Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

OECD
OCDE
PARIS

ECONOMICS DEPARTMENT

Cancels & replaces the same document:
distributed 06-Aug-1999

**TECHNOLOGY UPGRADING WITH LEARNING COST :
A SOLUTION FOR TWO "PRODUCTIVITY PUZZLES"**

ECONOMICS DEPARTMENT WORKING PAPERS N0. 220

by
Sanghoon Ahn

80355

ABSTRACT/RÉSUMÉ

As it takes time and effort to learn how to fully utilise new technology and realise its maximum potential productivity gain, adoption of new technology tends to reduce productivity temporarily, even though the potential productivity gain in the long run outweighs this short run loss. This paper points to such "learning cost" in technology upgrading as a potential explanation of the following two "productivity puzzles" reported in the Information Technology (IT) literature and in the studies of East Asian economic growth. First, in the 1980s, US companies made enormous IT investments, but little productivity gain was observed. Second, Total Factor Productivity (TFP) growth of "Newly Industrialising Countries" (NICs) in East Asia was mediocre in spite of the impressive investment drive in those countries.

A simple model of optimal intervals for technology upgrading with learning cost is developed. This model predicts that a company with higher frequency in technology upgrading will tend to have higher market value even with lower current profitability. An empirical study using unbalanced panel data of 1,031 US companies from 1986 to 1995 supports this prediction. Extending the scope from firm-level to industry-level, the paper estimates the magnitude of industry-wide learning-by-doing effects using annual data on 15 sub-industries in the Japanese machinery manufacturing sector from 1955 to 1990. The results show that industry-wide learning-by-doing was strong in low-tech industries where technological change was relatively slow, while it was insignificant in high-tech industries which experienced rapid technological evolution. It is also observed in the US and Japanese manufacturing industries that TFP growth tends to decrease with faster capital accumulation. This negative correlation is reproduced in simulations based on the extended model of learning cost.

JEL classification: O30; O47; D24; F43
Keywords: Technology; Learning; IT investment; Total factor productivity.

*****

Apprendre à pleinement utiliser une nouvelle technologie et achever son gain potentiel de productivité demandent du temps et des efforts. Pour ces raisons, l'adoption d'une nouvelle technologie tend à baisser temporairement la productivité même si le gain de productivité à long terme dépasse cette perte à court terme. Cette étude présente l'existence des "coûts d'apprentissage" comme une explication potentielle à deux paradoxes de la productivité relevés dans les études concernant les technologies informatiques ainsi que celles portant sur la croissance économique en Asie de l'Est. Premier paradoxe: en dépit des investissements énormes par les firmes américaines pendant les années 80, il en a résulté peu de gain de productivité. Second paradoxe: la croissance de la productivité totale des facteurs (TFP) dans les nouveaux pays industrialisés (NPI) de l'Asie de l'Est a été médiocre malgré leur taux d'investissements exceptionnellement élevé.

On modélise, dans cette étude, l'adoption de nouvelles technologies sous l'hypothèse des coûts d'apprentissage, et on obtient, dans ce cadre, son intervalle optimal. Le modèle prévoit que l'adoption technologique plus fréquente tend à augmenter la valeur de marché des firmes même si elle baisse la rentabilité courante. A l'aide des données de panel sur 1031 firmes américaines pour la période 1986-95, une étude économétrique nous permet de confirmer la prévision du modèle. Pour étendre l'analyse, on estime aussi l'amplitude des effets de l'apprentissage par pratique (learning by doing) au niveau des branches industrielles. Pour ce faire, on utilise les données annuelles (1955-1990) pour 15 branches japonaises fabriquant des machines. Les résultats montrent que l'apprentissage par pratique est plus important pour le gain de productivité dans les industries à basse technologie où les changements techniques sont relativement lents, alors qu'il occupe une place non significative dans les industries à haute technologie ayant connu une évolution technologique rapide. Par ailleurs, on trouve aussi que le taux de croissance de TFP tend à diminuer avec une accumulation de capital plus rapide dans les industries manufacturières américaines et japonaises. Cette corrélation négative est aussi reproduite par les simulations basées sur le modèle élargi des coûts d'apprentissage.

Classification : JEL: O30, O47, D24, F43
Mots-clés : technologie, apprentissage, investissement dans les technologies informatiques, productivité total des facteurs.

TABLE OF CONTENTS

Tables

Figures

TECHNOLOGY UPGRADING WITH LEARNING COST:
A SOLUTION FOR TWO "PRODUCTIVITY PUZZLES"

Sanghoon Ahn[1]

1. Introduction

Recently, two "productivity puzzles" drew the attention of researchers in the Information Technology (IT) literature and in the studies of East Asian economic growth. The one was the slowdown in the US productivity growth during the 1980s in spite of the massive investment boom in IT capital.[2] The other was the unimpressive productivity growth from the impressive investment drive of "Newly Industrialising Countries" (NICs) in East Asia.[3] These two cases seem to have interesting common aspects. In both cases, there was a massive investment in sectors related to more advanced technology, but the return on such investments was marginal in terms of productivity. This paper points to "learning cost" in technology upgrading as a potential explanation of these two "puzzles".

In adopting a new technology, one must acquire a set of necessary skills and know-how in order to fully utilise the new technology and realise its maximum potential productivity gain. This will involve acquiring a basic and then advanced understanding of the technology as well as gaining experience with the application of the technology in a particular business / industry context. Because no technology can be implemented instantaneously, it takes time and resources for the potential productivity of the new technology to be fully realised. Such time and effort spent in adopting a new technology should be included as part of the cost for technology adoption. I will call it "learning cost".

Consider, for example, upgrading from a typewriter to word processor. One should buy not only a personal computer and word processing software but also a printer, floppy diskettes, etc. One might have to re-arranging office in order to accommodate the personal computer and related equipment and may have to dispose of the old typewriter. After these initial efforts, one will have to acquire the basic skills necessary to operate a personal computer. And yet, one will still have to become familiar with the word

---

2. For example, see the beginning paragraph of Oliner and Sichel (1994):

> During the past 15 years, U.S. companies have poured billions of dollars into information technology. Yet, through the 1980s, many observers argue that these companies were not getting their money's worth. As hard as analysts scoured the numbers, they could not show that computing equipment contributed much to productivity growth, leading to Robert Solow's famous quip that "you can see the computer age everywhere but in the productivity statistics."

3. See Young (1995). He concludes:

> While the growth of output and manufacturing exports in the newly industrializing countries of East Asia is virtually unprecedented, the growth of total factor productivity in these economies is not.

processing software. During the frustrating transition period, one might even lose a whole day's work simply by making a computer error. It is not surprising that one's productivity with the new word processor could be initially lower than that with the old typewriter.

Learning cost is relevant at the organisation level as well. Consider a factory or a company which has adopted a new technology. Just as with individual learning, it takes time and effort for an organisation to learn how to fully utilise a new technology and realise any productivity growth. For example, in order to most effectively implement a new technology, a new organisational structure might be called for. Of course, restructuring an organisation requires tremendous time and effort. This restructuring can also be regarded as a part of organisational learning.

Figure 1. Learning cost in adopting a new technology
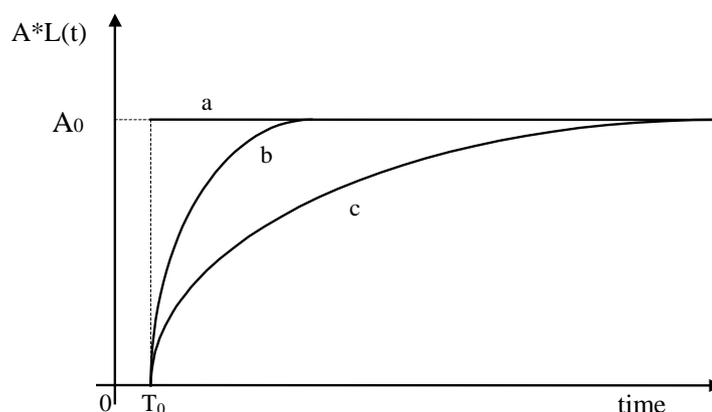
Figure 1 illustrates the idea of learning cost as it relates to adopting a new technology. In Figure 1, the horizontal axis shows time and the vertical axis productivity. Suppose that a new technology with a potential productivity A_0 was adopted at time T_0. Case a shows the case where there is no learning cost. In this case, the potential productivity of the new technology is fully realised at the moment when it is installed. Usually, however, it takes time and resources to assimilate a new technology and realise its benefits. Cases b and c illustrate situations that include such learning costs. In the figure, case c incurs higher learning costs than case b.

Because there are learning costs associated with the adoption of new technology, its implementation tends to reduce productivity in the short term, even though the potential productivity gain in the long run outweighs this short term loss. New technologies incur both tangible and intangible costs. The tangible costs are those associated with purchasing the new technology itself, e.g. capital equipment, computer hardware, software. However these specific costs are only one part of the total cost of purchasing a new technology. There are also significant learning and adjustment costs associated with implementing a new technology. For instance Compaq Computer Corporation, the world's largest personal computer manufacturer, estimates that the initial purchase price of a personal computer represents only about 20% of the total cost of owning and operating that PC in a corporate network environment. This is direct replacement cost, assuming no other environmental changes such as changing software, etc. With such substantial learning costs, upgrading technology will usually imply a short run loss and long run gain. Fundamentally, adopting new technology is similar to investing in physical capital in the sense that it requires short term expenditures that offer long term returns when the technology is appropriate and implemented and utilised successfully.

Figure 2. Effects of adopting new technology at time T_0
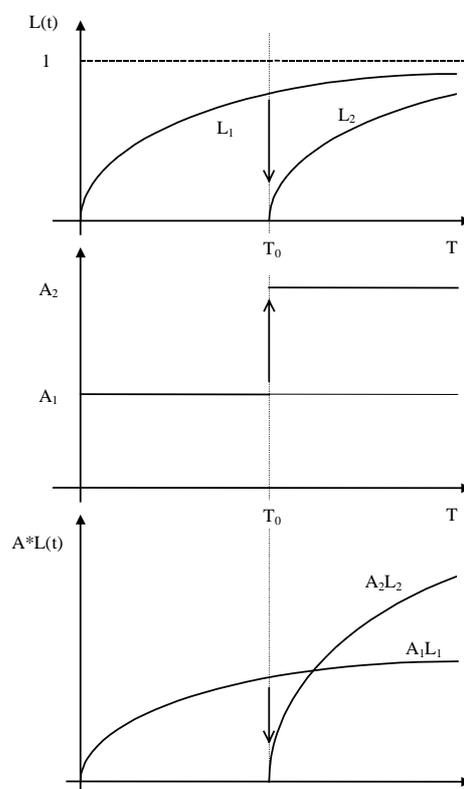
Figure 2 illustrates this idea in a simple context. In the first and the second panels, the effects of upgrading technology are separated into two factors: level effect and learning effect. The first panel shows two learning curves L_1 and L_2 with different starting points in time. Both start from 0 and converges to 1 over time. The second panel shows potential productivity levels of the old and the new technology as A_1 and A_2. Consider a technology upgrading at time t_0. This upgrade increases potential productivity from A_1 to A_2, but it also requires a shift to the new learning curve, L_2. The combined effect of the technology upgrade is depicted in the third panel. At the moment of the upgrade at time t_0, the productivity of the new technology $A_2 L_2$ is lower than that of the old one $A_1 L_1$.[4] But, even with this dramatic initial fall in productivity, it is not necessarily cost effective to continue utilising the old technology. The productivity of the new technology $A_2 L_2$ increases faster and catches up with that of the old technology $A_1 L_1$ over time.

The idea of learning cost in technology upgrading can also be extended and applied to issues affecting industries or economies in aggregate. Brezis, Krugman, and Tsiddon (1993), for example, tried to explain technological leapfrogging in international trade based on this idea:

---

4.   This idea of initial inferiority of a new technology with potential superiority was succinctly summarised in Young (1993) as follows:

> [Existing] models of invention make the surprising assumption that new technologies attain their full productive potential at the moment of their invention and are, at that point in time, superior (or at least equal) to the older technologies for which they substitute. The history of technical change suggests, however, that most new technologies are, in fact, initially broadly inferior to the older technologies they seek to replace and are actually competitive in only a narrow range of specialised functions. Incremental improvements over time, however, allow new technologies to ultimately dominate older systems of production across a wide variety of activities.

[...] at times of a new invention or a major technological breakthrough, economic leadership, since it also implies high wages, can deter the adoption of new ideas in the most advanced countries. A new technology may well seem initially inferior to older methods to those who have extensive experience with those older methods; yet that initially inferior technology may well have more potential for improvements and adaptation. When technological progress takes this form, economic leadership will tend to be the source of its own downfall.

As the authors admitted, however, their paper was based on a restrictive assumption that individuals are too myopic to consider the long run benefits of adopting a new technology when they choose between the new technology and the existing one. In this paper, I will explicitly consider optimisation behaviour that weighs short run costs of technology upgrading against its long run benefits.

