CIRJE-F-350

Innovation Versus Diffusion:
Determinants of Productivity Growth Among Japanese Firms

Kiyohiko G. Nishimura
Policy Board, Bank of Japan
Takanobu Nakajima
Keio University
Kozo Kiyota
Yokohama National University

June 2005

Innovation Versus Diffusion: Determinants of Productivity Growth Among Japanese Firms*

Kiyohiko G. Nishimura†
Takanobu Nakajima‡
Kozo Kiyota§

June 19, 2005 version

Abstract

This paper presents a model of firm-level productivity growth that distinguishes between innovation and technology diffusion, and then applies the model to a large-scale data set of Japanese manufacturing and non-manufacturing firms between 1994 and 2000. We find both innovation and diffusion are important factors in firm-level productivity growth. Results also suggest that innovation comes not only directly from R&D activities, but also indirectly from patent purchases and imports. Previously, patent purchases and imports were considered as sources of technology diffusion rather than innovation. In fact, we find patent purchases are more effective in this regard than R&D expenditure.

1 Introduction

Productivity growth involves both innovation and diffusion. Innovation resulting in new products and novel production processes has been considered as one of the most important determinants of growth for capitalist economies since the age of Schumpeter.[1] In particular, research and development (R&D) has received considerable attention as an economic activity that produces innovation

*This paper is a part of our ongoing joint research with Flora Bellone, Patrick Musso, and Michel Quéré of Institut de Droit et d'Economie de la Firme et de l'Industrie. We wish to thank Flora Bellone, Dale Jorgenson, Tsutomu Miyagawa, Patrick Musso, Lionel Nesta, and Michel Quéré for helpful comments on an earlier version of the paper. Views expressed here are strictly those of the authors and in no way represent those of the Bank of Japan, the University of Tokyo, Keio University, or Yokohama National University.

†Policy Board, Bank of Japan, and Manufacturing Management Research Center, Faculty of Economics, University of Tokyo.

‡Faculty of Business and Commerce, Keio University

§Faculty of Business Administration, Yokohama National University, and Manufacturing Management Research Center, Faculty of Economics, University of Tokyo.

[1] See Schumpeter (1934).

Figure 1: Sources of Productivity Growth: Traditional View

- Innovation
  - Research and Development (R&D)
- Diffusion
  - Explicit Emulation ("Active" Diffusion)
    (spillover through patent purchases, imports, etc.)
  - Implicit Emulation ("Passive" Diffusion)
    (autonomous productivity convergence through learning-by-doing)

and is, subsequently, viewed as an engine of growth. (Griliches, 1998; Hulten, Dean and Harper, eds, 2001). Indeed, possible links between R&D activities and productivity growth are central issues in recent strands of growth theory (Grossman and Helpman, 1991; Barro and Sala-i-Martin, 2004, chapter 8).

Technology diffusion is also an important determinant of productivity growth. If firms are quick to emulate the performance of the industry leader, we would expect faster productivity improvement in this industry than would otherwise occur, assuming that the performance leader adopts the most advanced technology and management systems. Industry productivity performance in a country which follows this path is likely to rapidly achieve results similar to the best world performer. If the leading firm in a country is on the cutting-edge of technology in the world, and if technology diffusion is fast in its country's industry, that country can outdistance other countries.

A traditional view about the sources of productivity growth is summarized in Figure 1. There, besides innovation, technology diffusion can be further divided into two groups. One is explicit emulation, which can be described as "active" technology diffusion to adopt new technology. Typical channels are patent purchases and the imitation of technology embodied in imports. The other is implicit emulation, which is depicted as a "passive" technology diffusion, or, in other words, "autonomous" productivity convergence. One typical mode of this type of diffusion is productivity catch-up through learning-by-doing.

A number of studies have examined the effects of explicit emulation, especially focusing on the role of imports.[2] The foreign knowledge embodied in the large variety of intermediate products

[2]For instance, Coe and Helpman (1995) examined the effects of R&D spillovers through imports among 21 OECD countries plus Israel. They found positive effects on productivity growth through the spillovers of international R&D

and capital equipment enables countries to boost their productivity growth. However, few studies incorporate both explicit and implicit emulation at the same time. None of these studies have combined the effects of innovation and explicit and implicit emulation in one coherent framework of firm-level productivity growth.

This paper examines the growth of productivity at the firm level, distinguishing between the effects of innovation and those of technology diffusion (explicit and implicit emulation) and then investigates possible determinants of innovation and diffusion, quantitatively in a large-scale data set for both manufacturing and non-manufacturing firms. The data used in this paper is the micro database of *Kigyou Katsudou Kihon Chousa Houkokusho* (*The Results of the Basic Survey of Japanese Business Structure and Activities*) prepared by the Research and Statistics Department, Ministry of Economy, Trade and Industry (METI) for the period 1994-2000.

We find that not only innovation but also technology diffusion is an important source of the productivity growth. The results indicate that, after controlling for the effects of innovation, there exists strong evidence of productivity convergence among firms in most industries. This clearly implies that the technological knowledge of the most advanced firm(s) spills over to other firms and that technological diffusion is one of major determinants of productivity growth.

As for innovation, R&D expenditure has a significantly positive effect on productivity growth as expected, but the source of innovative improvement of productivity is not limited to R&D activities. In fact, "the new impetus of thoughts and the effects" brought by patent purchases are shown to be more effective and stronger than R&D activities in producing innovation-related productivity growth. Imports are also found as an important innovation factor, though the effects of imports are weaker than those of R&D and patent purchases. With respect to technological diffusion, we find that there is strong evidence that imports speed up the productivity convergence process.

Finally, the productivity convergence in technology diffusion is stronger in information and communication technology (IT) industries than non-IT industries. We were not able to establish the difference of impacts of innovation on IT and non-IT industries. The results imply that the difference of the productivity growth between IT and non-IT industries results from "autonomous" productivity convergence through learning-by-doing.

The organization of this paper is as follows. In Section 2, we present a model of a firm's productivity growth, distinguishing between innovation and technology diffusion. The estimation results of the model are presented in Section 3. In Section 4, we discuss the implication of our

from imports. Similarly, Lee (1995) found positive impacts on per capita income from R&D spillovers through capital goods trade.

results. Concluding remarks are presented in Section 5.

2 Innovation versus Diffusion: A Model

The starting point for our discussion of firm-level productivity growth is a model of productivity convergence proposed by Bernard and Jones (1996), which is extensively examined in the literature including our own companion paper (Nishimura, Nakajima and Kiyota, 2005). Let us denote total factor productivity (TFP) for a firm i in year t as θ_{it}. Then, TFP growth is assumed to be described as:

$$\ln \theta_{it} = \gamma_i + \lambda \left\{ \ln \theta_{1t-1} - \ln \theta_{it-1} \right\} + \ln \theta_{it-1} + \ln \epsilon_{it}, \tag{1}$$

where $\ln \theta_{1t-1} - \ln \theta_{it-1}$ represents a catch-up variable, which represents the distance in productivity between the most productive firm, denoted by 1, and a firm in question, i. The speed of catch-up therefore is captured by λ while the asymptotic rate of productivity growth of firm i is denoted by γ_i. Finally, $\ln \epsilon_{it}$ represents a disturbance term.

This formulation captures "passive" technology diffusion, in which technological knowledge spreads out without costly efforts by firms trying to catch-up to the leader(s). The best example of this kind of technology diffusion may be learning-by-doing. Catching-up firms may improve their productivity by emulating the best practices of the most advanced firm without incurring significant costs.

There are, however, other conceivable determinants of productivity growth as suggested in the Introduction. Firstly, R&D is a particular effort to improve productivity in terms of product value as well as production cost. R&D activities "enlarge" the production possibility frontier of even the most advanced firm, and can be considered as a leapfrogging factor. Secondly, even technological diffusion or catch-up has an "active" form. For example, "active" catching-up factors are patent purchases from cutting-edge technology firms and/or to emulating advanced technology through imports.

