Center for Economic Institutions

Working Paper Series

CEI Working Paper Series, No. 2003-21

Technology Upgrading with Learning Cost

Sanghoon Ahn

Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitachi, Tokyo, 186-8603 JAPAN
Tel: +81-42-580-8405
Fax: +81-42-580-8333
e-mail: cei-info@ier.hit-u.ac.jp

TECHNOLOGY UPGRADING WITH LEARNING COST

Sanghoon Ahn*

Hong Kong University of Science and Technology
Division of Social Science
Clearwater Bay, Kowloon
Hong Kong

Phone: (85 2) 2358-0009
Fax: (85 2) 2335-0014
E-mail: sosa@ust.hk

Abstract

Adoption of new technology requires diversion of resources from direct production activities to learning/adjusting activities, which could reduce productivity temporarily. Focusing on the existence of such "learning cost", we derive a simple model on the optimal timing for technology upgrading. This model suggests that a firm perceived to have better learning ability will show more frequent technology upgrading and higher market value even with possibly lower current profitability. The model predictions are supported by regression results from a panel data set of more than 1,000 companies in the US during the late 1980s and the early 1990s. Simulations based on an extended model reproduce the negative correlation between investment growth and TFP growth.

Keywords: Technology; Learning; Total factor productivity (TFP); Market Value.

JEL classification: O30; O47; G10; D24.

* Special thanks go to Susanto Basu, John Fernald, John Haltiwanger, Boyan Jovanovic, Pierre Mohnen, Plutarchos Sakellaris, and other participants of "Productivity Growth and the Macroeconomy" workshop at the NBER 2001 Summer Institute and of the DRUID 2003 Summer Conference for their useful comments and suggestions on earlier versions of the paper. All remaining errors are the author's own.

1. Introduction

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. Therefore, it takes time and effort for the potential productivity of the new technology to be fully realised. Such resources spent in adopting a new technology should be considered as part of the cost for technology adoption. This paper will call it as "learning cost".

Consider, for example, technology upgrading from a typewriter to word processor. On the top of tangible costs for purchasing a personal computer, word processing software, a printer, floppy diskettes, etc., 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 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 such intangible learning costs could make one's productivity with the new word processor initially lower than that with the old typewriter.

The notion of learning cost is relevant at the organisation level as well. Consider a factory or a firm 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.[1] Of course, restructuring an organisation requires tremendous time and effort. This restructuring can also be regarded as a part of organisational learning.

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 Economies" (NIEs) in East Asia.[3] These two cases seem to have interesting

---

1 For this point, see Bresnahan, Brynjolfsson, and Hitt (1999), amongst others.

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.

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 focuses on "learning cost" in technology upgrading as a potential explanation of these two "puzzles".

In the literature on technology and productivity, aggregate productivity slowdown is often attributed to required huge adjustment/learning costs in transition to a new technology regime. For example, David [1990] pointed out the case of pronounced productivity slowdown during electrification of the US and the UK over the period 1890-1913. According to his comparison of the introduction of dynamos and that of computers, contemporary observers in 1900 might well have said that the electric dynamos were "everywhere but in the productivity statistics" like what Solow said about computers in 1987. Indeed, it took decades for large potential gains from factory electrification to be fully realised.[4] Greenwood and Yorukoglu [1997] offered a dynamic general equilibrium model based on the idea of vintage capital and investment in learning. Their model simulations supported their conclusion that "[a]t the dawn of an industrial revolution, the long-run advance in labor productivity temporarily pauses as economic agents undertake the (unmeasured) investment in information required to get new technologies operating closer to their full potential."[5]

This paper offers a simple model of optimal timing for technology upgrading, which considers both tangible upgrading cost and intangible learning cost.[6] While most previous studies focused on macroeconomic implications of a large-scale technological change (namely, "technology revolution" or advent of a new "General Purpose Technology (GPT)"), this paper pays attention to implications of firm-level heterogeneity under continuous introduction of newer frontier technologies with higher potential productivity. While previous studies based on a dynamic general equilibrium model had to rely on simulations, a minimalist model in this paper is very simple, and hence, basic questions can be answered analytically without relying on numerical methods. A deterministic learning equation hired in this paper, as well as the assumption of perfect foresight, might look too simplistic. In fact, however, such problem is common in the main body of the existing literature.[7]

4 More generalised version of David's insight from a historical perspective can be found in the notion of "General Purpose Technology (GPT)". Some technologies such as steam engine, electricity, railway, and microelectronics are regarded as GPTs in the sense that they are pervasively used as inputs by a wide and expanding range of sectors in the economy. In this approach, a technology revolution can be defined as the advent of a new GPT. See Bresnahan and Trajtenberg [1995], Helpman and Trajtenberg [1996], and Helpman and Rangel [1998], for example.

5 Using simulations on a vintage capital model with conventional capital and information technology (IT) capital, Yorukoglu [1998] concluded that the marginal return of IT capital is underestimated by around 20% of its actual size. This conclusion is based on the simulation results that IT capital is associated with a strong learning-by-doing effect and that IT capital investment is mostly concentrated at the replacement dates. For a prototypical vintage capital model on the decision to replace old technologies with new ones, see Cooley, Greenwood, and Yorukoglu [1997].

6 According to back-of-the-envelope calculations by Jovanovic [1996], the costs of adopting new technologies are 20 times as large as invention cost and may amount to 10% of GDP. In our model, technology adoption costs consist of two parts, namely, tangible upgrading cost and intangible learning cost.

7 Papers based on deterministic learning process includes Parente [1994], Greenwood and Yorukoglu [1996], Horstein and Krusell [1996], and Yorukoglu [1998]. Klenow [1998] incorporated uncertainty in demand side, but the learning process remained deterministic. Jovanovic and Nyarko [1995, 1996] based on Bayesian learning belong to a rare exception.

Predictions from models of deterministic learning process should be interpreted with special caution. In particular, the possibility of unsuccessful technology adoption cannot be easily incorporated in deterministic models or models with perfect foresight.

