RIETI Discussion Paper Series 09-E-027

Productivity and Characteristics of Firms: An application of a bootstrapped data envelopment analysis to Japanese firm-level data

KATO Atsuyuki
RIETI

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 09-E-027

Productivity and Characteristics of firms:
■ An Application of a Bootstrapped Data Envelopment Analysis to Japanese Firm-Level Data

Atsuyuki KATO †
Research Institute of Economy, Trade and Industry, IAA, Tokyo Japan
kato-atsuyuki@rieti.go.jp

Preliminary version

Abstract

This paper examines the relationships between productivity growth and characteristics of firms using Japanese firm-level data during the period 1995-2004. Applying a bootstrapped Malmquist index approach and weighted least squares (WLS) to two retail trade industries, we estimate the firm-level productivity growth rates and the effects of firms' characteristics on those growth rates. In addition, decomposing productivity growth into technical efficiency change and technical progress, we discuss mechanisms of productivity growth in detail. Our estimation reveals that productivity growth of department stores and supermarkets was stagnant during the sample period. It also indicates that positive technical efficiency changes are usually offset by technical regress and vice versa. Furthermore, effects of firms' characteristics on both productivity components are sometimes conflicting as well. In view of these findings, industrial policies should be carefully devised, based upon their efficiency distribution.

† Atsuyuki KATO, RIETI, 1-3-1 Kasumigaseki Chiyoda-ku Tokyo, 100-8901, Japan, Phone: +81-3-3501-8275, Fax: +81-3-3501-8416, E-mail: kato-atsuyuki@rieti.go.jp

1. Introduction

Recently, productivity analysis using micro (firm or establishment level) data has been employed often by economists and policy makers. This is because only productivity growth is considered as an engine to yield economic growth in the long run and productivity growth at the micro level results in productivity growth at the industrial and macro levels. So, in order to draw more desirable economic and industrial policies, further understanding of productivity at the micro level is quite important. For this reason, many papers have estimated the productivity of firms and establishments and its determinants using various approaches.

Reviewing this work, Biesebroeck (2000) discusses the advantages and disadvantages of five frequently applied methodologies, (1) index numbers, (2) data envelopment analysis (DEA), (3) stochastic frontier analysis (SFA), (4) instrumental variables (GMM), and (5) semiparametric estimation. He examines the robustness of their productivity estimates, introducing randomness via factor price heterogeneity, measurement errors and differences in production technology, respectively, and obtains the following results. For data with small measurement errors, index numbers are most desirable while parametric approaches are better if measurement or optimisation errors are not negligible. On the other hand, DEA outweighs the others where the production technology is heterogeneous and returns to scale are not constant.

This result is interesting because DEA is still relatively poorly applied compared from index numbers and parametric approaches in productivity empirics. Although those papers which apply index numbers or some parametric approaches provide many contributions to our understandings of productivity, the validity of their fundamental assumptions, such as constant returns to scale or homogeneous production technology is still controversial in studies using micro data. For such studies, estimation of productivity based upon DEA is considered to provide additional contributions to productivity empirics. In addition, productivity analysis based on DEA allows us to decompose productivity growth into efficiency improvement and technical progress.

In this paper, we also conduct productivity analysis using DEA for two retail trade industries, department stores and supermarkets in Japan. Using the obtained efficiency scores from DEA, we estimate the Malmquist productivity index of firms as the growth rate of their total factor productivity, and decompose it into technical efficiency change and technical progress. In order to make the obtained DEA scores available for statistical inference, we apply a bootstrap method proposed by Simar and Wilson (1999). It possibly corrects the upward bias of the DEA scores if our examined samples do not include the actual best practice firm as well. Regression analysis is also carried out to examine various determinants of productivity growth and obtain some industry-specific implications. Since the assumption of homoscedasticity is possibly violated, we apply weighted least squares (WLS) regression to our samples.

The layout of this paper is as follows. In section 2, we detail the methodologies of estimation in this model. Section 3 describes the data which we use. Section 4 discusses the result and implications of empirical analysis. And the last section draws conclusion from the above discussion.

2. Methodology

This section briefly describes our methodologies. First, we calculate the Malmquist productivity change index as a measure of productivity growth over time. The Malmquist index is constructed as ratios of distance functions which are estimated by

DEA[1]. The Malmquist productivity index approach is initially suggested by Caves, Christensen and Diewert (1982) and developed as an empirical index by Färe, Grosskopf, Norris and Zhang (1994), henceforth FGNZ. This index has been used in empirical analysis because of an advantage that it is decomposable into further components, technical progress and technical efficiency change without price information as Kumbharkar and Lovell (2000) mention.

The output distance function introduced by Shephard (1970) at time period t is defined as follows,

$$D_o^t\left(x^t,\ y^t\right) = \min_\theta \left\{\theta : \left(x^t,\ y^t/\theta\right) \in P^t\right\} = \left(\max\left\{\theta : \left(x^t,\ \theta y^t\right)\right\} \in P^t\right)^{-1} \qquad (1)$$

where x^t, y^t, and P^t represent inputs, outputs and the production possibility set, respectively. The subscript t represents time period, and $t = 1, \ldots, T$. Here P^t is defined that it transforms inputs $x^t \in R_+^k$ into outputs $y^t \in R_+^M$. That is described as $P^t = \left\{\left(x^t,\ y^t\right) : x^t\ can\ produce\ y^t\right\}$. In equation (1), $D_o^t\left(x^t,\ y^t\right) \le 1$ if $\left(x^t,\ y^t\right) \in P^t$. $D_o^t\left(x^t,\ y^t\right) = 1$ if and only if a firm manages its production activity on the technology frontier. Since this output distance function is the reciprocal of the output-oriented measure of efficiency proposed by Farrell (1957), $D_o^t\left(x^t,\ y^t\right)$ is also used as a measure of efficiency. Similarly, the output distance function as well as a measure of efficiency at time period $t+1$ is denoted as $D_o^{t+1}\left(x^{t+1},\ y^{t+1}\right)$.

