The returns to scale effect in labour productivity growth

Mizobuchi, Hideyuki
Ryukoku University

27. May 2011

The Returns to Scale Effect in Labour Productivity Growth

Hideyuki Mizobuchi[†]

May 6, 2011

Abstract

Labour productivity is defined as output per unit of labour input. Economists acknowledge that technical progress as well as growth in capital inputs increases labour productivity. However, little attention has been paid to the fact that changes in labour input alone could also impact labour productivity. Since this effect disappears for the constant returns to scale short-run production frontier, we call it the *returns to scale effect*. We decompose the growth in labour productivity into two components: 1) the joint effect of technical progress and capital input growth, and 2) the returns to scale effect. We propose theoretical measures for these two components and show that they coincide with the index number formulae consisting of prices and quantities of inputs and outputs. We then apply the results of our decomposition to U.S. industry data for 1987–2007. It is acknowledged that labour productivity in the services industries grows much more slowly than in the goods industries. We conclude that the returns to scale effect can explain a large part of the gap in labour productivity growth between the two industry groups.

Key Words: Labour productivity, index numbers, Malmquist index, Törnqvist index, output distance function, input distance function

JEL classification: C14, D24, O47, O51

[†] Faculty of Economics, Ryukoku University, 67 Fukakusa Tsukamoto-cho, Fushimi-ku, Kyoto 612-8577, Japan; mizobuchi@econ.ryukoku.ac.jp

1. Introduction

Economists broadly think of productivity as measuring the current state of the technology used in producing the firm's goods and services. The production frontier, consisting of inputs and the maximum output attainable from them, characterizes the prevailing state of technology. Productivity growth is often identified by the shift in the production frontier, reflecting changes in production technology.[1,2] However, productivity growth can also be driven by movement along the production frontier.

Even in the absence of changes in the production frontier, changes in the inputs used for production can lead to productivity growth, moving along the production frontier and making use of its curvature. Productivity growth that is induced by the movement along the production frontier is called the *returns to scale effect*. This effect does not reflect changes in the production frontier. Thus, in order to properly evaluate improvements in the underlying production technology reflecting the shift in the production frontier, we must disentangle the returns to scale effect from labour productivity.

Productivity measures can be classified into two types: total factor productivity (TFP) and partial factor productivity. The former index relates a bundle of total inputs to outputs, whereas the latter index relates a portion of total inputs to outputs. The present paper deals with labour productivity (LP) among several measures of partial factor productivity. LP is defined as output per labour input in the simple one-output one-labour-input case. Economy-wide LP is the critical determinant of a country's standard of living in the long-run. For example, U.S. history reveals that increases in LP have translated to nearly one-for-one increases in per capita income over a long period of time.[3] The importance of LP as a source for the progress of economic well-being prompts many researchers to investigate what determines LP growth.[4] Technical progress and capital input growth have been emphasized as the main determinants of a country's enormous LP growth over long periods (Jorgenson and Stiroh 2000, Jones 2002) as well as the wide differences in LP across countries (Hall and Jones 1999). The present paper adds one more explanatory factor to LP growth.[5]

---

[1] See Griliches (1987). The same interpretation is also found in Chambers (1988).

[2] In principle, productivity improvement can occur through technological progress and gains in technical efficiency. Technical efficiency is the distance between the production plan and the production frontier. The present paper assumes a firm's profit-maximizing behaviour, and in our model the current production plan is always on the current production frontier. The assumption of profit maximization is common in economic approaches to index numbers. See Caves, Christensen and Diewert (1982) and Diewert and Morrison (1986).

[3] See the 2010 Economic Report of the President.

[4] The LP growth and the capital input growth are the abbreviations for the growth rates of LP and capital input. In this paper, for example, the growth rate of LP between the current and previous periods is the ratio of LP in the current period to LP in the previous period.

[5] If the number of workers or the number of hours worked are adopted as the measure of labour input, changes in characteristics of labour input also affects LP. These authors also found an important role of labour quality growth (in other words, human capital accumulation) for explaining changes in their measure of LP that is defined by using the number of workers or the number of hours worked. Since we allow wages to vary across different types of labour input, the quality of each labour input is differentiated in our measure of labour input. Thus, we ignore the role of the labour quality growth for explaining LP growth, throughout this paper. See Footnote 6 for the unmeasured improvement in labour quality.

LP relates labour inputs to outputs, holding technology and capital inputs fixed. The short-run production frontier, which consists of labour inputs and the maximum output attainable from them, represents the capacity of current technology to translate labour inputs to outputs. Both technical progress and capital input growth, which have been identified as the sources of LP growth, induce LP growth throughout the shift in the short-run production frontier. However, the returns to scale effect, which is the extent of LP growth induced by movement along the short-run production frontier, has never been exposed.

We decompose LP growth into two components: 1) the joint effect of technical progress and capital input growth, 2) the returns to scale effect.[6] First, we propose theoretical measures representing the two effects by using distance functions. Second, we derive the index number formulae consisting of prices and quantities and show that they coincide with theoretical measures, assuming the translog functional form for the distance functions and the firm's profit-maximizing behaviour.

Our approach to implementing theoretical measures is drawn from Caves, Christensen and Diewert (1982) (hereafter, CCD). Using the distance function, CCD formulate the (theoretical) Malmquist productivity index, which measures the shift in the production frontier, and show that the Malmquist productivity index and the Törnqvist productivity index coincide, assuming the translog functional form for the distance functions and the firm's profit-maximizing behaviour.

The Törnqvist productivity index is a measure for the TFP growth calculated by the Törnqvist quantity indexes. It is an index number formula consisting of prices and quantities of inputs and outputs. Equivalence between the two indexes breaks down if the underlying technology does not exhibit constant returns to scale. CCD shows that its difference depends on the degree of returns to scale in the underlying technology, which captures the curvature of the production frontier. Thus, following Diewert and Nakamura (2007) and Diewert and Fox (2010), we can interpret that CCD decompose the TFP growth that is calculated by the Törnqvist quantity indexes into Malmquist productivity index and the returns to scale effect.[7] The former component captures TFP growth induced by the shift in the production frontier. The latter component, which is the difference between the Malmquist productivity index and the Törnqvist productivity index, captures TFP growth induced by the movement along the production frontier exploiting its curvature.

Many researchers have been concerned with the growth in TFP induced by movement along the underlying production frontier. For example, Lovell (2003) calls it the *scale effect*. In the literature of Data Envelopment Analysis (Balk, 2001, Coelli et al., 2003), the product of *scale efficiency change* and *input mix effect* or that of *scale efficiency change* and *output mix effect* summarizes the TFP growth induced by movement along the production frontier, and it can be interpreted as the returns to scale effect.[8]

Although scholars have recognized the significance of the returns to scale effect for TFP growth, its effect on LP growth has never been addressed even though it is more

---

[6] In case when our measure of labour inputs fails to capture the improvement in labour quality, TFP growth induced by that unmeasured improvement in labour quality is interpreted as that induced by technical progress. Thus, it is captured by the joint effect of technical progress and capital input growth.

[7] CCD use the word of 'scale factor' for the returns to scale.

[8] For the decomposition of Nemoto and Goto (2005), we interpret the product of 'scale change' and 'input and output mix effects' as the returns to scale effect. Their result identifies the combined effect of changes in the composition of inputs and that of outputs.

important in explaining LP growth than in explaining TFP growth. When the underlying technology exhibits constant returns to scale, the returns to scale effect disappears from TFP growth. However, it still plays a role for LP growth. This is because even if the underlying technology exhibits constant returns to scale, the short-run production frontier is likely not to exhibit constant returns to scale.

Triplett and Bosworth (2004, 2006) and Bosworth and Triplett (2007) observed that LP growth in the service industries was much less than in the goods industries in U.S. economy since the early 1970s. As we discussed above, there are two underlying factors to LP growth; therefore, possible explanations for the low LP growth in the services industries are as follows: 1) the joint effect of technical progress and increases in capital inputs is modest; 2) an increase in labour inputs induces negative returns to scale effects; 3) both 1) and 2). We apply our decomposition result to U.S. industry data to compare the relative contributions of the two effects.

Section 2 illustrates the two effects underlying LP growth graphically. Section 3 discusses the measure of the joint effect of technical progress and capital input growth in the multiple-inputs multiple-outputs case. Section 4 discusses the measure of the returns to scale effect in the multiple-inputs multiple-outputs case. We show that the product of the joint effect of technical progress and capital input growth, and the returns to scale effect coincides with LP growth. Section 5 includes the application to the U.S. industry data. Section 6 presents the conclusions.

2. Two Sources of Labour Productivity Growth

We display graphically what derives LP growth using a simple model of one output y and two inputs: labour input x_L and capital input x_K. Suppose that a firm produces outputs y^0 and y^1 using inputs (x_K^0, x_L^0) and (x_K^1, x_L^1). Period t production technology is described by the period t production function $y = f^t(x_K, x_L)$ for t = 0 and 1. Let us begin by considering how this joint effect of technical progress and capital input growth raises LP. Figure 1 illustrates the case when the joint effects of technical progress and capital input growth positively affect the productive capacity of labour. The lower curve represents the period 0 short-run production frontier. It indicates how much output can be produced by using a specified quantity of labour given the capital and technology available in period 0. Similarly, the higher curve represents the period 1 short-run production frontier. It indicates how much output can be produced by using a specified quantity of labour given the capital and technology available in period 1.

Since the short-run production frontier shifts upward, the output attainable from a given labour input x_L increases between the two periods such that $f^1(x_K^1, x_L) > f^0(x_K^0, x_L)$. The corresponding LP also grows such that $f^1(x_K^1, x_L)/x_L > f^0(x_K^0, x_L)/x_L$. Thus, the ratio $f^1(x_K^1, x_L)/f^0(x_K^0, x_L) = (f^1(x_K^1, x_L)/x_L)/(f^0(x_K^0, x_L)/x_L)$ captures the joint effect on LP growth of technical progress and capital input growth. Note that the ratio is also a measure of the distance between the short-run production frontiers of periods 0 and 1 in the direction of the y axis, evaluated at x_L. The ratio increases as the distance between the period 0 and the period 1 short-run production frontiers increases. Therefore, the joint effect of technical progress and capital input growth can be captured throughout by measuring the shift in the short-run production frontier.