In Section 2, I develop a simple model which helps determine optimal intervals for technology upgrading in this standard scenario that includes both tangible and intangible costs. The optimal timing for upgrading technology is considered in terms of maximising the net present value of future output. As an application of this model, I provide an explanation for the seemingly puzzling phenomena: the coexistence of high market value and low current profitability in IT intensive companies. This explanation is supported by empirical studies with a firm-level data set of 1,031 US companies from 1986 to 1995 in Section 3. The scope is extended from firm-level to industry-level in later sections. Section 4 estimates the magnitude of industry-wide learning-by-doing effects using annual data on 15 sub-industries in the Japanese machinery manufacturing sector from 1955 to 1990. The results show that industry-wide learning-by-doing was strong in low-tech industries where technological change was relatively slow, while it was insignificant in high-tech industries which were experiencing rapid technological evolution. Section 5 presents a finding that total factor productivity (TFP) growth tends to decrease when capital accumulation becomes faster. This negative correlation between the movement of the capital growth rate and that of the productivity growth rate is very widely observed. Moreover, this basic pattern is not affected by various methods for measuring capital stock, or varying levels of aggregation. This paper argues that this intriguing pattern is explained by the idea of learning cost in installing new capital better than by other competing stories. To support this argument, Section 6 extends the simple model of learning cost in technology upgrading, presented in Section 2, from the individual producer's level to the aggregated industry level. Simulations based on this extended model of learning cost reproduce the observed negative correlation between the capital growth rate and the TFP growth rate. This negative correlation, which is widely observed in industry-level data and reproduced in simulations based on the learning cost model, will make mediocre TFP growth in East Asian NICs look much less puzzling.

2.   Technology upgrading with learning cost

In this section, I build a model which formalises the idea of learning costs associated with adopting technology in a very simple setting. Predictions of this model will be tested with firm-level data in the next section.

2.1   Setup of the model

This is a minimalist model which focuses on the issue of productivity change caused by upgrading technology. The only choice variable in this simplistic model is the timing of technology upgrading. The producer is chooses the optimal timing for upgrading technology so as to maximise the net present value of the output flow.

Assumption 2.1 (Production) (1) Labour is the only input for production and it is fixed. Specifically, I normalise labour input as 1. By construction, output is equal to productivity. (2) The producer cannot

employ more than one type of technology at the same time. Therefore, the producer upgrades technology by discarding the current one and adopting a new technology with higher potential productivity, as better and better technology becomes available.

Improvements in technology allow the producer to adopt newer technologies with higher potential productivity. These upgrades, however, incur both tangible and intangible costs. Tangible costs are the costs associated with obtaining the new technology. Intangible costs arise from the fact that it takes time and resources to implement the new technology and to see its utilisation reach its full potential. I will label these "upgrade cost" and "learning cost", respectively. Following two assumptions with specific functional forms makes the model very simple and tractable without sacrificing the spirit of the model.

Assumption 2.2 (Upgrading technology and upgrade cost) (1) Upgrading technology allows the producer to adopt the most up-to-date technology that has the highest available productivity potential. When technology is upgraded, potential productivity jumps to the frontier level which increases at a constant growth rate, α. (2) Each upgrade incurs cost associated with the upgrade that is proportional to the potential productivity of the new technology with a ratio of β.[5]

Assumption 2.3 (Learning a new technology and learning cost) (1) When a technology is upgraded, experience with the existing technology is not transferred to the new technology.[6] (2) Within each generation of technology, productivity increases over time starting from zero and converging to its potential level with a decelerating growth rate. Specifically, the convergence of productivity to potential takes the following functional form with a parameter γ : $L(t) = 1 - e^{-\gamma t}$.[7]

From Assumption 2.1-2.3, output at time t under the n-th generation of technology is the following:

$$y_t = A_0 \, e^{\alpha t_{n-1}} \, (1 - e^{-\gamma(t-t_{n-1})}) \quad \text{where} \quad t \in [\, t_{n-1}, t_n \,) \tag{2.1}$$

Without loss of generality, we can change the unit so that A_0 equals 1.

$$y_t = e^{\alpha t_{n-1}} \, (1 - e^{-\gamma(t-t_{n-1})}) \quad \text{where} \quad t \in [\, t_{n-1}, t_n \,) \tag{2.1'}$$

Now the producer has to choose the sequence of optimal upgrade timings $\{t_1^*, t_2^*, t_3^*, t_4^*, ...\}$ which maximises the net present value. (See Figure 3)

---

5.   This proportional upgrading cost assumption allows us to have evenly spaced optimal timing and makes the calculation much simpler. (See Proposition 2.1) In case we have fixed upgrading cost, the optimal time interval of upgrading gets shorter and shorter as potential productivity increases.

6.   We can relax this strong assumption by allowing partial transfer of experience. But, again, this strong assumption makes the calculation much simpler without sacrificing the spirit of the model.

7.   Notice that large γ represents low learning cost while large β means high upgrading cost.

Figure 3. Repeated technology upgrades and productivity
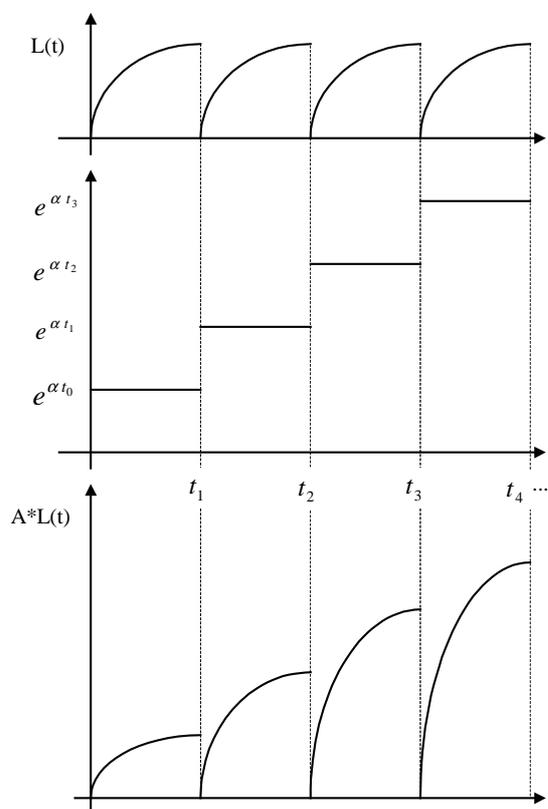
In calculating the net present value, we assume that the discount rate is given and constant for simplicity.

Assumption 2.4 (Discount rate) (1) The discount rate is given as ρ. (2) The discounted present value of output is well defined with this discount rate. To be specific, the growth rate of potential productivity from the evolution of frontier technology is lower than the discount rate: $\rho > \alpha$.

The net present value can be calculated in the following two steps. First, we calculate the discounted net present value of output flow under the n-th generation of technology evaluated at the beginning of that technology. Second, we calculate the discounted infinite sum of discounted present value of output flow in each generation (V_1, V_2, V_3, V_4,...) evaluated at time 0.

The following proposition makes the calculation of net present value even simpler.

Proposition 2.1 The optimal upgrade timing is evenly spaced. In other words, it is optimal to upgrade technology at a constant time interval.

Proof: See appendix. ∎

Proposition 2.1 allows us to solve the maximisation problem with respect to just one choice variable, the upgrading interval. We define this upgrading interval as follows. $\lambda \equiv t_n - t_{n-1}$ for any $n = 1, 2, 3, ...$

Now, output at time $(\tau + (n-1)\lambda)$ under the *n*-th generation of technology is:

$$y_{\tau+(n-1)\lambda} = e^{\alpha(n-1)\lambda}(1 - e^{-\gamma\tau}) \text{ where } \tau \in [0,\lambda].$$
(2.1'')

First, the discounted net present value of output flow under the *n*-th generation of technology evaluated at the moment when the technology is introduced is:

$$V_n = \int_0^\lambda y_{\tau+(n-1)\lambda} e^{-\rho\tau} d\tau - \beta e^{\alpha(n-1)\lambda} = e^{\alpha(n-1)\lambda}(\Psi - \beta)$$
(2.2)

where $\Psi \equiv \dfrac{1}{\rho}(1 - e^{-\rho\lambda}) - \dfrac{1}{\rho+\gamma}(1 - e^{-(\rho+\gamma)\lambda})$
(2.3)

Second, discounted infinite sum of net present value of output in each generation can be expressed as follows.

$$V = \sum_{n=1}^{\infty} V_n e^{-\rho(n-1)\lambda} = \frac{\Psi - \beta}{1 - e^{-\lambda(\rho-\alpha)}} \cong \frac{\Psi - \beta}{\lambda(\rho - \alpha)}.^8$$
(2.4)

2.2 *Optimal upgrading interval and its implications*

Now, the optimal timing for upgrading technology can be found by choosing λ so as to maximise the net present value of the output flow.

$$\max_{\lambda} V = \frac{\Psi - \beta}{\lambda(\rho - \alpha)}$$
(2.5)

From the first order condition,

$$\frac{\Psi}{\lambda} - \frac{\partial\Psi}{\partial\lambda} = \frac{\beta}{\lambda}$$
(2.6)

By inserting (2.3) in (2.6) and rearranging, the first order condition is expressed as follows.

$$\frac{1}{\rho\lambda}[1 - (1+\rho\lambda)e^{-\rho\lambda}] - \frac{1}{(\rho+\gamma)\lambda}[1 - \{1+(\rho+\gamma)\lambda\}e^{-(\rho+\gamma)\lambda}] = \frac{\beta}{\lambda}$$
(2.7)

From equation (2.7), we can study how the optimal upgrading interval and the maximised net present value of the output are affected by upgrading cost, learning cost, or exogenous technological progress. Proposition 2.2 and 2.3 answer this question.

---

8.   The approximation for the last term in equation (2.4) comes from the first order Taylor approximation, $e^{-x} \cong 1 - x$ when $x \cong 0$. So, my approximation is justifiable as long as α is close enough to ρ. I will use this approximation because it makes the following analyses much simpler without distorting the main results.

Proposition 2.2 The optimal upgrading interval, $\lambda^*(\beta, \gamma, \rho) \equiv \arg\max_\lambda V$, has the following properties:

$$\frac{\partial \lambda^*}{\partial \alpha} = 0, \quad \frac{\partial \lambda^*}{\partial \beta} > 0, \quad \frac{\partial \lambda^*}{\partial \gamma} < 0. \tag{2.8}$$

Proof: See appendix. ∎

Proposition 2.3 The maximised net present value, $V^*(\alpha, \beta, \gamma, \rho) \equiv \max_\lambda V$, has the following properties:

$$\frac{\partial V^*}{\partial \alpha} = \frac{\partial V}{\partial \alpha} > 0, \quad \frac{\partial V^*}{\partial \beta} = \frac{\partial V}{\partial \beta} < 0, \quad \frac{\partial V^*}{\partial \gamma} = \frac{\partial V}{\partial \gamma} > 0. \tag{2.9}$$

Proof: Envelope theorem. See appendix. ∎

According to Proposition 2.2 and 2.3, a reduction in the cost of an upgrade ($\beta\downarrow$) or faster learning ($\gamma\uparrow$) results in faster technology upgrades and higher net present value. These results are in line with intuition. The reduction in the cost of an upgrade or an improvement in the learning process makes it less expensive to adopt a new technology by reducing its tangible or intangible costs, and hence, makes upgrading more frequent and increases the net present value. For example, consider a case involving a software upgrade. If the price of the new software drops, or if it becomes easier to learn the new software, customers will be more willing to purchase it.

Proposition 2.2 and 2.3 also state that faster exogenous technological progress ($\alpha\uparrow$) does not affect the upgrading interval even though it increases net present value. This prediction is less immediately intuitive. If exogenous technological progress occurs more rapidly, the benefit of upgrading also increases. If this is the case, then how could the optimal upgrade interval be unaffected? The answer comes from the second part of Assumption 2.2, which states that the upgrade cost is proportional to potential productivity. Under this assumption, faster technological progress increases both potential productivity and upgrade cost at the same time. For this reason the assumption of a proportional upgrade cost is not totally benign. However, this assumption is very useful in this model and allows us to stipulate Proposition 2.1. Moreover, the strict assumption of proportional upgrade cost may provide a bench mark for other situations.

Using this bench mark case of proportional upgrade cost, for example, consider the case where the upgrade cost is fixed and independent of the potential productivity of the adopted technology. In this scenario, the upgrade interval will become shorter and shorter over time, because the ratio of upgrade cost to potential productivity will get smaller and smaller as the potential productivity of the frontier technology increases. In this case, faster technological progress will make technology upgrades more frequent, because it increases the benefit of new technology adoption without increasing the cost. For example, consider the case of personal computers. The computing power of personal computers has been increasing exponentially,[9] but the price of personal computers has remained almost constant regardless of the increase in computing power. My model predicts that, in this case, people will tend to upgrade their personal computers over decreasing intervals of time. Moreover, if the increase in computing power becomes more rapid, people will also respond by upgrading more quickly.

---

9.   For example, Moore's law says computer processing speed has been doubled every other year.

According to Proposition 2.2 and 2.3, all things being equal, a firm with lower upgrade cost or faster learning ability will have a shorter upgrade interval and higher net present value of their output. As long as the stock market reflects the fundamental value of the firm, the market value of a firm with lower upgrade cost or faster learning ability will tend to be higher. Then, will a firm with a shorter upgrade interval necessarily have a higher current profit rate as well? The answer from our model is: "No, not necessarily." In our simple model, the initial effect of technology upgrading on current net output is negative. So, higher net present value of future output does not necessarily mean higher current profit.

3.   IT capital, market value, and profitability

The model in Section 2 predicts that a firm having lower upgrade costs, tangible or intangible, will tend to upgrade technology more frequently and have a higher market value. But, it is uncertain whether a firm with more frequent technology upgrading will necessarily have a higher current profit rate. This section tests those predictions using a firm-level accounting data set and offers an explanation to the aforementioned seemingly contradictory phenomenon reported in the literature on Information Technology (IT).