In order to capture productivity change over time, we also define the following two

[1] Kordbacheh (2007) briefly reviews literature on the Malmquist index. Our explanation of this index largely relies on his work as well.

hypothetical distance functions which represent the transformation of input x^{t+1} (x^t) into outputs y^{t+1} (y^t) by technology P^t (P^{t+1}),

$$D_o^t\left(x^{t+1},\ y^{t+1}\right) = \min\left\{\theta : \left(x^{t+1},\ y^{t+1}/\theta\right) \in P^t\right\} \qquad (2)$$

$$D_o^{t+1}\left(x^t,\ y^t\right) = \min\left\{\theta : \left(x^t,\ y^t/\theta\right) \in P^{t+1}\right\} \qquad (3).$$

Using these equations, productivity change is measured as the ratio of the actual distance function to the hypothetical production function, $M_o^t = \left(\frac{D_o^t\left(x^{t+1},\ y^{t+1}\right)}{D_o^t\left(x^t,\ y^t\right)}\right)$ or $M_o^{t+1} = \left(\frac{D_o^{t+1}\left(x^{t+1},\ y^{t+1}\right)}{D_o^{t+1}\left(x^t,\ y^t\right)}\right)$. These two measures of productivity change are identical if and only if one input yields one output. Otherwise, $M_o^t \neq M_o^{t+1}$. In order to measure unbiased productivity change between periods t and $t+1$, the Malmquist productivity change index for multiple inputs and outputs are defined as the geometric mean of M_o^t and M_o^{t+1}. That is,

$$M_o\left(x^{t+1},\ y^{t+1},\ x^t,\ y^t\right) = \left[\left(\frac{D_o^t\left(x^{t+1},\ y^{t+1}\right)}{D_o^t\left(x^t,\ y^t\right)}\right)\left(\frac{D_o^{t+1}\left(x^{t+1},\ y^{t+1}\right)}{D_o^{t+1}\left(x^t,\ y^t\right)}\right)\right]^{1/2} \qquad (4).$$

If the estimated index is larger (smaller) than unity, the examined firm experienced productivity progress (regress). Using the above distance functions, we can decompose productivity change into technical efficiency change and technical progress. That is,

$$M_o\left(x^{t+1}, y^{t+1}, x^t, y^t\right) = \underbrace{\left(\frac{D_o^{t+1}\left(x^{t+1}, y^{t+1}\right)}{D_o^t\left(x^t, y^t\right)}\right)}_{(TEC)} \underbrace{\left[\left(\frac{D_o^t\left(x^{t+1}, y^{t+1}\right)}{D_o^{t+1}\left(x^{t+1}, y^{t+1}\right)}\right)\left(\frac{D_o^t\left(x^t, y^t\right)}{D_o^{t+1}\left(x^t, y^t\right)}\right)\right]^{1/2}}_{(TP)} \quad (5)$$

The first term in the right hand side of Equation (5), (*TEC*), is the ratio of distance functions between periods t and $t+1$, and represents technical efficiency change. On the other hand, the second term (*TP*) captures effects of shift of the production frontier, thus denotes change of technology. As well as the Malmquist index, the estimated values of TEC and TP are interpreted as progress (regress) if they are larger (smaller) than unity.

One shortcoming of this approach is that the estimated productivity change is possibly biased. In the above Malmquist index approach, the distance functions are obtained by DEA. According to the definition of DEA, our data should include at least one best practice firm (See Appendix 2). However, our data do not always include the real best practice in all existing firms. If not, the estimated distance functions have upward bias and productivity indices are biased as well. In addition, the standard DEA scores are sensitive to measurement errors. Furthermore, DEA scores are not directly applicable to statistical inference because they are deterministic values. In order to overcome these shortcomings of DEA, Simar and Wilson (1998, 2000, and 2007) propose to apply bootstrapping methods to DEA[2]. Their procedure constructs pseudo-data sets iteratively. Using those new data, DEA scores are re-estimated

[2] The approach proposed in Simar and Wilson (2007) is considerably time consuming for large samples.

respectively, and then, their procedure yields a good approximation of the true distribution of sampling[3]. Simar and Wilson (1999) also apply a consistent method to the Malmquist productivity change index[4]. Although their procedure has a shortcoming that the estimated bias corrected estimator may have a higher mean-square error, it is still considered useful and we follow it as well.

In our methodologies, secondly, we carry out regression of the estimated productivity index, technical efficiency change, and technical progress on various regressors which are considered as determinants of productivity dynamics in preceding literature, respectively. This analysis is expected to reveal the properties of industrial productivity dynamics. A significant problem of this regression analysis is that the errors of our estimation models may violate the assumption of homoscedasticity. If so, ordinary least squares (OLS) is not an appropriate approach. To discuss it, we carry out two tests of that assumption such as the Breusch-Pagan and the White tests and apply weighted least squares (WLS) to our data. Following Lee and Kang (2007), we formulate the estimation model below,

$$TFPG(TEC,\ TP)_{it} = \alpha + X'\beta + \delta \widetilde{Eff}_{it-1} + \varepsilon_i \qquad (6)$$

where *TFPG*, *TEC*, *TP* and $\widetilde{Eff}$ are TFP growth, technical efficiency change, technical progress and the bootstrapped DEA score, respectively. Since the initial level of productivity is thought to be associated with the following productivity growth, we add the initial level of technical efficiency as a proxy of the productivity level. The subscript

---

3 Brümmer (2001) cited in Balcombe, Davidova and Latruffe (2008)

4 We do not discuss details of their bootstrapping methods in this paper in order to avoid redundancy.

i denotes firm i and $i = 1, \ldots, N$. X is the vector of control variables, β is the vector of the coefficients, and δ is the coefficient of the initial efficiency. ε is an error term.

3. Data

In this study, we use firm-level data of department stores and supermarkets in Japan. The data are extracted from the annually compiled official statistics of firms' activities by the Ministry of Economy, Trade and Industry of Japan[5]. This statistics covers many activities of firms and is considered reliable[6]. Since these statistics do not identify which firms are department stores or supermarkets, we construct the lists of them for name identification using Nikkei Almanac of Retail and Wholesale Companies. From these data sources, we construct our own dataset composing of output, labour and capital inputs, and various control variables.