[Place Figure 1 appropriately here]

Any quantity of labour input can produce more output in period 1 than in period 0, reflecting the positive joint effect of technical progress and capital input growth. The firm increases its demand for labour input from x_L^0 to x_L^1, exploiting the increased productive capacity of labour input. Suppose that production takes place at *A* for period 0 and at *B* for period 1. The slope of the ray from the origin to *A* and *B* indicates the LP of each period. Since y^1/x_L^1 is smaller than y^0/x_L^0, LP declines between the two periods. The fact that LP can decline despite the outward shift in the short-run production frontier suggests that another factor contributes to LP growth.[9]

The path from *A* to *B* can be divided into two parts: the vertical jump from *A* to *A'* and the movement along the period 1 short-run production frontier from *A'* to *B*. Along the vertical jump from *A* to *A'*, the LP changes from y^0/x_L^0 to $f^1(x_K^1, x_L^0)/x_{L.}^1$ Its ratio $(y^1/x_L^1)/(f^1(x_K^1, x_L^0)/x_L^0)$ is considered to be the growth in LP induced by the shift in the short-run production frontier, which is the joint effect of technical progress and capital input growth. LP growth is offset by the change in labour input from x_L^0 to x_L^1. The movement along the period 1 short-run production frontier from *A'* to *B* reduces LP from $f^1(x_K^1, x_L^0)/x_L^0$ to y^1/x_L^1. We call the LP growth induced by movement along the short-run production frontier $(y^1/x_L^1)/(f^1(x_K^1, x_L^0)/x_L^0)$ the returns to scale effect.

In this section, we illustrate two sources of LP growth using the simple one-output two-inputs model. However, the division of the path from *A* to *B* into two steps from *A* to *A'* and from *A'* to *B* is merely an example. It is also possible to decompose the change from *A* to *B* into the movement along the period 0 short-run production function from *A* to *B'* and the vertical jump from *B'* to *B*. In this case, the former movement reflects the returns to scale effect, and the latter jump reflects the joint effect of technical progress and capital input growth.

For measuring the joint effect of technical progress and capital input growth, the important consideration is the quantity of labour input at which the distance between two short-run production frontiers is evaluated. For measuring the returns to scale effect, whether we consider the movement along the period 0 or 1 short-run production frontier matters. Hereafter, we generalize our discussion to the more general multiple-inputs multiple-outputs case and propose measures for the two effects that are immune from choosing the arbitrary benchmark.

3. Joint Effect of Technical Progress and Capital Input Growth

A firm is considered as a productive entity transforming inputs into outputs. We assume there are *M* (net) outputs, $\boldsymbol{y} = [y_1,\ldots, y_M]^{\mathrm{T}}$, and $P + Q$ inputs consisting of *P* types of capital inputs, $\boldsymbol{x}_K = [x_{K,1},\ldots, x_{K,P}]^{\mathrm{T}}$, and *Q* types of labour inputs, $\boldsymbol{x}_L = [x_{L,1},\ldots, x_{L,Q}]^{\mathrm{T}}$.[10] The period *t* production possibility set S^t consists of all feasible combinations of inputs and outputs, and it is defined as

$$(1)\ S^t \equiv \{(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) : (\boldsymbol{x}_K, \boldsymbol{x}_L)\ \text{can produce}\ \boldsymbol{y}\}.$$

[9] This is just an example of the fact that the shift in the short-run production frontier is not the only one contribution factor to LP growth. We do not exclude the case that LP increases under the outward shift in the short-run production frontier.

[10] Outputs include intermediate inputs. If output *m* is an intermediate input, then $y_m < 0$. Hence, the nominal value of total (net) outputs $\boldsymbol{p}\cdot\boldsymbol{y}$ is the value-added that a firm generates.

We assume S^t satisfies Färe and Primont's (1995) axioms that guarantee the existence of output and input distance functions. The period t production frontier, which is the boundary of S^t, is represented by the *period t input requirement function* G^t. It is defined as follows:

$$(2)\ F^t(\boldsymbol{y}, \boldsymbol{x}_{K,-1}, \boldsymbol{x}_L) \equiv \min_{x_{K,1}}\{x_{K,1} : (\boldsymbol{y}, x_{K,1}, \boldsymbol{x}_{K,-1}, \boldsymbol{x}_L) \in S^t\}.$$

It represents the minimum amount of the first capital input that a firm can use at period t, producing output quantities $\boldsymbol{y}$, holding other capital inputs $\boldsymbol{x}_{K,-1} = [x_{K,2},\ldots, x_{K,P}]^{\mathrm{T}}$ and labour inputs $\boldsymbol{x}_L$ fixed. This function, which is originally formulated for characterizing the period t production frontier, also can be used for characterizing the period t short-run production frontier. Given period t capital input $\boldsymbol{x}_K{}^t$, the set of labour inputs $\boldsymbol{x}_L$ and outputs $\boldsymbol{y}$ satisfying $x_{K,1}{}^t = F^t(\boldsymbol{y}, \boldsymbol{x}_{K,-1}{}^t, \boldsymbol{x}_L)$ forms the period t short-run production frontier.

CCD measure the shift in the production frontier by using the output distance function. Adjusting their approach, we also use the output distance function to measure the shift in the short-run production frontier. Using the input requirement function, the *period t output distance function* for t = 0 and 1 is defined as follows:

$$(3)\ D^t_O(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) \equiv \min_\delta \left\{\delta : F^t\left(\frac{\boldsymbol{y}}{\delta}, \boldsymbol{x}_{K,-1}, \boldsymbol{x}_L\right) \le x_{K,1}\right\}.$$

Given capital inputs $\boldsymbol{x}_K$ and labour inputs $\boldsymbol{x}_L$, $D_O{}^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is the minimum contraction of outputs $\boldsymbol{y}$ so that the contracted outputs $\boldsymbol{y}/D_O{}^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$, capital inputs $\boldsymbol{x}_K$ and labour inputs $\boldsymbol{x}_L$ are on the period t production frontier. If $(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is on the period t production frontier, $D_O{}^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ equals 1. Note that $D_L{}^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is linearly homogeneous in $\boldsymbol{y}$.

We also can relate it to the short-run production frontier. Given labour inputs $\boldsymbol{x}_L$, $D_O{}^t(\boldsymbol{y}, \boldsymbol{x}_K{}^t, \boldsymbol{x}_L)$ is the minimum contraction of outputs so that the contracted outputs $\boldsymbol{y}/D_O{}^t(\boldsymbol{y}, \boldsymbol{x}_K{}^t, \boldsymbol{x}_L)$ and labour input $\boldsymbol{x}_L$ are on the period t short-run production frontier. Thus, $D_O{}^t(\boldsymbol{y}, \boldsymbol{x}_K{}^t, \boldsymbol{x}_L)$ provides a radial measure of the distance of $\boldsymbol{y}$ to the period t short-run production frontier. We measure the shift in the short-run production frontier by comparing the radial distances of $\boldsymbol{y}$ to the short-run production frontiers of the periods 0 and 1. It is defined as follows:[11]

$$(4)\ SHIFT(\boldsymbol{y}, \boldsymbol{x}_L) \equiv \frac{D^0_O(\boldsymbol{y}, \boldsymbol{x}^0_K, \boldsymbol{x}_L)}{D^1_O(\boldsymbol{y}, \boldsymbol{x}^1_K, \boldsymbol{x}_L)}.$$

If technical progress and capital input growth have a positive effect on the productive capacity of labour between periods 0 and 1, the short-run production frontier shifts outward. Given labour inputs $\boldsymbol{x}_L$, more outputs can be produced. Thus, the minimum contraction factor for given outputs $\boldsymbol{y}$ declines such that $D_O{}^1(\boldsymbol{y}, \boldsymbol{x}_K{}^1, \boldsymbol{x}_L) \le D_O{}^0(\boldsymbol{y}, \boldsymbol{x}_K{}^0, \boldsymbol{x}_L)$, leading to $SHIFT\ (\boldsymbol{y}, \boldsymbol{x}_L) \ge 1$. Similarly, the negative joint effect of technical progress and capital input growth leads to $SHIFT\ (\boldsymbol{y}, \boldsymbol{x}_L) \le 1$.

Each choice of reference vectors $(\boldsymbol{y}, \boldsymbol{x}_L)$ might generate a different measure of the shift in the short-run production frontier from periods 0 to 1. We calculate two measures using different reference vectors $(\boldsymbol{y}^0, \boldsymbol{x}_L{}^0)$ and $(\boldsymbol{y}^1, \boldsymbol{x}_L{}^1)$. Since these reference outputs

[11] CCD and Färe et al (1994) introduce a measure of the shift in the production frontier by using the ratio of the output distance function. Given $(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$, Färe et al (1994) measure the shift in the production frontier by $D_O{}^0(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)/D_O{}^1(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$.

and labour inputs are, in fact, chosen in each period, they are equally reasonable. Following Fisher (1922) and CCD, we use the geometric mean of these measures as a theoretical measure of the joint effect of technical progress and capital input growth, *SHIFT*, as follows:[12]

$$(5)\ SHIFT \equiv \sqrt{SHIFT(\boldsymbol{y}^0, \boldsymbol{x}_L^0) \cdot SHIFT(\boldsymbol{y}^1, \boldsymbol{x}_L^1)}.$$

The case of one output and one labour input offers a graphical interpretation of *SHIFT*. In Figure 1, it is reduced to the following formula:

$$(6)\ SHIFT = \sqrt{(f^1(\boldsymbol{x}_L^1, \boldsymbol{x}_L^0)/\boldsymbol{y}^0)(\boldsymbol{y}^1/f^0(\boldsymbol{x}_L^0, \boldsymbol{x}_L^1))}.$$

Given a quantity of labour input, the ratio of the output attainable from such a labour input at period 1 to the output attainable at period 0 represents the extent to which the short-run production function expands. *SHIFT* is the geometric mean of those ratios conditional on $\boldsymbol{x}_L^0$ and $\boldsymbol{x}_L^1$.