To incorporate these leapfrogging factors, such as innovation, and "active" diffusion factors, such as patent purchases, we assume that TFP growth is described by

$$\ln \theta_{it} = \gamma_i + \lambda_{it} \left\{ \ln \theta^*_{it-1} - \ln \theta_{it-1} \right\} + \ln \theta_{it-1} + \ln \varepsilon_{it}. \tag{2}$$

There are two basic differences between equations (2) and (1). Firstly, the speed-of-convergence, λ_{it} , is now dependent on i and t, reflecting "active" catching-up activities of the firm. As suggested before, the firm can have influence on the speed-of-convergence, λ_{it}, so that this term now depends on time t and has the form

$$\lambda_{it} = \lambda_0 + \sum_j \lambda_j \, DiffusionFactor_{ij,t-1}, \tag{3}$$

where $DiffusionFactor$ denotes a determinant of the diffusion speed, which we will consider in the next section.

Secondly, the firm's technological distance from the most advanced firm (denoted by 1) $\ln \theta_{1t-1} - \ln \theta_{it-1}$ in equation (1) is replaced by $\ln \theta^*_{it-1} - \ln \theta_{it-1}$ in (2), where θ^*_{it-1} is the productivity level that the firm can achieve, in an ideal case in which the firm has all explicit and implicit know-hows, and technological expertise in producing its products so that there is no need for catching up. This is a "target" productivity level. In an environment of homogeneous products and homogeneous technology, this is the productivity level of the most productive firm, θ_{1t}. This is the case implicitly assumed in Bernard and Jones as well as others.

The conceivably best productivity level of the firm i in period t, $\ln \theta^*_{it}$, is assumed to follow a stochastic process

$$\ln \theta^*_{it} = \Theta^*_{it} + \ln \theta^*_{it-1} + \ln \varepsilon^*_{it}, \tag{4}$$

where technological advance Θ^*_{it} represents innovation that is determined by R&D activities and other determinants (denoted by $TechChangeFactor$) such that

$$\Theta^*_{it} = \zeta_0 + \sum_k \zeta_k \, TechChangeFactor_{ik,t-1}. \tag{5}$$

In the next section, we will consider these determinants of innovation that bring about technological advance.

This extended framework of productivity growth yields a qualitatively similar model to the baseline "passive" diffusion model of Bernard and Jones. Letting $\widehat{\theta}^*_{it} = \theta_{it}/\theta^*_{it}$ and $\widehat{\varepsilon}^*_{it} = \varepsilon_{it}/\varepsilon^*_{it}$ we have

$$\ln \widehat{\theta}^*_{it} = (\gamma_i - \Theta^*_{it}) + (1 - \lambda_{it}) \ln \widehat{\theta}^*_{it-1} + \ln \widehat{\varepsilon}^*_{it},$$

which implies the average TFP growth rate of firm i relative to the best level is

$$\begin{aligned} \frac{\ln \widehat{\theta}^*_{iT} - \ln \widehat{\theta}^*_{i0}}{T} &= -\frac{1}{T} \left\{ 1 - \left(\prod_{s=1}^{T} (1 - \lambda_{i,T+1-s}) \right) \right\} \ln \widehat{\theta}^*_{i0} \\ &\quad + \frac{1}{T} \sum_{\tau=0}^{T-1} \left(\prod_{s=1}^{\tau} (1 - \lambda_{i,T+1-s}) \right) \left(\gamma_i - \Theta^*_{i,T-\tau} + \ln \widehat{\varepsilon}^*_{i,T-\tau} \right) \end{aligned} \tag{6}$$

with a convention that $\Pi^{\tau}_{s=1} (1 - \lambda_{i,T+1-s}) = 1$ when $\tau = 0$.

Consequently, substituting $\ln \widehat{\theta}_{it}^{*} = \ln \theta_{it} - \ln \theta_{it}^{*}$ into (6) and then rearranging terms, we have

$$\Delta \ln \theta_{iT} \equiv \frac{\ln \theta_{iT} - \ln \theta_{i0}}{T} = \beta_0 + \beta_1 \ln \theta_{i0} + \mu_{iT},$$

where coefficients β_0 and β_1 are determined by the determinants of "active" diffusion ($DiffusionFactor$) and the determinants of innovation ($TechChangeFactor$), respectively.

Firstly, we have

$$\beta_1 = \frac{-\left[1 - \left\{1 - \overline{\lambda}_i\left(1, T\right)\right\}^T\right]}{T},$$

where $\overline{\lambda}_i\left(1, T\right)$ is the "average" rate of productivity convergence such that

$$\left\{1 - \overline{\lambda}_i\left(1, T\right)\right\}^T = \prod_{s=1}^{T}\left(1 - \lambda_{i,T+1-s}\right).$$

Because of (3), it is natural to assume that $\overline{\lambda}_i$ is determined by the time average of the determinants

$$\overline{\lambda}_i = \phi_0 + \sum_j \phi_j \overline{DiffusionFactor}_{ij}. \tag{7}$$

Taking the first-order approximation of β_1 with respect to $\overline{DiffusionFactor}_{ij}$ around $\overline{DiffusionFactor}_{ij} = 0$, we have

$$\beta_1 = \beta_{10} + \sum_j \beta_{11j} \overline{DiffusionFactor}_{ij},$$

where

$$\beta_{10} = \frac{-\left[1 - \left(1 - \phi_0\right)^T\right]}{T}$$

and

$$\beta_{11j} = -\left(1 - \phi_0\right)^{T-1} \phi_j.$$

Secondly, we have

$$\begin{aligned}\beta_0 = &\left(\frac{1}{T}\sum_{s=1}^{T}\Theta_{i,T+1-s}^{*}\right) \\ &\times\left\{1 + \sum_{\tau=0}^{T-1}\left\{1 - \overline{\lambda}_i\left(1, \tau\right)\right\}^{\tau}\frac{\left(\gamma_i - \Theta_{i,T-\tau}^{*}\right)}{\sum_{s=1}^{T}\Theta_{i,T+1-s}^{*}}\right. \\ &\left. + \left[1 - \left\{1 - \overline{\lambda}_i\left(1, T\right)\right\}^T\right]\frac{\ln \theta_{i0}^{*}}{\sum_{s=1}^{T}\Theta_{i,T+1-s}^{*}}\right\}.\end{aligned} \tag{8}$$

It should be noted that the rate of convergence $\overline{\lambda}_i$ influences β_0 indirectly, only through cross terms with the productivity difference $\gamma_i - \Theta_{i,T-\tau}^{*}$ and the initial best productivity level θ_{i0}^{*}. Thus, so long as productivity difference and the initial productivity level are small relative to accumulated productivity changes $\sum_{s=1}^{T} \Theta_{i,T+1-s}^{*}$, then we can approximate the term in the curly bracket in equation (8) as a constant. We hereafter use this approximation so that

$$\beta_0 = \psi_0 + \sum_k \psi_k \overline{TechChangeFactor}_{ik}.$$

Finally, the disturbance term μ_{iT} has the form:

$$\mu_{iT} = \sum_{s=1}^{T} \ln \varepsilon_{i,T+1-s}^{*} + \sum_{\tau=0}^{T-1} \left\{1 - \overline{\lambda}_i (1, \tau)\right\} \ln \widehat{\varepsilon}_{i,T-\tau}^{*}.$$

The above results show that an appropriate model of productivity, distinguishing between innovation and "active" as well as "passive" diffusion, is

$$\begin{aligned} \Delta \ln \theta_{iT} &\equiv \frac{\ln \theta_{iT} - \ln \theta_{i0}}{T} \\ &= \psi_0 + \sum_k \psi_k \overline{TechChangeFactor}_{ik} \\ &\quad + \left(\beta_{10} + \sum_j \beta_{11j} \overline{DiffusionFactor}_{ij}\right) \ln \theta_{i0} + \mu_{iT}. \end{aligned} \tag{9}$$

Here the implicit rate of "passive" technology convergence with no "active" diffusion, that is, λ when $\overline{DiffusionFactor}_{ij} = 0$, can be recovered from β_{10}, which is ϕ_0.