In our dataset, output is represented as total sales as well as many existing papers because the estimated TFP based on value added is biased using micro data as Basu and Fernand (1995) discuss. The proxy of accumulated capital is the tangible fixed assets. In some previous papers, land is excluded out of capital. However, we include it in the capital data because we believe that conditions of location play an important role in service production by department stores and supermarkets. Labour input is calculated as man-hours[7]. In addition, following Tokui, Inui and Kim (2007) and Kim, Kwon and Fukao (2007), the intermediate input is obtained as follows[8]:

---

5 This statistics is named as 'the Basic Survey of Business Structure and Activity'.
6 Kiyota and Matsuura (2004)
7 The data of working hours are available from Monthly Labour Survey.
8 In calculation of intermediate input, we slightly modify both Tokui et al. and Kim et al. The former does neither include tax and dues nor purchase in calculation of the intermediate inputs while the later does not include tax and dues.

$$Intermediate\ Input = COGS + SGA - (TW + Dep + T\ \&\ D + Purchase), \quad (7)$$

where *COGS*, *SGA*, *TW*, *Dep* and *T&D* are the cost of goods sold, the selling and general administrative expenses, the total wages, the depreciation and the tax and dues, respectively. Since data of output, intermediate input and capital are nominal values, we construct real series of them using deflators in JIP database[9]. In constructing our dataset, we rule out the firms which report zero or negative values for total sales, the number of regular workers, the tangible fixed assets, total wage, or intermediate inputs.

4. Empirical Results

This section discusses the empirical results and their implications. The model estimation of bootstrapped Malmquist productivity indices is conducted using the computer program FEAR on R[10]. Since Kato (2009) rejects the assumption of constant returns to scale for these two retail trade industries, the index is also estimated based upon the assumption of variable returns to scale. The number of iterations is 2000 following Balcombe, Davidova and Latruffe (2008). A drawback of DEA for panel data is to identify technical regress if the production level decreases due to recession.

Figure 1 presents TFP dynamics of both industries. The fluctuations of GDP in the retail trade industry are also drawn to capture the demand fluctuation as well[11]. In many

---

9 JIP database includes deflators for output and intermediate input. We construct capital deflator series following Kim et al.

10 R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

FEAR is a freely downloadable program to estimate DEA scores and conduct the bootstrap algorithm proposed by Simar and Wilson (1998, 1999).

11 The growth rate of GDP in the retail trade industry is calculated using national

papers, TFP growth is procyclical and it casts a serious doubt concerning whether we estimate TFP appropriately. Because of simultaneity of production and consumption in the service sector, it is difficult for us to tell reduction of demand from regress of technical ability under such procyclicality. Compared with them, our estimates of TFP growth do not always follow the demand fluctuation. It seems to reveal that our TFP estimates reasonably capture the dynamics of technical ability in the above two industries. Obviously, a hike in the consumption tax (from 3% to 5%) in 1997 gave a significant negative impact on GDP growth of the retail trade industry while not on TFP growth. The figure also implies that the recent recovery of demand in the retail trade sector does not lead the resurgence of TFP growth for supermarkets.

Figures 2A and B present the dynamics of TFP growth and its components such as TEC and TP[12]. Following FGNZ, positive estimates of TEC are interpreted as technical diffusion while those of TP are expansion of the technology frontiers. From these figures, it is obvious that effects of components are offset each other. When the technical frontier expanded, the efficiency gaps also expanded. On the other hand, the positive effects of technical diffusion on TFP growth were countered by technical regress. As a result, TFP growth in these two industries was stagnant during the examined period.

To discuss the productivity dynamics further, we carry out regression analysis of productivity on various control variables. This empirical analysis is thought to possibly give implications to draw more desirable industrial policies. Similar analysis has been conducted in the previous literature, using some different measures of TFP growth. For

---

account of Japan (2006).

[12] TEC and TP for department stores are dropped from Figure 2A during the period 2002-2004 because the estimates are extremely fluctuated.

example, Morikawa (2007) examines this issue (for TFP and value added) in Japanese manufacturing, wholesale, retail trade and service industries. In this paper, we provide additional contributions to those existing literature, examining the effects of characteristics of firms on their TFP components as well as TFP growth. The examined control variables are as follows[13]. First, the averaged wage (*wage*) is considered as a proxy of labour quality[14]. Second, the number of establishments per regular worker, (*establish*) denotes firms' strategic choice between the focus and market saturation approaches[15]. Third, ratios of part-time and temporary workers to regular workers (*part* and *temporary*) are also included in order to examine effects of shifts in employment structure on productivity[16]. The fourth control variable is the ratio of outsourcing to total sales (*outsource*). Fifth, as firms' history, ages (*age*) and original institutional forms of firms (= marged (*form1*), decomposed (*form2*), organisationally changed (*form3*), newly established (*form4*) and the others) are examined. As well as these factors, Information and Communication Technology (ICT) utilization is considered to provide positive contributions to productivity growth. Therefore, our estimation includes three different forms of the network systems, such as the Intranet (*within*), the network between limited firms (*between*), and the open network (*open*). In relation to them, share of information and communication cost to total sales (*infocost*) is tested as well. Finally, the bootstrapped estimate of the initial efficiency levels (*deab*) is also included in our estimation. We expect a negative coefficient on this variable for TFP and technical

---

13 All control variables but dummies are logs, and the names of variables are in parentheses.
14 Jorgenson, Gallop and Fraumeni (1987)
15 Following their business models, retail traders are thought to focus their business into the small number of commercial establishments (focus approach) or to widely distribute their shops through their business areas (market saturation approach).
16 The variable, *temporary*, consists of temporary and dispatched workers.

efficiency change estimation because initially more efficient firms have the smaller room to improve their efficiency levels. On the other hand, we examine neither of the share of R&D expenditure to total sales, the share of export to total sales, nor the share of foreign capital to capital stock while those factors are examined in many existing papers. This is because most of those variables have zero values in our data[17].

Tables 1 and 2 show the results of estimation for department stores and supermarkets, respectively. In all regressions, the null hypothesis of homoscedasticity is rejected by both the Breusch-Pagan and the White tests. Therefore, we apply WLS to our panel data. Adjusted R-squares indicate that the goodness of fit is relatively high in both industries. Since the results of both industries are not always consistent, their mechanisms of TFP growth seem to vary as well.