3.1   *Predictions to be tested*

One of the biggest challenges in testing the model is the fact that variables such as upgrade or learning cost coefficients and technology upgrade intervals are not directly observable from ordinary firm-level accounting data. To overcome this problem, this section takes an indirect approach of using an observable proxy which seems to be highly correlated with the unobservable variable in the model.

Even though a company's learning ability in technology upgrading is not directly observable from the company's accounting data, for example, IT capital intensity seems to be correlated with such learning speed in technology upgrading. It is not only because IT investment itself is a particular form of technology upgrading but also because IT capital is expected to improve learning ability for other forms of technology upgrading. In other words, a company with faster learning ability will be more active in accumulating IT capital, while installed IT capital will again help a company to reduce its intangible upgrade costs in subsequent technology adoptions. Similarly, the average frequency of technology upgrading is not directly observable from firm-level accounting data. But, more frequent technology upgrading will be accompanied by higher capital turnover rate, because technology upgrading usually requires the replacement of old capital by new one.

Having IT capital intensity and capital turnover rate as proxies for learning ability and upgrading frequency, respectively, we can derive the following three testable predictions from the model, especially from Proposition 2.2 and 2.3.

Prediction 3.1 The partial correlation between IT capital intensity and capital turnover rate will be positive.

Prediction 3.2 The partial correlation between IT capital intensity and market value of the company will be positive.

Prediction 3.3 The partial correlation between IT capital intensity and return on asset (ROA) can have any sign.

In the following subsections, I will describe the data set and report the results of my panel regressions which test these predictions.

3.2 Data

The data is an unbalanced panel data from 1,031 US firms for a 10-year period (1986-1995) constructed by combining two major sources, *Computer Intelligence* and *Compustat*.

- *The Computer Intelligence* panel has the total value of computer stock (including central processors, personal computers, peripherals, etc.) for *Fortune 1000* firms. Because the list of *Fortune 1000* firms has names that drop on or off the list from year to year, the number of firms covered in *Computer Intelligence* data set exceeds 1,000.[10]

- *The Compustat* data set was consulted to gather other information such as market value, physical capital, return on assets, debt-to-equity ratio, etc. Those two data sets were linked together by using the ticker symbol of each company.

Variables in the regressions are constructed in the following way.

- **Market Value (of a Company):** A company's market value is calculated by adding up each item of its liabilities evaluated in the market price. I added up common stock, preferred stock, and debt reported in the *Compustat* data base. In this data base, only common stock is evaluated in the market price. As preferred stock and debt are not evaluated in the market value, but in book value, the market value in my regressions is a very rough measure of a company's real market value. A much more detailed data set and additional assumptions (for pricing each type of debt) would be required to calculate the market value more precisely. But, what is more important in my regressions is not the exact level of market value but the trend of market value depending on other explanatory variables. The market value series in my regressions capture the response of each company's market value in the equity market.

- **IT Capital:** I used the total value of computer stock (central processors, personal computers, peripherals, etc.) reported in the *Computer Intelligence* data set.

- **Capital:** I used the series of "Property, Plant, & Equipment" reported in the *Compustat*. When IT capital intensity is calculated, I used the net value of "Property, Plant, & Equipment". When calculating the capital turnover rate, I used the gross value of "Property, Plant, & Equipment".

- **IT Capital Intensity:** IT-capital divided by capital. That is, the total value of computer stock divided by the net value of "Property, Plant, & Equipment"

- **Capital Turnover Rate:** I used the log growth rate of gross capital over the coming three years from the current year. By using the gross value of the capital rather than the net value, this capital turn over rate captures not only the increase in new capital but also the disposal of old capital.

---

10. Shin-Kyu Yang at MIT generously allowed me to use his data set on the computer capital stock reported by Computer Intelligence.

− **Return on Asset (ROA):** As a measure of profitability, I used the ROA reported in *Compustat*. In calculating this ROA, annual net profit is divided by the average of assets at the beginning of the year and at the end of the year.

3.3. *Regression results*

To test the three predictions in the previous subsection, I ran regressions for three equations explaining capital turnover rate, market value, and return on asset, respectively.

3.3.1. *Capital Turnover Rate and IT Capital Intensity*

To test Prediction 3.1, I regressed the capital turnover rate on IT capital intensity, ROA, and the ratio of the market value to assets. All variables were transformed into natural logs. It is expected that a company with higher ROA or with a higher "*q*-ratio" (market value divided by total assets) will tend to invest more, and hence, to have a higher capital turnover rate. These two variables were added to see the pure contribution of IT capital intensity to the capital turnover rate, after controlling other factors which affect investment. Firm specific effects were also controlled either by fixed effect panel regression or by very detailed (SIC 4-digit) sub-industry dummies.

Table 1 shows the results of the (fixed effect) panel regression and the pooled regression with SIC 4-digit industry dummies. Estimated coefficients for all the right-hand-side variables had expected signs and were statistically significant. This was especially true for the coefficient for IT capital intensity which was significantly positive. This means the partial correlation between IT capital intensity and capital turnover rate is positive. These empirical results are consistent with Prediction 3.1. To summarise, a company with higher IT intensity tends to upgrade technology more frequently.

Table 1. **Regressions on capital turnover rate**

Dependent Variable: ln (Capital Turnover Rate)   (*t*-ratio in parenthesis)

	(1)	(2)
ln (IT Capital Intensity)	0.116 (7.461)	0.038 (3.540)
ln (ROA)	0.00885 (4.965)	0.017 (9.741)
ln (Market Value / Asset)	0.295 (7.800)	0.264 (10.419)
Number of Observations	3328	3328
R-squared	0.081	0.319
Regression Method	Panel Regression (fixed effect)	Pooled Regression (with sub-industry dummies)

3.3.2. *Market Value and IT Capital Level*

To test Prediction 3.2, I regressed the market value on IT capital, total assets, ROA, R&D spending, and advertisement spending. All variables were transformed into natural logs. In this regression, I had to use the IT capital level instead of the IT capital intensity because the left-hand-side variable (market value) is a level not a ratio. I added other controlling variables such as total assets, ROA, R&D spending, and advertisement spending, which are expected to affect market value positively. It is obvious that the market value of a firm is an increasing function of the company's size and profitability. One could also suspect that high-tech or big-name companies would tend to be valued higher in the stock market. The last two variables, R&D spending and advertisement spending, were added to control for this. Again, possible firm specific effects were controlled either by fixed effect panel regression or by very detailed (SIC 4-digit) sub-industry dummies.

Table 2 shows that the estimated coefficients for all the right-hand-side variables had the expected signs and were statistically significant. Especially, the coefficient for the IT capital level is significantly positive, which means that the partial correlation between IT capital and market value is positive even when the total assets have been controlled for. These empirical results are consistent with Prediction 3.2. To summarise, a company with higher IT capital but with the same total assets tends to have higher market value, if all the other factors are the same.

Table 2. **Regressions on market value**

Dependent Variable: ln (Market Value)		(*t*-ratio in parenthesis)
	(1)	(2)
ln (IT Capital)	0.067 (11.134)	0.044 (7.673)
ln (Asset)	0.704 (35.576)	0.812 (55.865)
ln (ROA)	0.012 (18.384)	0.023 (29.921)
ln (R&D)	0.065 (4.253)	0.056 (6.467)
ln (Advertisement)	0.032 (2.665)	0.069 (6.070)
Number of Observations	5406	5406
R-squared	0.790	0.925
Regression Method	Panel Regression (fixed effect)	Pooled Regression (with sub-industry dummies)

3.3.3. *Return on Asset (ROA) and IT Capital Intensity*

For Prediction 3.3, I regressed ROA on IT capital intensity and two control variables, the sales-to-asset ratio, and the debt-to-equity ratio. All variables were transformed into natural logs. The sales-to-asset ratio is expected to affect ROA positively, while the debt-to-equity ratio is expected to affect ROA

negatively by increasing interest spending. As in previous cases, possible firm specific effects were also controlled either by fixed effect panel regression or by very detailed (SIC 4-digit) sub-industry dummies.

The estimated coefficients for the sales-to-asset ratio and the debt-to-equity ratio had correct signs and were statistically significant (Table 3). Interestingly, the coefficient for IT capital intensity is significantly negative. That is, the partial correlation between IT capital intensity and ROA is negative. These empirical results are consistent with Prediction 2.3. In other words, a company with higher IT capital intensity tends to upgrade technology more rapidly and have a higher market value, but its profitability is not necessarily higher.[11]

Table 3. **Regressions on ROA**

Dependent Variable: ln (ROA)		(*t*-ratio in parenthesis)
	(1)	(2)
ln (IT Capital Intensity)	-0.758 (-8.237)	-0.646 (-8.709)
ln (Sales / Asset)	2.805 (7.276)	1.622 (6.594)
ln (Debt / Equity)	-1.132 (-10.498)	-1.605 (-20.281)
Number of Observations	5769	5769
R-squared	0.117	0.379
Regression Method	Panel Regression (fixed effect)	Pooled Regression (with sub-industry dummies)

4.   Industry-wide learning and technological change

When learning costs are incurred as a part of upgrading technology, the adoption of new technology tends to reduce productivity and hence profitability in the short term, even though the potential productivity gain in the long run outweighs this short run loss. This idea of learning costs in technology upgrading can also be extended and applied to issues affecting industries or economies in aggregate. From this section, I will extend the scope of the study from firm-level to industry-level and develop my argument further both empirically and theoretically.

---

11.   To summarise, the regression results are consistent with the model's predictions as follows.

Partial correlation	Model's predictions	Regression results
	IT capital intensity	IT capital intensity/level
Turnover rate of capital	+ (Proposition 2.2)	+
Market value	+ (Proposition 2.3)	+
ROA	-/+	-

4.1. *Learning-by-doing at the industry level*

An employee's skill, or the productivity of a production line, increases with experience and repetition. This is the basic concept underlying so-called "learning-by-doing." Learning-by-doing effects have been observed both at the personal level and in larger work units such as factories. It should be emphasised, however, that productivity growth at the factory level is not the simple sum of the learning-by-doing effects of individual workers. This is evident from the fact that the productivity growth of a factory is observed even when there is high worker turnover. That is to say, there exists "organisational learning" which cannot be reduced to the sum of "individual learning." For example, co-ordination among different divisions in a factory becomes increasingly efficient as the factory gains experience through repetition. We can also consider learning-by-doing at the industry level, which cannot be reduced to the simple sum of individual firms comprising that industry. This is because there are various types of externalities or spillover effects among firms which belong to the same industry. A new entrant in the auto industry would accrue benefits from the experience of existing firms, for example, in training workers or in constructing a system of auto-part suppliers.

Many theoretical models for international trade and economic growth simply assume the existence of industry-wide learning-by-doing. Based on this assumption, these models try to explain the success of the export-led growth strategy of Japan and the Newly Industrialising Countries (NICs) in East Asia. In trade models, the concept of industry-wide learning-by-doing is used to show that protective trade policies could be effective under certain conditions. For example, Krugman (1984, 1987) showed that Japanese style protective industrial policies could create a comparative advantage if there exists dynamic economies of scale due to industry-wide learning-by-doing. In growth models, learning-by-doing effects are a significant component of endogenous productivity growth, especially in developing countries where imitation rather than invention is the dominant form of technological progress. In this context, Lucas (1993) pointed to learning-by-doing as the source of the rapid growth in East Asian NICs. Surprisingly, however, there are very few previous empirical works which tried to measure the size of learning-by-doing effects at the industry-level.[12]

In this section, I estimate the relative size of industry-wide learning-by-doing effects using annual data of Japanese machinery manufacturing from 1955 to 1990. I chose the Japanese machinery manufacturing data as a starting point, considering the following facts. First, the aforementioned "strategic trade policy" models, which assume industry-wide learning-by-doing, are motivated by the model of Japanese industrial policies and export-led growth. Secondly, Japan seems to have the best industry-level data among the East Asian countries that pursued the export-led growth strategy based primarily on imitation rather than invention. Thirdly, those 15 sub-industries,[13] whose primary business involved machine manufacturing, make the majority of the products which have been associated in the literature with strong factory-level learning-by-doing (for example, shipbuilding, aircraft, motor vehicle, electronic parts, etc.).

---

12.   Sheshinski (1967) and Bahk and Gort (1993) belong to rare exceptions, but their estimations were basically cross-sectional estimations: Sheshinski did cross-state and cross-country regressions, while Bahk and Gort did cross-firm regressions. Using cross-sectional data instead of time-series data is not the best way of measuring the size of learning-by-doing which occurs over time. Naturally, conventional learning curve estimations at individual or factory level usually use time series data. Argote and Epple (1990) provides a succinct survey on learning curve estimations.

13.   The categories of these sub-industries are comparable to three digit SIC codes.

4.2.   Estimating industry-wide learning-by-doing

In the literature which discusses learning-by-doing at the factory or firm level, the most typical way of measuring the learning-by-doing effect is to estimate the "learning curve" which relates a productivity measure (or, unit cost as the reciprocal of a productivity measure) to an experience measure (e.g., cumulative output) in a log-linear form as follows.[14]

$$\ln UC_t = c_0 - c_1 \ln Q_t \qquad\qquad (4.1)$$

where UC_t: unit cost, Q_t: cumulative output, c_1: learning coefficient.

