $f_x(m,x) \gtreqless 0$ depending on whether $m \gtreqless \bar{m}$.*

*Proof* Function $f(m,x)$ may be rewritten as follows:

$$f(m,x) = \frac{(\rho - x)(m^{\rho} - 1) - \rho(m^{\rho-x} - 1)}{x(m^{\rho-x} - 1)(m^{\rho} - 1)}.$$

Taking logarithmic derivative of this function, we obtain, after some arragement,

$$\frac{f_x(m,x)}{f(m,x)} = \frac{m^{\rho-x}[x(\rho - x)(m^{\rho} - 1)\log m - \rho(m^{x} - 1)]}{x[(\rho - x)(m^{\rho} - 1) - \rho(m^{\rho-x} - 1)]} \tag{(B.4)}$$

The denominator of the right hand side of this equation is positve. Hence, this expression is positive or negative depending on whether the numerator is positive or negative. We thus obtain

$$\frac{f_x(m,x)}{f(m,x)} \gtreqless 0 \ \ depending\ on\ whether\ \ \frac{(m^{\rho} - 1)\log m}{m^{x} - 1} \gtreqless \frac{\rho}{x(\rho - x)}. \tag{(B.5)}$$

Let us put

$$g(m) \equiv \frac{(m^{\rho} - 1)\log m}{m^{x} - 1} \tag{(B.6)}$$

Then, since $\rho > x$, this function is shown to have the following properties:

$$g'(m) > 0 \ \ for\ m > 1, \quad \lim_{m \to 1} g(m) = 0, \quad \lim_{m \to \infty} g(m) = \infty.$$

Therefore, there exists one and only one value of $m$, denoted by $\bar{m}$, such that

$$\frac{(\bar{m}^{\rho} - 1)\log \bar{m}}{\bar{m}^{x} - 1} = \frac{\rho}{x(\rho - x)}. \tag{(B.7)}$$

Then, in view of (B.4), we have $f_x(m,x) \gtreqless 0$ depending on whether $m \gtreqless \bar{m}$. This completes the proof.

**Proposition B1.** *There exists some value of $m$ denoted by $\bar{m}$, which is larger than 1, such that $D(m) \gtreqless 0$ depending on whether $m \gtreqless \bar{m}$.*

*Proof* By the above lemma, if $m > \bar{m}$ holds, $f(m,x)$ is an increasing function with respect to $m$. Hence, $f(m,1) > f(m,\lambda)$, since $\lambda < 1$. Then, we have $D(m) > 0$, in view of (B.1). The opposite holds if $m < \bar{m}$. This concludes the proof.

**A Note on Proposition B1.** An important problem is whether $D(m)$ is positive or negative for the economically range of $m$. To examine this problem, let us consider the case $\rho = 1.2$ and $\lambda = 0.8$, which appropriate for Japanese manufacturing industry as a whole. Then, it is caluclated that $\bar{m} \fallingdotseq 18$. In other words, if the ratio of the largest firm's output to the smallest firm's output exceeds 18, then we have $D(m) > 0$. Obviously, this number of $\bar{m}$ is quite small compared to the actually observed number. Therefore, we may safely assume that $D(m) > 0$ holds in reality.

# References


Aghion, P. and P. Howitt (1992), "A Model of Growth through Creative Destruction," *Econometrica*, Vol.60, pp.325-51.

Aghion, P. and P. Howitt (1998), *Endogenous Growth Theory*, Cambridge, Mass.: The MIT Press.

Champernowne, D. (1953), "A Model Income Distribution," *Economic Journal*, Vol.63, pp.318-51.

Ijiri, Y. and H.A. Simon (1964), "Business Firm Growth and Size," *American Economic Review*, Vol.54, pp.77-89.

Ijiri, Y. and H.A. Simon (1971), "Effects of Mergers and Acquisitions on Business Firm Concentration," *Journal of Political Economy*, Vol.79. pp.314-22.

Ijiri, Y. and H.A. Simon (1974), "Interpretation of Departure from Pareto Curve Firm-Size Distribution," *Journal of Political Economy*, Vol.82, pp.315-32.

Lydall, H. (1971), "A Theory of Distribution and Growth with Economies of Scale," *Economic Journal*, Vol.81, pp.91-112.

Lydall, H. (1979), *A Theory of Income Distribution*, Oxford: Clarendon Press.

Lydall, H. (1998), *A Critique of the Orthodox Economics*, London: Macmillan.

Mandelbrot, B. (1961), "Stable Paretian Random Functions and the Multiplicative Variation of Income," Econometrica, Vol.29, pp. 517-43.

Romer, P.M. (1986), "Increasing Returns and the Long-Run Growth," *Journal of Political Economy*, Vol.94, pp.1002-37.

Romer, P.M. (1990), "Endogenous Technological Change," *Journal of Political Economy*, Vol.98, pp.71-102.