As for department stores, Table 1 gives negative estimates on *wage* in all three regressions. It seems that labour quality has a negative impact on TFP growth through both efficiency deterioration and technical regress under our assumption. However, these results should be carefully interpreted. During the examined period, the average wage levels might not reflect the labour quality of firms well. Because of the long lasted recession, firms could obtain well-educated or highly skilled workers with relatively lower salaries[18]. If so, it is difficult to draw any conclusion from the negative estimates with respect to labour quality. In order to further discuss this issue, we need additional information related to labour quality, for example education levels.

The focus approach could raise TFP growth rate through improvement of technical efficiency. This result is consistent with the intensifying elimination and consolidation of commercial establishments during the examined period. The dependency on the

---

17 All of these variables are insignificantly estimated if included.

18 Morikawa (2007)

part-time workers is negatively associated with TP. At the 10 percent significance level, it is also negatively related to TFP growth and TEC. It implies that increasing in part-time workers is not advantageous in this industry. The negative coefficient of *temporary* in the TEC and TP regressions also seems that the higher share of the dispatched and temporary workers harms efficiency improvement and technical progress. However, this result should be carefully discussed as well. Our data sources cover only a small amount of dispatched and temporary workers because of the definition and does not always capture the actual increase in them[19]. The estimated results of *outsource* indicate that outsourcing can provide positive contributions to TFP growth through technical progress. On the other hand, it deteriorates firms' technical efficiency. However, we should also carefully study it because of data problems[20].

Firms' age is positively related to all of TFP growth, TEC and TP while none of the original institutional forms of firms is significantly estimated. Thus, it does not matter whatever the starting forms of firms are, but experiences are helpful to raise TFP growth as well as efficiency improvement and technical progress. Among three network systems, *within* is significantly and positively estimated. It suggests that adopting the Intranet system is helpful for improving efficiency and fuelling technical progress while the other two network systems do not have significant effects on productivity dynamics. The resultant estimates of *infocost* are contradistinctive between TEC and TP. The negative coefficient in the TEC regression reveals that the higher ratio of information and communication cost to total sales might distort efficient use of production resources. On the other hand, it possibly facilitates technical progress. The positive estimate on

---

[19] Official statistics of firms report only the workers whose salaries are recorded as the labour cost in the head offices. However, those workers are hired through various paths.
[20] This variable includes many zero values.

deab in the TFP regression implies that the productivity gaps between firms tend to increase in this industry.

The empirical results for supermarkets are considerably different from those for department stores as Table 2 presents. In this industry, *wage* is not significantly estimated in any regression at the 1 or 5 percent level. However, it is positively estimated in the TFP regression at the 10 percent level. It implies that labour quality represented as the average wage level might have a positive connection with TFP growth. Unlike department stores, supermarkets prefer the market saturation approach in terms of efficiency improvement and technical progress. The difference in two industries possibly reflects the difference in their business models. In general, department stores work as malls for luxury goods. In such business, the focus approach is reasonable. On the other hand, supermarkets focus on consumers’ daily demand. In order to follow the localised demand, the market saturation approach is helpful.

The dependency on the part-time workers is positively related to TFP growth although negatively related to TEC. This positive estimate implies that the recently increasing in the ratio of the part-time workers to total employees is reasonable in terms of productivity growth. It is also supportive to adopt the distributing strategy for firms. However, increasing in the part-time workers does not contribute to efficiency improvement. It suggests that it is not always reasonable to highly rely on part-time workers, for the firms which remain far behind the technical frontier. Although data problem should be kept in mind, the positively estimated coefficient on *temporary* implies that further utilising dispatched and temporary workers possibly helps TFP growth through technical progress. As well as department stores, *outsource* is positively estimated in the TFP and TP regressions while negative in the TEC regression. It indicates that outsourcing fuels technical progress and provides positive contributions to TFP growth. On the other hand, it also reveals that less efficient firms should revise their current internal production process before outsourcing it.

The coefficient on *age* is significantly positive in the TFP regression but negative in the TEC regression. Unlike department stores, long-established supermarkets do not always achieve larger improvement of technical efficiency[21]. Among four original institutional forms, only *form 2* has a significant and positive estimate. A possible interpretation of this finding is that firms sometimes make specialised sectors independent and such specialised firms achieve higher performance. In order to discuss this issue in details, we need further analysis from business studies.

As to the ICT, utilisation of the Intranet and the Internet provides positive contribution to TFP growth of supermarkets. These results indicate that ICT can play positive roles to yield higher TFP growth in this industry as the US experience shows. A possible interpretation of these findings is that supermarkets can construct better supply chains using those ICTs. On the other hand, the network system between limited firms is negatively related to productivity growth in any path. It implies that the ICT reinforcing the closed business relations deteriorates firms' productivity performance.

These empirical results give us various implications related to productivity performance for these two industries. In addition to considerable differences between industries, our study also finds that effects of various factors on TEC and TP within an industry are not always consistent, but also sometimes conflicting. This finding implies that we need to draw different industrial policies based on distribution of firms'

---

[21] Since the correlation coefficient between firms' age and their efficiency scores is 0.2, it is not always true that long-established supermarkets already reached higher efficiency levels and had small room for further improvement.

efficiency scores for achieving better productivity performance. If an industry has huge and consistent efficiency gaps between firms, industrial policies should put larger weight on reinforcing technical diffusion from the best practice to the followers. On the other hand, if large amount of firms in an industry already reached higher efficiency levels, expanding the technical frontier is strongly required.

In relation to this issue, Figures 3A and B respectively present the distributions of firms' efficiency scores for each industry in 1995, 2000 and 2004. Both figures indicate that these industries seem to achieve significant improvement in their technical efficiency, in particular during 1995-2000. However, it is not always consistent with findings in Figures 2A and B that there is little evidence of efficiency improvement on average. It is thought that such discrepancy stems from the fact that the above figures cover slightly different data. Figures 2A and B cover only firms appeared in two consecutive years while Figures 3A and B do the whole firms appeared in each year. It implies that the exit of less efficient firms yields significant improvement in distribution of efficiency scores without technical diffusion[22]. On the other hand, for department stores, it is difficult to detect any evidence of efficiency improvement since millennium. Rather, large amounts of them already reached relatively higher efficiency levels in 2000. It implies that industrial policies to facilitate technical progress are advantageous for further productivity growth. As to supermarkets, significant efficiency gaps still remain although we find considerable catch-up. It reveals that desirable industrial policies for supermarkets should reinforce technical diffusion from the best practice to the followers[23]. Our empirical analysis provides guidelines of those industrial policies[24].