The overview of the paper is as follows. **Section 2** presents the basic model. The model suggests that an optimising company with more frequent technology upgrading will tend to have higher market value even with lower current profitability. An empirical study using unbalanced panel data of more than 1,000 companies in the US from 1986 to 1995 supports this prediction (**Section 3**). Extending the scope from firm-level to industry-level, **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 experienced rapid technological changes. In **Section 5**, it is widely observed in the US and Japanese manufacturing industries that TFP growth tends to decrease with faster capital accumulation. This basic pattern is not affected by various methods for measuring capital stock, or varying levels of aggregation. This intriguing pattern can be explained by the idea of learning cost in installing technology-embodied new capital better than by other competing stories. To support this argument, **Section 6** extends the basic model to incorporate capital and to take aggregation from the individual producer's level to the aggregated industry level. Simulations based on this extended model reproduce the observed negative correlation between the capital growth rate and the TFP growth rate. **Section 7** concludes the paper.

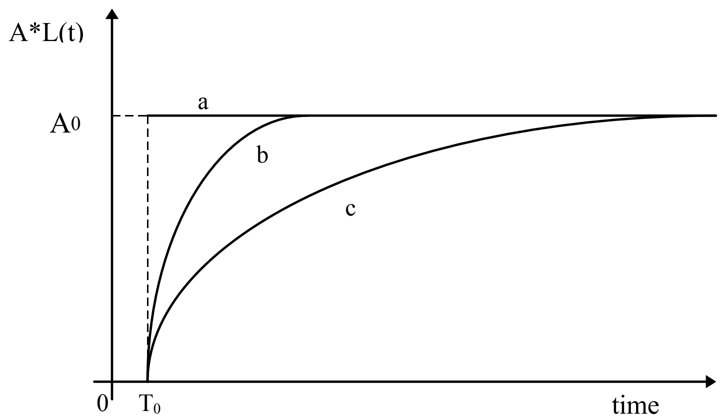
2. Technology Upgrading with Learning Cost

This section 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 Learning Cost

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 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*.

Figure 1. Learning cost in adopting a new technology



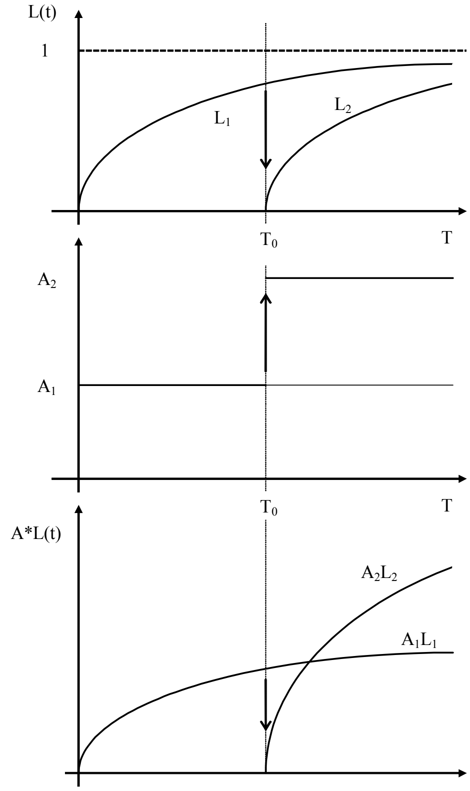
Because there are learning costs associated with the adoption of new technology, its implementation tends to reduce productivity in the short term, even though the potential productivity gain in the long run outweighs this short term loss. Fundamentally, adopting new technology is similar to investing 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 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 converge 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.[8] 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.

[8] 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.

Figure 2. Effects of adopting new technology at time T_0



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:

> [...] 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 section, we will explicitly consider optimisation behaviour that weighs short run costs of technology upgrading against its long run benefits.

2.2 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.

Assumptions 2.1 (Production) (1) *Labour is the only input for production and it is fixed. Specifically, labour input is normalised as* 1. *By construction, output is equal to productivity.*[9] (2) *The producer cannot employ more than one type of technology at the same time.*[10] *The producer can upgrade technology by discarding the current one and adopting a new technology with higher potential productivity, as better and better technology becomes available.*

Expanding available technology set allows 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.

Assumptions 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 [11]. *To avoid the corner solution case where the upgrade cost is prohibitively high, it is also assumed that* $\beta < \rho^{-1} - (\rho + \gamma)^{-1}$.

Assumptions 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.*[12] (2)

9 In this case, one can assume that technology-specific experience is embodied in labour as in Zeckhauser [1968] and Helpman and Rangel [1998].

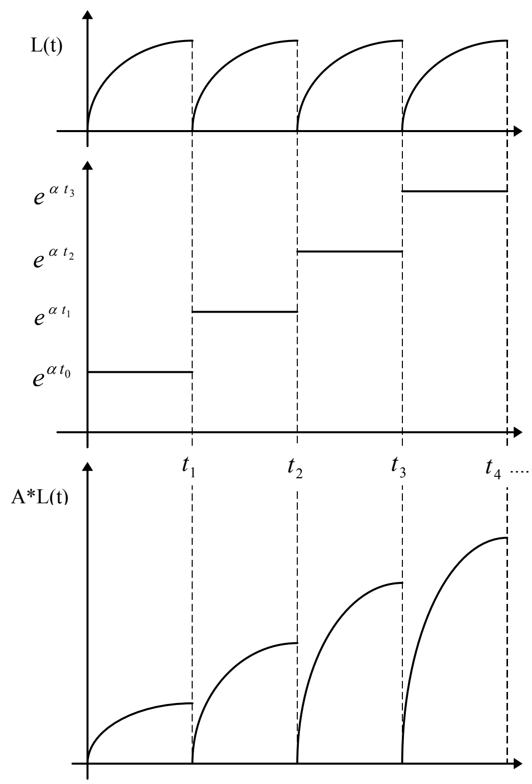
10 This is a usual assumption in the exiting models. For example, Cooley, Greenwood, and Yorukoglu [1997] or Klenow [1997] assumes that a plant can use only one technology at a time, which raises the question of when to replace the old technology with the new one.

11 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.