3 Specification and Estimation Results

3.1 Data and Measurement of Productivity

In this section, we apply the model described in Section 2 to a large-scale data set of Japanese firms in order to examine determinants of both innovation and diffusion. We use a micro database of *Kigyou Katsudou Kihon Chousa Houkokusho* (*The Results of the Basic Survey of Japanese Business Structure and Activities*) prepared by METI (1996-2002). The survey covers both manufacturing and non-manufacturing firms with more than 50 employees and with capital of more than 30 million yen. Classification of industries is at a 3-digit level. From this survey, we develop a longitudinal data set of firms for the years from 1994 to 2000. Using this data set, we construct each firm's TFP level by using a multilateral index method developed by Caves, Christensen and Diewert (1982) and extended by Good, Nadiri, Roeller and Sickles (1983).[3] Detailed information

[3]There is an alternative method that is based on the econometric estimation of gross production functions, which is proposed by Olley and Pakes (1996) and extended by Levinsohn and Petrin (2003). However, this framework has to specify a production function, although we do not have any reliable information about the specific functional form of a true production function. Moreover, because of the limited availability of intermediate inputs, their method was not feasible in practice. Consequently, we employ a multilateral index method described in this present study.

on this procedure is found in a companion paper (Nishimura et al., 2005), which itself follows a procedure taken by Nishimura, Nakajima and Kiyota (forthcoming).

3.2 Determinants and Controls

Let us now consider determinants of innovation ($TechChangeFactor$) and those of "active" diffusion ($DiffusionFactor$) in the equation (9). We first deal with innovation producing an outward shift of the production possibility frontier and then consider "active" diffusion speeding up productivity convergence. Finally, we examine other control variables such as adjustment costs, which are not explicitly considered in the past literature of productivity growth, but appear to be important in practice.

3.2.1 Determinants of Innovation

The most obvious source of innovation is the firm's own R&D activities ($R\&D$). Specifically, we measure the level of R&D activities by R&D expenditure scaled by sales. If innovation by R&D has positive effects on productivity growth, the coefficient of $R\&D$ should become significantly positive.

However, there are other possible determinants of innovation. It is often argued that a new combination of even old thoughts stimulates new ideas and thus are innovation-enhancing. Thus, patent purchases (involving "old ideas") might bring new ideas to develop the firm's own products and/or production processes. Moreover, imports of parts and equipment may reveal new approaches from foreign sources, which would enable the firm to innovate products and production processes. Even in non-manufacturing industries such as retail trade, imports of foreign merchandise may stimulate R&D activities to develop new lines of domestic merchandise.

From this perspective, both patent purchases (PAT) and imports (IMP) may not only speed up technology diffusion but also enhance innovation. Hence we modify the traditional view of innovation (Figure 1) as in Figure 2 and explore these possibilities as well.

We also evaluate synergetic effects between R&D and patent purchases. R&D activities might be more effective if they are combined with related technologies. However, firms do not necessarily have the related technologies by themselves. In that case, the R&D activities work more effectively when the related technologies are introduced through the purchase of technologies. To investigate the synergistic effects, we introduce the cross-term between R&D and patent purchases.

Figure 2: Sources of Productivity Growth: This Paper's View

Innovation — R&D, **Patent Purchases, Imports, etc.**

Diffusion
- Explicit Emulation ("Active" Diffusion)
  (spillover through patent purchases, imports, etc.)
- Implicit Emulation ("Passive" Diffusion)
  (autonomous productivity convergence through learning-by-doing)

3.2.2 Determinants of Explicit Emulation ("Active" Diffusion)

Let us now turn to explicit emulation. In past literature, two variables were considered to capture "active" diffusion thus speeding up technology diffusion. The first variable is patent purchases (PAT). The importance of patent purchases is fairly straightforward in technology diffusion because the use of patents means the direct purchase of technology through market transaction. Indeed, several studies such as Branstetter (2000) focused on the role of patents as a channel of technology diffusion. We measure the patent variable as patent payments scaled by sales.

The second variable is imports (IMP). Many economists believe that imports are one of the most important channels in international technology spillovers (e.g., Coe and Helpman, 1995; Lee, 1995). As was discussed in the Introduction, advanced foreign technology can be embodied in a wide variety of products, such as capital equipment. For example, the importation of these products enable countries to boost their productivity growth through emulation achieved by reverse-engineering. Based on this argument, we utilize a firm's imports (scaled by sales) as an "active" diffusion variable.

In addition, we introduce R&D expenditure as the third variable. Although R&D is traditionally regarded as an innovation factor, it can also be a catch-up factor. For instance, the follower firms have to invest in R&D to catch up to a leading firm if the imitation through imports is difficult. Similarly, when patent purchases generate costs higher than R&D activities, follower firms tend to conduct R&D for themselves. It is also not surprising that the introduction of new technology requires some efforts in R&D to enhance the capacity of the firm. We thus include that R&D expenditures are not only an innovation factor but also an "active" diffusion factor.

Another important channel might be direct foreign investment (FDI). For instance, Aitken and

Harrison (1999), Haskel, Pereira and Slaughter (2002), Keller and Yeaple (2004), and Javorcik (2004) focused on the role of FDI as a channel of explicit emulation, utilizing firm-level and/or establishment-level data. However, the effects of explicit emulation through FDI on productivity growth are ambiguous. While Keller and Yeaple (2004) and Javorcik (2004) confirmed that FDI led to substantial productivity gains for domestic firms, Aitken and Harrison (1999) and Haskel et al. (2002) did not find any evidence to support the spillovers from foreign-owned firms to domestic firms. This ambiguity results from the difficulty in the proper measurement of multinational activities. As Keller and Yeaple (2004) pointed out, the measurement makes a big difference in the estimation results. Given these considerations, and since there is no consensus on the proper measurement of FDI as a channel of emulation, this study focuses only on the effects of patents and imports.

3.2.3 Scale Effects and Adjustment Costs

In addition to innovation and diffusion, we controlled for scale effects and adjustment costs. Scale effects are captured by an employment scale (natural log, $\ln L$). Capital stocks may be another possible variable to control for the scale. However, the correlation between the scale of employment and that of capital stocks is high (0.71),[4] which causes multicollinearity. Hence we use the employment scale only. The coefficient of $\ln L$ is expected to be positive if scale effects exist.

We also control for the effects of adjustment costs. From a dynamic perspective, it is not easy for a firm to instantaneously allocate its inputs optimally because some of the inputs are quasi-fixed. This, in turn, implies the existence of adjustment costs, which drags productivity growth. Following Nakamura (1993), we employ the quadratic form of adjustment costs. Adjustment costs are measured by the square of the capital and employment growth ($\{\Delta \ln L\}^2$ and $\{\Delta \ln K\}^2$).

3.2.4 Industry Differences

To examine industry differences in the speed-of-convergence across industries, we include industry dummies in both constants and initial TFP levels. Therefore, industry j's speed-of-convergence is measured as the difference from a reference industry (say, industry 0). Constants and initial TFP levels are now represented as:

$$\beta_0^0 + \beta_0^1 D_{i0}^1 + ... + \beta_0^j D_{i0}^j$$

and

$$\beta_1^0 \ln \hat{\theta}_{i0} + \beta_1^1 D_{i0}^1 \ln \hat{\theta}_{i0} + ... + \beta_1^j D_{i0}^j \ln \hat{\theta}_{i0},$$

[4]For the correlation matrix of variables, see Table 3.

where D^j_{i0} is a dummy variable that takes the value of unity if firm i belongs to industry j. The "passive" speed-of-convergence of an industry j, ϕ^j_0, is represented as:

$$\frac{-[1-(1-\phi^j_0)^T]}{T} = \begin{cases} \beta^0_1 + \beta^j_1 & \text{if } \beta^j_1 \text{ is statistically significant}^5; \\ \beta^0_1 & \text{otherwise.} \end{cases}$$

In addition to industry specific effects, we control for IT industry effects. Recent studies on the TFP growth confirmed that IT products and industries strongly affect the national-level TFP growth. In the United States, Jorgenson (2001, Table 7) showed that the IT products contributed 0.5 percentage points to TFP growth in the latter half of the 1990s. In Japan, Nishimura et al. (2005) provided detailed analysis on the sectoral difference of the speed-of-convergence and found that part of such a difference was attributed to the difference between IT and non-IT industries. Although we control for industry effects by industry dummies, the effects of innovation we consider in this paper might also be different between IT and non-IT industries. Thus, we introduce cross-term of the IT industry dummy (D^{IT}) and innovation ($TechChangeFactor$).[6]