---

[22] This view is not consistent with some preceding work such as Nishimura, Nakajima and Kiyota (2003).

[23] It does not mean that technical progress is not important for supermarkets.

[24] We do not list effective control variables repeatedly, in order to avoid redundancy.

5. Concluding Remarks

In this paper, we estimate the productivity growth and its components and examine the relationships between productivity performance and characteristics of firms for department stores and supermarkets the period 1995-2004. From the results of a bootstrapped Malmquist approach, we find that productivity growth of department stores and supermarkets were stagnant during the examined period. It also indicates that the positive technical efficiency changes are usually offset by technical regress, vice versa. In addition, regression analysis brings the relations between productivity performance and firms' characteristics into daylight. It reveals that higher dependence on the part-time workers hurts technical progress for department stores and efficiency improvement for supermarkets although it is positively associated with TFP growth of supermarkets. The intranet is useful for department stores while other network technology has insignificant effects. On the other hand, the network system between limited firms provides significant negative effects on TFP growth of supermarkets through efficiency deterioration and technical regress while the Intranet and the Internet have significant positive coefficients. As a whole, various characteristics of firms sometimes have conflicting effects between productivity components as well as industries. It suggests that the Malmquist index approach to productivity empirics at the firm-level data provides a great deal of implication to design effective industrial policies.

In future research, we should examine the validity of our interpretation of the results, using various other approaches. It provides many implications for economists and policy makers. In addition, expanding this research to unexamined industries would allow us to better understand industry-specific features of them. Furthermore, we also need to examine the relationships of characteristics between complementarily related industries.

Appendix 1: Some Descriptive Statistics

	Department	**Super**
DEA Max	0.9946	0.9906
DEA Min	0.3068	0.6441
DEA adv	0.8954	0.8651
DEA skw	-2.7019	-0.4051
Sales adv	86368.81	76356.77
Labour adv	1482.75	2692.46
wage adv	4584120	2614779
Establish adv	8.63	42.02
Part adv	0.3124	0.7130
temporary adv	0.0451	0.0478
outsource adv	0.0016	0.0054
age adv	45.26	35.91
infocost adv	0.0495	0.0296
Sample	1067	2424

Note: DEA means a bootstrapped DEA
Estab denotes the number of Establishments
Part represents the part-time worker ratio
Max, Min, adv and skw are maximum, minimum, average and skew, respectively

Appendix2: Data Envelopment Analysis[25]

DEA is a linear programming approach which constructs a nonparametric piecewise surface from data and then measures the efficiency of each firm by comparing its location from the frontier using the Farrell index[26]. Charnes, Cooper and Rhodes (1978) (henceforth CCR) proposes the efficiency measure which maximises the ratio of the weighted sum of outputs to that of inputs where the measured efficiency scores are fallen into the range, $0 < Eff \leq 1$. For the case that each firm employs m different inputs to produce s different outputs, CCR formulates the estimation of the efficiency score for firm j as follows,

$$
\begin{aligned}
&\max_{\mu_k\,,\,\omega_k} \left(\sum_{r=1}^{s} \mu_{rk} y_{rk}\right) \\
&\text{subject to} \\
&\qquad \sum_{i=1}^{m} \omega_{ik} x_{ik} = 1 \\
&\qquad \sum_{r=1}^{s} \mu_{rk} y_{rj} - \sum_{i=1}^{m} \omega_{ik} x_{ij} \leq 0 \ \text{ for } j = 1,\ldots,n \qquad \text{(A1)} \\
&\qquad \mu_{rk}, \omega_{ik} \geq 0
\end{aligned}
$$

where x_{ik} and ω_{ik} are i-th input and its weight for firm k. On the other hand, y_{rk} and μ_{rk} are r-th output and its weight for firm k as well. Firm k is the reference firm. The duality theory in linear programming says that an equivalent form of this problem is described as follows,

---

25 We largely rely on Cracolici, Nijkamp and Cuffaro (2006) for this appendix.

26 Debreu (1951) and Farrell (1957)

$$\min_{\lambda_j, \theta_k} \quad \theta_k$$

subject to

$$\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta_k x_{ik} \quad i = 1,\ldots,m; \qquad \text{(A2)}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{rk} \quad r = 1,\ldots,s;$$

$$\lambda_j \ge 0 \quad j = 1,\ldots,n..$$

where θ_k is the dual variable corresponding to the equality constraint that normalised the weighted sum of inputs, and λ is a vector of the dual variables corresponding to the inequality constraints of the original problem indicating the intensity variables. In equation 2, firm j is efficient if $\theta^* = 1$, otherwise inefficient[27].

The above CCR model assumes constant returns to scale. Banker, Charnes and Cooper (1984) (henceforth BCC) expand it to the assumption of variable returns to scale (VRS), adding $\sum_{j=1}^{n} \lambda_j = 1$ to the CCR model. In this paper, we apply the BCC model because the assumption of CRS might not be reasonable in analysing firm-level production.

The CCR and the BCC models above reveal that DEA does not rely on information of prices. Since price data of inputs or outputs are hardly available in actual empirical study, this is an important advantage. In addition, DEA doesn't need to specify the form of production or cost function[28]. On the other hand, DEA scores are considerably sensitive to the extreme samples and measurement errors.

---

27 An asterisk to a variable denotes its optimal solution.

28 It is also considered as an advantage that DEA manages to deal with multi-outputs and multi-inputs simultaneously because firms provide various products or services in their buisness.

References

Balcombe K., S. Davidova, and L. Latruffe (2008), “The Use of Bootstrapped Malmquist Indices to Reassess Productivity Change Findings: An Application to a Sample of Polish Farms”, Applied Economics, 40 (16), 2055-2061.

Banker R. D., A. Charnes and W. W. Cooper (1984), “Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis”, *Management Science*, 30, 1078-1092.

Basu S. and J. G. Fernald (1995), “Are Apparent Productive Spillovers a Figment of Specification Error”, *Journal of Monetary Economics*, 36 (1), 165-188.