12 In other words, experience is technology-specific. An empirical justification of this assumption is found in Irwin and Klenow [1994]. Using firm-level data on seven generations of DRAM semiconductor chips over 1974-92, they showed that intergenerational learning spillovers are weak. One can also consider partial transfer of experience, but this strong assumption of zero-transferability makes the calculation much simpler without sacrificing the spirit of the

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 : $L(t) = 1 - e^{-\gamma t}$.[13,14]

Figure 3. Repeated technology upgrades and productivity



model. One can easily consider partial transfer of experience, under a more general assumption that experience is transferred to the next generation technology with the ratio of ($1 - \delta$) and that $0 \leq \delta \leq 1$. By defining β' as ($\beta + c\delta$) and substituting it for β, all the following equations can easily include a more general case allowing for partial transferability. In this case, a change in δ has the qualitatively same effect as a change in β.

13 This is a typical functional form of deterministic learning process. See Parente [1994] and Jovanovic and Nyarko [1995] for the differential equation version of this representation. Its discreet time version is found in Hornstein and Krusell [1996] and in Yorukoglu [1998]. A major limitation of this type of learning process is that it is given as an exogenous process. Learning process can be made endogenous by introducing investment in learning (as in Greenwood and Yorukoglu [1996]) or by making it dependent upon cumulative output (as in Klenow [1998]).

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

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**)

In calculating the net present value, we assume that the discount rate is given and constant for simplicity.

Assumptions 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.*

Intuitively, a constant time interval seems consistent with the exponential and multiplicative functional form of the optimisation problem which makes the choice of the (n + 1)-th upgrading timing at the moment of n-th upgrading look alike for any n. For more formal proof,

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:

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

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:

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

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

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

$$V = \sum_{n=1}^{\infty} V_n e^{-\rho(n-1)\lambda} = \frac{\Psi-\beta}{1-e^{-\lambda(\rho-\alpha)}} \cong \frac{\Psi-\beta}{\lambda(\rho-\alpha)}.^{15} \tag{2.4}$$

2.3 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.

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

From the first order condition,

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

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

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

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.

---

15 The approximation for the last term in equation (2.4) comes from the first order Taylor approximation, $e^{-x} \cong 1-x$ when x is close enough to 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,[16] 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.

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, 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

[16] For example, Moore's law says computer processing speed has been doubled every 18 months.

IT capital, while installed IT capital will again help a company to reduce its intangible upgrade costs in subsequent technology adoptions. One can also consider the skilled-worker share instead of the IT-capital share as a proxy variable for a company's learning ability. For example, the learning equation in the model by Greenwood and Yorukoglu [1997] is based on the idea that the amount of skilled labour hired by the plant would facilitate the adoption of a new technology. In this paper, IT capital rather than skilled labour was used as a proxy for the learning ability in the following regressions, simply because this information was available from the obtained data set. Empirical evidence of skill-technology complementarity is found in Bartel and Lichtenberg [1987] and Autor, Katz, and Krueger [1997], among others.

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. According to the notion of investment-specific technological progress, technological progress is embodied in the form of new capital goods and investment is a central measure of technology upgrading. Of course, some investment is obviously unrelated with technological progress. Investment for simply expanding the existing capacity will be a good example. However, observations that technology use is virtually uncorrelated with the plant age and that investment has a distinct spiked pattern suggest that incumbent plants also update their technologies by installing new capital goods.[17]

With IT capital intensity and capital turnover rate as proxies for learning ability and upgrading frequency, respectively, the following three testable predictions are derived 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*.

---

17 See Cooley, Greenwood, and Yorukoglu [1997] and Klenow [1998].

- *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.

- *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.

- **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.

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)

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)

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.

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

The notion 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.

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.[18]

[18] 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

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[19], 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.).

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.[20]

$$\ln UC_t = c_0 - c_1 \ln Q_t \tag{4.1}$$

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.

---

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.

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

20 For example, see Argote and Epple [1990] and references there.

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

(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:

$$\begin{cases} \ln Y_{1t} = c_1 + \gamma_1 \ln Q_{1t} + \delta_1 \cdot t + \alpha_1 \ln K_{1t} + \beta_1 \ln L_{1t} + \varepsilon_{1t} \\ \ln Y_{2t} = c_2 + \gamma_2 \ln Q_{2t} + \delta_2 \cdot t + \alpha_2 \ln K_{2t} + \beta_2 \ln L_{2t} + \varepsilon_{2t} \\ \cdots\cdots \\ \ln Y_{15t} = c_{15} + \gamma_{15} \ln Q_{15t} + \delta_{15} \cdot t + \alpha_{15} \ln K_{15t} + \beta_{15} \ln L_{15t} + \varepsilon_{15t} \end{cases} \tag{4.3}$$

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.

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 Coordination Agency.[21]

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.

