ISSN 1881-6436

Discussion Paper Series

No. 11-03

Returns to scale effect in labour productivity growth

Hideyuki Mizobuchi

March 2012

Faculty of Economics,
Ryukoku University

67 Tsukamoto-cho, Fukakusa, Fushimi-ku,
Kyoto, Japan
612-8577

Returns to scale effect in labour productivity growth

Hideyuki Mizobuchi

March 2012

Abstract

Labour productivity is defined as output per unit of labour input. Economists acknowledge that technical progress and growth in capital inputs increase labour productivity. However, less focus is given to the fact that changes in labour input alone could also affect labour productivity. Because this effect disappears for the constant returns to scale short-run production frontier, we call it the returns to scale effect. We decompose 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 labour inputs and outputs. We then apply the results of our decomposition to US industry data for 1987–2009. Labour productivity in the services sector is acknowledged to grow much more slowly than in the goods sector during the productivity slowdown period. 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.

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

JEL classification: C14, D24, O47, O51

Acknowledgement: A part of this paper was circulated under the title, "New Indexes of Labour Productivity Growth: Baumol's Disease Revisited". We are grateful to Bert Balk for his helpful comments and suggestions. We also wish to thank Erwin Diewert, Jiro Nemoto, Takanobu Nakajima, Toshiyuki Matsuura, Mitsuru Sunada and seminar participants in the 12th European Workshop on Efficiency and Productivity Analysis, June 2011, Verona, Italy and the annual meeting of the Japanese Economic Association, Kumamoto Gakuen University, May, 2011, Tohoku University, Keio University, Chuo University, Tokyo University, Kyoto University. All remaining errors are the author's responsibility.

Corresponding author: Department of Economics, Ryukoku University, 67 Fukakusa Tsukamoto-cho, Fushimi-ku, Kyoto 612-8577, Japan. +81-75-645-8439, 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, characterises the prevailing state of the technology. Productivity growth is often identified by a shift in the production frontier, reflecting changes in production technology.[1] However, movement along the production frontier also derives productivity growth.[2]

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 induced by movement along the production frontier is called the *returns to scale effect*. This effect does not reflect changes in the production frontier. Thus, to properly evaluate improvements in the underlying production technology reflecting a shift in the production frontier, we must disentangle the returns to scale effect from overall productivity growth.

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, US history reveals that increases in LP have translated to nearly one-for-one increases in per capita income over a long period.[3] The importance of LP as a source for the progress of economic well-being prompts many researchers to investigate the determinants of LP growth. Technical progress and capital input growth have been emphasised 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).[4] The present paper adds one more explanatory factor to LP growth.

---

[1] See Griliches (1987). Moreover, the same interpretation is found in Chambers (1988).

[2] In principle, productivity improvement can also occur through 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-maximising behaviour, and in our model, the current production plan is always on the current production frontier. The assumption of profit maximisation 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] In addition, these authors found that improvements in the quality of labour inputs (in other words, human capital accumulation) play an important role for explaining changes in LP. If we adopt the method of these authors; that is the number of workers or the number of hours worked are adopted as the measure of labour input, changes in the quality of labour input raise the amount of output attainable from a given number of workers or a given hours worked, leading to an outward shift in the short-run production frontier. However, because we differentiate qualities of different labour inputs, allowing wages to vary among them, our measure of changes in the total labour input reflects changes in labour qualities among varieties of labour inputs. Thus, improvements in labour qualities do not affect the short-run production frontier itself and we can ignore the role of the labour quality growth for explaining LP growth throughout this paper. See Footnote 5 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 into 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 and 2) the returns to scale effect.[5] First, we propose theoretical measures representing the two effects by using the short-run 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 short-run distance functions and the firm's profit-maximising behaviour.

Our approach to implementing theoretical measures is drawn from Caves, Christensen and Diewert (1982) (CCD). Using the distance functions, CCD formulated the (theoretical) Malmquist productivity index that measures the shift in the production frontier, and show that the Malmquist productivity and the Törnqvist productivity indexes coincide, assuming the translog functional form for the distance functions and the firm's profit-maximising behaviour.[6]

The Törnqvist productivity index is a measure of 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 showed 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 decomposed the TFP growth calculated by the Törnqvist quantity indexes into the 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 and the Törnqvist productivity indexes, captures TFP growth induced by the movement along the production frontier exploiting its curvature.