This conventional formulation of learning-by-doing, which means that the unit cost decreases by c_1 % as the cumulative output increases by 1 %, is valid only in ceteris paribus cases. In other words, this formulation assumes that there have been no changes either in capital stock or in technology. Under such assumptions, all the productivity gains are attributed to the learning-by-doing effects. These assumptions are reasonable for the estimation of the learning coefficient of a production line over a relatively short period.

In measuring industry-wide learning-by-doing with data spanning over several decades, however, it is unrealistic to assume that the capital stock and the technology level remain unchanged. Therefore, changes in the capital stock and in the technology level should be explicitly taken into account when attempting to identify industry-wide learning-by-doing. For this purpose, consider an aggregate production function having the productivity factor, A_t, as a function of both cumulative output and time trend.

$$Y_t = A_t f(K_t, L_t) \quad \text{where} \quad A_t = A_0 \cdot Q_t^{\gamma} \cdot Exp(\delta \cdot t) \ , \quad Q_t = \sum_{s=1}^{t-1} Y_s \ . \qquad (4.2)$$

(*Y*: output, *K*: capital input, *L*: labour input, *Q*: cumulative output)

Here, γ is the learning coefficient, and δ represents the speed of exogenous technological progress. From equation (4.2), we can estimate the learning coefficient after the changes in inputs and in the technology level are controlled for.

For estimation, I specified production functions for 15 sub-industries in a log-linear form as follows:

$$
\begin{cases}
\ln Y_{1t} = c_1 + \gamma_1 \ln Q_{1t} + \delta_1 \cdot t + \alpha_1 \ln K_{1t} + \beta_1 \ln L_{1t} + \varepsilon_{1t} \\
\ln Y_{2t} = c_2 + \gamma_2 \ln Q_{2t} + \delta_{21} \cdot t + \alpha_2 \ln K_{2t} + \beta_2 \ln L_{2t} + \varepsilon_{2t} \\
\ldots\ldots \\
\ln Y_{15t} = c_{15} + \gamma_{15} \ln Q_{15t} + \delta_{15} \cdot t + \alpha_{15} \ln K_{15t} + \beta_{15} \ln L_{15t} + \varepsilon_{15t}
\end{cases}
\qquad (4.3)
$$

where $E[\varepsilon_{it}] = 0$, but $E[\varepsilon_{it} \cdot \varepsilon_{jt}] \neq 0$, $E[\varepsilon_{it} \cdot \varepsilon_{is}] \neq 0$.

I did not impose the constant-returns-to-scale assumption (i.e., $\alpha + \beta = 1$) in order to prevent the possibility of wrongly attributing static scale effects to learning-by-doing.

---

14.   For example, see Argote and Epple (1990) and references there.

In (4.3), I assumed that all the equations in this system are connected with one another through the correlation of the error terms. It is reasonable to assume that a shock to a certain sub-industry propagates to related ones. For example, communication equipment industry or computer industry is very likely to be affected by a major upheaval in the electronic parts industry. Also, a common shock can affect related sub-industries at the same time. In such cases, where error terms are correlated with one another, we can get a more efficient estimate by using so-called "Seemingly Unrelated Regression (SUR)." Before running regressions on the system equation (4.3), endogeneity problem in some of the right-hand-side variables should be considered. For instance, if there is a positive shock to a sub-industry, whereby it increases not only its output but also the demand for and utilisation of its inputs. This means that capital and labour in our model are endogenous and therefore correlated with the error terms. In this case, "Ordinary Least Square (OLS)" estimation or its system estimation equivalent, "Seemingly Unrelated Regression (SUR)", will yield inconsistent estimates. To solve this problem, we need to use instrumental variables. In the context of "full information system estimation," it is just "Three Stage Least Square Estimation (3SLS)." Real wages, the real interest rate, the relative price of capital goods, and their lagged variables are used as instruments.

I used data from 15 sub-industries of the Japanese machinery industry from 1955 to 1990. The data for net output, capital, and labour in each sub-industry were obtained from *Census of Manufactures*, Vol. 2., *Report by industries* (each year). This annual report is published by the Research and Statistics Department in the Ministry of International Trade and Industry (MITI). The current format based on the Japanese Standard Industrial Classification (JSIC) was established in 1955. The data on price, wage, and interest rate were obtained from the *Japan Statistical Yearbook* published by the Statistics Bureau in Management and Co-ordination Agency.[15]

Each variable in the regression was constructed in the following way.

- **Net output:** I deflated the value-added numbers from the *Census of Manufactures* using the wholesale price indexes from the *Japan Statistical Yearbook*. The definition of "value-added" in the *Census of Manufactures* is: (Value-added) = (Value of gross output) - (Domestic consumption tax) - (Value of intermediate input) - (Depreciation). Industry-specific wholesale price indexes (i.e., wholesale price index for general machinery, that for electrical machinery, and that for transport equipment) were used as deflators for each sub-industry depending on the industry to which it belongs.

- **Capital:** The current stock of the fixed tangible assets reported in the *Census of Manufactures* was deflated by the wholesale price index of capital goods.

- **Labour:** I used the number of workers reported in the *Census of Manufactures*. This number includes self-employed workers and unpaid family workers as well as employees.

- **Real wage:** The "Average monthly cash earnings per regular worker by industry" as reported in the *Japan Statistical Yearbook* were deflated by the consumer price index. Cash earnings are the sum of contracted earnings and extra payments.

---

15. This yearbook refers to following original sources: Research and Statistics Department, The Bank of Japan, Annual Report on Price Indexes; Statistics Bureau, Management and Co-ordination Agency, Annual Report on the Consumer Price Index; Research and Statistics Department, The Bank of Japan, Annual Report on Economic Statistics; Policy Planning and Research Department, Ministry of Labour, Monthly Labour Survey.

− **Real interest rate:** The "Averages of agreed interest rates on loans" of all banks were used after being corrected by wholesale price inflation.

− **Relative Price of Capital Goods:** The wholesale price index of capital goods was divided by the weighted-average wholesale price index of all commodities.

4.3. *Imitation vs. innovation*

The main result of the estimation is summarised in Figure 4. It shows an interesting relation between the learning coefficient and the speed of technological change. Industry-wide learning-by-doing effects were most pronounced in low-tech industries where technological change was relatively slow. In contrast, the learning-by-doing effect was insignificant or even negative in high-tech industries which were experiencing rapid change in their technology.

These findings, on one hand, seem to confirm the validity of theoretical models which assume industry-wide learning-by-doing. But, on the other hand, they reveal the limit of the applicability of those models by showing that the industry-wide learning-by-doing effect is strong only in low-tech industries. Already questionable effectiveness of protective trade/industry policies, namely, policies of nurturing an infant industry via learning-by-doing, becomes even weaker in high-tech industries.

Figure 4. Speed of technological change and learning-by-doing
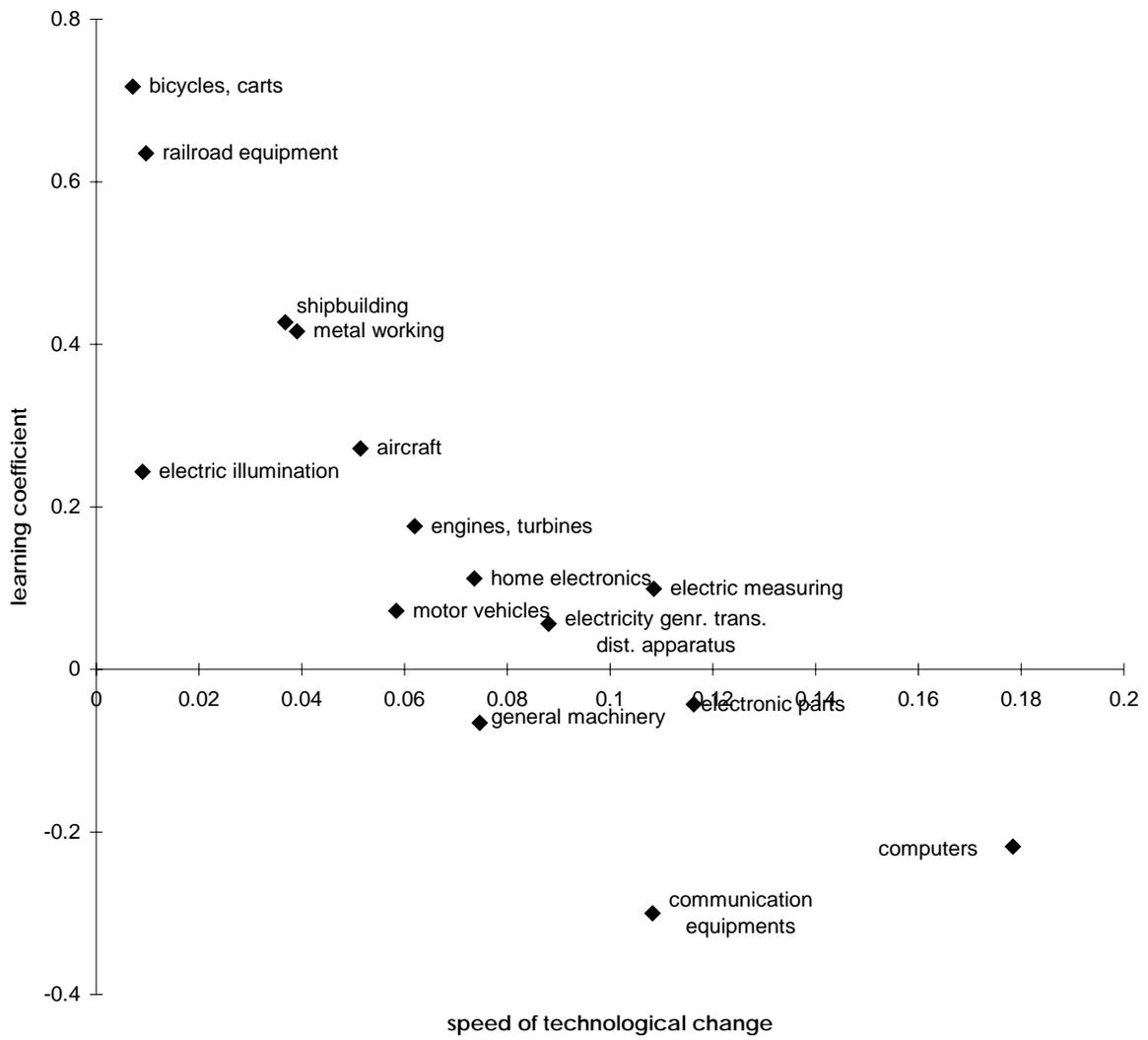
This negative correlation between learning-by-doing and exogenous technological progress can be explained in both directions of causality: (1) from technology to learning; and (2) from learning to technology.

− *From Technology to Learning:* In low-tech industries, production methods do not change frequently and it is relatively easy to improve productivity simply by imitating existing technology and best practise in using it. As the production method evolves more rapidly, however, the opportunity to improve productivity through imitation decreases. In high-tech industries, invention rather than imitation is the dominant form of technological change, and hence, R&D rather than learning-by-doing becomes the primary factor that determines productivity growth.

− *From Learning to Technology:* The technology upgrading model in Section 2 delineates another aspect of the story explaining the negative correlation between technological change and learning-by-doing. According to the model (especially, see Proposition 2.2), the more significant learning becomes in upgrading technology (i.e., the smaller γ is), the slower the pace for technology upgrades. This prediction of the model can be explained intuitively as follows. In an industry where productivity growth from learning-by-doing lasts for a long time, the incentive for frequent innovation will be relatively small. On the contrary, in an industry where the productivity potential for a specific technology is fully realised within a short period of learning, the room for productivity growth through learning-by-doing will disappear quickly. In this case, technology upgrading will play a much more significant role than learning-by-doing with existing technology.

5.   Investment and total factor productivity (TFP)

Two "productivity puzzles" reported in the IT literature and in the studies of East Asian economic growth have interesting common aspects. In both cases, there was a massive investment in advanced technology sectors, but the return on such investments did not materialise immediately in terms of profitability or productivity. The idea of learning costs provides, at least, a partial explanation for these observations. Whether it is adopting new information technology or industrialising a pre-modern economy, the initial learning cost related to how to manage these new modes of production will be enormous. Therefore, it is not surprising that the immediate productivity gain in such massive-scale "technology upgrades" was small.

This section presents a finding that total factor productivity (TFP) growth tends to decrease with faster capital accumulation. This negative correlation between the movement of the capital growth rate and that of the productivity growth rate is very widely observed. Moreover, this basic pattern is not affected by various methods for measuring capital stock, or varying levels of aggregation. This section argues that this intriguing pattern conforms better with the idea of learning cost in installing new capital than with other competing stories.

5.1.   *Empirics: accelerating investment decelerates TFP growth[16]*

Comparing the TFP growth rate and the capital growth rate reveals a very intriguing pattern: the TFP growth rate tends to decrease [increase] when capital accumulation becomes more rapid [slower]. This pattern is repeatedly observed in various data with various aggregation levels.