- **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 Findings

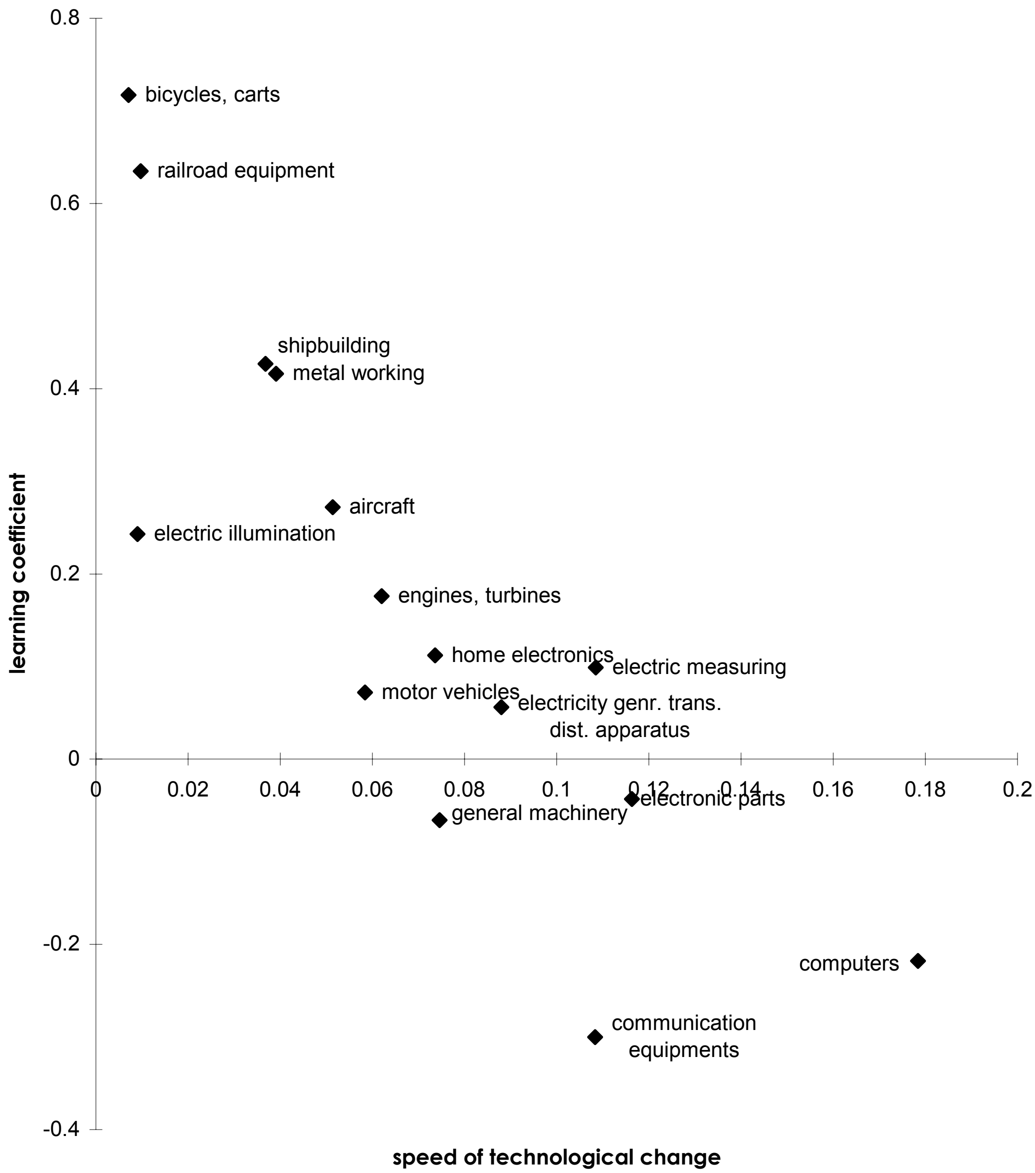
The main result of the estimation is summarised in **Figure 4**.[22] It shows an interesting relation between the learning coefficient and the speed of technological change. Industry-wide learning-by-doing effects were most pronounced in low-tech industries where technological change

[21] 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*.

[22] Due to the lack of space, more detailed tables for regression results are not reported here but available upon request.

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.

Figure 4. Speed of Technological Change and Learning-by-Doing



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.

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)

Data suggest 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 across different sectors. 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

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.[23]

Figure 5. TFP Growth Rate and Capital Growth Rate

(General Machinery Industry, Japan)

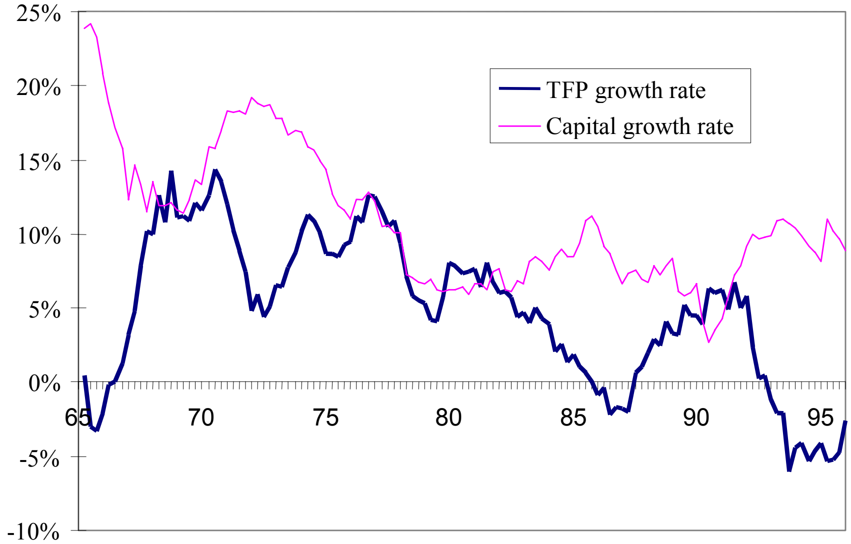
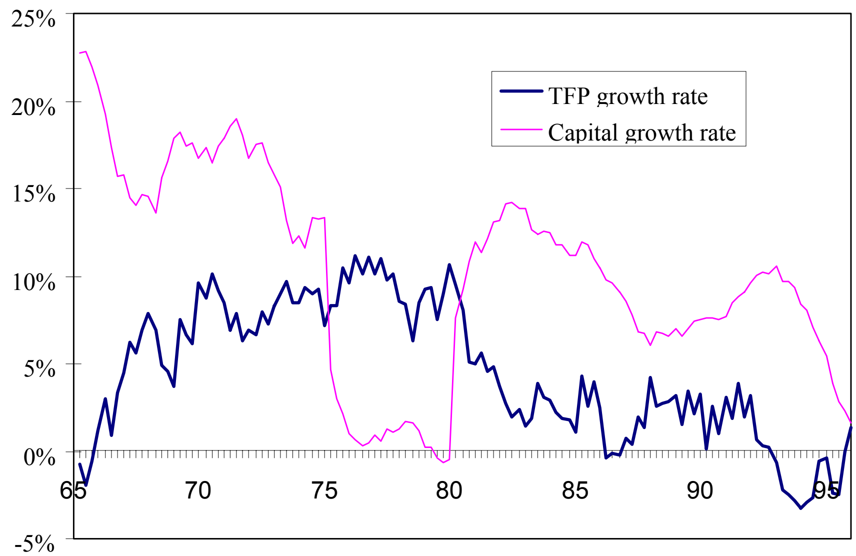


Figure 6. TFP Growth Rate and Capital Growth Rate

(Transportation Equipment Industry, Japan)



23 See Appendix 2, for methodology and data sources.

Figure 7. TFP Growth Rate and Capital Growth Rate

(Boilers, Engines, and Turbines, Japan)