SHIFT is a theoretical measure defined by the unknown distance functions, and there are alternative ways of implementing it. We show that the theoretical measure coincides with a formula of price and quantity observations under the assumption of a firm's profit-maximizing behaviour and a translog functional form for the output distance function.[13] Our approach is drawn from CCD, which implements the Malmquist productivity index, a theoretical measure of the shift in the production frontier. They show that the Malmquist productivity index coincides with a different index number formula of price and quantity observations, called the Törnqvist productivity index, under similar assumptions.[14]

CCD also show that the first-order derivatives of the distance function D^t with respect to quantities at the period t actual production plan ($\boldsymbol{y}^t$, $\boldsymbol{x}_K^t$, $\boldsymbol{x}_L^t$) are computable from price and quantity observations. Their equivalence result between the Malmquist and Törnqvist productivity indexes relies on these relationships. Utilizing the same relationships, we also show that *SHIFT* coincides with an index number formula of price and quantity observations. Equations (7)–(16) already have been derived by CCD, but for completeness of discussion we outline below how to compute the first-order derivatives of the distance functions from price and quantity observations. The implicit function theorem is applied to the input requirement function $F^t(\boldsymbol{y}/\delta, \boldsymbol{x}_{K,-1}, \boldsymbol{x}_L) = x_{K,1}$ to solve for $\delta = D_O^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ around ($\boldsymbol{y}^t$, $\boldsymbol{x}_K^t$, $\boldsymbol{x}_L^t$).[15] Its derivatives are

[12] Since the firm's profit maximization is assumed, it is possible to adopt a different formulation for the measure of the shift in the short-run production frontier: $SHIFT = (D_O^0(\boldsymbol{y}^1, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)/D_O^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0))^{1/2}(D_O^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1)/D_O^1(\boldsymbol{y}^0, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0))^{1/2}$. This formulation is closer to the Malmquist productivity index introduced by CCD.

[13] Alternative approaches involve estimating the underlying distance function by econometric or linear programming approaches. Either approach requires sufficient empirical observations. Our approach, originated by CCD, is applicable so long as price and quantity observations are available for the current and the reference periods. See Nishimizu and Page (1982) for the application of the econometric technique, and see Färe et al. (1994) for the application of the linear programming technique.

[14] CCD justify the use of the Törqvist productivity index, which is the Törqvist output quantity index divided by the Törqvist input quantity index. In addition to the translog functional form for the output distance function, CCD assume the firm's cost-minimizing and revenue-maximizing functions. Their approach is known as the exact index number approach, which constructs a formula of price and quantity observations that approximate theoretical measures.

[15] We assume the following three conditions are satisfied for t = 0 and 1: F^t is differentiable at the point ($\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t$), $\boldsymbol{y}^t >> 0_M$ and $\boldsymbol{y}^t \cdot \nabla_y F(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) > 0$.

represented by the derivatives of $F^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$. We have the following equations for t = 0 and 1:

$$(7)\ \nabla_y D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{\boldsymbol{y}^t \cdot \nabla_y F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t)} \nabla_y F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t),$$

$$(8)\ \nabla_{x_K} D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{\boldsymbol{y}^t \cdot \nabla_y F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t)} \begin{bmatrix} -1 \\ \nabla_{x_{K,-1}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t) \end{bmatrix},$$

$$(9)\ \nabla_{x_L} D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{\boldsymbol{y}^t \cdot \nabla_y F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t)} \nabla_{x_L} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t).$$

We assume the firm's profit-maximizing behaviour. Thus, $(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) >> 0_{N+P+Q}$ is a solution to the following period t profit maximization problem for t = 0 and 1[16]:

$$(10)\ \max\{\boldsymbol{p}^t \cdot \boldsymbol{y} - r_1^t F^t(\boldsymbol{y}, \boldsymbol{x}_{K,-1}, \boldsymbol{x}_L) - \boldsymbol{r}_{-1}^t \cdot \boldsymbol{x}_{K,-1} - \boldsymbol{w}^t \cdot \boldsymbol{x}_L\}.$$

Outputs are sold at the positive producer prices $\boldsymbol{p} = [p_1,\ldots, p_M]^T >> 0$, capital inputs are purchased at the positive rental prices $\boldsymbol{r} = [r_1,\ldots, r_P]^T >> 0$ and labour inputs are purchased at the positive wages $\boldsymbol{w} = [w_1,\ldots, w_Q]^T >> 0$. Note that $\boldsymbol{r}_{-1} = [r_{-1},\ldots, r_P]^T$. The period t profit maximization problem yields the following first-order conditions for t = 0 and 1:

$$(11)\ \boldsymbol{p}^t = r_1^t \nabla_y F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t),$$

$$(12)\ \boldsymbol{r}_{-1}^t = -r_1^t \nabla_{x_{K,-1}} F^s(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t),$$

$$(13)\ \boldsymbol{w}^t = -r_1^t \nabla_{x_L} F^s(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t).$$

By substituting (11)–(13) into (7)–(9), we obtain the following equations for t = 0 and 1:

$$(14)\ \nabla_y D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \boldsymbol{p}^t / \boldsymbol{p}^t \cdot \boldsymbol{y}^t,$$

$$(15)\ \nabla_{x_K} D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = [r_1^t / \boldsymbol{p}^t \cdot \boldsymbol{y}^t] \begin{bmatrix} -1 \\ \nabla_{x_{K,-1}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t) \end{bmatrix} = [1/ \boldsymbol{p}^t \cdot \boldsymbol{y}^t] \begin{bmatrix} -r_1 \\ -\boldsymbol{r}_{-1} \end{bmatrix},$$

$$(16)\ \nabla_{x_L} D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = -\boldsymbol{w}^t / \boldsymbol{p}^t \cdot \boldsymbol{y}.$$

Equations (14)–(16) allow us to compute derivatives of the distance function without knowing the output distance function itself. Information concerning the derivatives is useful for calculating values of the output distance functions. However, one disadvantage is that the derivatives of the period t output distance function need to be evaluated at the period t actual production plan $(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t)$ in equations (14)–(16) for t = 0 and 1. The distance functions evaluated at the hypothetical production plan such as $(\boldsymbol{y}^1, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $(\boldsymbol{y}^0, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ also constitute *SHIFT*. Hence, the above equations are insufficient for implementing *SHIFT*. In addition to a firm's profit maximization, we further assume a following translog functional form for the period t output distance function for t = 0 and 1. It is defined as

[16] We assume that there always exists a solution to the firm's profit maximization problem. Thus, we implicitly exclude the case that the underlying technology exhibits increasing returns to scale. However, it is possible that the underlying technology exhibits constant or decreasing returns to scale.

$$
\begin{aligned}
\ln D_O^t(\boldsymbol{y},\boldsymbol{x}_K,\boldsymbol{x}_L) \equiv\ & \alpha_0^t + \sum_{m=1}^{M} \alpha_m^t \ln y_m + (1/2)\sum_{i=1}^{M}\sum_{j=1}^{M} \alpha_{i,j} \ln y_i \ln y_j \\
& + \sum_{p=1}^{P} \beta_p^t \ln x_{K,p} + (1/2)\sum_{i=1}^{P}\sum_{j=1}^{P} \beta_{i,j} \ln x_{K,i} \ln x_{K,j} \\
& + \sum_{q=1}^{Q} \chi_q^t \ln x_{L,q} + (1/2)\sum_{i=1}^{Q}\sum_{j=1}^{Q} \chi_{i,j} \ln x_{L,i} \ln x_{L,j} \\
& + \sum_{m=1}^{M}\sum_{p=1}^{P} \delta_{m,p} \ln y_m \ln x_{K,p} + \sum_{m=1}^{M}\sum_{q=1}^{Q} \varepsilon_{m,q} \ln y_m \ln x_{L,q} \\
& + \sum_{p=1}^{P}\sum_{q=1}^{Q} \varphi_{p,q} \ln x_{K,p} \ln x_{L,q}
\end{aligned} \tag{17}
$$

where the parameters satisfy the following restrictions:

(18) $\alpha_{i,j} = \alpha_{j,i}$ for all i and j such that $1 \le i < j \le M$;

(19) $\beta_{i,j} = \beta_{j,i}$ for all i and j such that $1 \le i < j \le P$;

(20) $\chi_{i,j} = \chi_{j,i}$ for i and j such that $1 \le i < j \le Q$;

(21) $\sum_{n=1}^{N} \alpha_n^t = 1$;

(22) $\sum_{i=1}^{M} \alpha_{i,m} = 0$ for $m = 1,\ldots,M$;

(23) $\sum_{m=1}^{M} \delta_{m,p} = 0$ for $p = 1,\ldots,P$;

(24) $\sum_{i=1}^{M} \varepsilon_{m,q} = 0$ for $q = 1,\ldots,Q$.

Restrictions (21)–(24) guarantee the linear homogeneity in $\boldsymbol{y}$. The translog functional form characterized in (17)–(24) is a flexible functional form so that it can approximate an arbitrary output distance function to the second order at an arbitrary point. Thus, the assumption of this functional form does not harm any generality of the output distance function. Note that the coefficients for the linear terms and the constant term are allowed to vary across periods. Thus, technical progress under the translog distance function is by no means limited to Hicks neutral and is able to represent a variety of technical progress. Under the assumptions of the profit-maximizing behaviour and the translog functional form, a theoretical measure *SHIFT* is computed from price and quantity observations.

Proposition 1: Assume that the output distance functions D_O^0 and D_O^1 have the translog functional form defined by (17)–(24) and that a firm follows competitive profit-maximizing behaviour in periods t = 0 and 1. Then, the joint effect of technical progress and capital input growth, *SHIFT*, can be computed from observed prices and quantities as follows:

$$
\ln SHIFT = \sum_{m=1}^{M} s_m \ln\left(\frac{y_m^1}{y_m^0}\right) - \sum_{q=1}^{Q} s_{L,q} \ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right), \tag{25}
$$

where s_m and $s_{L,q}$ are the average value-added shares of output m and labour input q between periods 0 and 1 such that:

$$
s_m = \frac{1}{2}\left(\frac{p_m^0 y_m^0}{\boldsymbol{p}^0 \cdot \boldsymbol{y}^0} + \frac{p_m^1 y_m^1}{\boldsymbol{p}^1 \cdot \boldsymbol{y}^1}\right) \text{ and } s_{L,q} = \frac{1}{2}\left(\frac{w_q^0 x_{L,q}^0}{\boldsymbol{p}^0 \cdot \boldsymbol{y}^0} + \frac{w_q^1 x_{L,q}^1}{\boldsymbol{p}^1 \cdot \boldsymbol{y}^1}\right).
$$

The index number formula in (25) can be interpreted as the ratio of a quantity index of output to a quantity index of labour input. Note that no data on price and quantity of capital inputs appear in this formula. Although the shift in the short-run production frontier reflects technical progress as well as the change in capital input, we can measure its shift without resort to capital input data explicitly.