---

16.   See Appendix 2, for methodology and data sources.

Figure 5. TFP growth rate and capital growth rate

(general machinery industry, Japan)
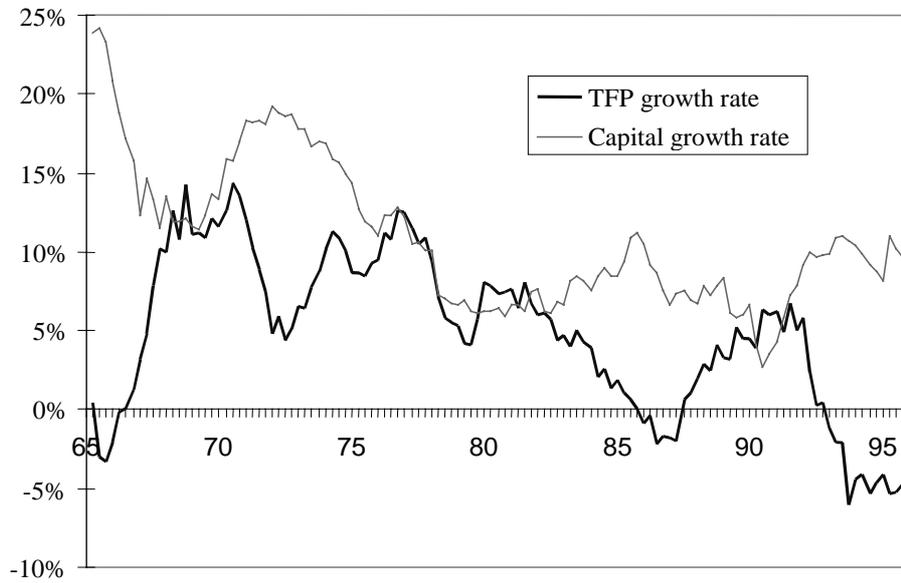
Figure 6. TFP growth rate and capital growth rate

(transportation equipment industry, Japan)
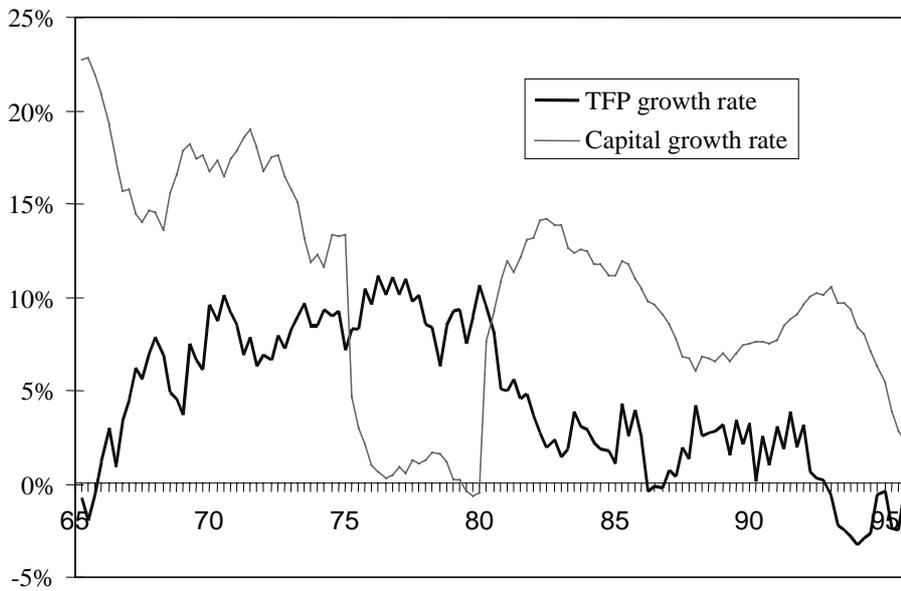
**Figure 7. TFP growth rate and capital growth rate
(boilers, engines, and turbines, Japan)**
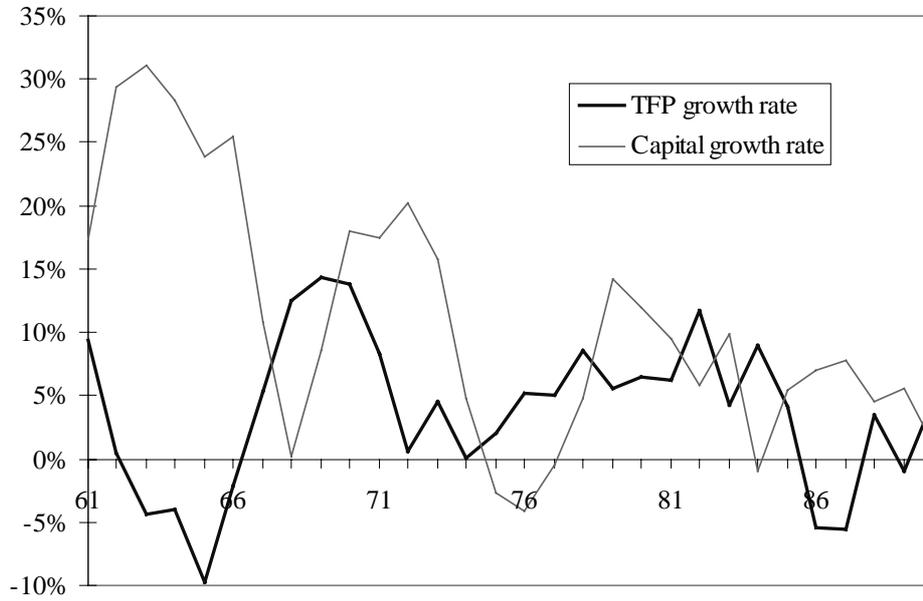
**Figure 8. TFP growth rate and capital growth rate
(motor vehicles and equipment, Japan)**

Figure 9. TFP growth rate and capital growth rate

(engines and turbines: SIC 3510, U.S.)
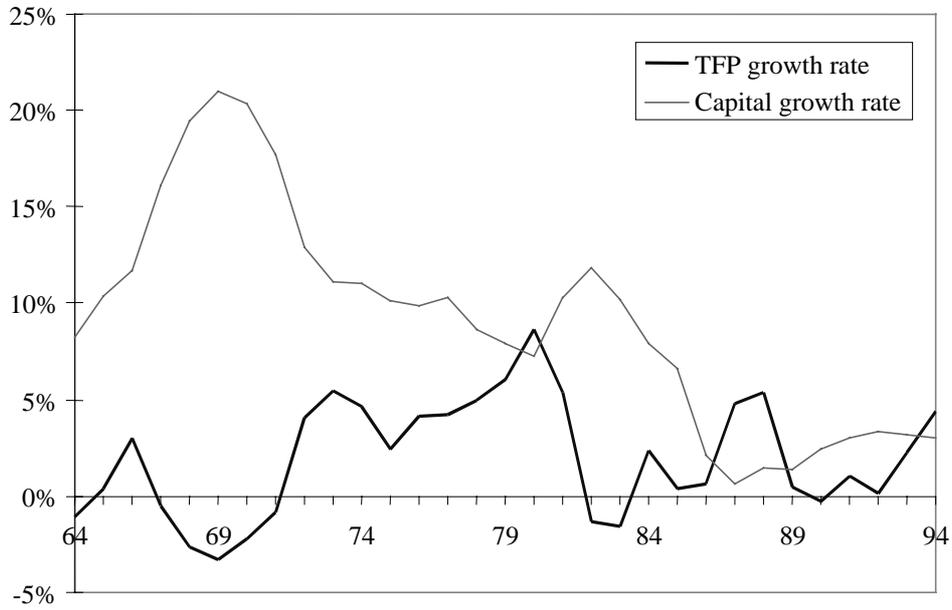
Figure 10. TFP growth rate and capital growth rate

(motor vehicles and car bodies: SIC 3711, U.S.)
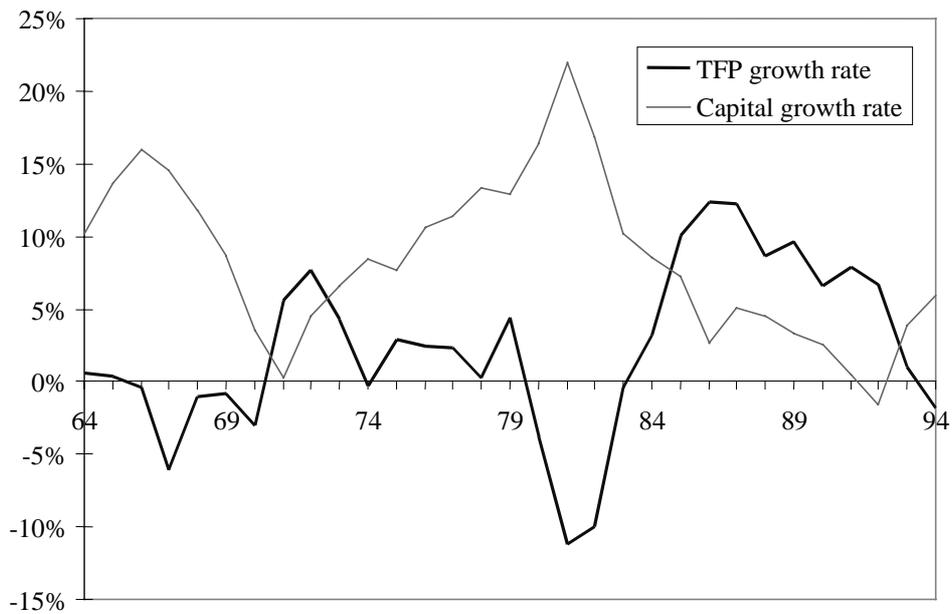
Table 4. **Capital accumulation and productivity growth
(major industries, Japan)**

*: significantly different from 0 at the level of 95%   (*t*-ratio from simple regression in parenthesis)

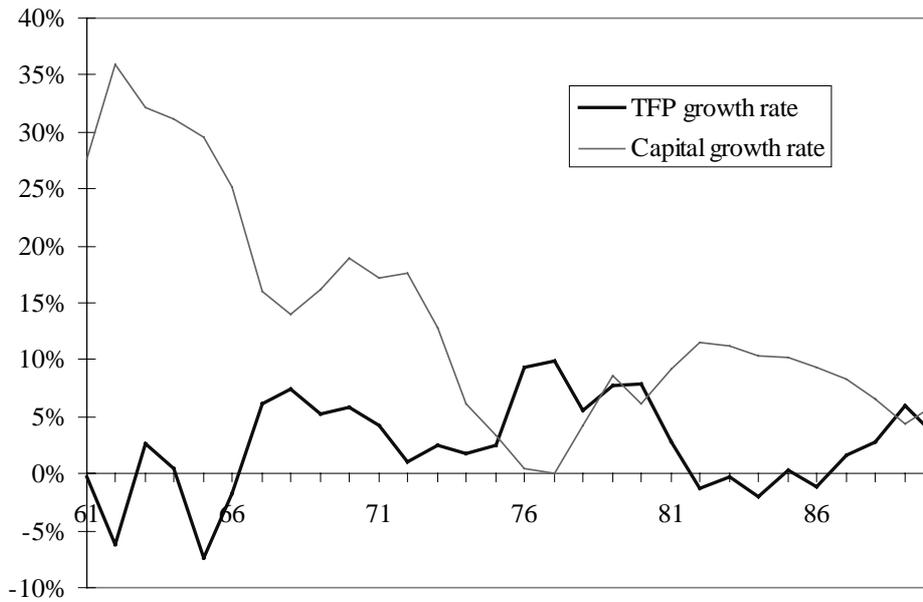|	Correlation $[\Delta^2 \ln(TFP) , \Delta^2 \ln(K)]$
Total Industry Aggregate	-0.357* (-5.790)
Manufacturing Aggregate	-0.478* (-4.547)
Food, Tobacco	-0.185* (-5.140)
Textile	-0.120* (-10.614)
Apparel	0.343 (0.138)
Lumber, Wood	0.202 (0.333)
Pulp, Paper	-0.195* (-11.077)
Publishing, Printing	0.114 (0.307)
Chemical	-0.387* (-5.501)
Petroleum, Coal	0.343 (0.555)
Ceramic	0.131 (0.215)
Iron, Steel	-0.434* (-8.267)
Non-ferrous metals	-0.229* (-8.489)
Fabricated metal	-0.113* (-8.827)
General machinery	-0.247* (-5.290)
Electrical machinery	-0.422* (-2.546)
Transportation equipment	-0.450 (-1.457)
Precision machinery	0.418 (0.042)
Shipbuilding	-0.154 (-0.353)

Table 5. **Capital accumulation and productivity growth**
(machinery industries, Japan)

(*t*-ratio from simple regression in parenthesis)

	Correlation $[\Delta^2 \ln(TFP), \Delta^2 \ln(K)]$
Boilers, engines, and turbines	-0.480 (-1.708)
Metal working machinery	-0.517 (-2.196)
General industrial machinery and equipment	-0.376 (-3.288)
Electrical generating, transmission, distribution appara	-0.572 (-2.327)
Household electric appliances	-0.537 (-1.817)
Electric illumination appliances	-0.202 (-6.552)
Communication equipment and related products	-0.291 (-3.016)
Computers	-0.586 (-1.226)
Electronic parts	-0.347 (-2.513)
Electric measuring instrument	-0.367 (-3.381)
Motor vehicles and equipment	-0.609 (-1.877)
Railroad equipment and parts	-0.477 (-2.006)
Bicycles, carts and parts	-0.064 (-23.835)
Ship and boat building and repairing	-0.395 (-3.880)
Aircraft	-0.153 (-8.276)

Table 6. **Capital accumulation and productivity growth**
(3 sub-industries, US)

(*t*-ratio from simple regression in parenthesis)

	Correlation $[\Delta^2 \ln(TFP), \Delta^2 \ln(K)]$
Engines and turbines (SIC 3510)	-0.525 (-2.431)
Motor vehicles and car bodies (SIC 3711)	-0.587 (-2.302)
Aircraft (SIC 3721)	-0.306 (-2.144)

Figure 5 through Figure 10 demonstrate this point graphically. They have different data sources, aggregation levels, industries, and different countries, but still they show the same pattern. The capital growth rate and the TFP growth rate tend to move in opposite directions under most circumstances. In other words, one curve tends to go up when the other one goes down, and *vice versa*.