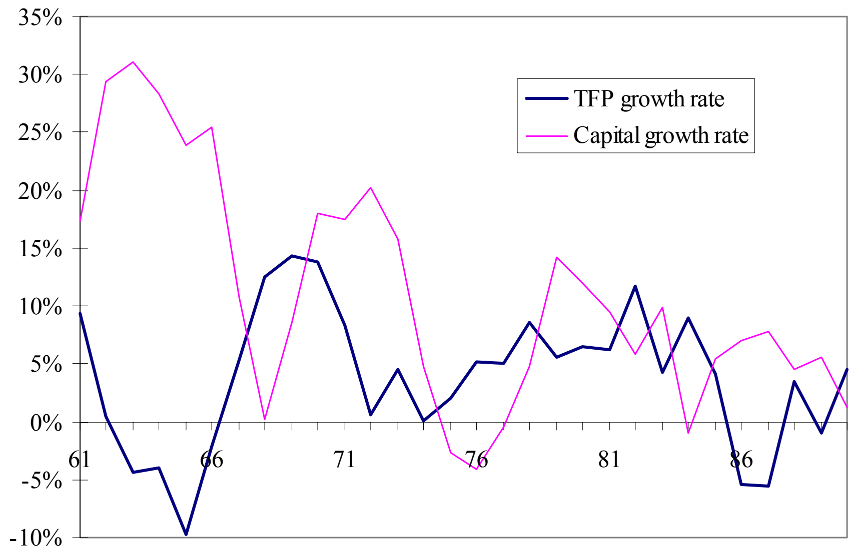


Figure 8. TFP Growth Rate and Capital Growth Rate

(Motor Vehicles and Equipment, Japan)

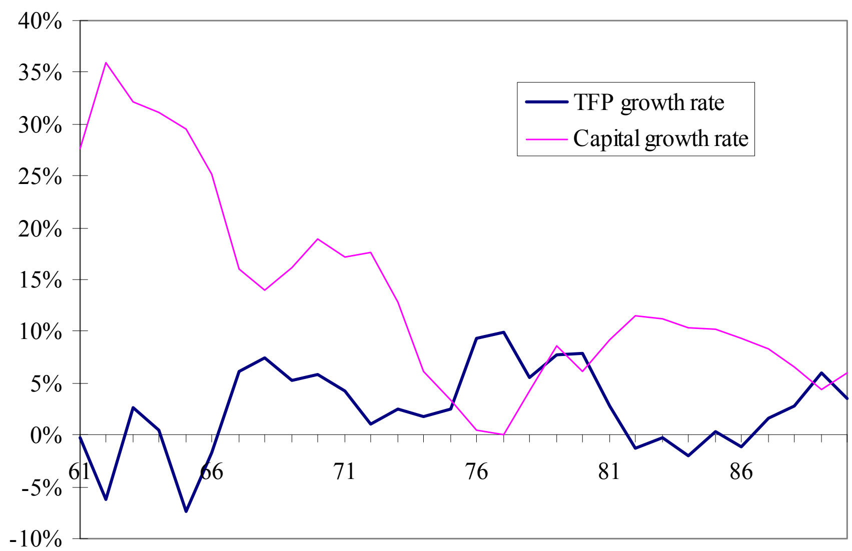


Figure 9. TFP Growth Rate and Capital Growth Rate
(Engines and Turbines: SIC 3510, US)

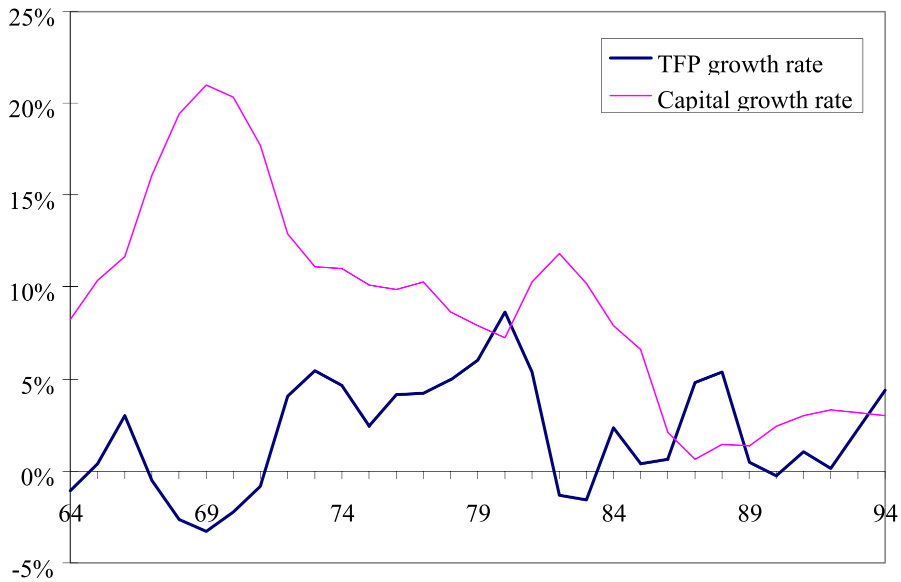


Figure 10. TFP Growth Rate and Capital Growth Rate
(Motor Vehicles and Car Bodies: SIC 3711, US)