4. Returns to Scale Effect

As shown in Figure 1, the shift in the short-run production frontier is not the only factor contributing to the growth in LP. Even when there is no change in the short-run production frontier, the movement along the frontier could raise LP, exploiting the curvature of the short-run production frontier. We refer to LP growth induced by the movement along the short-run production frontier as the *returns to scale effect*. In the simple model consisting of one output and one labour input, LP is defined as output per one unit of labour input. Therefore, LP growth, which is the growth rate of LP from the previous period to the current period, coincides with the growth rate of output divided by the growth rate of labour input. Since the returns to scale effect is the LP growth induced by the movement along the short-run production frontier, it is computed by the growth rates of output and labour input between the two end points of the movement. Figure 2 shows how the movement along the period t short-run production frontier from point C to D affects LP. Comparing points C and D, the growth rate of output is $f^t(\boldsymbol{x}_K^t, \boldsymbol{x}_L^1)/f^t(\boldsymbol{x}_K^t, \boldsymbol{x}_L^0)$, and the growth ratio of labour input is $\boldsymbol{x}_L^1/\boldsymbol{x}_L^0$. The growth rate of LP between the two points coincides with the growth rate of output divided by that of labour input, so that $(f^t(\boldsymbol{x}_K^t, \boldsymbol{x}_L^1)/f^t(\boldsymbol{x}_K^t, \boldsymbol{x}_L^0))/(\boldsymbol{x}_L^1/\boldsymbol{x}_L^0) = (f^t(\boldsymbol{x}_K^t, \boldsymbol{x}_L^1)/\boldsymbol{x}_L^1)/(f^t(\boldsymbol{x}_K^t, \boldsymbol{x}_L^0)/\boldsymbol{x}_L^0)$.

[Place Figure 2 appropriately here]

We generalize the growth rates of labour input and output between two points on the period t short-run production frontier in order to measure the returns to scale effect in the multiple-inputs multiple-outputs case. First, we investigate the counterpart of the growth rate of labour inputs in the multiple-inputs multiple-outputs case. CCD define the input quantity index, which is the counterpart of the growth rate of total inputs, by comparing the radial distance between the two input vectors and the period t production frontier. The input distance function is used for the radial scaling of total inputs. Adapting the input distance function used by CCD, we introduce the labour input distance function that measures the radial distance of labour inputs $\boldsymbol{x}_L$ to the period t production frontier. The *period t labour input distance function* for $t = 0$ and 1 is defined as follows:

$$(26)\ D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) \equiv \max_\delta \left\{ \delta : F^t\left(\boldsymbol{y}, \boldsymbol{x}_{K,-1}, \frac{\boldsymbol{x}_L}{\delta} \right) \le x_{K,1} \right\}.$$

Given outputs $\boldsymbol{y}$ and capital inputs $\boldsymbol{x}_K$, $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is the maximum contraction of labour inputs $\boldsymbol{x}_L$ so that the contracted labour inputs $\boldsymbol{x}_L/D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ and capital inputs $\boldsymbol{x}_K$ with outputs $\boldsymbol{y}$ are on the period t production frontier. If $(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is on the period t production frontier, $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ equals 1. Note that $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is linearly homogeneous in $\boldsymbol{x}_K$.

We can also relate it to the short-run production frontier. Given outputs $\boldsymbol{y}$, $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L)$ is the maximum contraction of labour inputs so that the contracted labour inputs

$\boldsymbol{x}_L/D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L)$ and outputs $\boldsymbol{y}$ are on the period t short-run production frontier. Thus, $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L)$ provides a radial measure of the distance of $\boldsymbol{x}_L$ to the period t short-run production frontier conditional on $\boldsymbol{y}$. We construct the counterpart of the growth rate of labour input by comparing two labour inputs $\boldsymbol{x}_L^0$ and $\boldsymbol{x}_L^1$ to the period t short-run production frontier conditional on $\boldsymbol{y}$. It is defined as follows:

$$(27)\ LABOUR(t, \boldsymbol{y}) \equiv D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1) / D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0)$$

If labour inputs increase between two periods, $\boldsymbol{x}_L^1$ moves further away from the origin than $\boldsymbol{x}_L^0$, meaning that the labour input vector $\boldsymbol{x}_L^1$ is larger than the labour input vector $\boldsymbol{x}_L^0$. The maximum contraction factor for producing outputs $\boldsymbol{y}$ with the period t capital inputs $\boldsymbol{x}_K^t$ using the period t technology increases such that $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0) \leq D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1)$, leading to $LABOUR(t,\boldsymbol{y}) \geq 1$. Similarly, if labour input shrinks between two periods, $\boldsymbol{x}_L^1$ moves closer to the origin than does $\boldsymbol{x}_L^0$, leading to $LABOUR\ (t,\boldsymbol{y}) \leq 1$.

Second, we generalize the growth rate of outputs between two points on the period t short-run production frontier. In the multiple-inputs multiple-outputs case, outputs attainable from given capital inputs $\boldsymbol{x}_K$ and labour inputs $\boldsymbol{x}_L$ are not uniquely determined by the short-run production frontier. Let $P^t(\boldsymbol{x}_K,\boldsymbol{x}_L)$ be the portion of the period t production frontier that is conditional on capital inputs $\boldsymbol{x}_K$ and labour inputs $\boldsymbol{x}_L$, consisting of the set of maximum outputs that are attainable from $\boldsymbol{x}_K$ and $\boldsymbol{x}_L$ using technology available at period t. It is defined as follows:

$$(28)\ P^t(\boldsymbol{x}_K, \boldsymbol{x}_L) \equiv \{\boldsymbol{y} : F^t(\boldsymbol{y}, \boldsymbol{x}_{K,-1}, \boldsymbol{x}_L) = x_{K,1}\}.$$

We can also relate it to the short-run production frontier. The portion of the period t short-run production frontier that is conditional on $\boldsymbol{x}_L$ is represented by $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L)$. Since $D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L)$ provides a radial measure of the distance of $\boldsymbol{y}$ to the period t short-run production frontier conditional on $\boldsymbol{x}_L$, it can also be interpreted as a radial measure of the distance of $\boldsymbol{y}$ to $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L)$. We construct the counterpart of the growth rate of outputs between two points on the period t short-run production frontier by measuring the distance between $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L^0)$ and $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L^1)$. We start with the reference outputs vector $\boldsymbol{y}$. We measure the distance between $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L^0)$ and $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L^1)$, comparing the radial distances from outputs $\boldsymbol{y}$ to $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L^0)$ and $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L^1)$. It is defined as follows:

$$(29)\ OUTPUT(t, \boldsymbol{y}) \equiv D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0) / D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1).$$

If the labour input growth makes it possible to produce more outputs while holding capital input fixed and using the same technology, the set of outputs attainable from $\boldsymbol{x}_L^1$, $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L^1)$, shifts outward to that of outputs attainable from $\boldsymbol{x}_L^0$, $P^t(\boldsymbol{x}_K^t,\boldsymbol{x}_L^0)$. Thus, the minimum contraction factor for given outputs $\boldsymbol{y}$ declines such that $D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1) \leq D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0)$, leading to $OUTPUT\ (t, \boldsymbol{y}) \geq 1$. Similarly, if the change in labour inputs reduces outputs attainable from given capital and labour, inputs leads to $OUTPUT\ (t, \boldsymbol{y}) \leq 1$.

Using the counterparts of the growth rate of outputs and labour inputs between two points on the period t short-run production frontier, we can propose a measure for the LP growth between these two points. When we consider the movement along the period t short-run production and use outputs $\boldsymbol{y}$ as reference, the returns to scale effect is defined as follows[17]:

$$SCALE(t,\boldsymbol{y}) \equiv OUTPUT(t,\boldsymbol{y}) / LABOUR(t,\boldsymbol{y})$$

(30)
$$= \left(\frac{D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0)}{D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1)} \right) \Bigg/ \left(\frac{D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1)}{D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0)} \right).$$

Each choice of reference short-run production frontier and reference output vector $\boldsymbol{y}$ might generate a different measure of the returns to scale effect going from period 0 to period 1. We calculate two measures by using short-run production frontiers and output vectors that are available at the same period: period 0 short-run production frontier and period 0 output vector $\boldsymbol{y}^0$; period 1 short-run production frontier and period 1 output vector $\boldsymbol{y}^1$. Since these sets of short-run production frontiers and output vectors are equally reasonable, we use the geometric mean of these measures as a theoretical index of the returns to scale effect, *SCALE*, as follows:

(31) $$SCALE \equiv \sqrt{SCALE(0,\boldsymbol{y}^0) \cdot SCALE(1,\boldsymbol{y}^1)} .$$

The case of one output and one labour input offers us a graphical interpretation of *SCALE*. In Figure 1, (30) can be reduced to the following formula:

(32) $$SCALE = \sqrt{\left(\frac{f^0(x_K^0, x_L^1)}{y^0} \Bigg/ \frac{x_L^1}{x_L^0} \right) \left(\frac{y^1}{f^1(x_K^1, x_L^0)} \Bigg/ \frac{x_L^1}{x_L^0} \right)} .$$

Given the period t short-run production frontier, the ratio of the LP associated with x_L^1 to the LP associated with x_L^0 represents the LP growth induced by the change in labour input. *SCALE* is the geometric mean of those ratios conditional on the period 0 and 1 short-run production frontiers.

CCD apply the implicit function theorem to the input requirement function with the output distance function such as $F^t(\boldsymbol{y}^t/\delta, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t) = x_{K,1}^t$ where $\delta = D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t)$. Accordingly, we apply the implicit function theorem to the input requirement function with the labour input distance function such as $F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t/\delta) = x_{K,1}^t$ where $\delta = D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t)$. In this case, $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is differentiable around the point $(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t)$.[18] Its derivatives are represented by the derivatives of $F^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$. We have the following equations for $t = 0$ and 1:

(33) $$\nabla_y D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{\boldsymbol{x}_L^t \cdot \nabla_{x_L} F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t)} \nabla_y F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t) ,$$

(34) $$\nabla_{x_K} D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{\boldsymbol{x}_L^t \cdot \nabla_{x_L} F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t)} \begin{bmatrix} -1 \\ \nabla_{x_{K,-1}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t) \end{bmatrix},$$

(35) $$\nabla_{x_L} D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{\boldsymbol{x}_L^t \cdot \nabla_{x_L} F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t)} \nabla_{x_L} F^t(\boldsymbol{y}^t, \boldsymbol{x}_{K,-1}^t, \boldsymbol{x}_L^t) .$$

---

[17] This formulation is a counterpart of the scale effect on TFP growth that is proposed by Lovell (2003; 450). Lovell's definition is based on the input distance function instead of the labour input distance function.