Biesebroeck J. van (2007), “Robustness of Productivity Estimates”, *The Journal of Industrial Economics*, 55 (3), 529-569.

Brümmer B., “Estimating Confidence Intervals for Technical Efficiency: The Case of Private Farms in Slovenia”, European Review of Agricultural Economics, 28 (3), 285-306.

Caves D. W., L. R. Christensen and W. E. Diewert (1982), “The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity”, *Econometrica*, 50(6), 1393-1411.

Charnes A., W. W. Cooper and E. Rhodes (1978), “Measuring the Efficiency of Decision Making Units”, *European Journal of Operational Research*, 2, 429-444.

Cracolici M. F., P. Nijkamp and M.Cuffaro (2006), “Efficiency and Productivity of Italian Tourist Destinations: A Quantitative Estimation based on Data Envelopment Analysis and the Malmquist Method”, *Tinbergen Institute Discussion Paper*, TI 2006-096/3.

Debreu G. (1951), “The Coefficient of Resource Utilization”, *Econometrica*, 19, 273-292.

Färe R., S. Grosskopf, M. Norris and Z. Zhang (1994), “Productivity Growth, Technical Progress, and Efficiency Change in Industrialised Countries”, *American Economic Review*, 64, 66-83.

Farrell M. J. (1957), “The Measurement of Productive Efficiency”, *Journal of the Royal Statistical Society*, Series A 120 (3), 253-290.

Jorgenson D., F. Gallop and B. Fraumeni (1987), *Productivity and U.S. Economic Growth*, Harvard University Press, MA.

Kato A. (2009), “Productivity, Returns to Scale and Product Differentiation: An Empirical Analysis using Japanese Firm-Level Data”, *RIETI Discussion Paper*, 90-E-009.

Kim Y. G., H. U. Kwon and K. Fukao (2007), “Entry and Exit of Companies and Establishments, and Productivity at the Industry Level (in Japanese)”, *RIETI*

Discussion Paper, 07-J-022.

Kiyota K. and T. Matsuura (2004), "Construction and Usage of A Panel Data of 'the Basic Survey of the Business Structure and Activity': Problems in Application to Economic Analysis and Arrangement of Data (in Japanese)", *RIETI Policy Discussion Paper*, 04-P-004.

Kordbacheh H. (2007), *Evaluating the Performance of International Airlines: Non-Parametric versus Parametric Methods*, PhD Thesis, University of Essex.

Kumbharkar S. C. and C. A. K. Lovell (2000), *Stochastic Frontier Analysis*, Cambridge University Press.

Lee K-J and S-M Kang (2007), "Innovation Types and Productivity Growth: Evidence from Korean Manufacturing Firms", *Global Economic Review*, 36(4), 343-359.

Morikawa M. (2007), "What Kinds of Company have High Productivity?: Company Characteristics and TFP (in Japanese)", *RIETI Discussion Paper*, 07-J-049.

Nishimura G. K., T. Nakajima and K. Kiyota (2003), "What happened to Japanese Industries in the 90s, the lost decade?: Firms' entry/exit and productivity (in Japanese)", *RIETI Discussion Paper*, 03-J-002.

Shephard R. W. (1970), *Theory of Cost and Production Frontiers*, Princeton University

Press, New Jersey.

Simar L. and P. W. Wilson (1998), “Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models”, *Management Science*, 44, 49-61.

--------------------, (1999), “”Estimating and Bootstrapping Malmquist Indices”, *European Journal of Operating Research*, *115*, 459-471.

--------------------, (2000), “A General Methodology for Bootstrapping in Nonparametric Frontier Models”, *Journal of Applied Statistics*, 27, 779-802.

--------------------, (2007), “Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Process”, *Journal of Econometrics*, 136 (1), 31-64.

Tokui J., T. Inui and Y. G. Kim (2007), “The Embodied Technical Progress and the Average Vintage of Capital (in Japanese)”, *RIETI Discussion Paper*, 07-J-035.

Wilson P. W. (2006), “FEAR: A Software Package for Frontier Efficiency Analysis with R”, mimeo.

Table 1: Department Stores

Variables	TFP	P>t	TEC	P>t	TP	P>t
wage	-0.0715**	0.000	-0.0354**	0.001	-0.1817**	0.000
establish	0.0042*	0.031	-0.0088**	0.003	0.0160	0.446
part	-0.0133	0.054	-0.0065	0.078	-0.3257**	0.000
temporary	-0.0025	0.194	-0.0068*	0.000	-0.1960**	0.000
outsource	0.0025**	0.000	-0.0028*	0.015	0.0689**	0.000
age	0.0062**	0.008	0.0390*	0.000	0.1207**	0.000
form1	-0.0287	0.509	-0.0336	0.481	0.5483	0.559
form2	-0.0655	0.108	-0.0290	0.499	0.3727	0.692
form3	0.0234	0.967	-0.0289	0.769	-0.0854	0.938
form4	-0.0221	0.597	-0.0196	0.641	0.0505	0.957
within	0.0202	0.112	0.0182*	0.026	0.2934**	0.000
between	0.0161	0.224	-0.0212	0.118	0.0550	0.826
open	0.0253	0.399	0.0061	0.895	-0.1322	0.898
infocost	0.0073**	0.000	-0.0080**	0.000	0.0627**	0.000
deab	0.1626**	0.000	-0.0456	0.177	-1.2616**	0.000
t	-0.0132**	0.000	-0.0296**	0.000	0.0127	0.159
Constant	1.1049**	0.000	0.4467**	0.006	2.4012*	0.035
$adj\ R^2$	0.9506		0.5118		0.9848	
RMSE	0.0064		0.0476		0.0402	
$\sigma_i^2 = \sigma^2$	Stats	P-value	Stats	P-value	Stats	P-value
B-P	818.13	0.000	463.31	0.000	227.10	0.000
White	341.11	0.000	394.36	0.000	285.27	0.000

Note: RMSE = Root Mean Square Error

B-P = Breusch Pagan

* and ** represent significantly estimated coefficients at the five and one percent levels