CCD's formula for the returns to scale effect appeared as the residual of two indexes and CCD did not explicitly model the returns to scale effect using the underlying production frontier.[8] On the other hand, other studies model the

[5] In case when our measure of labour inputs fails to capture the improvement in labour quality, the unmeasured improvement in labour quality shifts the short-run production frontier. Thus, its effect on LP growth is captured by the joint effect of technical progress and capital input growth.

[6] Since CCD are concerned with measurement of TFP, they deal with the underlying production frontier that consists of total inputs (capital and labour inputs) and the maximum output attainable from them, indicating the capacity of current technology to translate total inputs into outputs. From this point forward, 'the underlying production frontier' or simply 'the production frontier' means this type of the underlying production frontier, in distinction from the short-run production frontier.

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

[8] In the present paper, we show that our index number formula for the returns to scale effect coincides with the growth in LP induced by movement along the short-run production frontier. As

growth in TFP induced by the movement along the underlying production frontier but adopt different approaches to estimating the modelled returns to scale effect rather than relying on index number formulae. Lovell (2003) modelled the returns to scale effect by using input and output distance functions and calls it the *scale effect* or *activity effect*. In Balk's (2001) decomposition of TFP growth, the product of *scale efficiency change* and *input mix effect* or that of *scale efficiency change* and *output mix effect* summarised the TFP growth induced by movement along the production frontier, and it can be interpreted as the returns to scale effect.[9]

Although scholars have recognised the significance of the returns to scale effect for TFP growth, its effect on LP growth has never been addressed even though it plays a more important role 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 in LP growth 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 services sector of the US economy has been much slower than in the goods sector since the early 1970s. As we discussed above, there are two underlying factors to LP growth. Thus, different explanations are possible for its stagnated LP growth depending on the factor emphasised. We apply our decomposition result to US industry data to compare the relative contributions of the two effects.

Section 2 graphically illustrates the two effects underlying LP growth. 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 includes the main result. It 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 US industry data. Section 6 presents the conclusions.

2. Two Sources of Labour Productivity Growth

We graphically display the drivers of 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^1 and y^0 using inputs (x_K^0, x_L^0) and (x_K^1, x_L^1). The period t production technology is described by the period t production frontier $y = f^t(x_K, x_L)$ 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 in which the joint effect of technical progress and capital input growth

---

will be emphasised further on, our result implies that CCD's formula for the returns to scale effect coincides with the TFP growth induced by the movement along the underlying production frontier.

[9] 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 identified the combined effect of changes in the composition of inputs and of outputs.

positively affects the productive capacity of labour. The lower curve represents the period 0 short-run production frontier and indicates how much output can be produced using a specified quantity of labour given capital input and technology available in period 0. Similarly, the higher curve represents the period 1 short-run production frontier and indicates how much output can be produced using a specified quantity of labour given capital input 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 $f^1(x_K^1, x_L) > f^0(x_K^0, x_L)$ for all x_L. Moreover, the corresponding LP 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. Additionally, note that the ratio is 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. That LP can decline despite the outward shift in the short-run production frontier suggests that another factor contributes to LP growth.[10] 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 $f^1(x_K^1, x_L^0)/x_L^0$. 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. However, 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.

However, the division of the path from A to B into two steps from A to A' and from A' to B is only an example. Decomposing the path from A to B into the movement along the period 0 short-run production frontier from A to B' and the vertical jump from B' to B is also possible. In this case, the former

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

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 is significant. Hereafter, we generalise our discussion to the more general multiple-inputs multiple-outputs case and propose measures for the two effects that are immune from the selection of 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, $\boldsymbol{y} \equiv (y_1, \dots, y_M)$ and $P + Q$ inputs consisting of P types of capital inputs, $\boldsymbol{x}_K \equiv \left(x_{K,1}, \dots, x_{K,P}\right)$; and Q types of labour inputs, $\boldsymbol{x}_L \equiv \left(x_{L,1}, \dots, x_{L,Q}\right)$. Outputs include intermediate inputs. If output m is an intermediate input, then $y_m < 0$. If output is not an intermediate input but a (gross) output, $y_m > 0$. We also assume that any outputs and labour inputs are non-zero such that $y_m \neq 0$ for all m and $x_{L,q} \neq 0$ for all q.[11] The period t production possibility set S^t consists of all feasible combinations of inputs and outputs, and it is defined as