[18] We also assume the following three conditions are satisfied for $t = 0$ and 1: F^t is differentiable at the point $(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t)$, $x_L^t >> 0_M$ and $x_L^t \cdot \nabla_{x_L} F(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) > 0$.

We assume that $(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) >> 0_{N+P+Q}$ is a solution to the period t profit maximization problem (10) for t = 0 and 1. By substituting (11)–(13) obtained from the profit maximization into (33)–(35), we obtain the following equations for t = 0 and 1:

(36) $\nabla_y D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = -\boldsymbol{p}^t / \boldsymbol{w}^t \cdot \boldsymbol{x}_L^t$,

(37) $\nabla_{x_K} D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = [1/\boldsymbol{w}^t \cdot \boldsymbol{x}_L^t]\begin{bmatrix} r_1 \\ \boldsymbol{r}_{-1} \end{bmatrix}$,

(38) $\nabla_{x_L} D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \boldsymbol{w}^t / \boldsymbol{w}^t \cdot \boldsymbol{x}_L^t$.

Equations (36)–(38) allow us to compute the derivatives of the labour input distance function without knowing the labour input distance function itself. Information concerning the derivatives is useful for calculating the values of *SCALE*, which is defined by the distance functions. However, one disadvantage is that the derivatives of the period t distance function need to be evaluated at the period t actual production plan $(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t)$ in equations (36)–(38) for t = 0 and 1. The distance functions evaluated at the hypothetical production plan such as $(\boldsymbol{y}^1, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $(\boldsymbol{y}^0, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ also constitute *SCALE*. Hence, the above equations are insufficient for obtaining *SCALE*. In addition to the firm's profit maximization, we further assume the translog functional form for the period t labour input distance function for t = 0 and 1. It is defined as follows:

(39)
$$\begin{aligned} \ln D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) \equiv\ & \tau_0^t + \sum_{m=1}^M \tau_m^t \ln y_m + (1/2)\sum_{i=1}^M \sum_{j=1}^M \tau_{i,j} \ln y_i \ln y_j \\ & + \sum_{p=1}^P \upsilon_p^t \ln x_{K,p} + (1/2)\sum_{i=1}^P \sum_{j=1}^P \upsilon_{i,j} \ln x_{K,i} \ln x_{K,j} \\ & + \sum_{q=1}^Q \omega_q^t \ln x_{L,q} + (1/2)\sum_{i=1}^Q \sum_{j=1}^Q \omega_{i,j} \ln x_{L,i} \ln x_{L,j} \\ & + \sum_{m=1}^M \sum_{p=1}^P \xi_{m,p} \ln y_m \ln x_{K,p} + \sum_{m=1}^M \sum_{q=1}^Q \psi_{m,q} \ln y_m \ln x_{L,q} \\ & + \sum_{p=1}^P \sum_{q=1}^Q \zeta_{p,q} \ln x_{K,p} \ln x_{L,q} \end{aligned}$$

where the parameters satisfy the following restrictions:

(40) $\tau_{i,j} = \tau_{j,i}$ for all i and j such as $1 \le i < j \le M$;

(41) $\upsilon_{i,j} = \upsilon_{j,i}$ for all i and j such as $1 \le i < j \le P$;

(42) $\omega_{i,j} = \omega_{j,i}$ for i and j such as $1 \le i < j \le Q$;

(43) $\sum_{q=1}^Q \omega_q^t = 1$;

(44) $\sum_{i=1}^Q \omega_{i,q} = 0$ for $q = 1,\ldots, Q$;

(45) $\sum_{q=1}^Q \psi_{m,q} = 0$ for $m = 1,\ldots, M$;

(46) $\sum_{q=1}^Q \zeta_{p,q} = 0$ for $p = 1,\ldots, P$.

Equation (39) is the same functional form defined by (17) that we assumed for the output distance function in the discussion of *SHIFT*. However, the restrictions on parameters on the labour input distance function differ from those on the output distance function. We replace restrictions (21)–(24) with (43)–(46). While restrictions (21)–(24) guarantee the linear homogeneity in outputs $\boldsymbol{y}$ for the output distance function, restrictions (43)–(46) guarantee the linear homogeneity in labour inputs $\boldsymbol{x}_L$ for the labour input distance function.

The translog functional form characterized by (39)–(46) is a flexible functional form so that it can approximate an arbitrary labour input distance function to the second order at an arbitrary point. Thus, the assumption of this functional form does not harm any generality of the labour input distance function. Note that the coefficients for the linear terms and the constant term are allowed to vary across periods. Thus, technical progress under the translog distance function is by no means limited to Hicks neutral, and a variety of technical progress is possible. Under the assumptions of profit-maximizing behaviour and the translog functional form, a theoretical index of the returns to scale, *SCALE*, is computable from price and quantity observations.

Proposition 2: Assume the following: output distance functions $D_O{}^0$ and $D_O{}^1$ have the translog functional form defined by (17)–(24); labour input distance functions $D_L{}^0$ and $D_L{}^1$ have the translog functional form defined by (39)–(46) and a firm follows competitive profit-maximizing behaviour in periods t = 0 and 1. Then, the returns to scale effect, *SCALE*, can be computed from observable prices and quantities as follows:

$$(47)\ \ln SCALE = \sum\nolimits_{q=1}^{Q} s_{L,q} \ln\left(\frac{x_{L.q}^1}{x_{L,q}^0}\right) - \sum\nolimits_{q=1}^{Q} \bar{s}_{L,q} \ln\left(\frac{x_{L.q}^1}{x_{L,q}^0}\right).$$

where $s_{L,q}$ is the average value-added shares of labour input q and $\bar{s}_{L,q}$ is the average labour-compensation share of labour input q between periods 0 and 1 such that

$$s_{L,q} = \frac{1}{2}\left(\frac{w_q^0 x_{L,q}^0}{\boldsymbol{p}^0 \cdot \boldsymbol{y}^0} + \frac{w_q^1 x_{L,q}^1}{\boldsymbol{p}^1 \cdot \boldsymbol{y}^1}\right) \text{ and } \bar{s}_{L,q} = \frac{1}{2}\left(\frac{w_q^0 x_{L,q}^0}{\boldsymbol{w}^0 \cdot \boldsymbol{x}_L^0} + \frac{w_q^1 x_{L,q}^1}{\boldsymbol{w}^1 \cdot \boldsymbol{x}_L^1}\right).$$

The index number formula on the right-hand side of (47) can be interpreted as the ratio of the quantity indexes of labour inputs. Both terms in (47) are the weighted geometric average of the growth rates for labour inputs. The first term uses the ratio of labour compensation for a particular type of labour input to the value-added as weight, and the second term uses the ratio of labour compensation for a particular type of labour input to the total labour compensation as weight. Thus, if labour income share, which is the ratio of the total labour compensation to the value-added, is large, the difference between two terms (47) becomes small, making the magnitude of *SCALE* smaller. Conversely, if labour income share is small, the magnitude of *SCALE* becomes larger.

Starting from the understanding that the two contribution factors exist for the LP growth, we reached the index number formula for these factors independently. Our result, however, does not deny the possibility that other unknown factors explain LP growth. Fortunately, two factors of *SHIFT* and *SCALE* can fully explain the LP growth. The product of *SHIFT* and *SCALE* coincides with the index of the LP growth, as follows

Corollary 1: Assume the following: output distance functions $D_O{}^0$ and $D_O{}^1$ have the translog functional form defined by (17)–(24); labour input distance functions $D_L{}^0$ and $D_L{}^1$ have the translog functional form defined by (39)–(46) and a firm follows competitive profit-maximizing behaviour periods t = 0 and 1. Then, the product of *SHIFT* and *SCALE* can be computed from observed prices and quantities as follows:

$$(48)\ \ln SHIFT + \ln SCALE = \sum_{m=1}^{M} s_m \ln\left(\frac{y_m^1}{y_m^0}\right) - \sum_{q=1}^{Q} \bar{s}_{L,q} \ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right),$$

where s_m is the average value-added shares of output m and $\bar{s}_{L,q}$ is the average labour-compensation share of labour input q between periods 0 and 1 such that

$$s_m = \frac{1}{2}\left(\frac{p_m^0 y_m^0}{\boldsymbol{p}^0 \cdot \boldsymbol{y}^0} + \frac{p_m^1 y_m^1}{\boldsymbol{p}^1 \cdot \boldsymbol{y}^1}\right) \text{ and } \bar{s}_{L,q} = \frac{1}{2}\left(\frac{w_q^0 x_{L,q}^0}{\boldsymbol{w}^0 \cdot \boldsymbol{x}_L^0} + \frac{w_q^1 x_{L,q}^1}{\boldsymbol{w}^1 \cdot \boldsymbol{x}_L^1}\right).$$

The right side of equation (48) represents the logarithm of LP growth. The first term of the right-hand side coincides with the Törnqvist quantity index of outputs, and the second term is the Törnqvist quantity index of labour inputs. Equation (48) allows us to decompose LP growth fully into two components, *SHIFT* and *SCALE*, when multiple inputs and outputs are employed. This decomposition is justifiable as a generalization of the one-input one-output case in which LP growth is induced by the shift in the production frontier and the movement along the production frontier.

5. An Application to U.S. Industry Data

Having discussed the theory of the decomposition, we now explore its empirical significance with industry data. The industry data covering the period 1987–2007 is taken from the Bureau of Labour Statistics (BLS) multifactor productivity data. We use gross output, intermediate input, and labour input at current and constant prices by 58 industries, which constitute the non-farm private business sector. Labour input at constant prices measures the number of hours worked.[19] These industries are categorized either as goods-producing industries (goods industries, thereafter) or services-providing industries (services industries, thereafter).