3.3 Estimation and Results

In a companion paper (Nishimura et al., 2005), we found that industry-level estimates of the speed of productivity convergence might be biased if one fails to take account of the effects of exits. In this paper, we have considered this possible bias explicitly.[7] To obtain a consistent estimator, we employ a sample selection model that is based on two equations. One is a selection equation describing which firms are exiting or surviving between years 0 and T. The other is a productivity equation, equation (9), to estimate productivity movement using information about surviving firms only. These two equations are unified into one likelihood function and estimated by a maximum likelihood (ML) method.[8]

3.3.1 Estimation Issues: Selection Equation

The selection equation captures the effects of exiting decisions by the exiting firms. Dunne, Roberts and Samuelson (1989) found that plant size, age, and ownership type (single-plant firm

[6] For the definition of IT industries, we follow the definition of US Department of Commerce (1999). Our IT industries include the following seven industries: 1) Office, service industry and household machines; 2) Electronic data processing machines, digital and analog computers, equipment and accessories; 3) Non-ferrous metal-worked products; 4) Electronic parts and devices; 5) Communication equipment and related products; 6) Miscellaneous electrical machinery, equipment and supplies; 7) Miscellaneous precision instruments and machinery.

[7] For a detailed discussion on econometric issues of firm-level productivity convergence regression, including endogeneity and sensitivity, see Nishimura et al. (2005).

[8] An alternative method to correct this type of selection bias is Heckman's two-step estimation procedure (Heckit). However, the ML (one-step) procedure is generally more efficient than the Heckit estimation (Johnston and DiNardo, 1997, p. 450; Davidson and MacKinnon, 1993, p. 545). Hence this paper employs ML rather than Heckit.

or multi-plant firm) were statistically significant determinants of plant growth and failure. Following the findings of Dunne, Roberts and Samuelson, we assume that the exit of a firm depends on three factors: firm age (AGE), employment scale (L), and multi-plant dummy (D^{multi} that takes the value of unity if a firm has multi-plants and zero for otherwise). In addition, we assume that the natural selection mechanism works: firms with lower productivity exit from the market. The selection equation for the exiting firms is represented as follows:

$$s_{iT} = \begin{cases} 1 & \text{if } \gamma_0 + \gamma_1 \ln \hat{\theta}_{i0} + \gamma_2 \ln AGE_{i0} + \gamma_3 \ln L_{i0} + \gamma_4 D^{multi}_{i0} + \nu_{iT} \geq 0; \\ 0 & \text{otherwise,} \end{cases}$$

where s_{iT} is the selection indicator that takes the value of unity if a firm exists before year T and zero otherwise.

3.3.2 Estimation Issues: Productivity Equation

The productivity equation, (9), captures the effects of innovation and diffusion. We estimate four models, which are summarized in Table 1. Summary statistics and a correlation matrix are presented in Tables 2 and 3, respectively. As indicated before, there exists a high correlation between two scale variables: labor ($\ln L$) and capital ($\ln K$). If we included these two variables at the same time, the regression equation would have multicollinearity. Thus, we have only used $\ln L$ to control for scale effects.

Table 1: Summary of the Models Estimated

Determinants/Variables	Model 1	Model 2	Model 3	Model 4
Innovation (ψ) $R\&D$, PAT, IMP	○	○	○	○
Explicit emulation (ξ) $\ln\theta \times R\&D$, $\ln\theta \times PAT$, $\ln\theta \times IMP$	○	○	○	○
Implicit emulation (β_1) $\ln\theta$	○	○	○	○
Scale effects (ω) $\ln L$	○	○	○	○
Adjustment costs (ω) $\{\Delta \ln L\}^2$, $\{\Delta \ln K\}^2$	○	○	○	○
IT industry dummy (ψ^{IT}) $R\&D \times D^{IT}$, $PAT \times D^{IT}$, $IMP_{i0} \times D^{IT}$	×	○	×	○
Cross-effect (ψ^{CROSS}) $R\&D \times PAT$	×	×	○	○

Table 2: Summary Statistics of Explanatory Variables

		N	Mean	S.E.
$\Delta \ln \theta$	TFP growth	12851	0.016	0.112
$\ln \theta$	TFP	12851	-0.007	0.561
$\ln \theta \times R\&D$	TFP $\times$ R&D-sales ratio	12851	0.001	0.009
$\ln \theta \times PAT$	TFP $\times$ Patent-sales ratio	12851	0.000	0.001
$\ln \theta \times IMP$	TFP $\times$ Import-sales ratio	12851	0.005	0.050
$R\&D$	R&D-sales ratio	12851	0.006	0.016
PAT	Patent-sales ratio	12851	0.000	0.002
IMP	Imports-sales ratio	12851	0.014	0.056
$R\&D \times PAT$	R&D-sales ratio $\times$ Patent-sales ratio	12851	0.000007	0.000
$\ln L$	Employment	12851	5.327	1.016
$\ln K$	Capital	12851	7.077	1.642
$\{\Delta \ln L\}^2$	(Labor growth)^2	12851	0.012	0.051
$\{\Delta \ln K\}^2$	(Capital growth)^2	12851	9.094	0.366

Note: For the definition of variables, see Section 4.

Table 3: Correlation Matrix of Variables

	$\Delta \ln \theta$	$\ln \theta$	$\ln \theta$ $\times R\&D$	$\ln \theta$ $\times PAT$	$\ln \theta$ $\times IMP$	$R\&D$	PAT	IMP	$R\&D$ $\times PAT$	$\ln L$	$\ln K$	$\{\Delta \ln L\}^2$	$\{\Delta \ln K\}^2$
$\Delta \ln \theta$	1.00												
$\ln \theta$	-0.38	1.00											
$\ln \theta \times R\&D$	-0.09	0.32	1.00										
$\ln \theta \times PAT$	-0.03	0.11	0.24	1.00									
$\ln \theta \times IMP$	-0.07	0.27	0.09	0.07	1.00								
$R\&D$	0.05	0.12	0.49	0.12	0.01	1.00							
PAT	0.03	0.06	0.15	0.60	0.04	0.20	1.00						
IMP	0.02	0.16	0.05	0.05	0.60	0.06	0.08	1.00					
$R\&D \times PAT$	0.02	0.05	0.23	0.53	0.03	0.31	0.74	0.06	1.00				
$\ln L$	0.03	0.09	0.20	0.04	0.02	0.27	0.07	0.04	0.09	1.00			
$\ln K$	0.02	0.00	0.15	0.04	0.00	0.24	0.07	0.04	0.08	0.71	1.00		
$\{\Delta \ln L\}^2$	-0.04	0.09	0.01	0.00	0.02	-0.02	-0.01	0.01	-0.01	-0.03	-0.23	1.00	
$\{\Delta \ln K\}^2$	-0.11	0.04	0.00	0.01	-0.01	0.02	0.01	-0.04	0.01	0.05	-0.03	0.08	1.00

In sum, the regression equation is specified as follows.