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

We assume S^t satisfies convexity and Färe and Primont's (1995) axioms that guarantee the existence of distance functions.[12] The period t production frontier, which is the boundary of S^t, is represented by the *period t input requirement function F^t* and it is defined as follows:

$$F^t(\boldsymbol{y}, \boldsymbol{x}_K . \boldsymbol{x}_L) \equiv \min\left\{x_{L,1} : \left(\boldsymbol{y}, \boldsymbol{x}_K, x_{L,1}, \boldsymbol{x}_{L-1}\right) \in S^t\right\}. \quad (2)$$

F^t represents the minimum amount of the first labour input that a firm can use at period t, producing output quantities $\boldsymbol{y}$ and holding capital inputs $\boldsymbol{x}_K$ and other labour inputs $\boldsymbol{x}_{L,-1} \equiv \left(x_{L,2}, \dots, x_{L,Q}\right)$ fixed. This function, originally formulated for characterising the period t production frontier, can also be used for characterising 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 $x_{L,1}^t = F^t\left(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}\right)$ forms the period t short-run production frontier.

[11] Excluding zero quantities for outputs and labour inputs is a crucial condition for deriving index number formulae by aggregating the growth rates of input and output. Since capital input quantities do not appear in the index number formula of the present paper, we do not impose the condition of non-zero quantities for the capital inputs. There is an alternative approach in index number theory called the 'difference approach'. As emphasised by Diewert and Mizobuchi (2009), the difference approach can apply to the situation that there are certain inputs or outputs whose quantities are zero.

[12] Originally, Färe and Primont's (1995) axioms are for input and output distance function. Moreover, they guarantee the existence of the labour input distance function that we introduce in this paper.

We assume that $F^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_{L,-1})$ is differentiable at $(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^{\mathrm{t}})$ with respect to $\boldsymbol{y}$ and $\boldsymbol{x}_L$ for $t = 0$ and 1 and satisfies the following conditions.[13] These conditions are necessary for discussing the relationship between the input requirement function and the distance functions, which we introduce later.

$$\boldsymbol{y}^t \cdot \nabla_{\boldsymbol{y}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t) \neq 0, \tag{3}$$

$$x_{L,1}^t - \boldsymbol{x}_{L,-1}^t \cdot \nabla_{\boldsymbol{x}_{L,-1}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t) \neq 0. \tag{4}$$

CCD measured the shift in the production frontier 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:

$$D_O^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) \equiv \min_\delta \left\{\delta : F^t\left(\frac{\boldsymbol{y}}{\delta}, \boldsymbol{x}_K, \boldsymbol{x}_{L,-1}\right) \leq x_{L,1}\right\}. \tag{5}$$

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 $\boldsymbol{y}$ enabling the contracted outputs $\boldsymbol{y}/D_O^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$, capital inputs $\boldsymbol{x}_K$ and labour inputs $\boldsymbol{x}_L$ to fall 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 $D_O^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is linearly homogeneous in $\boldsymbol{y}$.

Furthermore, we can relate the *period t output distance function* to the *period t* 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 $\boldsymbol{y}$ causing the contracted outputs $\boldsymbol{y}/D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L)$ and labour inputs $\boldsymbol{x}_L$ to fall on the period t short-run production frontier. Thus, $D_O^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ 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 from $\boldsymbol{y}$ to the short-run production frontiers of the periods 0 and 1, which is defined as follows:[14]

$$SHIFT(\boldsymbol{y}, \boldsymbol{x}_L) \equiv \frac{D_O^0(\boldsymbol{y}, \boldsymbol{x}_K^0, \boldsymbol{x}_L)}{D_O^1(\boldsymbol{y}\ \boldsymbol{x}_K^1, \boldsymbol{x}_L)}. \tag{6}$$

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 $D_O^1(\boldsymbol{y}, \boldsymbol{x}_K^1, \boldsymbol{x}_L) \leq D_O^0(\boldsymbol{y}, \boldsymbol{x}_K^0, \boldsymbol{x}_L)$, leading to $SHIFT(\boldsymbol{y}, \boldsymbol{x}_L) \geq 1$. Similarly, the negative joint effect of technical progress and capital input growth leads to $SHIFT(\boldsymbol{y}, \boldsymbol{x}_L) \leq 1$.