As a quantitative measure for the pattern observed in the figures listed above, consider the correlation between the change in TFP growth rate (i.e. $\Delta^2 \ln(TFP)$) and the change in the capital growth rate (i.e. $\Delta^2 \ln(K)$). If the TFP growth rate decreases [increases] when the capital growth rate increases [decreases], this correlation will be negative. Table 4 through Table 6 report the values of *Correlation*$[\Delta^2 \ln(TFP), \Delta^2 \ln(K)]$. To check the statistical significance of the sign of the correlation, I added the *t*-ratio from the simple regression of $\Delta^2 \ln(TFP)$ on $\Delta^2 \ln(K)$ in the parenthesis below the correlation.

As shown in Table 4, the correlation is -0.357 for total industry aggregates, and -0.478 for manufacturing aggregates. The negative sign of either number is statistically significant. Among the 17 industries in manufacturing, 9 industries showed a significant, negative correlation. 6 showed positive correlation, but none of them were statistically significant. Table 5 and Table 6 show the correlation values at the level of three or four digit SIC sub-industries in the machinery industry. For all of the 18 sub-industries in Japan and the US, the correlation was negative. All in all, the correlation values confirm the observed pattern of the capital growth rate and the TFP growth rate's tendency to move to opposite directions. In other words, there is a tendency for productivity growth to slow down when investment is accelerated.

5.2. *Possible explanations*

Investment reduces productivity? How to solve this "puzzle"? If the TFP growth rate decreases when the observed capital growth rate increases, there are two possibilities: (1) it is an accounting phenomenon (discrepancy between reported capital stock and real capital stock); or, (2) it is a real phenomenon (some lag effect between increasing investment and output increase).

The "time-to-build" argument can be a good explanation for the former case. Usually, there is a lag between the time when the investment is first received and the time when the capital equipment is finally built and utilised in production. When we have a discrepancy between the time when the investment

expenditure is made and the time when new equipment started operating on the production floor, increased capital stock as it shows on the ledger will not increase real output immediately. In this sense, "time-to-build" can be regarded as an accounting problem. Or, one might suspect that it is problematic to regard the book value of fixed assets as capital stock.

But, this negative correlation between investment and productivity is observed even after controlling this time-to-build factor. In Figure 5 and Figure 6, for example, the capital series was calculated based on fixed tangible assets excluding construction in process. In other words, the "time-to-build" factor cannot affect capital stock in this case, because capital expenditures on unfinished projects were not included in the value of capital stock. If one points to the problem of using the book value in determining capital stock series, Figure 9 and Figure 10 are good counter examples because the capital series was constructed instead of using the book value of fixed assets. Therefore, we need to explain this negative relation between investment growth and productivity growth as a real phenomenon.

One can think of several real factors that could produce a delay in the full realisation of increased productivity from newly acquired capital equipment. For a good example of what causes the "gestation lag" when a company is procuring new capital equipment, one can think about "(intangible) adjustment cost" in investment. This cost is intangible in the sense that it is not captured as an explicit cost in accounting. Once a new investment project is initiated, human resources as well as physical resources are diverted to this investment project. As those diverted resources are employed for this new priority, their contribution to output is diminished or may be negligible during the transition period. If such adjustment costs are high, productivity will temporarily decrease when new investment increases. Following Section 2, I will call such intangible cost in starting a new investment project the "learning cost". In this terminology, I understand adjustment of an organisation to a new circumstance as "organisational learning".

6.   Simulations: learning cost and productivity

In this section, I show that the observed pattern between investment and productivity is explained quite well by the idea of learning cost associated with acquiring new capital equipment. This is an extension of the technology upgrading model discussed in Section 2. As it takes time and resources to realise the maximum benefit from new technology, so it will take time and effort for any company to realise the potential productivity gain from newly acquired capital equipment. When the number of producers installing new capital equipment increases, the capital growth rate will increase but the TFP growth rate will initially decrease due to the increased intangible cost (i.e., learning cost).

At first, I will extend the technology upgrading model presented in Section 2, from the individual producer's level to the aggregated industry level. Using this extended model, I will do some simulations and show that the learning cost model successfully reproduces the observed negative correlation between the investment trend and the productivity growth trend.

6.1.   Model

I am going to build a simple model for investment and productivity in which there is learning cost associated with acquiring new capital equipment. In this model, productivity growth is the result of either learning-by-doing effects from utilising existing technology or higher productivity from adopting a new technology with higher productivity potential. To adopt a new technology, however, one requires new capital. Acquiring new capital equipment requires not only capital expenditures but also incurs learning cost, because it takes time and effort for the potential productivity of the newly installed equipment to be fully realised. This model has the same basic structure as the technology upgrading model in Section 2.

For simplicity, I will specify the production function as follows. The output of an individual producer at time t under the current technology employed at time t_n is:

$$Y_t = A_0 \, e^{\alpha t_n} \cdot (1 - e^{-\gamma(t - t_n)}) \cdot \overline{F} \quad \text{where} \quad t \in [t_n, t_{n+1}) \qquad (6.1)$$

This specification is based on the following assumptions.

Assumption 6.1 (Production) Capital functions as the container of technology. Labour is fixed for each producer. By construction, the output level is determined by productivity, and the productivity growth rate is equal to the output growth rate. [$\overline{F}$]

Assumption 5.2 (Productivity) (1) Productivity is determined by two factors: the potential productivity of the current technology and the learning effect through accumulated experience with the currently employed technology. (2) The potential productivity level depends on the time when the current technology is employed, and the more recently employed technology has higher potential productivity with the growth rate of α, reflecting exogenous technological progress. [$A_0 \, e^{\alpha t_n}$] (3) The potential productivity is gradually realised as experience with the current technology increases over time. [$(1 - e^{-\gamma(t - t_n)})$]

With another simplifying assumption as follows, the model has exactly the same basic structure as the technology upgrading model in Section 2.

Assumption 5.3 (Capital) The adoption of new technology requires the acquisition of new capital equipment as the container of the technology. The amount of the required capital is proportional to the potential productivity of the contained technology.[17] This technology-container capital exists without depreciation as long as the contained technology is employed. In other words, the capital stock at time t is equal to the amount of capital expenditure at the time of the most recent investment:

$$K_t = \beta \, A_0 \, e^{\alpha t_n} \quad \text{where} \quad t \in [t_n, t_{n+1}) \qquad (6.2)$$

Given the assumptions listed above, an individual producer who tries to maximise the net present value of output flow will acquire new capital at the optimal interval.[18]

Now, we are ready to extend the model from the level of individual producers to the industry level. The following assumption simplifies this extension.

Assumption 6.4 (Industry) (1) The industry consists of a continuum of identical producers. Individual producers have the same production function and the same time interval for upgrading, λ. They are different from one another only in the time frame in which they choose to upgrade. (2) Index x, $x \in [0, \lambda]$, represents a producer who made the most recent upgrade at x-units of time ago. The distribution of individual producers with index x at time t follows a probability density function, f(x,t).

Once the probability density function, $f(x,t)$, is known, one can determine the industry aggregates for output and for capital stock from equation (6.3) and equation (6.4), respectively.

---

17.   In this case, capital expenditure for new technology adoption is just like the upgrading cost in the technology upgrading model in Section 2.

18.   For the properties of this optimal interval, see Section 2.

$$\textit{Aggregate output:} \, Y_t^I = \int_0^\lambda A_0 \, e^{\alpha(t-x)} \, (1 - e^{-\gamma x}) \, \overline{F} f(x,t) \, dx \qquad (6.3)$$

$$\textit{Aggregate capital stock:} \, K_t^I = \int_0^\lambda \beta \, A_0 \, e^{\alpha(t-x)} \, f(x,t) \, dx \qquad (6.4)$$

From equation (6.3) and equation (6.4), the industry-level productivity growth rate (which is the same as the output growth rate)[19] and the industry-level capital growth rate can be calculated. Now, once parameters $[\lambda, \alpha, \gamma]$ and the functional form of the probability density function $[f(x,t)]$ are given, the correlation between the investment trend and productivity growth trend can be calculated immediately. That is, we are ready for running simulations.

6.2.   Simulations

At first, we need to specify the probability density function for the distribution of individual producers in the industry. As a simple way of mimicking fluctuations in the aggregate output and in the aggregate capital stock observed in actual data, I will specify the probability density function using a sine (or cosine) function. Consider the following specification.

$$f(x,t) = \frac{1}{2\lambda} \left(2 - \cos\left(\frac{2\pi}{\lambda}(t-x) \right) \right) \text{ where } x \in [0, \lambda] \qquad (6.5)$$

This $f(x,t)$ in equation (5.5) fluctuates around $1/\lambda$ with the wave length λ and the wave height $1/\lambda$. (Remember that λ is the interval at which a producer chooses to upgrade.)[20]

Before running simulations, let me explain again what is meant by this specification of the distribution of the producers following a sine curve. This specification simply captures a situation where aggregate investment moves up and down along the business cycle. In this case, the business cycle is created by the fact that there is a fluctuation in the number of producers who start a new investment project.

Suppose that each producer initiates a new investment project every five years ($\lambda = 5$) and that the exogenous technological progress enables frontier technologies to increase potential productivity by a 10% growth rate ($\alpha = 0.1$). The third parameter γ is supposed to be 1.[21] From equation (6.3) and equation (6.4), with equation (6.5), the productivity growth rate (which is equal to the output growth rate) and the capital growth rate can be calculated. Then, by repeating random drawing of time t, one can get

---

19.   See Assumption 6.1.

20.   Of course, $f(x,t)$ in equation (5.5) satisfies the following necessary condition to be a probability density function.

$$\int_0^\lambda f(x,t) \, dx = 1 \text{ for any } t.$$

21.   In this case, it takes 1.6 years until newly installed capital equipment [technology] realises 80% of its potential productivity through the learning process. 99.3% of the potential productivity is realised in 5 years.

Correlation $[\Delta^2 \ln(Y), \Delta^2 \ln(K)]$. With given parameter values, such stochastic simulation has obtained *Correlation* $[\Delta^2 \ln(Y), \Delta^2 \ln(K)] = -0.791$. Table 7 shows that this correlation is negative most of the cases.

Table 7. **Simulation results**

Correlation $[\Delta^2 \ln(TFP), \Delta^2 \ln(K)]$

	α =0.1	α =0.2
γ =0.5	-0979	-0.719
γ =1	-0.791	-0.200
γ =2	-0.257	0.392

I repeated this simulation changing the parameter values for α and γ, and plotted the correlation as a function of α and γ in Figure 11. Unless the learning cost is very small (i.e. γ is very big) or the exogenous productivity growth rate α is so large as to dominate the learning cost effects, *Correlation* $[\Delta^2 \ln(Y), \Delta^2 \ln(K)]$ is negative.

Figure 11. Simulation results

Correlation $[\Delta^2 \ln(TFP), \Delta^2 \ln(K)]$

33
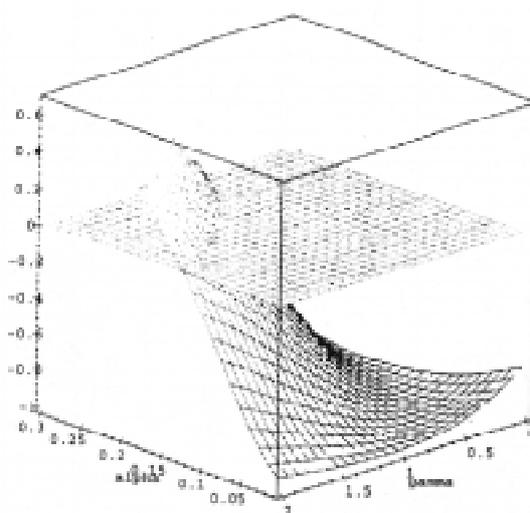
7.   Conclusion

As it takes time and effort to learn how to fully utilise new technology and realise its maximum potential productivity gain, adoption of new technology tends to reduce productivity temporarily, even though the potential productivity gain in the long run outweighs this short run loss. This paper points to such "learning cost" in technology upgrading as a potential explanation of the following two "productivity puzzles" reported in the IT literature and in the studies of East Asian economic growth. First, in the 1980s, US companies made enormous IT investments, but little productivity gain was observed. Second, TFP growth of East Asian NICs was mediocre in spite of the impressive investment drive in those countries.

A simple model of optimal intervals for technology upgrading with learning cost is developed. This model predicts that a company with higher frequency in technology upgrading will tend to have higher market value even with lower current profitability. An empirical study using unbalanced panel data of 1,031 US companies from 1986 to 1995 supports this prediction. Extending the scope from firm-level to industry-level, the paper estimates the magnitude of industry-wide learning-by-doing effects using annual data on 15 sub-industries in the Japanese machinery manufacturing sector from 1955 to 1990. The results show that industry-wide learning-by-doing was strong in low-tech industries where technological change was relatively slow, while it was insignificant in high-tech industries which experienced rapid technological evolution. It is also observed in the US and Japanese manufacturing industries that TFP growth tends to decrease with faster capital accumulation. Simulations based on an extended model of learning cost reproduce the observed negative correlation between the capital growth rate and the TFP growth rate. This negative correlation, which is widely observed in industry-level data and reproduced in simulations based on the learning cost model, makes mediocre TFP growth in East Asian NICs look much less puzzling.