[Place Table 1 appropriately here]

Table 1 compares LP growth and its components across the non-farm private business sector, the goods industries and the services industries. For the entire sample period 1987–2007, the returns to scale effect had a negative impact on LP growth of 2.17 percent per year in the non-farm private business sector. While the average rate of the joint effect of technical progress and capital input growth was 2.53 percent, it was largely offset by the returns to scale effect of −0.36 percent.

Triplett and Bosworth (2004, 2006) and Bosworth and Triplett (2007) found, in the U.S. industry data, that LP growth in the services industries was stagnant and lower than LP growth in the goods industries.[20] That difference in LP growth between the goods and the services industries is also documented in our dataset. During 1987–2007, the average growth rate of the goods industries LP of 2.76 percent was almost

[19] Thus, this measure of labour input does not appropriately capture changes in labour quality. The joint effect of technical progress and capital input growth includes the LP growth that is induced by changes in the characteristics of labour input.

[20] Triplett and Bosworth (2006) call the situation that LP growth in the services industries is likely to stagnate *Baumol's disease*. They argue that this disease has been cured in the middle 1990s.

twice that of the services industries LP of 1.95 percent. Although the returns to scale effect was subtle and negligible in the goods industries for the period, a significant negative returns to scale effect appeared in the services industries, and LP growth in the services industries rose significantly when the returns to scale effect was excluded. More than half the difference in LP growth between the goods and the services industries can be explained by the difference in the returns to scale effect. The joint effects of technical progress and capital input growth in both industry groups were, on average, very close over the period 1987–2007: 2.71 percent for the goods industries and 2.47 percent for the services industries. This reflects that the increase in labour input was occurring primarily in the services industries.

[Place Table 2 appropriately here]

Table 2 summarizes the growth in labour input for the non-farm private business sector, the goods industries and the services industries. Both the weighted and the un-weighted average of the detailed industries show labour input in the services industries increases more rapidly than that in the goods industries by at least 1.9 percent per year.

It is useful to divide the entire sample period 1987–2007 into two periods: the ‘productivity slowdown’ period 1987–1995 and the ‘productivity resurgence’ period 1995–2007. A productivity slowdown in U.S. economy started in the early 1970s, with an average annual growth rate of 1.39 percent for the non-farm private business sector during the period 1987–1995. Productivity growth surged after 1995, with an average annual growth rate of 2.69 percent during the period 1995–2005. As Triplett and Bosworth (2004, 2006) and Bosworth and Triplett (2007) pointed out, the services industries LP grew slowly, especially during the productivity slowdown period, with an average growth rate of 1.22 percent during the period 1970–1995. During the same period, the goods industries grew at an average annual rate of 1.82 percent. However, once we control for the returns to scale effect and consider only the joint effect of technical progress and capital input growth, the services industries, with the average annual rate of 1.83 percent, come very close to the goods industries with an average annual rate of 1.88 percent.

Thus, although LP growth was lesser in the services industries than it was in the goods industries, the productive capacity of labour, which is the output attainable from given labour inputs, increased in the services industries as much as it did in the goods industries. It reflects that the large increase in labour input in the services industries restrained LP from increasing significantly. While labour input in the goods industries slightly increased during the period 1987–1995, it even fell during the period 1995–2007, leading to the positive returns to scale effect. On the other hand, labour input in the services industries increased throughout the sample period, leading to the negative returns to scale effect. Thus, the part of the gap in LP growth explained by the gap in the returns to scale effect between the goods and the services industries became larger in the period 1995–2007.

[Place Tables 3 and 4 appropriately here]

Tables 3 and 4 show LP growth and its components, and growth in labour input and labour income share by industry during the periods 1987–1995 and 1995–2007. The pattern found in the aggregate study based on the sector data in Table 1 is also documented in the detailed industries. Most industries in the services industries show the negative returns to scale effect. It is especially significant during the period 1987–

1995. On the other hand, many industries in the goods industries show a modest returns to scale effect during the period 1987–1995, and a positive returns to scale effect during the period 1995–2007.

There are exceptional industries in both the goods and the services industries. Labour input grew largely with the average growth rate of more than 2 percent in three industries within the goods industries, leading a significantly negative returns to scale effect: *plastics and rubber products* industry during the period 1987–1995, *support activities for mining* industry and *construction* industry during the period 1995–2007. On the other hand, labour input fell for few industries in the services industries. Especially, the average growth rates of labour input in both the periods 1987–1995 and 1995–2007 are negative for three industries: *utilities* industry, *rail transportation* industry, and *pipeline transportation* industry. In these industries, there is a trend of decrease in labour input for the entire sample period. The accumulated positive returns to scale effects was significant.

The returns to scale effect depends on labour income share as well as growth in labour input. The detailed industry study reveals cases when the returns to scale effect induced by labour input growth is mitigated by the small labour income share. The largest average growth rate of labour input is 6.36 percent per year in *other transportation and support activities* industry during the period 1987–1995. However, its returns to scale effect is –1.47, which is not the largest among all the industries. The largest and the second largest returns to scale effect during the period 1987–1995 are found in *information and data processing services* industry and *rental and leasing services and lessors of intangible assets* industry. The average annual rate is −2.43 percent for the former industry and −2.29 percent for the latter industry. The average growth rate of labour input of *other transportation and support activities* industry exceeds that in the above two industries by at least 2 percent. This is because the impact of the rapid increase in labour input of *other transportation and support activities* industry is mitigated by its large labour income share of 77.18 percent.[21]

6. Conclusion

In this paper, we distinguished two effects on LP growth by examining the short-run production frontier. The joint effect of technical progress and capital input growth appears as the growth in LP that is induced by the shift in the short-run production frontier. The returns to scale effect appears as the LP growth induced by the movement along the short-run production frontier. The LP growth calculated by Törnqvist quantity indexes is fully decomposed into the product of these two effects. We applied this decomposition result to U.S. industry data for the period 1987-2007. It is shown that a large part of the difference in LP growth between the goods industries and the services industries can be explained by the returns to scale effect.

In this paper, we assumed the firm’s profit-maximizing behaviour and ruled out inefficient production processes. If we relax the firm’s profit-maximizing behaviour, another factor—*technical efficiency change*—appears in the decomposition of LP growth. Even with no change in the short-run production frontier and no change in labour input, a firm can approach closer to the short-run production frontier by

[21] The impact of the rapid increase in labour input is also mitigated by the large labour income share in *Computer systems design and related services* industry.

improving technical efficiency. For example, a firm improves technical efficiency by increasing output up to the maximum level attainable from given labour inputs under current technology. For the implementation of the decomposition of the LP growth without assuming a firm's profit-maximizing behaviour, we can estimate the distance function by using econometric techniques or Data Envelopment Analysis's linear programming technique. However, we leave this exercise to the future research.

Appendix A

Proof of Proposition 1

$$\ln SHIFT = \left(\frac{1}{2}\right)\ln\left(\frac{D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{D_O^1(\boldsymbol{y}^0,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}\right)+\left(\frac{1}{2}\right)\ln\left(\frac{D_O^0(\boldsymbol{y}^1,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}\right)$$

$$=\left(\frac{1}{2}\right)\ln\left(\frac{D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{D_O^1(\boldsymbol{y}^0,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}\right)+\left(\frac{1}{2}\right)\ln\left(\frac{D_O^0(\boldsymbol{y}^1,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}\right)$$

Since the firm's profit maximization is assumed, the period t production plan is on the period t production frontier for t = 0 and 1.

$$=\left(\frac{1}{4}\right)\sum\nolimits_{m=1}^{M}\left(\frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln y_m}+\frac{\partial \ln D_O^1(\boldsymbol{y}^0,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{\ln y_m}+\frac{\partial \ln D_O^0(\boldsymbol{y}^1,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{\ln y_m}+\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln y_m}\right)\ln\left(\frac{y_m^1}{y_m^0}\right)$$

$$+\left(\frac{1}{4}\right)\sum\nolimits_{m=1}^{M}\left(\frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}+\frac{\partial \ln D_O^1(\boldsymbol{y}^0,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{\ln x_{L,q}}+\frac{\partial \ln D_O^0(\boldsymbol{y}^1,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{\ln x_{L,q}}+\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

using the translog identity in CCD

$$=\left(\frac{1}{2}\right)\sum\nolimits_{m=1}^{M}\left(\frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln y_m}+\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln y_m}\right)\ln\left(\frac{y_m^1}{y_m^0}\right)$$

$$+\left(\frac{1}{2}\right)\sum\nolimits_{q=1}^{Q}\left(\frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}+\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

$$+\left(\frac{1}{4}\right)\sum\nolimits_{m=1}^{M}\left(-\frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln y_m}+\frac{\partial \ln D_O^1(\boldsymbol{y}^0,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{\ln y_m}+\frac{\partial \ln D_O^0(\boldsymbol{y}^1,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{\ln y_m}-\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln y_m}\right)\ln\left(\frac{y_m^1}{y_m^0}\right)$$

$$+\left(\frac{1}{4}\right)\sum\nolimits_{m=1}^{M}\left(-\frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}+\frac{\partial \ln D_O^1(\boldsymbol{y}^0,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{\ln x_{L,q}}+\frac{\partial \ln D_O^0(\boldsymbol{y}^1,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{\ln x_{L,q}}-\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

$$=\left(\frac{1}{2}\right)\sum\nolimits_{m=1}^{M}\left(\frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln y_m}+\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln y_m}\right)\ln\left(\frac{y_m^1}{y_m^0}\right)$$

$$+\left(\frac{1}{2}\right)\sum\nolimits_{q=1}^{Q}\left(\frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}+\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

from the equation (17).

$$=\left(\frac{1}{2}\right)\sum\nolimits_{m=1}^{M}\left(\frac{p_m^0 y_m^0}{\boldsymbol{p}^0\cdot\boldsymbol{y}^0}+\frac{p_m^1 y_m^1}{\boldsymbol{p}^1\cdot\boldsymbol{y}^1}\right)\ln\left(\frac{y_m^1}{y_m^0}\right)-\left(\frac{1}{2}\right)\sum\nolimits_{q=1}^{Q}\left(\frac{w^0 x_{L,q}^0}{\boldsymbol{p}^0\cdot\boldsymbol{y}^0}+\frac{w^1 x_{L,q}^1}{\boldsymbol{p}^1\cdot\boldsymbol{y}^1}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

substituting equations (14) and (16).