Table 2: Supermarkets

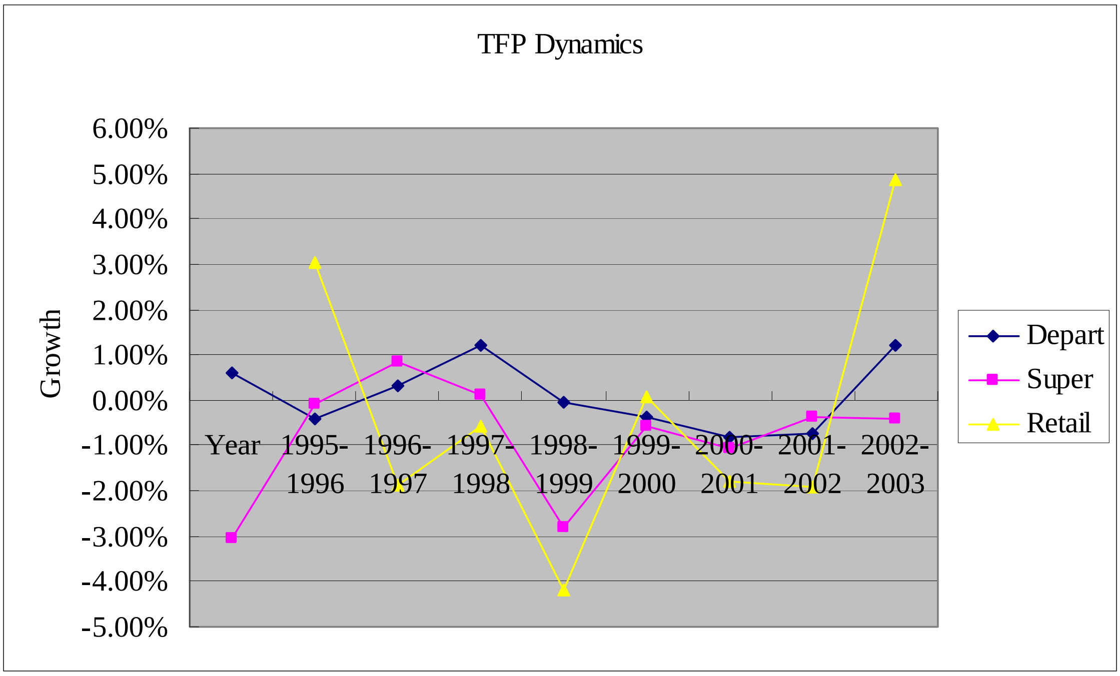
Variables	TFP	P>t	TEC	P>t	TP	P>t
wage	0.0640	0.083	-0.0018	0.958	0.0014	0.957
establish	-0.0012	0.743	-0.0045**	0.006	-0.0177**	0.000
part	0.0895**	0.000	-0.0247*	0.045	-0.0040	0.752
temporary	0.0051**	0.000	-0.0003	0.851	0.0036**	0.000
outsource	0.0023**	0.000	-0.0021*	0.001	0.0028**	0.000
age	0.0342**	0.000	-0.0108*	0.024	-0.0100	0.389
form1	-0.0211	0.122	0.0401	0.231	0.0099	0.914
form2	0.1825*	0.000	-0.0150	0.678	-0.0384	0.680
form3	0.0036	0.944	0.0053	0.904	-0.0095	0.934
form4	-0.0218	0.200	0.0176	0.607	-0.0336	0.716
within	0.0143**	0.000	0.0104	0.137	-0.0182**	0.000
between	-0.0267**	0.000	-0.0286*	0.000	-0.0212**	0.000
open	0.0216**	0.000	-0.0053	0.387	0.0022	0.914
infocost	-0.0114**	0.000	-0.0034	0.156	0.0074**	0.000
deab	-0.0772**	0.009	0.0049	0.857	0.1010	0.086
t	0.0017	0.341	0.0028**	0.000	-0.0003	0.903
Constant	-1.0737	0.059	-0.0004	0.999	-0.0261	0.950
$adj\ R^2$	0.9414		0.3242		0.4786	
RMSE	0.0066		0.0008		0.0092	
$\sigma_i^2 = \sigma^2$	Stats	P-value	Stats	P-value	Stats	P-value
B-P	3375.59	0.000	31.86	0.000	2488.04	0.000
White	258.21	0.000	245.40	0.000	269.22	0.000

Note: RMSE = Root Mean Square Error

B-P = Breusch Pagan

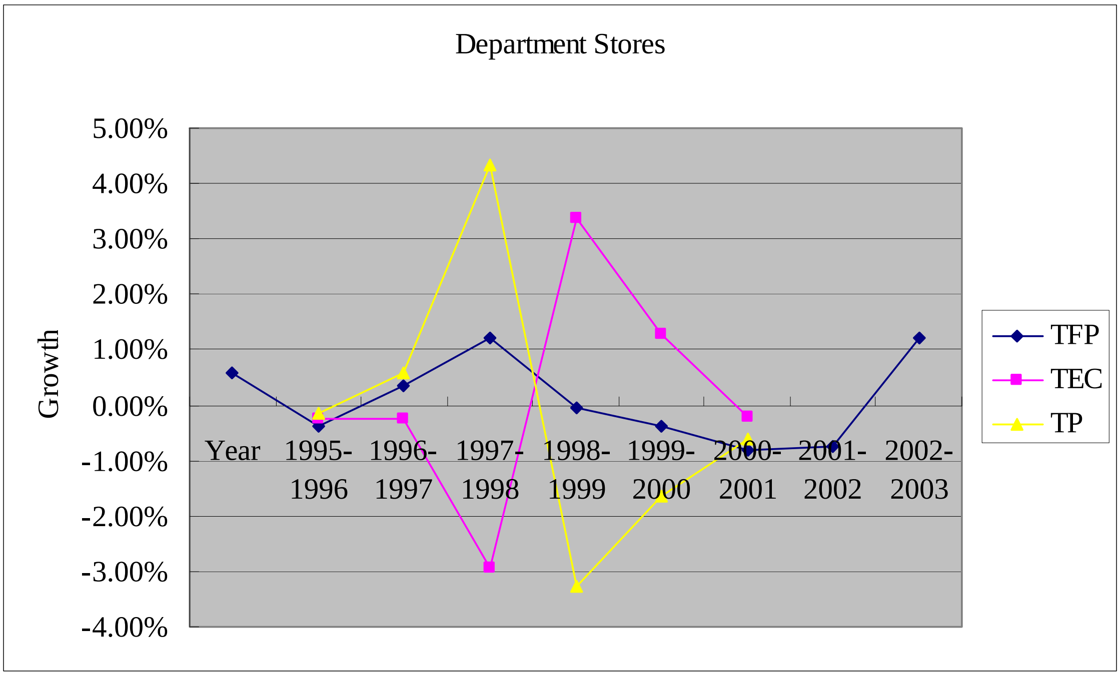
* and ** represent significantly estimated coefficients at the five and one percent levels