Model 1 (baseline model)

$$\begin{aligned}\Delta \ln \theta_{iT} &= \beta_0^0 + \beta_0^1 D_{i0}^1 + \ldots + \beta_0^j D_{i0}^j \\ &+ \beta_1^0 \ln \theta_{i0} + \beta_1^1 D_{i0}^1 \ln \theta_{i0} + \ldots + \beta_j^j D_{i0}^j \ln \theta_{i0} \\ &+ \xi_1 \ln \theta_{i0} \times R\&D_{i0T} + \xi_2 \ln \theta_{i0} \times PAT_{i0T} + \xi_3 \ln \theta_{i0} \times IMP_{i0T} \\ &+ \psi_1 R\&D_{i0T} + \psi_2 PAT_{i0T} + \psi_3 IMP_{i0T} \\ &+ \omega_1 \ln L_{i0T} + \omega_2 \left\{\Delta \ln L_{iT}\right\}^2 + \omega_3 \left\{\Delta \ln K_{iT}\right\}^2 + \mu_{iT}.\end{aligned} \quad (10)$$

Model 2

$$\begin{aligned}\Delta \ln \theta_{iT} &= \beta_0^0 + \beta_0^1 D_{i0}^1 + \ldots + \beta_0^j D_{i0}^j \\ &+ \beta_1^0 \ln \theta_{i0} + \beta_1^1 D_{i0}^1 \ln \theta_{i0} + \ldots + \beta_j^j D_{i0}^j \ln \theta_{i0} \\ &+ \xi_1 \ln \theta_{i0} \times R\&D_{i0T} + \xi_2 \ln \theta_{i0} \times PAT_{i0T} + \xi_3 \ln \theta_{i0} \times IMP_{i0T} \\ &+ \psi_1 R\&D_{i0T} + \psi_2 PAT_{i0T} + \psi_3 IMP_{i0T} \\ &+ \psi_1^{IT} R\&D_{i0T} \times D_{i0}^{IT} + \psi_2^{IT} PAT_{i0T} \times D_{i0}^{IT} + \psi_3^{IT} IMP_{i0T} \times D_{i0}^{IT} \\ &+ \omega_1 \ln L_{i0T} + \omega_2 \left\{\Delta \ln L_{iT}\right\}^2 + \omega_3 \left\{\Delta \ln K_{iT}\right\}^2 + \mu_{iT}.\end{aligned} \quad (11)$$

Model 3

$$\begin{aligned}\Delta \ln \theta_{iT} &= \beta_0^0 + \beta_0^1 D_{i0}^1 + \ldots + \beta_0^j D_{i0}^j \\ &+ \beta_1^0 \ln \theta_{i0} + \beta_1^1 D_{i0}^1 \ln \theta_{i0} + \ldots + \beta_j^j D_{i0}^j \ln \theta_{i0} \\ &+ \xi_1 \ln \theta_{i0} \times R\&D_{i0T} + \xi_2 \ln \theta_{i0} \times PAT_{i0T} + \xi_3 \ln \theta_{i0} \times IMP_{i0T} \\ &+ \psi_1 R\&D_{i0T} + \psi_2 PAT_{i0T} + \psi_3 IMP_{i0T} + \psi_4^{CROSS} R\&D_{i0T} \times PAT_{i0T} \\ &+ \omega_1 \ln L_{i0T} + \omega_2 \left\{\Delta \ln L_{iT}\right\}^2 + \omega_3 \left\{\Delta \ln K_{iT}\right\}^2 + \mu_{iT}.\end{aligned} \quad (12)$$

Model 4

$$\begin{aligned}\Delta \ln \theta_{iT} &= \beta_0^0 + \beta_0^1 D_{i0}^1 + \ldots + \beta_0^j D_{i0}^j \\ &+ \beta_1^0 \ln \theta_{i0} + \beta_1^1 D_{i0}^1 \ln \theta_{i0} + \ldots + \beta_j^j D_{i0}^j \ln \theta_{i0} \\ &+ \xi_1 \ln \theta_{i0} \times R\&D_{i0T} + \xi_2 \ln \theta_{i0} \times PAT_{i0T} + \xi_3 \ln \theta_{i0} \times IMP_{i0T} \\ &+ \psi_1 R\&D_{i0T} + \psi_2 PAT_{i0T} + \psi_3 IMP_{i0T} + \psi_4^{CROSS} R\&D_{i0T} \times PAT_{i0T} \\ &+ \psi_1^{IT} R\&D_{i0T} \times D_{i0}^{IT} + \psi_2^{IT} PAT_{i0T} \times D_{i0}^{IT} + \psi_3^{IT} IMP_{i0T} \times D_{i0}^{IT} \\ &+ \omega_1 \ln L_{i0T} + \omega_2 \left\{\Delta \ln L_{iT}\right\}^2 + \omega_3 \left\{\Delta \ln K_{iT}\right\}^2 + \mu_{iT}.\end{aligned} \quad (13)$$

As we summarized in Table 1, the coefficients ($\psi_1, ...$) capture the effects of innovation ($TechChangeFactor$) on productivity growth while (ξ_1, ξ_2, ξ_3) captures those of "active" diffusion, or explicit emulation ($DiffusionFactor$). The "passive" diffusion, or implicit emulation, are measured by ($\beta_1^0, \beta_1^1, ..., \beta_1^j$), which is used to calculate the "passive" speed of productivity convergence, ϕ_0^j. The effects of other control variables are represented by $(\omega_1, ...)$. $R\&D$, PAT, IMP, and $\ln L$ is the average of the period from 1995-2000 while $\Delta \ln \theta_{iT}$, $\Delta \ln L_{iT}$, and $\Delta \ln K_{iT}$ is the annual average growth rate between 1995 and 2000.

Here we distinguish the difference between innovation and "active" diffusion. The coefficients of innovation ($\psi_1, ...$) should be significantly positive if innovation contributes to the productivity growth. On the other hand, if "active" diffusion positively affects the productivity growth, the coefficients of "active" diffusion (ξ_1, ξ_2, ξ_3) should be significantly *negative* since negative coefficients mean speeding up the speed of productivity convergence.

3.3.3 Implicit Emulation ("Passive" Diffusion)

Tables 4-6 present the regression results for Models 1-4, which are generated by ML estimation. Table 4 reports the distribution of the speed of "passive" diffusion, ϕ_0^j. There are two notable findings in this table. Firstly, even after we control for explicit emulation as well as innovation, we obtain quite similar results in the speed of "passive" diffusion among different types of models. Out of 70 industries, nearly two-thirds of the industries report less than 10 percent, and less than or equal to 10 industries report more than 20 percent, regardless of the type of model.[9]

Secondly, there are large differences in the "passive" diffusion among industries. While most industries are concentrated in less than five percent of the speed-of-convergence, some industries show more than 20 percent of the speed of productivity convergence. These industry differences seem relatively robust: we observe them in all models.

There seems to be several reasons for these large industry differences in the speed of productivity convergence. As was discussed in Nishimura et al. (2005), one of the most important reasons may be the difference between IT and non-IT industries. Table 5 presents the distribution of the

[9]The speed of productivity convergence is significantly faster than the speed reported in the previous country-level studies. For instance, Dorwick and Nguyen (1989) reported that the speed-of-convergence among countries was 2.5 percent annually. At first glance, this seems to be a very high rate, but it is not so high if one looks at its order of magnitude. Suppose that the productivity level of firm i is 10 while that of the most productive firm is 100. If the speed-of-convergence is 10 percent (i.e., $\lambda = 0.10$), it still takes about 24 years for firm i to catch up the most productive firm. Note that whether or not a firm can survive for more than 24 years is an important issue since (Nishimura et al., forthcoming, Table 3) confirmed that about half of new firms in Japan exited from the market within five years of start up. Similarly, Bellone, Musso and Quéré (2003) found that about 70 percent of new firms exited from the market within 10 years in France.

Table 4: Distribution of the “Passive” Diffusion (Number of Industries)

	“Passive” Speed-of-convergence (ϕ_0^j)					
	$< 3\%$	$< 5\%$	$< 10\%$	$< 15\%$	$< 20\%$	$20\% \leq$
Model 1	2	0	43	8	8	9
Model 2	2	0	43	7	8	10
Model 3	2	0	43	8	8	9
Model 4	2	0	43	7	8	10

Table 5: Difference between IT and Non-IT industries (Number of Industries)

		“Passive” Speed-of-convergence (ϕ_0^j)					
		$< 3\%$	$< 5\%$	$< 10\%$	$< 15\%$	$< 20\%$	$20\% \leq$
Model 1	IT industries	0	0	0	2	1	4
	Non-IT industries	2	0	43	6	7	5
Model 2	IT industries	0	0	0	1	1	5
	Non-IT industries	2	0	43	6	7	5
Model 3	IT industries	0	0	0	2	1	4
	Non-IT industries	2	0	43	6	7	5
Model 4	IT industries	0	0	0	1	1	5
	Non-IT industries	2	0	43	6	7	5

speed of productivity convergence for IT and non-IT industries (the sum of IT and non-IT industries corresponds to the results of Table 4). Table 5 indicates that IT industries are more likely to have faster convergence speed. Out of seven IT industries, no industries indicate a rate of convergence of less than 5 percent and five industries show more than a 20 percent rate of convergence.