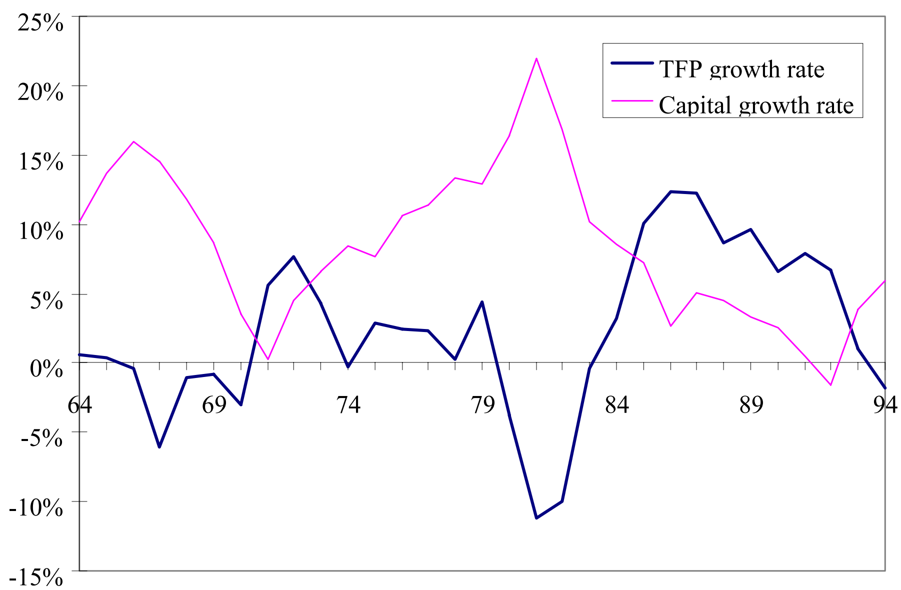


Table 4. Capital Accumulation and Productivity Growth (Major Industries, Japan)

*: 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 apparatus	-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

The immediate impact of increasing investment growth on productivity growth appears to be negative. Similarly, Power (1998) also found that there was little evidence of a robust positive correlation between recent investment spike and productivity level, from examining the relationship between investment age, plant age, and productivity using the Longitudinal Research Database (LRD) in the US covering 14,000 plants over 18-year period. Then, how to explain this finding that investment drive could decelerate productivity growth? 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 this kind of "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 try to 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.

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 effect from utilising the existing technology or technology upgrading effect from adopting a new technology with higher productivity potential. To adopt a new technology, however, one must buy new capital goods which embody the new technology.[24] 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:

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

This specification is based on the following assumptions.

Assumptions 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.*

[24] In this respect, this simple model shares some main features of a vintage capital model in Colley, Greenwood, and Yorukoglu [1997]: i) Technological progress is embodied in the form of new capital goods; ii) The firm should decide when to replace its existing capital with a new vintage; and therefore iii) Investment is a lumpy decision and capital only disappears because of replacement.

Assumptions 6.2 (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.

Assumptions 6.3 (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.*[25] *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:*

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

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.[26]

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

Assumptions 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.

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

25. A machine replacement model by Cooper, Haltiwanger, and Power [1999], for example, makes very similar assumptions. In our case, capital expenditure for new technology adoption is just like the upgrading cost in the technology upgrading model in Section 2.

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

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

From equation (6.3) and equation (6.4), the industry-level productivity growth rate (which is the same as the output growth rate)[27] and the industry-level capital growth rate can be calculated. Now, once parameters [λ,α,γ] 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.

Now, we need to specify the probability density function for the distribution of individual producers in the industry. As a simple way of mimicking cyclical 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.[28] 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.)[29] This specification simply captures a situation where aggregate investment moves up and down along the cycle. In this case, the cyclical movement 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 (λ = 5) and that the exogenous technological progress enables frontier technologies to increase potential productivity by a 10% growth rate (α = 0.1). The third parameter γ is supposed to be 1.[30] 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 *Correlation* [$\Delta^2 \ln(Y)$, $\Delta^2 \ln(K)$]. With given parameter values, such

---

27 See Assumption 6.1.

28 The links between investment-specific technological changes and the business cycle have been analysed in many studies. Recent studies include: Andolfatto and MacDonald [1998]; Campbell [1998]; Cooper and Haltiwanger [1993]; Cooper, Haltiwanger, and Power [1999]; Jovanovic and Lach [1997]; and Klenow [1998].

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

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

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

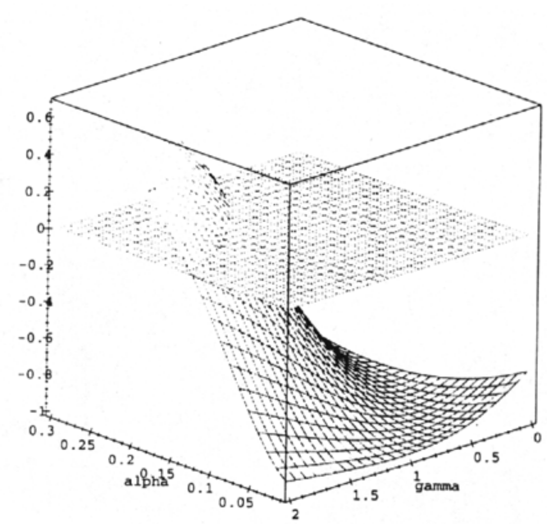
stochastic simulation has obtained *Correlation* [$\Delta^2 \ln(Y)$, $\Delta^2 \ln(K)$] = -0.791. **Table 7** shows that this correlation is negative most of the cases.

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



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. However, it should be emphasised that such seemingly optimistic conclusion from a deterministic learning model masks the existence of huge uncertainty and underlying heterogeneity accompanying technology upgrading. Adopting a new technology is a kind of investment activity in the sense that it pays short term costs expecting long term returns. Like any other investment activities, adopting a new technology involves risks and is sensitive to investors' perceptions regarding such risks. One important direction for further studies would be incorporating uncertainty into the model. Our model provides an interesting perspective on the large swing in the stock market observed in the late 1990s, in comparison with recent studies: (1) The IT revolution and the stock market (Greenwood and Jovanovic, 1999; Hobjin and Jovanovic, 2001); and (2) IT investment and intangible assets, namely, organizational capital (Brynjolfsson *et al*., 2002; Atkeson and Kehoe, 2002) or "e-capital" (Hall 2000, 2001b). Following the spirit of Hall [2001a], a follow-up study will offer an explanation for the recent boom and bust in the stock market as a rational response to new information affecting perceived learning cost for technology upgrading.

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:

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

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

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.

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

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

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

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.