Figure 1: TFP Dynamics and Demand Fluctuation



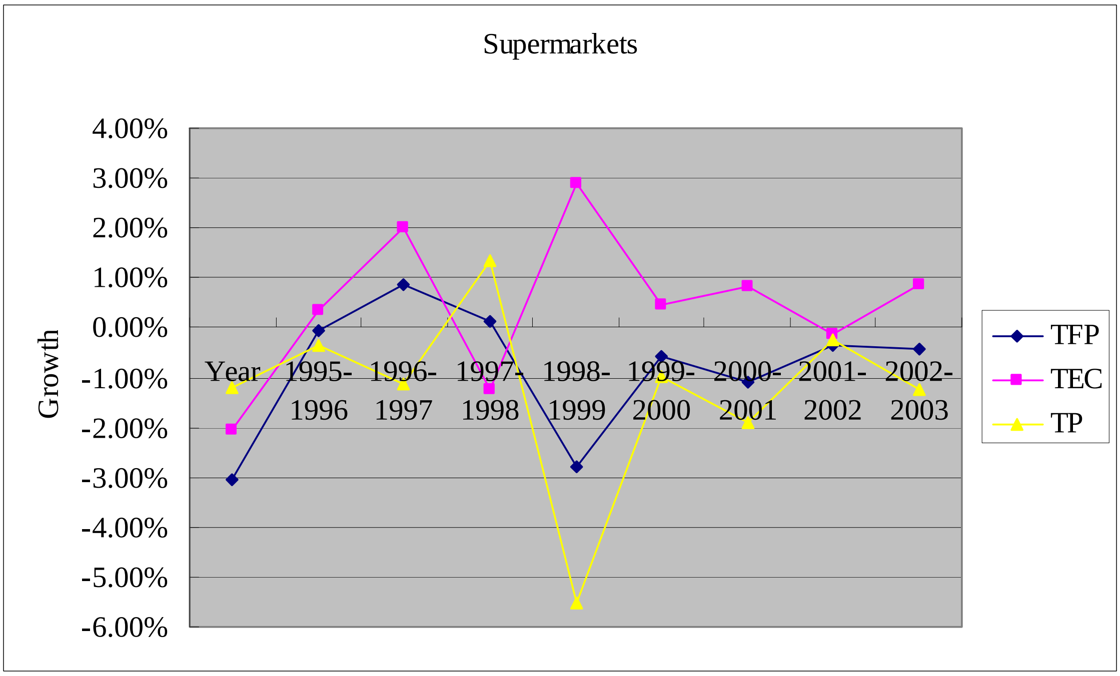
Note: TFP growth rates are calculated as the weighted averages in both industries.

Figure 2A: TFP, TEC and TP Dynamics of Department Stores



Note: TFP growth rates are calculated as the weighted averages in both industries.

Figure 2B: TFP, TEC and TP Dynamics of Supermarkets



Note: TFP growth rates are calculated as the weighted averages in both industries.

Figure 3A: Distribution of Efficiency Scores for Department Stores

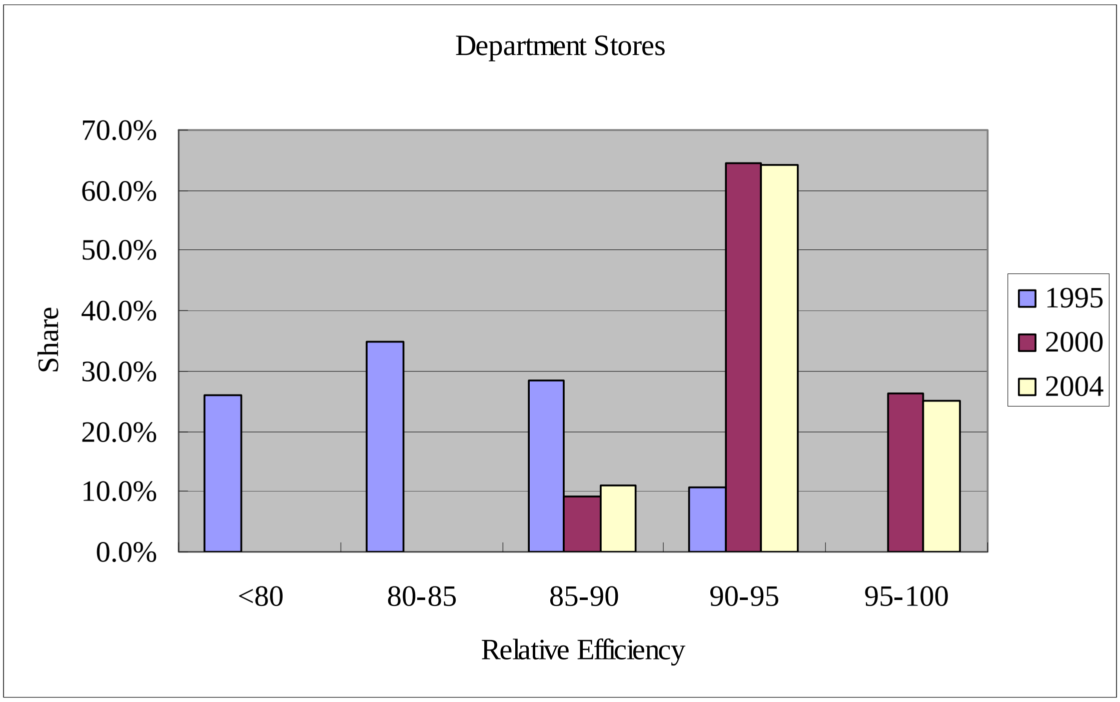


Figure 3B: Distribution of Efficiency Scores for Supermarkets