3.3.4 Innovation and Explicit Emulation (“Active” Diffusion)

Table 6 reports the coefficients of innovation and those of explicit emulation. Three notable findings stand out from this table. Firstly, both innovation and explicit emulation are important sources of productivity growth. All the coefficients of innovation ($R\&D$, PAT, IMP) present positive and significant signs. The coefficients of explicit emulation through imports and patents ($\ln\theta \times IMP$ and $\ln\theta \times PAT$) are large and negative, and the coefficients of imports are statistically significant (though those of patents are not statistically significant).

Secondly, in contrast with the difference in the speed-of-convergence between IT and non-IT industries in Table 5, the effects of innovation are not limited to IT industries. Models 2 and 4 illus-

Table 6: Innovation versus Diffusion

		Model 1 baseline model	Model 2	Model 3	Model 4
$\ln\theta$	TFP	0.122	0.127	0.124	0.130
$\times R\&D$	$\times$ R&D-sales ratio	(0.132)	(0.133)	(0.132)	(0.133)
$\ln\theta$	TFP	-1.302	-1.352	-1.190	-1.231
$\times PAT$	$\times$ Patent-sales ratio	(0.906)	(0.907)	(0.919)	(0.920)
$\ln\theta$	TFP	-0.058**	-0.061***	-0.058**	-0.061***
$\times IMP$	$\times$ Import-sales ratio	(0.023)	(0.023)	(0.023)	(0.023)
$R\&D$	R&D-sales ratio	0.234***	0.168*	0.244***	0.179**
		(0.074)	(0.086)	(0.075)	(0.087)
PAT	Patent-sales ratio	2.008***	1.768***	2.336***	2.128***
		(0.610)	(0.644)	(0.765)	(0.791)
IMP	Import-sales ratio	0.103***	0.109***	0.103***	0.109***
		(0.019)	(0.020)	(0.019)	(0.020)
$R\&D$	R&D-sales ratio			-10.366	-11.382
$\times PAT$	$\times$ Patent-sales ratio			(14.517)	(14.522)
$R\&D$	R&D-sales ratio		0.223		0.225
$\times D^{IT}$	$\times$ IT industry dummy		(0.147)		(0.147)
PAT	Patent-sales ratio		2.009		2.021
$\times D^{IT}$	$\times$ IT industry dummy		(1.442)		(1.442)
IMP	Import-sales ratio		-0.081		-0.083
$\times D^{IT}$	$\times$ IT industry dummy		(0.064)		(0.064)
$\ln L$	Employment scale	0.008***	0.008***	0.008***	0.008***
		(0.001)	(0.001)	(0.001)	(0.001)
$\{\Delta\ln L\}^2$	(Employment growth)^2	-0.026***	-0.026***	-0.026***	-0.026***
		(0.002)	(0.002)	(0.002)	(0.002)
$\{\Delta\ln K\}^2$	(Capital growth)^2	-0.010	-0.010	-0.010	-0.010
		(0.017)	(0.017)	(0.017)	(0.017)
ρ		0.119***	0.119***	0.119***	0.119***
		(0.045)	(0.045)	(0.045)	(0.045)
Log-likelihood		-3539.7	-3536.8	-3539.4	-3536.5
Akaike Information Criterion		7667.4	7667.6	7668.9	7669.0
N		23852	23852	23852	23852
LR test for $H_0:\rho=0$		5.26**	5.31**	5.29**	5.33**

Notes:

1) Standard errors are in parentheses.

2) ***, **, * indicate significance level at 1%, 5%, and 10%, respectively.

3) Constant and initial TFP level are included (but not reported).

trate the results that included IT industry dummies as a cross-term of innovation ($TechChangeFactor$). None of the coefficients of cross-term show significant signs. The results indicate that the positive effects of innovation on productivity growth exist and that the effects are widely observed in both IT and non-IT industries.

Thirdly, patent purchases are more effective than R&D activities in terms of the impacts on productivity growth *through innovation*. Table 6 indicates that the coefficients of patents (PAT) are larger than those of R&D ($R\&D$) in all models. On the other hand, although imports (IMP) have positive effects on productivity growth, the coefficients of imports are smaller than those of R&D. Since these three variables are in the same dimension, this result implies that the patents have stronger impacts on productivity growth through innovation than R&D, and R&D has much stronger effects than imports. The implication of these differences in impacts will be discussed in more detail in Section 4.

Finally, we cannot confirm any synergistic effects of R&D and the other factors. R&D has little impact on the speed of productivity convergence ($\ln\theta \times IMP$'s coefficient is insignificant). There is no synergy between R&D and patent purchases. The coefficients of cross-term between R&D and patent purchases ($R\&D \times PAT$) are neither positive nor statistically significant. Our results thus do not support the importance of the synergistic effects involving R&D.

3.3.5 Scale Effects and Adjustment Costs

Scale effects and adjustment costs are also important factors in explaining productivity growth. Let us first consider scale effects. The coefficients of the employment scale ($\ln L$) show significantly positive signs. This means that larger firms are more likely to grow faster than smaller firms, in terms of productivity.

Now, let us turn to adjustment costs. The results indicate that there are negative signs in $\{\Delta \ln L\}^2$ and $\{\Delta \ln K\}^2$. Besides, the coefficients of $\{\Delta \ln L\}^2$ are statistically significant. The results imply that adjustment costs exist, straining productivity growth. The rapid increases in inputs, in particular labor inputs, require large adjustment for firms, which result in negative effects on productivity growth.

4 Discussion

4.1 R&D, Patent Purchases, and Imports

We have found that both R&D and patent purchases are important sources of productivity growth. In particular, patent purchases are shown not only to speed up catch-up as expected (though somewhat weakly), but also to "wake up" firms' innovative activities to increase productivity further. However, in previous literature, innovation is considered to be primarily the product of R&D activities and researchers have paid little attention to the importance of patent purchases on productivity improvement through this innovation.

Thus, our results about the effects of R&D and patent purchases have important implications for management and academic research. Firms do not have to innovate everything from scratch by themselves. In fact, patent purchases are on average more effective than R&D to productivity increases *through innovation* as coefficients of these two variables reveal in Table 6. Consequently, it is important for managers and researchers to recognize that not only R&D but also the introduction of new ideas, such as patent purchases, could be an important source of innovation that leads to productivity improvement.

Imports have significantly positive effects on productivity growth as driving forces of innovation, although the effects of imports are not strong *vis-à-vis* R&D and patent purchases (the coefficients of imports are smaller than those of R&D and patent purchases). However, we find another important role of imports. The coefficients of cross-term between initial TFP level and imports show significantly negative signs, implying that imports contribute to accelerating the speed of catch-up.

The importance of international technology diffusion is discussed extensively in the recent studies in international economics.[10] Our results support the view emphasizing its importance. Even firms in developed countries like Japan can benefit from imports. Firms in developing countries are likely to obtain substantial productivity gains from international trade.

We have so far focused on the effects of imports rather than exports. Policy makers have shown a tendency to focus on the effects of exports on productivity growth (e.g., World Bank, 1993, pp. 316–326). Our results have revealed that imports are also an important source of productivity growth. Thus, growth strategies of developing countries should give sufficient attention to the role of imports. The protection on imports may result in preventing domestic firms from having

[10] See Barba Navaretti and Venables (2004, chapters 7 and 9) and Keller (2004), for an extensive survey of the related literature.

opportunities to obtain the state-of-art technology.

4.2 Importance of "Active" and "Passive" Diffusion

This paper has confirmed the importance of technology diffusion as well as innovation. "Active" diffusion helps firms with low productivity catch up to the most advanced firms. There are several strategies for secondary firms to catch up to the leading firms. Our analysis suggests that patent purchases and, especially, the imports of embedded technologies will likely benefit low-productivity firms more in catching up to the best technology than high-productivity firms.