Whether it is adopting new information technology or industrialising a pre-modern economy, the initial learning cost related to how to manage these new modes of production will be immense. Therefore, it is not surprising at all that the immediate productivity gain in such massive-scale "technology upgrades" was small. Interestingly, more recent data suggest that the productivity slowdown in the aforementioned two cases might be simply transitory. For example, Brynjolfsson and Hitt (1996) concludes that "the productivity paradox [of information systems spending] disappeared by 1991, at least in our sample of firms." My own back-of-the-envelop calculation of the annual TFP growth rate show that productivity growth has become significantly faster since 1980s both in South Korea and in Taiwan. Further study is needed to confirm this new tendency.

Appendix 1. Proofs

Proof of Proposition 2.1

Consider a sequence of upgrading timing, $\{t_0(=0), t_1, t_2, t_3, ...\}$. The n-th upgrading interval is defined as $\lambda_n \equiv t_n - t_{n-1}$.

Output at time t under the n-th generation of technology is:

$$y_t = e^{\alpha t_{n-1}}(1 - e^{-\gamma(t-t_{n-1})}) \quad \text{where} \quad t \in [t_{n-1}, t_n). \tag{A1.1}$$

From equation (A1.1), discounted net present value of output flow under the n-th generation of technology evaluated at the moment when the technology is introduced is:

$$V_n = \int_0^{\lambda_n} e^{\alpha t_{n-1}}(1 - e^{-\gamma \tau})e^{-\rho \tau} d\tau - \beta e^{\alpha t_{n-1}} = e^{\alpha t_{n-1}}(\Psi(\lambda_n) - \beta) \tag{A1.2}$$

where $\Psi(\lambda_n) \equiv \dfrac{1}{\rho}(1 - e^{-\rho \lambda_n}) - \dfrac{1}{\rho + \gamma}(1 - e^{-(\rho + \gamma)\lambda_n}). \tag{A1.3}$

Total net present value of output is the discounted infinite sum of net present value of output in each generation. It can be expressed as follows.

$$V = \sum_{n=1}^{\infty} V_n e^{-\alpha t_{n-1}} = \sum_{n=1}^{\infty}(\Psi(\lambda_n) - \beta)e^{-\alpha t_{n-1}} \tag{A1.4}$$

Consider the sequence of optimal upgrading intervals, $\{\lambda_1^*, \lambda_2^*, \lambda_3^*, ...\}$. This sequence is the solution to the following maximisation problem.

$$\max_{\{\lambda_n, n=1,2,3...\}} V = (\Psi(\lambda_1) - \beta) + (\Psi(\lambda_2) - \beta)e^{(\alpha-\rho)\lambda_1} + (\Psi(\lambda_3) - \beta)e^{(\alpha-\rho)(\lambda_1+\lambda_2)} + ... \tag{A1.5}$$

Now suppose that you choose the optimal intervals once again at time $t_1^*(= \lambda^*)$, and call the solution $\{\lambda_2^{**}, \lambda_3^{**}, \lambda_4^{**} ...\}$. This is the solution for the following problem.

$$\max_{\{\lambda_n, n=2,3,4...\}} V' = (\Psi(\lambda_2) - \beta) + (\Psi(\lambda_3) - \beta)e^{(\alpha-\rho)\lambda_2} + (\Psi(\lambda_4) - \beta)e^{(\alpha-\rho)(\lambda_2+\lambda_3)} + ... \tag{A1.6}$$

Compare equation (A1.5) and equation (A1.6). They are basically the same problem. Therefore, the first element of the optimal sequence should be the same in those two problems. That is,

$$\lambda_1^* = \lambda_2^{**}. \tag{A1.7}$$

But, the first problem in equation (A1.5) can be rewritten as follows.

$$\max_{\{\lambda_n, n=1,2,3...\}} \left\{ (\Psi(\lambda_1) - \beta) + (\Psi(\lambda_2) - \beta)e^{(\alpha-\rho)\lambda_1} + (\Psi(\lambda_3) - \beta)e^{(\alpha-\rho)(\lambda_1+\lambda_2)} + ... \right\}$$

$$= \max_{\{\lambda_1\}} \left[(\Psi(\lambda_1) - \beta) + \max_{\{\lambda_n, n=2,3,...\}} \left\{ (\Psi(\lambda_2) - \beta)e^{(\alpha-\rho)\lambda_1} + (\Psi(\lambda_3) - \beta)e^{(\alpha-\rho)(\lambda_1+\lambda_2)} + ... \right\} \right]$$

$$= \max_{\{\lambda_1\}} \left[(\Psi(\lambda_1) - \beta) + e^{(\alpha-\rho)\lambda_1} \max_{\{\lambda_n, n=2,3,...\}} \left\{ (\Psi(\lambda_2) - \beta) + (\Psi(\lambda_3) - \beta)e^{(\alpha-\rho)\lambda_2} + ... \right\} \right] \quad \text{(A1.8)}$$

Equation (A1.8) shows that the optimal choice of $\{\lambda_2^*, \lambda_3^*, ...\}$ does not depend on the choice of λ_1. Compare equation (A1.8) and equation (A1.6). The optimal choice of $\{\lambda_2^*, \lambda_3^*, ...\}$ in equation (A1.8) is the exactly same problem as the choice of $\{\lambda_2^{**}, \lambda_3^{**}, \lambda_4^{**} ...\}$ in equation (A1.6). So, the first element of the optimal sequence should be the same in those two problems:

$$\lambda_2^* = \lambda_2^{**}. \quad \text{(A1.9)}$$

From equation (A1.7) and equation (A1.9),

$$\lambda_1^* = \lambda_2^*. \quad \text{(A1.10)}$$

By repeating the reasoning, we get $\lambda_1^* = \lambda_2^* = \lambda_3^* = \lambda_4^* = ...$ ∎

Proof of Proposition 2.2

Consider the first order condition of the optimisation problem in equation (2.7).

$$\frac{1}{\rho\lambda}[1 - (1 + \rho\lambda)e^{-\rho\lambda}] - \frac{1}{(\rho+\gamma)\lambda}[1 - \{1 + (\rho+\gamma)\lambda\}e^{-(\rho+\gamma)\lambda}] = \frac{\beta}{\lambda} \quad \text{(A1.11)}$$

Define a function:

$$f(x) = \frac{1}{x}(1 - (1 + x)e^{-x}).$$

Notice that this is a hump shape function with $f(0) = 0, f(\infty) = 0$. Using this function, the first order condition can be rewritten as follows.

$$f(\rho\lambda) - f((\rho+\gamma)\lambda) = \frac{\beta}{\lambda}. \quad \text{(A1.12)}$$

In Figure A1.1, left hand side (*LHS*) and right hand side (*RHS*) of equation (A1.12) are drawn as a function of λ. It is trivial that $LHS(0) = 0, LHS(\infty) = 0$. Figure A1.1 shows why *LHS* intersects the horizontal axis from the below just once. As *RHS* is a hyperbola, *LHS* intersects *RHS* from the below at most once. This intersecting point gives the optimal interval, λ^*. When the upgrading cost β is too big [if

$\beta \geq \rho^{-1} - (\rho + \gamma)^{-1}$], they don't intersect. In this case, the optimal interval $\lambda^* = \infty$, which means there is no upgrading.

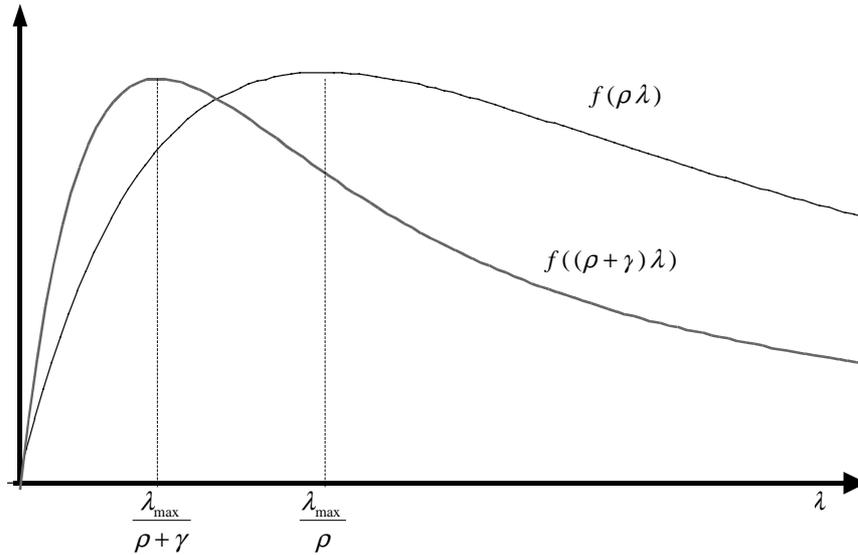1. As α does not affect either *LHS* or *RHS*, a change in α does not change λ^*.

2. When the upgrading cost β increases, the hyperbola *RHS* shifts away from the origin. Then, the intersecting point moves to the right. That is, upgrading interval increases upgrading becomes slower.

3. When the learning costs become less important (i.e., when γ becomes bigger), *LHS* shrinks to the left. (See Figure A1.2) Then, the intersecting point moves to the right. That is, upgrading interval decreases and upgrading becomes faster. ∎

Figure A1.1 Optimal upgrading interval: $\lambda^*(\beta,\gamma,\rho)$

$RHS(\lambda): \dfrac{\beta}{\lambda}$

β

γ

$\lambda^*(\beta,\gamma,\rho)$

λ

$LHS(\lambda): f((\rho + \gamma)\lambda) - f(\rho\lambda)$
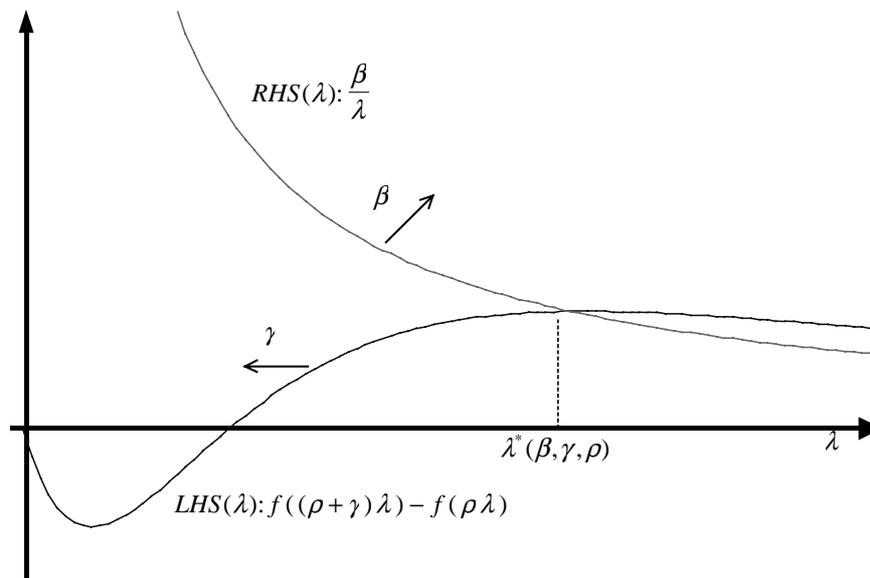
Figure A1.2 $f(\rho\lambda)$ *and* $f((\rho+\gamma)\lambda)$

Proof of Proposition 2.3

Take differentiation on equation (2.4). The first and second inequalities are trivial. The third one is proved as follows.

$$\frac{\partial V}{\partial \gamma} = \frac{1}{\lambda(\rho-\alpha)}\frac{\partial \Psi}{\partial \gamma} = \frac{1}{\lambda(\rho-\alpha)}\frac{1}{(\rho+\gamma)^2}\left(1-\left(1+(\rho+\gamma)\lambda\right)e^{-(\rho+\gamma)\lambda}\right) \qquad \text{(A1.13)}$$

Consider the following function:

$$g(x) = 1-(1+x)e^{-x}.$$

$g(0) = 0, g(\infty) = 1$, and $g'(x) = xe^{-x} > 0$ if $x > 0$. Therefore, $g(x) > 0$ if $x > 0$.

Therefore, in equation (A1.13),

$$\frac{\partial V}{\partial \gamma} = \frac{1}{\lambda(\rho-\alpha)}\frac{1}{(\rho+\gamma)^2}g\big((\rho+\gamma)\lambda\big) > 0. \quad \blacksquare$$

Appendix 2. Calculating total factor productivity growth

Methodology

The TFP growth rate indicates how quickly output increases as a result of productivity growth after controlling the growth in production factors. Consider the following production function.

$$Y_t = F(K_t, L_t, t) \tag{A2.1}$$

where Y_t: output at time t, K_t: capital input at time t, and L_t: labour input at time t.

Take the differential with respect to time on equation (A2.1) and then divide it by equation (A2.1). As a result, we get:

$$\hat{Y}_t = \frac{K_t F_K}{F} \hat{K}_t + \frac{L_t F_L}{F} \hat{L}_t + \frac{F_t}{F} \tag{A2.2}$$

where $\hat{X} \equiv \dfrac{dX/dt}{X}$ (i.e., growth rate of X) and $F_X \equiv \dfrac{\bar{o} F}{\partial X}$.