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

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.

$$\begin{aligned}
&\max_{\{\lambda_n, n=1,2,3...\}} \left\{(\Psi(\lambda_1)-\beta)+(\Psi(\lambda_2)-\beta)e^{(\alpha-\rho)\lambda_1}+(\Psi(\lambda_3)-\beta)e^{(\alpha-\rho)(\lambda_1+\lambda_2)}+......\right\} \\
&=\max_{\{\lambda_1\}}\left[(\Psi(\lambda_1)-\beta)+\max_{\{\lambda_n, n=2,3,...\}}\left\{(\Psi(\lambda_2)-\beta)e^{(\alpha-\rho)\lambda_1}+(\Psi(\lambda_3)-\beta)e^{(\alpha-\rho)(\lambda_1+\lambda_2)}+......\right\}\right] \\
&=\max_{\{\lambda_1\}}\left[(\Psi(\lambda_1)-\beta)+e^{(\alpha-\rho)\lambda_1}\max_{\{\lambda_n, n=2,3,...\}}\left\{(\Psi(\lambda_2)-\beta)+(\Psi(\lambda_3)-\beta)e^{(\alpha-\rho)\lambda_2}+......\right\}\right]
\end{aligned} \tag{A1.8}$$

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:

$$\lambda_2^* = \lambda_2^{**}. \tag{A1.9}$$

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

$$\lambda_1^* = \lambda_2^*. \tag{A1.10}$$

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).

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

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.

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

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 $LHS(0)=0, LHS(\infty)=0$. *LHS* in **Figure A1.1** is derived from **Figure A1.2** and shows that *LHS* intersects the horizontal axis (i.e., *RHS* when β is 0) from the below just once. When the upgrade cost is positive, *RHS* is represented as a hyperbola. In this case, *LHS* still intersects *RHS* from the below just once, under the second part of **Assumption 2.2** that upgrade cost β is not prohibitively high (i.e., $\beta < \rho^{-1} - (\rho+\gamma)^{-1}$]).

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

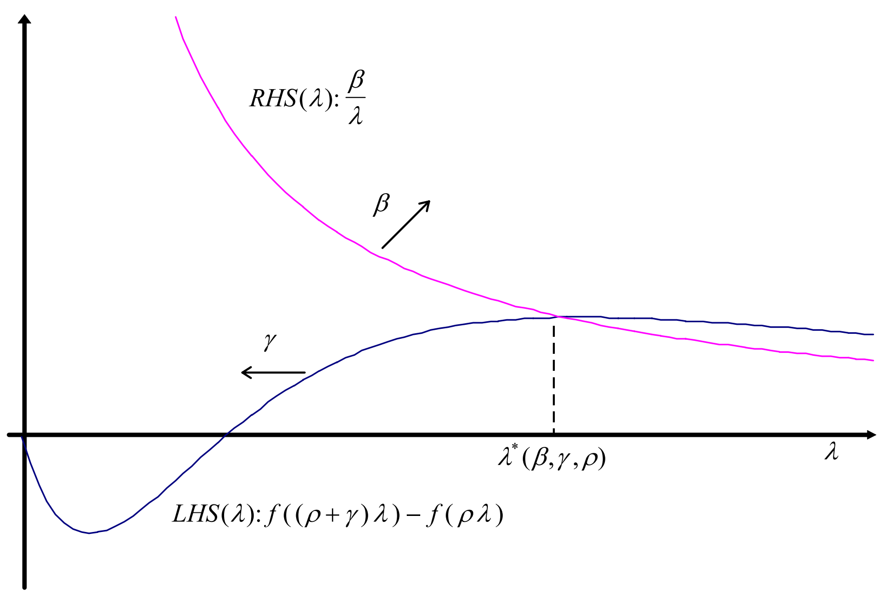
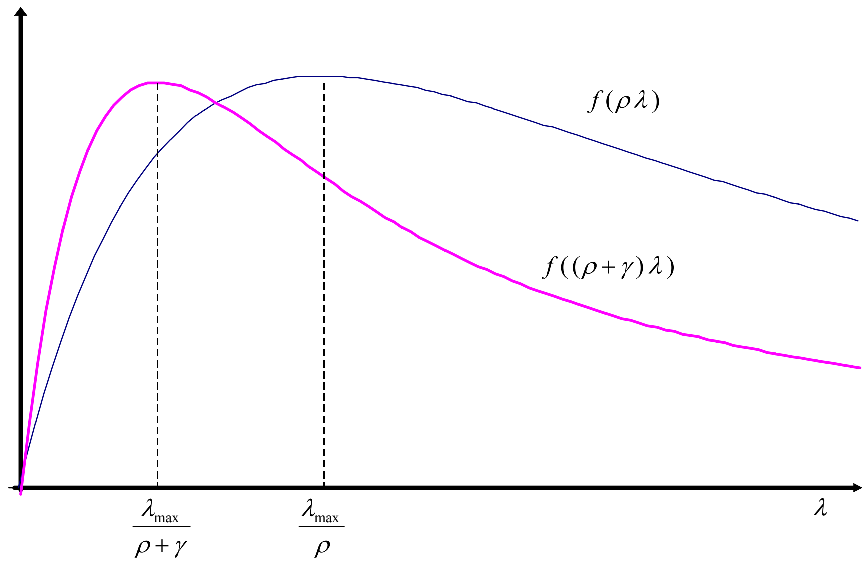


Figure A1.2 $f(\rho\lambda)$ *and* $f((\rho+\gamma)\lambda)$



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 $g'(x) = x e^{-x} > 0$ if $x > 0$. Therefore, $g(x) > 0$ if $x > 0$.

Therefore, in equation (A1.13),

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

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.

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

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:

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

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

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.

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

Under the assumption that the factor markets have perfect competition,

$$p F_K = r, \quad p F_L = w \qquad \text{(A2.3)}$$

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,

$$K F_K + L F_L = F \,. \qquad \text{(A2.4)}$$

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-1) can be calculated from the second order translog approximation of the production function.[31] 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

[31] For more detail, see Young [1995] and references there.

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 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

Andolfatto, David and Glenn M. MacDonald [1998], "Technology Diffusion and Aggregate Dynamics," *Review of Economic Dynamics*, 1, 338-70.

Atkeson, Andrew and Patrick J. Kehoe [2001], "The Transition to a New Economy after the Second Industrial Revolution", *NBER Working Paper Series*, No.8676.

________________________________ [2002], "Measuring Organizational Capital", *NBER Working Paper Series*, No.8676.

Autor, David H., Lawrence F. Katz, and Alan B. Krueger [1997], "Computing Inequality: Have Computers Changed the Labor Market?" *NBER Working Paper Series*, No.5956.