The importance of "passive" diffusion is worth mentioning as well. Although innovation and "active" diffusion are important determinants of productivity growth, these activities involve substantial costs of investment and/or patent purchase. This implies that small and medium-sized firms that cannot afford to conduct R&D activities and/or to purchase technologies are virtually prevented from achieving productivity growth.

With "passive" diffusion, however, firms can achieve high productivity growth without incurring substantial costs. In this paper, we found that the importance of "passive" diffusion implied that learning-by-doing is also an important factor of productivity growth. Thus, even firms that cannot afford to conduct innovation or purchase technologies can achieve productivity growth through learning-by-doing. Consequently, policies to encourage learning-by-doing activities, especially those of workers on shop floors should be given much more attention in the discussion to improve firms' and countries' productivity. We need bottoming-up efforts of shop floors to improve productivity as much as top-down activities of R&D and technology purchases.

Finally, we should note that instant technology diffusion causes an additional, different problem. If technology diffused easily, no firms would have an incentive to conduct R&D investment. However, our results clearly indicate that technology diffusion without costly efforts, or implicit emulation, is not instantaneous but rather takes a long time. Thus, there is still room to maintain a technological advantage for a long time, which gives firms enough incentives to innovate technologies.

4.3 Robustness

Our main results in Tables 4, 5 and 6 are robust with respect to specification changes. The detailed results of a robustness check are reported in the Appendix below. The main conclusions are summarized as follows. Firstly, our results are robust to the choice of base year. We re-estimate the baseline model, changing the base year from 1995 to 1994. The estimation results indicate that all innovation variables show significantly positive signs. Employment scale also has positive and significant effects. Imports accelerate the catch-up process. The labor adjustment has negative effects, implying that the adjustment costs exist.

Secondly, our results are not sensitive to the threshold of the data. Since our data does not cover the firms with less than 50 workers, firms whose employment dropped to less than 50 workers are regarded as exit firms. Thus, one may concern that our results might be sensitive to this artificial threshold. In order to check the sensitivity to the threshold, we re-estimate the model for firms with more than 55 workers. In spite of the reduction in the sample size, the results are almost the same as those obtained from the baseline model.

Finally, the endogeneity problem is not very serious enough to change the implication of our results. Once we introduce the instrumental variable (IV) in estimating the speed-of-convergence equation, the rate of the speed-of-convergence declines, implying that the implicit emulation become weak. However, innovation and explicit diffusion still have strong effects. We thus conclude that our results are relatively robust even when the endogeneity problems exist.

5 Concluding Remarks

This paper has examined the determinants of productivity growth at the firm level, incorporating both effects of innovation and those of diffusion (explicit and implicit emulation). We have developed a model of firm-level productivity growth distinguishing between innovation and technology diffusion. We have then applied the model to a large-scale data of Japanese manufacturing and non-manufacturing firms for the period 1994-2000. We have focused on R&D activities as a driver of innovation, as well as patent purchases and imports (of capital equipment and others) as sources of "new ideas" triggering innovation. Further, two types of diffusion has been considered. One is explicit emulation that is facilitated by patent purchases and imports. The other is implicit emulation that is achieved by learning by doing.

Major findings are summarized as follows. Firstly, the innovation is an important determinant of productivity growth. As expected, R&D expenditure has a positive effect on productivity growth but the positive effects of innovation are not limited to R&D activities. Patent purchases and imports also contribute to the productivity growth as innovation factors.

Secondly, not only innovation but also technology diffusion is an engine of productivity growth. The significantly positive effects of "active" technology diffusion, or explicit emulation, are confirmed in imports. Patent purchases can be another determinants of "active" technology diffusion but it is not strong enough to have a statistically significant effect. “Passive” technology diffusion, or implicit emulation, contributes to the productivity growth as well. Even after controlling for innovation and explicit emulation, a strong evidence of the “passive” diffusion is found in almost all industries in all specifications of the model.

References

Aitken, B. and A. Harrison, “Do Domestic Firms Benefit from Foreign Direct Investment? Evidence from Venezuela,” *American Economic Review*, 1999, *89* (3), 605–618.

Barba Navaretti, G. and A.J. Venables, *Multinational Firms in the World Economy*, Princeton, NJ: Princeton University Press, 2004.

Barro, R.J. and X.X. Sala-i-Martin, *Economic Growth*, 2 ed., Cambridge, MA: MIT Press, 2004.

Bellone, F., P. Musso, and M. Quéré, “Exiting Firms’ Patterns: Evidence from a French Panel Data Set,” *Paper presented at the International Industrial Organization Conference at Boston*, April 4-5 2003.

Bernard, A.B. and C.I. Jones, “Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries,” *American Economic Review*, 1996, *86* (5), 1216–1239.

Branstetter, L.G., “Is Foreign Direct Investment A Channel of Knowledge Spillovers? Evidence from Japan’s FDI in the United States,” *NBER Working Paper*, 2000, (8015).

Caves, D., L. Christensen, and W. Diewert, “Output, Input and Productivity Using Superlative Index Numbers,” *Economic Journal*, 1982, *92* (5), 73–96.

Coe, D.T. and E. Helpman, “International R&D Spillovers,” *European Economic Review*, 1995, *39* (6), 859–887.

Davidson, R. and J.G. MacKinnon, *Estimation and Inference in Econometrics*, Oxford: Oxford University Press, 1993.

Dorwick, S. and Duc-Tho Nguyen, “OECD Comparative Economic Growth 1950-85: Catch-Up and Convergence,” *American Economic Review*, 1989, *79* (5), 1010–1031.

Dunne, T., M.J. Roberts, and L. Samuelson, “The Growth and Failure of U.S. Manufacturing Plants,” *Quarterly Journal of Economics*, 1989, *104* (4), 671–698.

Good, D., I. Nadiri, L.H. Roeller, and R.C. Sickles, “Efficiency and Productivity Growth Comparisons of European and U.S/ Air Carriers: A First Look at the Data,” *Journal of Productivity Analysis*, 1983, *4*, 115–125.

Griliches, Z., *R&D and Productivity*, Chicago, IL: University of Chicago Press, 1998.

Grossman, G.M. and E. Helpman, *Innovation and Growth in the Global Economy*, Cambridge, MA: MIT Press, 1991.

Haskel, J.E., S.C. Pereira, and M.J. Slaughter, “Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?,” *NBER Working Paper*, 2002, (8724).

Hulten, C.R., E.R. Dean, and M.J. Harper, eds, *New Developments in Productivity Analysis*, Chicago, IL: University of Chicago Press, 2001.

Javorcik, B.S., "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages," *American Economic Review*, 2004, *94* (3), 605–627.

Johnston, J. and J. DiNardo, *Econometric Methods*, 4 ed., New York: McGraw-Hill, 1997.

Jorgenson, D.W., "Information Technology and the U.S. Economy," *American Economic Review*, 2001, *91* (1), 1–32.

Keller, W., "International Technology Diffusion," *Journal of Economic Literature*, 2004, *42* (3), 752–804.

__ **and S.R. Yeaple**, "Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States," 2004. manuscript.

Lee, J-W., "Capital Goods Imports and Long-run Growth," *Journal of Development Economics*, 1995, *48* (1), 91–110.

Levinsohn, J. and A. Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–341.

Nakamura, S., "An Adjustment Cost Model of Long-Term Employment in Japan," *Journal of Applied Econometrics*, 1993, *8* (2), 175–194.

Nishimura, K.G, T. Nakajima, and K. Kiyota, "Productivity Convergence at the Firm Level," *CIRJE Discussion Paper*, May 2005, (CIRJE-F-341).

Nishimura, K.G., T. Nakajima, and K. Kiyota, "Does the Natural Selection Mechanism Still Work in Severe Recessions? Examination of the Japanese Economy in the 1990s," *Journal of Economic Behavior and Organization*, forthcoming.

Olley, G.S. and A. Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Schumpeter, J.A., *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, Cambridge, MA: Harvard University Press, 1934.

Sørensen, A., "Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries: Comment," *American Economic Review*, 2001, *91* (4), 1160–1167.