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 $(\boldsymbol{y}^1, \boldsymbol{x}_L^1)$. Because these reference outputs and labour inputs are, in fact, chosen in each period, they

[13] Notation: $\nabla_x f(\boldsymbol{x}, \boldsymbol{y}) \equiv [\partial f(\boldsymbol{x}, \boldsymbol{y})/\partial x_1, \ldots, \partial f(\boldsymbol{x}, \boldsymbol{y})/\partial x_N]^\top$ is a column vector of the partial derivative of f with respect to the vector $\boldsymbol{x} \in \Re^N$, and $\boldsymbol{x} \cdot \boldsymbol{z} = \sum_{n=1}^N x_n z_n$.

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

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:[15]

$$SHIFT \equiv \sqrt{SHIFT(\boldsymbol{y}^0, \boldsymbol{x}_L^0) \cdot SHIFT(\boldsymbol{y}^1, \boldsymbol{x}_L^1)}. \tag{7}$$

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:

$$SHIFT \; = \; \sqrt{(f^1(x_K^1, x_L^0)/y^0) \cdot (y^1/f^1(x_K^1, x_L^0)\,)}\,. \tag{8}$$

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 frontier 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 several methods 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 short-run profit-maximising behaviour and a translog functional form for the output distance function.[16] Our approach is drawn from CCD, which deal with the Malmquist productivity index, a theoretical measure of the shift in the production frontier.

CCD showed that the first-order derivatives of the output distance function 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 under the assumption of a firm's profit-maximising behaviour and a translog functional form for the output distance function. Then, using these relationships, CCD showed that the Malmquist productivity index coincides with a different index number formula of price and quantity observations, the Törnqvist productivity index.[17] The following equations (14) and (15), already derived by CCD, allow us to compute the first-order derivatives of the output distance function from price and quantity observations. Moreover, they can be derived under our assumption in the same way as CCD. For completeness of discussion, we outline below how to obtain these equations.

The implicit function theorem is applied to the input requirement function $F^t(\boldsymbol{y}/\delta\,, \boldsymbol{x}_K, \boldsymbol{x}_{L,-1}) \; = x_{L,1}$ 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] Since the firm's short-run profit maximisation is assumed, it is possible to adopt a different formulation for the measure of the shift in the short-run production frontier:

$$SHIFT = \sqrt{\left(\frac{D_O^0(y^1,x_K^0,x_L^1)}{D_O^0(y^0,x_K^0,x_L^0)}\right)\left(\frac{D_O^1(y^1,x_K^1,x_L^1)}{D_O^1(y^0,x_K^1,x_L^0)}\right)}.$$

This above formulation is closer to the Malmquist productivity index introduced by CCD.

[16] 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 approach, and see Färe et al. (1994) for the application of the linear programming approach.

[17] CCD justified the use of the Törnqvist productivity index, which is the Törnqvist output quantity index divided by the Törnqvist input quantity index.

Its derivatives are represented by the derivatives of $F^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_{L,-1})$. We have the following equations for $t = 0$ and 1:[18]

$$\nabla_{\boldsymbol{y}} D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{\boldsymbol{y}^t \cdot \nabla_{\boldsymbol{y}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t)} \nabla_{\boldsymbol{y}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t), \tag{9}$$

$$\nabla_{\boldsymbol{x}_L} D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{\boldsymbol{y}^t \cdot \nabla_{\boldsymbol{y}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L-1}^t)} \begin{bmatrix} -1 \\ \nabla_{\boldsymbol{x}_{L,-1}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t) \end{bmatrix}. \tag{10}$$

We assume the firm's short-run profit-maximising behaviour. Thus, $(\boldsymbol{y}^t, \boldsymbol{x}_L^t)$ is a solution to the following period t short-run profit- maximisation problem for $t = 0$ and 1:[19]

$$\max_{\boldsymbol{y}, \boldsymbol{x}_{L,-1}} \{\boldsymbol{p}^t \cdot \boldsymbol{y} - w_1^t F^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_{L,-1}) - \boldsymbol{w}_{-1}^t \cdot \boldsymbol{x}_{L,-1}\}. \tag{11}$$