Sato, K. (1970), "Size, Growth, and Skew Distribution," *Discussion Paper*, No.145, SUNNY at Baffalo

Simon, H. A. (1955), "On a Class of Skew Distribution," *Biometrica*, Vol.82, pp.425-40, reprinted in Simon, Models of Man (New York: John Wiley, 1957), pp.145-64.

Simon, H.A. and Bonini, C.P. (1958), "The Size Distribution of Business Firms," *American Economic Review*, Vol.48, pp.607-17.

Solow, R.M. (2000), *Growth Theory: An Exposition*, Second Edition, Oxford: Oxford University Press.

Steindl, J. (1965), *Random Processes and the Growth of Firms*, London: Griffin.

Wold, O.A. and P. Whittle (1957), "A Model of Explaining the Pareto Distribution of Wealth," *Econometrica*, Vol. 25, pp.591-95.

Figure 1. Perfectly Competitive Equilibrium of the Firm

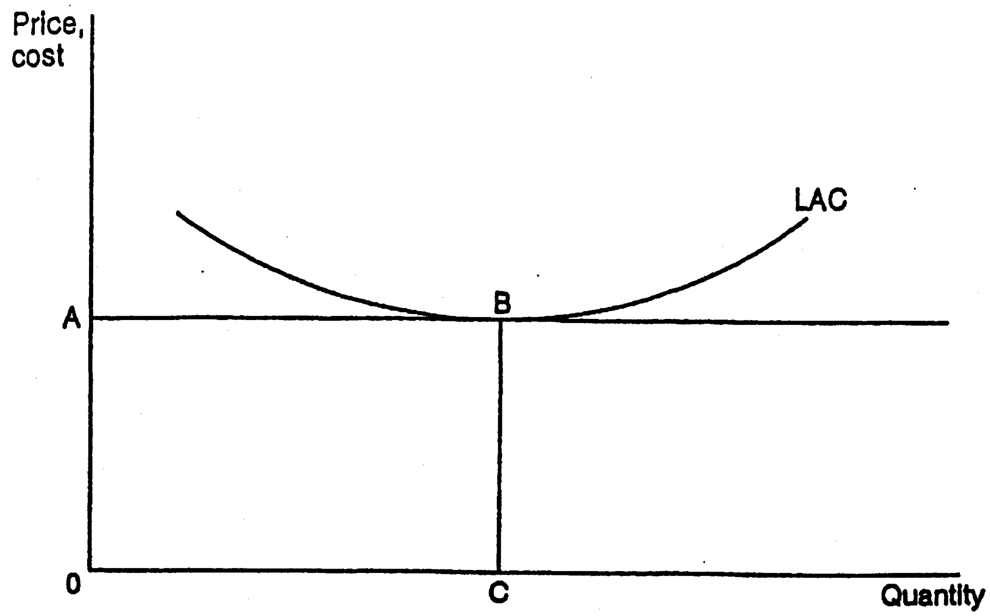
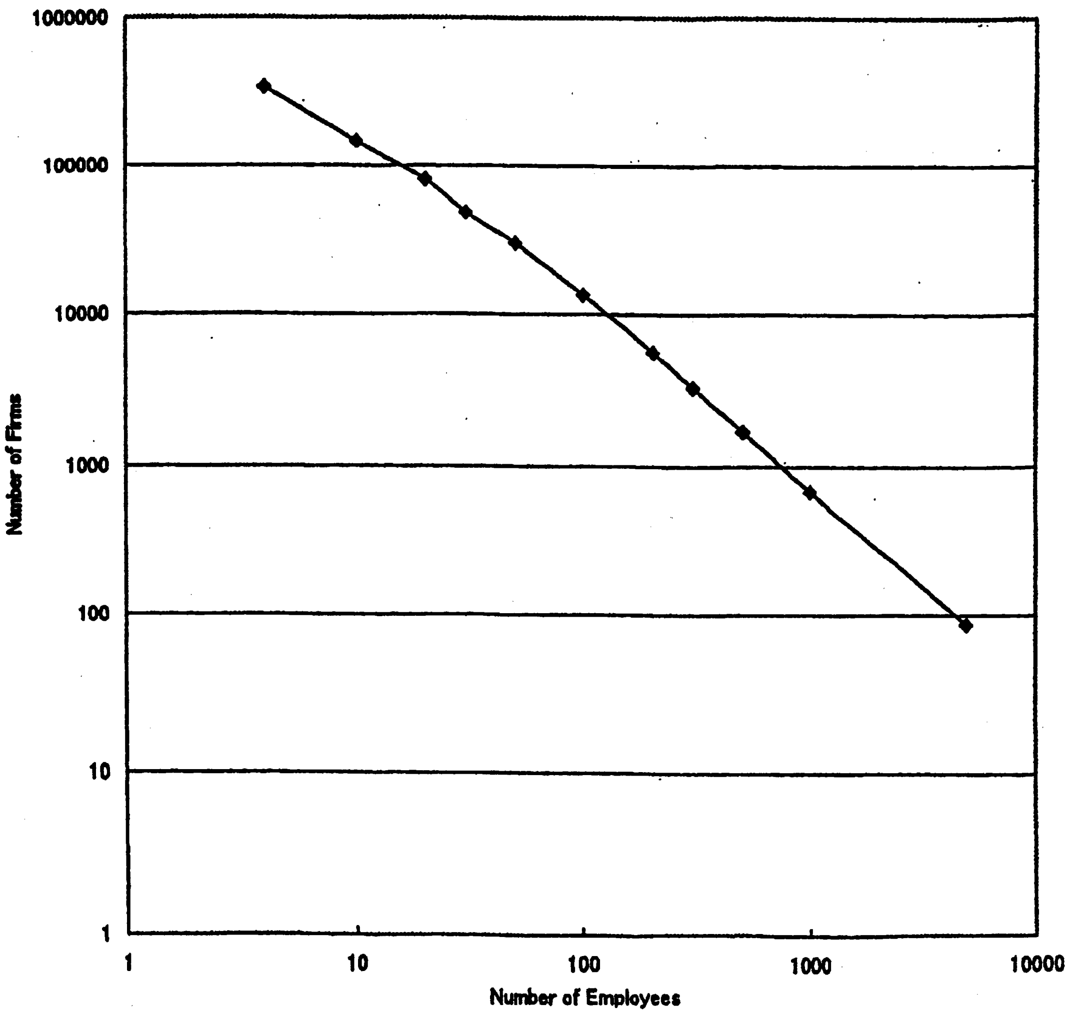