Proof of Proposition 2

$$\ln SCALE = \left(\frac{1}{2}\right)\ln\left(\left(\frac{D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}\right)\Bigg/\left(\frac{D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}\right)\right) + \left(\frac{1}{2}\right)\ln\left(\left(\frac{D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}\right)\Bigg/\left(\frac{D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}\right)\right)$$

$$= \left(\frac{1}{2}\right)\ln\left(\left(\frac{D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}\right)\left(\frac{D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}\right)\right) - \left(\frac{1}{2}\right)\ln\left(\left(\frac{D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}\right)\left(\frac{D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}\right)\right)$$

Since the firm's profit maximization is assumed, the period t production plan is on the period t production frontier for t = 0 and 1.

$$= \left(\frac{1}{4}\right)\sum_{m=1}^{M}\left(\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}} + \frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{\ln x_{L,q}} + \frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{\ln x_{L,q}} + \frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^0}{x_{L,q}^1}\right)$$

$$- \left(\frac{1}{4}\right)\sum_{m=1}^{M}\left(\frac{\partial \ln D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{\ln x_{L,q}} + \frac{\partial \ln D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}} + \frac{\partial \ln D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}} + \frac{\partial \ln D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

using the translog identity in CCD

$$= \left(\frac{1}{2}\right)\sum_{m=1}^{M}\left(\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}} + \frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^0}{x_{L,q}^1}\right)$$

$$- \left(\frac{1}{2}\right)\sum_{q=1}^{Q}\left(\frac{\partial \ln D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}} + \frac{\partial \ln D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

$$+ \left(\frac{1}{4}\right)\sum_{m=1}^{M}\left(-\frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}} + \frac{\partial \ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{\ln x_{L,q}} + \frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{\ln x_{L,q}} - \frac{\partial \ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^0}{x_{L,q}^1}\right)$$

$$- \left(\frac{1}{4}\right)\sum_{m=1}^{M}\left(\frac{\partial \ln D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^1)}{\ln x_{L,q}} - \frac{\partial \ln D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}} - \frac{\partial \ln D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}} + \frac{\partial \ln D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^0)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^0}{x_{L,q}^1}\right)$$

$$=\left(\frac{1}{2}\right)\sum_{m=1}^{M}\left(\frac{\partial\ln D_O^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}}+\frac{\partial\ln D_O^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^0}{x_{L,q}^1}\right)$$

$$-\left(\frac{1}{2}\right)\sum_{q=1}^{Q}\left(\frac{\partial\ln D_L^0(\boldsymbol{y}^0,\boldsymbol{x}_K^0,\boldsymbol{x}_L^0)}{\ln x_{L,q}}+\frac{\partial\ln D_L^1(\boldsymbol{y}^1,\boldsymbol{x}_K^1,\boldsymbol{x}_L^1)}{\ln x_{L,q}}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

from the equation (17) and (39).

$$=\left(\frac{1}{2}\right)\sum_{q=1}^{Q}\left(\frac{w^0x_{L,q}^0}{\boldsymbol{p}^0\cdot\boldsymbol{y}^0}+\frac{w^1x_{L,q}^1}{\boldsymbol{p}^1\cdot\boldsymbol{y}^1}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)-\left(\frac{1}{2}\right)\sum_{q=1}^{Q}\left(\frac{w^0x_{L,q}^0}{\boldsymbol{w}^0\cdot\boldsymbol{x}_L^0}+\frac{w^1x_{L,q}^1}{\boldsymbol{w}^1\cdot\boldsymbol{x}_L^1}\right)\ln\left(\frac{x_{L,q}^1}{x_{L,q}^0}\right)$$

substituting equations (16) and equations (38).

References

Balk, B.M. (2001), "Scale Efficiency and Productivity Change", *Journal of Productivity Analysis* 15, 159-183.

Bosworth, B.P. and J.E. Triplett (2007), "The Early 21st Century U.S. Productivity Expansion is Still in Services", *International Productivity Monitor* 14, 3-19.

Caves, D.W., L. Christensen and W.E. Diewert (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity", *Econometrica* 50, 1393-1414.

Chambers, R.G. (1988), *Applied Production Analysis: A Dual Approach*, New York: Cambridge University Press.

Council of Economic Advisors (2010), *Economic Report of the President*, Washington, D.C.: U.S. Government Printing Office.

Diewert, W.E. and C.J. Morrison (1986), "Adjusting Output and Productivity Indexes for Changes in the Terms of Trade", *Economic Journal* 96, 659-679.

Diewert, W.E. and A.O. Nakamura (2007), "The Measurement of Productivity for Nations", in J.J. Heckman and E.E. Leamer (ed.), *Handbook of Econometrics*, Vol. 6, Chapter 66, Amsterdam: Elsevier, pp. 4501-4586.

Fare, R., S. Grosskopf, M. Norris and Z. Zhang (1994), "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries", *American Economic Review* 84, 66-83.

Fisher, I. (1922), *The Making of Index Numbers*, Boston: Houghton-Mifflin.

Griliches, Z. (1987), "Productivity: Measurement Problems", in J. Eatwell, M. Milgate and P. Newman (ed.), *The New Palgrave: A Dictionary of Economics.* New York: McMillan, pp. 1010-1013.

Hall, R.E., and C.I. Jones (1999). "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics* 114, 83-116.

Hotelling, H. (1932), "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions", *Journal of Political Economy* 40, 577-616.

Jones, C.I. (2002), "Sources of U.S. Economic Growth in a World of Idea", *American Economic Review* 92, 220-239.

Jorgenson, D.W. and K.J. Stiroh (2000), "Raising the Speed Limit: U.S. Economic Growth in the Information Age", *Brookings Papers on Economic Activity* 2, 125-211.

Lovell, C. A. K. (2003), "The Decomposition of Malmquist Productivity Indexes", *Journal of Productivity Analysis* 20, 437-458.

Nemoto, J., and M. Goto (2005), "Productivity, Efficiency, Scale Economies and Technical Change: A New Decomposition Analysis of TFP Applied to the Japanese Prefectures", *Journal of Japanese and International Economies* 19, 617-634.

Nishimizu, M., and J.M. Page (1982), "Total Factor Productivity Growth, Technical Progress and Technical Efficiency Change: Dimensions of Productivity Change in Yugoslavia, 1965-78", *Economic Journal* 92, 920-936.

Triplett, J.E. and B.P. Bosworth (2004), *Services Productivity in the United States: New Sources of Economic Growth*, Washington, D.C.: Brookings Institution Press.

Triplett, J.E. and B.P. Bosworth (2006), "'Baumol's Disease' Has Been Cured: IT and Multi-factor Productivity in U.S. Services Industries", *The New Economy and Beyond: Past, Present, and Future*. Dennis W. Jansen, (eds.), Cheltenham: Edgar Elgar, pp. 34-71.

Figure 1: Sources of Sectoral Labour Productivity Growth

	1987-2007	*1987-1995*	*1995-2007*
Non-farm private business sector			
Labour productivity growth	2.17	1.39	2.69
Technical progress and capital input growth	2.53	1.82	3.00
Returns to scale effect	-0.36	-0.43	-0.31
Goods industries			
Labour productivity growth	2.76	1.82	3.38
Technical progress and capital input growth	2.71	1.88	3.26
Returns to scale effect	0.05	-0.06	0.12
Services industries			
Labour productivity growth	1.95	1.22	2.43
Technical progress and capital input growth	2.47	1.83	2.89
Returns to scale effect	-0.52	-0.61	-0.46

Note : All figures are average annual percentages.

Figure 2: Labour Input Growth and Labour Income Share

	1987-2007	*1987-1995*	*1995-2007*
Labour input growth (un-weighted average)			
Private non-farm business	0.62	1.22	0.22
Goods industries	-1.06	-0.26	-1.59
Services industries	1.63	2.09	1.33
Labour input growth (weighted average)			
Private non-farm business	1.13	1.41	0.95
Goods industries	-0.20	0.21	-0.48
Services industries	1.70	1.99	1.50
Labour income share			
Private non-farm business	69.18	69.57	68.92
Goods industries	67.44	69.29	66.21
Services industries	69.88	69.69	70.01

Note : All figures are average annual percentages. The un-weighted average growth rate of labour productivity is the arithmetic mean of the growth rate of industry labour productivity. The weighted average growth rate of labour productivity is calculated using labour income in each industry divided by the sum of industry labor