In equation (A2.2), the last term, F_t / F, represents the TFP growth rate: this term refers to the growth rate of output due to the productivity change independent of changes in factor inputs. In other words, TFP growth rate can be represented as the following.

$$\frac{F_t}{F} = \hat{Y}_t - \frac{K F_K}{F} \hat{K} - \frac{L F_L}{F} \hat{L} \tag{A2.2'}$$

Under the assumption that the factor markets have perfect competition,

$$p F_K = r, \quad p F_L = w \tag{A2.3}$$

where p: output price, r: factor price of capital, and w: factor price of labour.

The assumption of constant returns to scale makes this production function linearly homogeneous with respect to capital and labour. From the properties of linearly homogenous functions,

$$K F_K + L F_L = F . \tag{A2.4}$$

Therefore, under the assumption of perfect competition in the factor markets and constant returns to scale, capital income share and labour income share are well defined as follows.

$$\alpha_K \equiv \frac{K r}{p Y} = \frac{K p F_K}{p Y} = \frac{K F_K}{F}, \quad \alpha_L \equiv \frac{L w}{p Y} = \frac{L p F_L}{p Y} = \frac{L F_L}{F}, \quad \alpha_K + \alpha_L = 1. \tag{A2.5}$$

Now, using equation (A2.5), TFP growth rate can be rewritten as follows.

$$\frac{F_t}{F} = \hat{Y} - \alpha_K \hat{K} - \alpha_L \hat{L} = (\hat{Y} - \hat{L}) - \alpha_K (\hat{K} - \hat{L}). \tag{A2.6}$$

In growth accounting, the output growth rate subtracted by the weighted average of each input's growth rate (weighted by each inputs income share) is called the "Solow residual." Equation (A2.6) shows that the growth rate for "total factor productivity (TFP)" is equal to the Solow residual under the assumption of constant returns to scale and perfect competition in the factor markets. In this case, the TFP growth rate can be immediately calculated from output growth rate, input growth rate, and factor income share.

In order to calculate the TFP growth rate from actual discrete time data, however, we need a discrete time version of equation (A2.6). It is well known that the log of the growth rate of total factor productivity between time t and time $(t\text{-}1)$ can be calculated from the second order translog approximation of the production function.[22] That is,

$$\Delta \ln(TFP_t) \equiv \Delta \ln(Y_t) - \overline{\alpha}_K \Delta \ln(K_t) - \overline{\alpha}_L \Delta \ln(L_t) \tag{A2.7}$$

where $\Delta X_t \equiv X_t - X_{t-1}$, $\quad \overline{\alpha}_K \equiv \dfrac{\alpha_K(t) + \alpha_K(t-1)}{2}$, $\quad \overline{\alpha}_L \equiv \dfrac{\alpha_L(t) + \alpha_L(t-1)}{2}$.

With this equation, I calculated the TFP growth rate for various aggregation levels from four-digit SIC code sub-industry level to country level. I used value added, the value of capital stock, and number of employees for Y, K, and L, respectively. In general, the Solow residual as a TFP growth rate measure is very vulnerable to noise factors because this measure is basically a residual which includes any number of measurement errors. In order to reduce the effects from such noise, I used a five-year moving average in calculating both the TFP growth rate and the capital growth rate.

Data

Figure 5 and Figure 6 are drawn using the data from the *Quarterly Survey of Corporations* published by the Ministry of Finance in Japan. This survey covers Japanese corporations with capital of no less than 10 million yen. Aggregated numbers for balance sheet items and income statement items are reported by industry and by firm size. For the capital series used to calculate the TFP growth rate and the capital growth rate, I used the value of tangible fixed assets excluding land and construction in process. I reported the annual growth rate while the data is reported quarterly. As with all the other values in this section, the growth rates represent the moving average over the last five years.

In Figure 7 and Figure 8, I delved into more detailed categories by moving down from "general machinery" to "boilers, engines, and turbines" and from "transportation equipment" to "motor vehicles and equipment". Here, I used the *Survey of Manufactures* published by the Ministry of International Trade and Industry (MITI) in Japan. This is the same data that I used for the learning-by-doing regressions in Section 4.

In Figure 9 and Figure 10, I show the case of comparable US industries. For this purpose, I used the *Annual Survey of Manufactures* published by the US Department of Commerce. This particular survey offer numbers for new capital expenditures, but not capital stock. I constructed the capital stock series by

---

22.   For more detail, see Young (1992) or Young (1995) and references there.

using the perpetual inventory method. In constructing the capital stock series, I used the *Compustat* industry aggregation data for bench marking. The same pattern is confirmed in the US data as well.

Table 4, Table 5, and Table 6 are based on the same data sources for the above figures: *Quarterly Survey of Corporations* (Ministry of Finance, Japan), *Survey of Manufactures* (Ministry of International Trade and Industry, Japan), and *Annual Survey of Manufactures* (Department of Commerce, US).

Bibliography

ARGOTE, L. and D. Epple (1990), "Learning curves in manufacturing." *Science*, Vol. 247, February, 920-24.

ARROW, K.J. (1962), "The economic implications of learning by doing", *Review of Economic Studies*, 29, June, 155-73.

BAHK, B.-H., and M. Gort (1993), "Decomposing learning by doing in new plants", *Journal of Political Economy*, 101, August, 561-583.

BERNDT, E.R. and C.J. Morrison (1995), "High-tech capital formation and economic performance in U.S. manufacturing industries: an exploratory analysis", *Journal of Econometrics*, 65, 9-43.

BREZIS, K. and Tsiddon (1993), "Leapfrogging in international competition: a theory of cycles in national technological leadership", *American Economic Review*, 83(5).

BRYNJOLFSSON and Hitt (1996), "Productivity, business profitability, and consumer surplus: three different measures of information technology value", *MIS Quarterly*, June.

FAIR, R.C. (1970), "The estimation of simultaneous equation models with lagged endogenous variables and first order serially correlated errors", *Econometrica*, 38, May, 507-516.

GHEMAWAT, P. (1985), "Building strategy on the experience curve", *Harvard Business Review*, 63:2.

HORNSTEIN and Krusell (1996), "Can technology improvements cause productivity slowdowns?", *NBER Macroeconomics Annual*, 209-59.

HULTEN, C.R. (ed.) (1990), "Productivity growth in Japan and the United States", *NBER Studies in Income and Wealth*, Vol. 53, The University of Chicago Press, Chicago and London.

KRUGMAN, P.R. (1984), "Import protection as export promotion: international competition in the presence of oligopoly and economies of scale" in H. Kierzkowski (ed.), *Monopolistic Competition and International Trade*, Oxford University Press.

_____ (1987), "The narrow moving band, the dutch disease, and the competitive consequences of Mrs. Thatcher: notes on trade in the presence of dynamic scale economies", *Journal of Development Economics*, 27, 41-55.

LAU, L.J. and J.-I Kim (1994), "The sources of economic growth of the East Asian newly industrializing countries", *Journal of the Japanese and International Economies*, 8, 235-71.

LUCAS, R.E. Jr. (1993), "Making a miracle", *Econometrica*, 61, March, 251-272.

OLINER, S.D. and D.E. Sichel (1994), "Computers and output growth revisited: how big is the puzzle?", *Brookings Papers on Economic Activity*, 2, 273-334.

RESEARCH AND STATISTICS DEPARTMENT (in Ministry of International Trade and Industry), *Census of Manufactures*, Vol. 2., *Report by Industries*(in Japanese), each year.

ROMER, P.M. (1990), "Endogenous technical change", *Journal of Political Economy*, 98, October, 71-102.

SHESHINSKI, E. (1967), "Tests of the 'learning by doing' hypothesis", *Review of Economics and Statistics*, 49, November, 568-78.

SOLOW, R.M. (1956), "A contribution to the theory of economic growth", *Quarterly Journal of Economics*, 70, 65-94.

_____ (1957), "Technical change and the aggregate production function", *Review of Economics and Statistics*, Vol. 39, No. 3, August, 312-320.

_____ (1987), "We'd better watch out", *New York Times Book Review*, July, p. 36.

_____ (1988), "Growth theory and after", *American Economic Review*, 78, June, 307-17.

STATISTICS BUREAU (in Management and Coordination Agency), *Japan Statistical Yearbook* (in Japanese), each year.

YOUNG, A. (1991), "Learning by doing and the dynamic effects of international trade", *Quarterly Journal of Economics*, 106, May, 369-405.

_____ (1993), "Invention and bounded learning by doing", *Journal of Political Economy*, 101, June, 443-472.

_____ (1994), "Lessons from the East Asian NICs: a contrarian view", *European Economic Review*, 38, 964-973.

_____ (1995), "The tyranny of numbers: confronting the statistical realities of the East Asian growth experience", *Quarterly Journal of Economics*, August, 641-680.

ECONOMICS DEPARTMENT

WORKING PAPERS

219.   Testing for a Common OECD Phillips Curve
         (July 1999) Dave Turner and Elena Seghezza

218.   Sustainable Economic Growth: Natural Resources and the Environment
         (July 1999) Paul van den Noord and Ann Vourc'h

217.   Coping with Population Ageing in Australia
         (July 1999) David Carey

216.   Estimating Prudent Budgetary Margins for 11 EU Countries : A Simulated SVAR Model Approach
         (July 1999) Thomas Dalsgaard and Alain de Serres

215.   The Problems and Prospects Faced by Pay-As-You-Go Pension Systems : A Case Study of Greece
         (June 1999) Paul Mylonas and Christine de la Maisonneuve

214.   Greek Public Enterprises : Challenges for Reform
         (May 1999) Paul Mylonas and Isabelle Joumard

213.   *The Levels and Cyclical Behaviour of Mark-Ups Across Countries and Market Structures*
         (May 1999) Joaquim Oliveira Martins and Stefano Scarpetta

212.   Poverty Dynamics in Four OECD Countries
         (April 1999) Pablo Antolín, Thai-Thanh Dang and Howard Oxley
         Assisted by Ross Finnie and Roger Sceviour

211.   The Recent Experience with Capital Flows to Emerging Market Economies
         (February 1999) Sveinbjörn Blöndal and Hans Christiansen

210.   Foreign Portfolio Investors Before and During a Crisis
         (February 1999) Woochan Kim and Shang-Jin Wei

209.   *Towards More Efficient Government : Reforming Federal Fiscal Relations in Germany*
         (February 1999) Eckhard Wurzel

208.   *Stock Market Fluctuations and Consumption Behaviour : Some Recent Evidence*
         (December 1998) Laurence Boone, Claude Giorno and Pete Richardson

207.   *Microeconometric analysis of the retirement decision: The Netherlands*
         (June 1998) Maarten Lindeboom

206.   *Microeconometric analysis of the retirement decision: United Kingdom*
         (June 1998) Raffaele Miniaci and Elena Stancanelli

205.   *Microeconometric analysis of the retirement decision: Italy*
         (June 1998) Raffaele Miniaci

204.   *Microeconometric analysis of the retirement decision: Germany*
         (June 1998) Pablo Antolin and Stefano Scarpetta

203. *Microeconometric analysis of the retirement decision: United States*
     (June 1998) Joseph Quinn, Richard Burkhauser, Kevin Cahill and Robert Weathers

202. *The retirement decision in OECD countries*
     (June 1998) Sveinbjörn Blöndal and Stefano Scarpetta

201. *The macroeconomic effects of pension reforms in the context of ageing populations:*
     overlapping generations model simulations for seven OECD countries
     (June 1998) Ketil Hviding and Marcel Mérette

200. *The macroeconomics of ageing, pensions and savings: a survey*
     (June 1998) Richard Kohl and Paul O'Brien

199. *Marginal Effective Tax Rates on Physical, Human and R&D Capital*
     (May 1998) Kathryn Gordon and Harry Tchilinguirian

198. *The Norwegian Health Care System*
     (May 1998) Paul van den Noord, Terje Hagen and Tor Iversen

197. *APEC Trade Liberalisation : Its Implications*
     (May 1998) Seunghee Han and Inkyo Cheong

196. *The OECD Jobs Strategy : Progress Report on Implementation of Country Specific Recommendations*
     (May 1998)

196 *La Strategie de l'OCDE pour l'emploi : rapport sur l'état d'avancement de la mise en oeuvre des*
     recommandations par pays
     (May 1998)

195. *Trends in OECD Countries' International Competitiveness*
     (April 1998) Martine Durand, Christophe Madashi and Flavia Terribile

194. *The European Union's Trade Policies and their Economic Effects*
     (April 1998) Peter Hoeller, Nathalie Girouard and Alessandra Colecchia

193. *The Macroeconomic Implications of Ageing in a Global Context*
     (March 1998) Dave Turner, Claude Giorno, Alain De Serres, Ann Vourc'h and Pete Richardson

192. *Efficiency and Distribution in Computable Models of Carbon Emission Abatement*
     (March 1998) Joaquim Oliveira Martins and Peter Sturm

191. *Monetary Policy when Inflation is Low*
     (March 1998) Charles Pigott and Hans Christiansen

190. *Submission by the OECD to the G8 Growth, Employability and Inclusion Conference*
     (March 1998)

189. *Income Distribution and Poverty in Selected OECD Countries*
     (March 1998) Jean-Marc Burniaux, Thai-Thanh Dang, Douglas Fore, Michael Förster,
     Marco Mira d'Ercole and Howard Oxley

188. *Asset Prices and Monetary Policy*
     (February 1998) Mike Kennedy, Angel Palerm, Charles Pigott and Flavia Terribile