Figure 2. The Size Distribution of Firms in Japanese Manufacturing Industry




Source: Census of Manufactures, 1998.
(Ministry of International Trade and Industry)

Figure 3. Instability of the Steady Growth Equilibrium

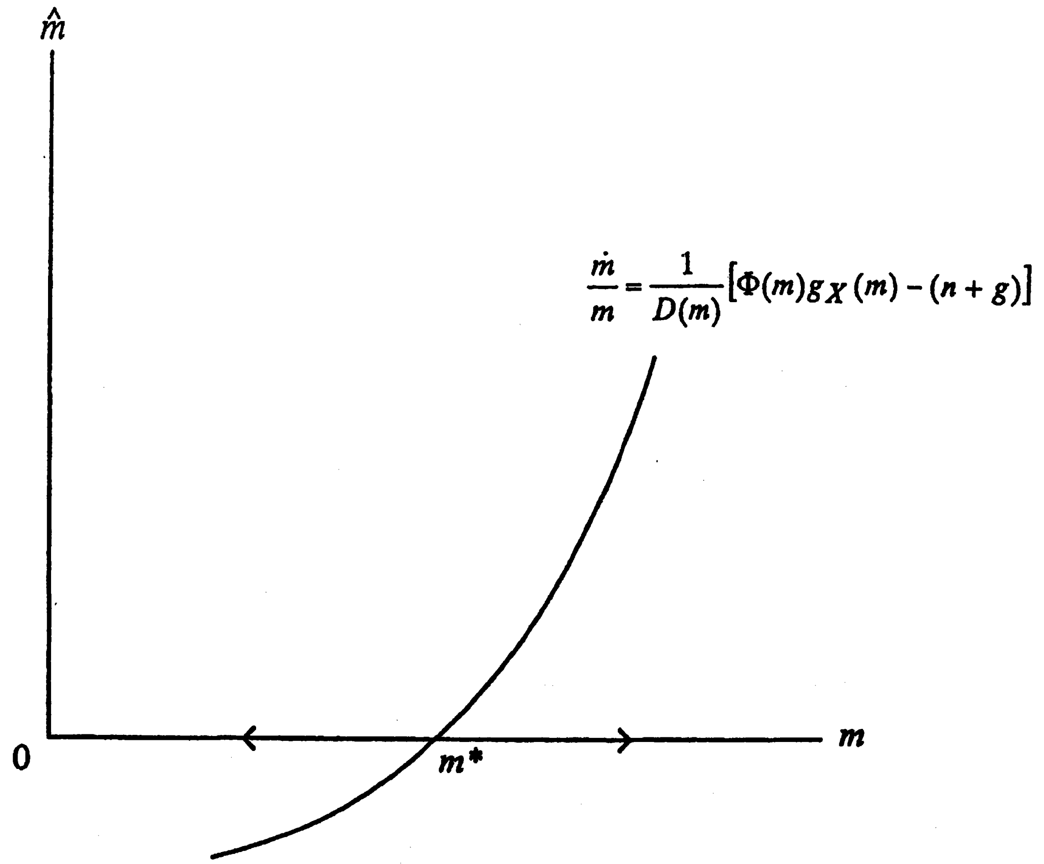



**Figure 4. Stability of the Steady Growth Equilibrium**