US Department of Commerce, *The Emerging Digital Economy II*, Washington, D.C.: Department of Commerce, 1999.

World Bank, *The East Asian Miracle: Economic Growth and Public Policy*, Oxford: Oxford University Press/The World Bank, 1993.

Appendix. Robustness Check

Appendix examines the robustness of our results. We address three issues. Firstly, we examine whether or not our results are sensitive to the choice of the base year. Secondly, we check how our results change when we use different threshold level of employment. Finally, we present the estimation results, controlling for possible endogeneity in the convergence equation.

The Choice of the Base Year

One of the major criticisms on the convergence studies is that the results are sensitive to the choice of base year. For instance, in his comments on Bernard and Jones (1996), Sørensen (2001) finds that whether or not we observe convergence depends crucially on the choice of the base year. To check the sensitivity of our results to the choice of the base year, we changed the base year from 1995 to 1994, examined the productivity growth between 1994 and 2000, and re-estimated the baseline model.

Tables A1 and A2 present the estimation results of the baseline model in Tables 4 and 6 respectively, changing the base year from 1994 to 1995. The results indicate that the coefficients are not very sensitive to the choice of the base year. The distribution of "passive" diffusion, or the speed-of-convergence, in Table A1 is not exactly the same as but similar to the distribution presented in Table 4. Similarly, in Table A2, All innovation factors ($R\&D, PAT, IMP$) show positive and significant effects on productivity growth. Employment scale ($\ln L$) has significantly positive effects. Imports speed up the productivity convergence process. The labor adjustment costs ($\{\Delta \ln L\}^2$) have negative effects. These results are the same as the results obtained in the baseline model. Thus, the speed-of-convergence may be affected by the choice of the base year but our main conclusion does not change for the choice to base year.

However, there are two differences between 1995 and 1994. One is that the capital adjustment costs ($\{\Delta \ln K\}^2$) in the case of 1994 show significantly negative signs, implying that both labor and capital adjustment costs exist. The other is that R&D expenditure now has *negative* effects on

Table A1: Robustness Check: Distribution of the "Passive" Diffusion (Number of Industries)

	"Passive" Speed-of-convergence (ϕ_0^j)					
	$<$ 3%	$<$ 5%	$<$ 10%	$<$ 15%	$<$ 20%	20% $\leq$
Base year = 1994	5	0	39	10	14	2
$L \geq 55$	2	0	42	8	9	9

Table A2: Robustness Check: Base Year and Threshold

		Base year = 1994	$L \geq 55$
		Model 1	Model 1
$\ln\theta$	TFP	0.300**	0.119
$\times R\&D$	$\times$ R&D-sales ratio	(0.117)	(0.132)
$\ln\theta$	TFP	-1.057	-1.303
$\times PAT$	$\times$ Patent-sales ratio	(0.953)	(0.905)
$\ln\theta$	TFP	-0.049**	-0.058**
$\times IMP$	$\times$ Import-sales ratio	(0.025)	(0.023)
$R\&D$	R&D-sales ratio	0.227***	0.234***
		(0.065)	(0.074)
PAT	Patent-sales ratio	1.697***	2.006***
		(0.597)	(0.610)
IMP	Imports-sales ratio	0.108***	0.105***
		(0.019)	(0.019)
$\ln L$	Employment	0.009***	0.008***
		(0.001)	(0.001)
$\{\Delta \ln L\}^2$	(Labor growth)^2	-0.025***	-0.027***
		(0.003)	(0.002)
$\{\Delta \ln K\}^2$	(Capital growth)^2	-0.049***	-0.013
		(0.012)	(0.017)
N		22077	23054
Log-liklihood		-1903.9	-3166.7
Akaike Information Criterion		4395.9	6921.3
LR test for $H_0 : \rho = 0$		4.59***	4.91**

Note: Standard errors are in parenthesis. *** and ** indicate significance level at 1% and 5%, respectively.

the convergence process, as the coefficient of the cross-term between TFP and R&D expenditure ($\ln\theta \times R\&D$) indicates positive sign. However, these two results seem dependent on the choice of this particular year (1994). Thus we base our argument on more robust results in the text.

Threshold

One may be concerned with the truncation based on the threshold of 50 workers. In our data, firms with less than 50 workers are not covered in the survey, and thus a firm whose employment is reduced below this level is regarded as an exiting firm.[11] To check the effects of threshold, we re-estimate the baseline model for firms with 55 workers.

The results in Tables A1 and A2 indicate that all innovation factors and employment scale have significantly positive effects on productivity growth while the labor adjustment costs have significantly negative effects. Imports help to speed up productivity convergence process. Despite the fact that 798 firms are eliminated from our sample, the results are quite similar to the results obtained from the baseline model. Thus, we can conclude that our results are not sensitive to the threshold of truncation.

Endogeneity

Finally, one may raise the issue of endogeneity: independent variables and μ_{iT} in the baseline model might be correlated. To examine possible effects of this endogeneity, we applied IV methods to the productivity equation (10).

In obtaining IV estimators, we employed Heckman's two-step estimation procedure (Heckit) rather than ML. In Section 3, we have used the ML method since this one-step procedure is generally more efficient than the two-step method of so-called Heckit estimation.[12] However, here we use the Heckit framework since it provides a straightforward extension to the case of endogeneity, which is unfortunately not the case in the ML method.

We first estimated the Mills ratio using a probit model and the Mills ratio is used as an additional variable to estimate the productivity equation. Instruments utilized are the lag of all independent variables. Because of the difficulty in obtaining the proper instruments, we run regressions without including industry dummy variables.

[11] There is also a truncation based on the amount of paid capital. However, since paid capital is usually not a good indicator of firm size in practice, this truncation is considered not as serious as the truncation based on the number of employees.

[12] See Johnston and DiNardo (1997, p. 450) and Davidson and MacKinnon (1993, p. 545).

Table A3 presents the estimation results of ML, Heckit and IV estimators. The estimated results indicate that the speed-of-convergence generated by Heckit is slower than ML but the speed generated by IV is much slower than Heckit. However, innovation and explicit emulation variables are quantitatively similar in ML, Heckit, and IV. Therefore, the endogeneity might have some effects on the speed-of-convergence estimates, but our major findings and implications are unchanged.

Table A3: Robustness Check: Endogeneity

		Model 1	Model 1	Model 1
Estimation method		ML	Heckit	IV
$\ln\theta$	TFP	-0.075	-0.065	-0.023
		(0.002)	(0.003)	(0.004)
$\ln\theta$	TFP	-0.356	-0.438	-0.622
$\times R\&D$	$\times$ R&D-sales ratio	(0.133)	(0.160)	(0.176)
$\ln\theta$	TFP	-2.146	-2.123	-3.615
$\times PAT$	$\times$ Patent-sales ratio	(0.973)	(1.098)	(1.296)
$\ln\theta$	TFP	-0.032	-0.042	-0.044
$\times IMP$	$\times$ Import-sales ratio	(0.028)	(0.032)	(0.035)
$R\&D$	R&D-sales ratio	0.696	0.734	0.723
		(0.073)	(0.085)	(0.081)
PAT	Patent-sales ratio	2.892	2.809	3.144
		(0.671)	(0.769)	(0.762)
IMP	Imports-sales ratio	0.177	0.174	0.123
		(0.023)	(0.026)	(0.026)
$\ln L$	Employment	0.007	0.026	0.055
		(0.001)	(0.004)	(0.004)
$\{\Delta\ln L\}^2$	(Labor growth)^2	-0.031	-0.040	-0.055
		(0.003)	(0.004)	(0.003)
$\{\Delta\ln K\}^2$	(Capital growth)^2	0.022	0.009	-0.038
		(0.022)	(0.024)	(0.022)
Mills ratio			0.393	0.392
			(0.026)	(0.026)
λ	Speed-of-convergence	9.0%	7.6%	2.4%
N		16138	16138	16138

Notes:
1) Standard errors are in parentheses.
2) ML: Maximum likelihood estimation method is used for the estimation.
3) Heckit: Heckman's two-step estimation method is used for the estimation.
4) IV: Instrumental variable method is used for the estimation.
5) For the IV results, standard errors are not adjusted.