Bartel, Ann P. and Frank Lichtenberg [1987], "The Comparative Advantage of Educated Workers in Implementing New Technologies," *Review of Economics and Statistics*, 69, 1-11.

Basu, Susanto, John G. Fernald, and Mathew D. Shapiro [2001], "Productivity Growth in the 1990s: Technology, Utilization, or Adjustment?" *Carnegie-Rochester Conference Series on Public Policy*, 55, 117-165.

Bresnahan, Timothy F. and Manuel Trajtenberg [1995], "General Purpose Technologies: 'Engine of Growth'?" *Journal of Econometrics*, 65, 83-108.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt [1999], "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence", *NBER Working Paper Series*, No.7136.

Brezis, Krugman, and Tsiddon [1993], "Leapfrogging in International Competition: A Theory of Cycles in National Technological Leadership," *American Economic Review*, 83(5).

Brynjolfsson, Erik, Lorin Hitt [1996], "Productivity, Business Profitability, and Consumer Surplus: Three Different Measures of Information Technology Value," *MIS Quarterly* (June 1996).

Brynjolfsson, Erik, Lorin Hitt, and Shinkyu Yang [2002], "Intangible Assets: Computers and Organizational Capital," *Brookings Papers on Economic Activity*, 1: 2002, 137-98.

Campbell, Jeffrey R. [1998], "Entry, Exit, Embodied Technology, and Business Cycles," *Review of Economic Dynamics*, 1, 371-408.

Cooley, Thomas F., Jeremy Greenwood, and Mehmet Yorukoglu [1997], "The Replacement Problem," *Journal of Monetary Economics*, 40, 457-99.

Cooper, Russell and John Haltiwanger [1993], "The Aggregate Implications of Machine Replacement: Theory and Evidence", *American Economic Review*, 83(3), 360-82.

Cooper, Russell, John Haltiwanger, and Laura Power [1999], "Machine Replacement and the Business Cycle: Lumps and Bumps," *American Economic Review*, 89(4), 921-46.

David, Paul A. [1990], "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *American Economic Review*, 80(2), 355-61.

Gordon, Robert J. [2000], “Does the New Economy Measure Up to the Great Invention of the Past?” *Journal of Economic Perspectives*, 14(4), 49-74.

______________ [2002], “High-tech Innovation and Productivity Growth: Does Supply Create Its Own Demand?” *mimeo*.

Greenwood, Jeremy and Boyan Jovanovic [1999], “The Information-Technology Revolution and the Stock Market,” *American Economic Review Papers and Proceedings*, 89(2), 116-22.

Greenwood, Jeremy and Mehmet Yorukoglu [1997], “1974,” *Carnegie-Rochester Conference Series on Public Policy*, 46, 49-95.

Hall, Bronwyn H. [1993], “The Stock Market's Valuation of R&D Investment during the 1980s,” *American Economic Review*, 83(2), 259-64.

_______________ [2000], “Innovation and Market Value,” in Barrell, Ray, Geoffrey Mason, and Mary O’Mahoney (eds.), *Productivity, Innovation and Economic Performance*, Cambridge: Cambridge University Press, 177-98.

Hall, Robert E. [2000], “E-Capital: The Link between the Stock Market and the Labor Market in the 1990s, *Brookings Papers on Economic Activity*, 2:2000, 73-102.

___________ [2001a], “Struggling to Understand the Stock Market,” *American Economic Review Papers and Proceedings*, 91(2), 1-11.

___________ [2001b], “The Stock Market and Capital Accumulation,” *American Economic Review*, 91(5), 1185-1202.

Helpman, Elhanan and Antonio Rangel [1998], “Adjusting to a New Technology: Experience and Training,” *NBER Working Paper Series*, No.6551.

Helpman, Elhanan and Manuel Trajtenberg [1996], “Diffusion of General Purpose Technologies,” *NBER Working Paper Series*, No.5773.

Hobjin, Bart and Boyan Jovanovic [2001], “The Information-Technology Revolution and the Stock Market: Evidence,” *American Economic Review*, 91(5), 1203-20.

Hornstein and Krusell [1996], “Can Technology Improvements Cause Productivity Slowdowns?”, *NBER macroeconomics annual 1996*, 209-59.

Irwin, Douglas A. and Peter J. Klenow [1994], “Learning-by-Dong Spillovers in the Semiconductor Industry,” *Journal of Political Economy*, 102(6), 1200-1227.

Jovanovic, Boyan [1997], “Learning and Growth,” *Advances in Economics*, D. Kreps and K.F. Wallis (eds.), Cambridge University Press.

Jovanovic, Boyan and Saul Lach [1997], “Product Innovation and the Business Cycle,” *International Economic Review*, 38(1), 3-22.

Jovanovic, Boyan and Yaw Nyarko [1995], “A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves,” *Brookings Papers on Economic Activity: Microeconomics*, 247-99.

____________________________ [1996], “Learning by Doing and the Choice of Technology,” *Econometrica*, 64(6), 1299-1310.

Klenow, Peter J. [1998], "Learning Curves and the Cyclical Behavior of Manufacturing Industries," *Review of Economic Dynamics*, 1, 531-50.

Oliner, Stephen D., and Daniel E. Sichel [1994], "Computers and Output Growth Revisited: How Big Is the Puzzle?" *Brookings Papers on Economic Activity, 2: 1994*, 273-334.

Parente, Stephen L. [1994], "Technology Adoption, Learning-by-Doing, and Economic Growth, " *Journal of Economic Theory*, 63, 346-69.

Power, Laura [1998], "The Missing Link: Technology, Investment, and Productivity," *Review of Economics and Statistics*, 300-13.

Solow, Robert M. [1987], "We'd Better Watch Out," *New York Times Book Review*, (July 12, 1987), p.36.

Yorukoglu, Mehmet [1998], "The Information Technology Productivity Paradox," *Review of Economic Dynamics*, 1, 551-92.

Zeckhauser, Richard [1968], "Optimality in a World of Progress and Learning," *Review of Economic Studies*, 35, 363-65.