Figure 3: Sources of Industry Labour Productivity Growth, 1987–1995

Industry	*Labour productivity growth*	*Technical progress and capital input growth*	*Returns to scale effect*	*Labour input growth*	*Labour income share*
	Goods industries				
Oil and gas extraction	4.08	2.10	1.98	-2.93	30.52
Mining, except oil and gas	4.75	4.11	0.64	-1.81	63.15
Support activities for mining	1.86	1.27	0.60	-1.97	68.22
Construction	-0.27	-0.11	-0.16	0.85	83.30
Food and Beverage and Tobacco Products	1.88	2.19	-0.32	0.72	53.61
Textile Mills and Textile Product Mills	2.90	2.71	0.19	-0.94	76.23
Apparel and Leather and Applied Products	4.68	3.92	0.76	-2.67	71.66
Paper Products	0.42	0.51	-0.10	0.26	60.61
Printing and Related Support Activities	0.46	0.76	-0.30	1.41	78.73
Petroleum and Coal Products	3.68	2.68	1.00	-1.64	37.32
Chemical Products	-0.39	-0.15	-0.24	0.53	48.20
Plastics and Rubber Products	1.94	2.65	-0.71	2.02	65.77
Wood Products	-1.53	-1.40	-0.12	0.34	68.46
Nonmetallic Mineral Products	1.27	1.37	-0.10	0.12	68.93
Primary Metal Products	1.29	1.18	0.11	-0.35	73.63
Fabricated Metal Products	1.47	1.71	-0.24	0.79	72.55
Machinery	-0.82	-0.52	-0.29	1.33	75.22
Computer and Electronic Products	19.63	19.33	0.29	-1.40	77.36
Electrical Equipment, Appliances, and Components	-3.80	-4.22	0.43	-1.09	60.59
Transportation Equipment	-1.70	-1.73	0.03	-0.49	82.27
Furniture and Related Products	0.64	0.65	-0.01	-0.08	79.99
Miscellaneous Manufacturing	2.41	2.80	-0.39	1.38	70.08
	Services industries				
Utilities	4.07	3.78	0.29	-0.39	24.97
Wholesale trade	2.32	2.66	-0.34	1.15	72.49
Retail trade	2.88	3.22	-0.34	1.41	76.28
Air transportation	1.50	2.79	-1.29	5.12	76.57
Rail transportation	5.43	4.63	0.80	-2.50	65.42
Water transportation	9.15	8.76	0.39	-0.80	51.35
Truck transportation	4.23	4.52	-0.28	1.17	74.84
Transit and ground passenger transportation	-2.32	-1.53	-0.79	3.24	76.08
Pipeline transportation	3.45	2.80	0.64	-1.27	49.26
Other transportation and support activities	-3.13	-1.65	-1.47	6.36	77.18
Warehousing and storage	5.16	5.35	-0.20	0.88	79.46
Publishing industries (includes software)	3.00	3.54	-0.55	1.70	67.89
Motion picture and sound recording industries	-2.97	-1.65	-1.33	5.22	74.50
Broadcasting and telecommunications	5.38	5.54	-0.17	0.24	41.89
Information and data processing services	0.22	2.65	-2.43	4.36	44.31
Federal Reserve banks, credit intermediation, and related activities	0.54	0.26	0.28	-0.69	58.39
Securities, commodity contracts, and investments	6.66	6.95	-0.29	1.70	79.04
Insurance carriers and related activities	2.43	2.63	-0.20	1.21	82.77
Funds, trusts, and other financial vehicles	-0.87	0.47	-1.35	1.87	26.56
Real estate	-0.06	0.73	-0.79	0.96	18.58
Rental and leasing services and lessors of intangible assets	1.61	3.90	-2.29	3.18	27.81
Legal services	-0.27	-0.18	-0.09	1.46	93.50
Computer systems design and related services	2.26	2.92	-0.66	5.46	87.60
Miscellaneous professional, scientific, and technical services	0.40	0.63	-0.23	2.50	90.73
Management of companies and enterprises	1.35	1.37	-0.01	0.08	84.11
Administrative and support services	0.41	1.04	-0.64	4.75	86.53
Waste management and remediation services	1.07	1.91	-0.83	2.02	59.54
Educational services	0.05	0.73	-0.69	2.58	71.51
Ambulatory health care services	-1.75	-1.01	-0.74	3.69	79.75
Hospitals and nursing and residential care facilities	-3.12	-2.56	-0.57	3.33	82.61
Social assistance	3.01	3.14	-0.13	2.11	93.74
Performing arts, spectator sports, museums, and related activities	2.08	2.33	-0.25	1.57	84.47
Amusements, gambling, and recreation industries	-0.89	0.72	-1.61	5.49	70.19
Accommodation	1.81	2.40	-0.59	1.87	69.55
Food services and drinking places	-1.20	-0.84	-0.35	1.94	82.12
Other services, except government	-0.73	-0.25	-0.48	2.25	78.75

Note: All figures are average annual percentages.

Figure 4: Sources of Industry Labour Productivity Growth, 1995–2007

Industry	*Labour productivity growth*	*Technical progress and capital input growth*	*Returns to scale effect*	*Labour input growth*	*Labour income share*
Goods-producing industries					
Oil and gas extraction	-2.91	-3.11	0.20	-0.62	21.03
Mining, except oil and gas	1.94	1.83	0.10	-0.54	52.91
Support activities for mining	-4.75	-2.02	-2.73	4.86	57.17
Construction	-2.89	-2.48	-0.41	2.45	84.53
Food and Beverage and Tobacco Products	0.83	0.78	0.05	-0.11	51.22
Textile Mills and Textile Product Mills	5.16	3.60	1.56	-5.95	72.96
Apparel and Leather and Applied Products	-0.20	-2.54	2.34	-9.55	75.38
Paper Products	3.81	2.60	1.20	-2.97	58.62
Printing and Related Support Activities	2.39	2.06	0.33	-2.53	85.17
Petroleum and Coal Products	5.77	4.80	0.97	-1.37	21.92
Chemical Products	6.69	5.87	0.82	-1.52	45.08
Plastics and Rubber Products	3.68	3.20	0.47	-1.34	62.25
Wood Products	2.31	2.12	0.19	-1.43	77.91
Nonmetallic Mineral Products	1.39	1.36	0.03	-0.18	63.23
Primary Metal Products	1.77	1.00	0.77	-2.88	67.13
Fabricated Metal Products	1.40	1.36	0.04	-0.28	69.41
Machinery	2.41	1.95	0.46	-2.00	73.55
Computer and Electronic Products	27.71	27.56	0.15	-2.79	79.11
Electrical Equipment, Appliances, and Components	1.89	0.91	0.98	-2.90	65.78
Transportation Equipment	5.90	5.44	0.47	-1.56	72.52
Furniture and Related Products	2.52	2.26	0.26	-1.17	76.67
Miscellaneous Manufacturing	5.20	4.97	0.23	-0.69	66.62
Services -producing industries					
Utilities	2.31	1.00	1.31	-1.76	25.20
Wholesale trade	3.81	4.00	-0.19	0.61	71.05
Retail trade	5.22	5.32	-0.10	0.41	76.42
Air transportation	10.25	9.80	0.45	-2.33	73.46
Rail transportation	2.59	2.06	0.53	-1.54	59.15
Water transportation	-1.15	0.40	-1.56	2.93	47.98
Truck transportation	1.15	1.42	-0.27	1.05	75.74
Transit and ground passenger transportation	1.02	1.26	-0.24	1.12	75.72
Pipeline transportation	5.68	4.87	0.81	-1.69	47.69
Other transportation and support activities	2.07	2.25	-0.18	0.66	78.02
Warehousing and storage	2.37	3.12	-0.75	3.13	76.67
Publishing industries (includes software)	3.88	3.80	0.08	-0.27	68.25
Motion picture and sound recording industries	0.53	0.85	-0.32	1.35	74.14
Broadcasting and telecommunications	6.67	6.68	-0.01	-0.10	44.64
Information and data processing services	6.42	7.51	-1.10	2.18	66.10
Federal Reserve banks, credit intermediation, and related activities	1.03	1.71	-0.69	1.56	55.56
Securities, commodity contracts, and investments	12.64	13.01	-0.37	2.89	87.06
Insurance carriers and related activities	0.21	0.36	-0.14	0.63	78.22
Funds, trusts, and other financial vehicles	-0.04	2.06	-2.11	2.62	18.77
Real estate	0.46	1.45	-0.99	1.29	24.12
Rental and leasing services and lessors of intangible assets	1.98	2.79	-0.80	1.08	25.45
Legal services	-0.29	-0.23	-0.06	0.80	92.26
Computer systems design and related services	2.91	3.59	-0.67	5.93	87.07
Miscellaneous professional, scientific, and technical services	3.40	3.72	-0.33	2.37	85.85
Management of companies and enterprises	-0.47	-0.27	-0.20	0.95	79.73
Administrative and support services	0.37	0.69	-0.32	2.75	88.10
Waste management and remediation services	-0.34	0.52	-0.86	2.23	61.92
Educational services	0.27	1.24	-0.97	3.57	72.01
Ambulatory health care services	0.35	1.09	-0.73	3.39	78.64
Hospitals and nursing and residential care facilities	-1.25	-0.80	-0.44	2.89	84.83
Social assistance	3.89	3.97	-0.08	1.38	93.38
Performing arts, spectator sports, museums, and related activities	1.61	1.77	-0.16	0.95	85.44
Amusements, gambling, and recreation industries	1.96	2.37	-0.41	1.32	67.78
Accommodation	-0.46	-0.11	-0.35	0.86	62.28
Food services and drinking places	0.86	1.20	-0.34	1.74	79.96
Other services, except government	-0.27	-0.12	-0.15	0.77	82.06

Note: All figures are average annual percentages.

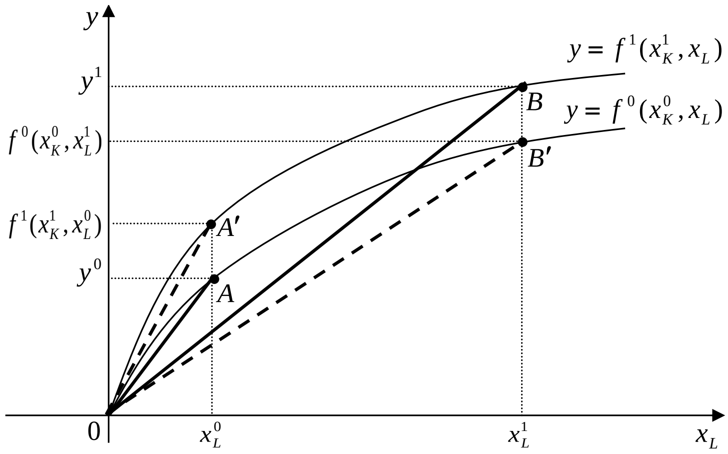

Figure 1: Labour Productivity Growth and Shift in the Short-run Production Frontier

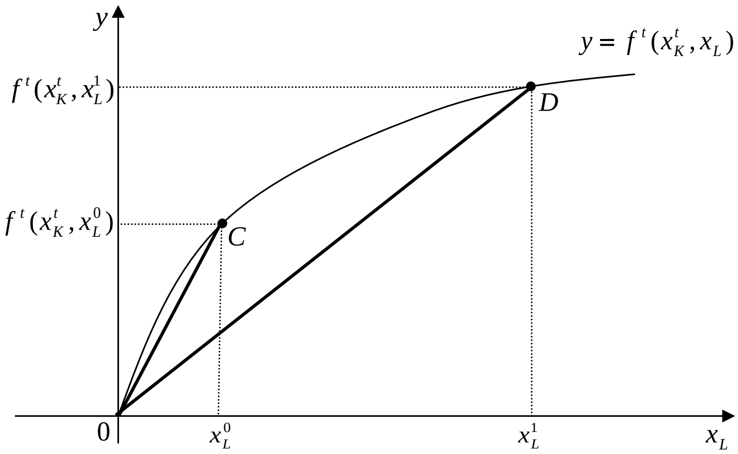

Figure 2: Returns to Scale Effect and Movement along the Short-run Production Frontier