Outputs are sold at the positive producer prices $\boldsymbol{p} \equiv (p_1, \dots, p_M) \in \Re_{++}^M$, capital inputs are purchased at the positive rental prices $\boldsymbol{r} \equiv (r_1, \dots, r_P) \in \Re_{++}^P$ and labour inputs are purchased at the positive wages $\boldsymbol{w} \equiv (w_1, \dots, w_Q) \in \Re_{++}^Q$. Note that $\boldsymbol{w}_{-1} \equiv (w_2, \dots, w_Q)$. The period t short-run profit-maximisation problem yields the following first-order conditions for $t = 0$ and 1:

$$\boldsymbol{p}^t = w_1^t \nabla_{\boldsymbol{y}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t), \tag{12}$$

$$\boldsymbol{w}_{-1}^t = -w_1^t \nabla_{\boldsymbol{x}_{L,-1}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t). \tag{13}$$

By substituting equations (12) and (13) into equations (9) and (10), we obtain the following equations (14) and (15) for $t = 0$ and 1:[20]

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

$$\begin{aligned} \nabla_{\boldsymbol{x}_L} D_O^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) &= [w_1^t / \boldsymbol{p}^t \cdot \boldsymbol{y}^t] \begin{bmatrix} -1 \\ \nabla_{\boldsymbol{x}_{L,-1}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t) \end{bmatrix} \\ &= [1 / \boldsymbol{p}^t \cdot \boldsymbol{y}^t] \begin{bmatrix} -w_1^t \\ -\boldsymbol{w}_{-1}^t \end{bmatrix}. \end{aligned} \tag{15}$$

Equations (14) and (15) 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) and (15) for $t = 0$ and 1. The output distance functions evaluated at the production plan in different period such as $D_O^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $D_O^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ also constitute $SHIFT$. Hence the above equations are insufficient for implementing $SHIFT$. In addition to a firm's short-run profit maximisation, we further assume a translog functional form with time-invariant

---

[18] Equation (3) implies that equations (9) and (10) are well defined.

[19] We assume a firm's short-run profit-maximising behaviour, unlike CCD's assumption that even capital inputs are optimally chosen. Thus, under our assumption, we cannot compute the first derivatives of the output distance function with respect to capital inputs from price and quantity observations. However, it is unnecessary to use the capital input counterpart to equations (14) and (15).

[20] Equations (3) and (4) implies $\boldsymbol{p} \cdot \boldsymbol{y} \neq 0$, meaning that equations (14) and (15) are well defined.

second-order coefficients for the period t output distance function for $t = 0$ and 1, which is defined as following:

$$\begin{aligned}
&\ln D_O^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) \equiv \\
&\alpha_0^t + \sum_{m=1}^{M} \alpha_m^t \ln y_m + \left(\frac{1}{2}\right) \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_{i,j} \ln y_i \ln y_j + \\
&\sum_{p=1}^{P} \beta_p^t \ln x_{K,q} + \left(\frac{1}{2}\right) \sum_{i=1}^{P} \sum_{j=1}^{P} \beta_{i,j} \ln x_{K,i} \ln x_{K,j} + \\
&\sum_{q=1}^{Q} \chi_q^t \ln x_{L,q} + \left(\frac{1}{2}\right) \sum_{i=1}^{Q} \sum_{j=1}^{Q} \chi_{i,j} \ln x_{L,i} \ln x_{L,j} + \\
&\sum_{m=1}^{M} \sum_{p=1}^{P} \delta_{m,p} \ln y_m \ln x_{K,p} + \sum_{m=1}^{M} \sum_{q=1}^{Q} \varepsilon_{m,q} \ln y_m \ln x_{L,q} + \\
&\sum_{p=1}^{P} \sum_{q=1}^{Q} \varphi_{p,q} \ln x_{K,p} \ln x_{L,q}
\end{aligned} \tag{16}$$

where the parameters satisfy the following restrictions:

$$\alpha_{i,j} = \alpha_{j,i} \text{ for all } i \text{ and } j \text{ such as } 1 \leq i < j \leq M; \tag{17}$$

$$\beta_{i,j} = \beta_{j,i} \text{ for all } i \text{ and } j \text{ such as } 1 \leq i < j \leq P; \tag{18}$$

$$\chi_{i,j} = \chi_{j,i} \text{ for all } i \text{ and } j \text{ such as } 1 \leq i < j \leq Q; \tag{19}$$

$$\sum_{n=1}^{N} \alpha_n^t = 1; \tag{20}$$

$$\sum_{i=1}^{M} \alpha_{i,m} = 0 \text{ for } m = 1, \dots, M; \tag{21}$$

$$\sum_{m=1}^{M} \delta_{m,p} = 0 \text{ for } p = 1, \dots, P; \text{ and} \tag{22}$$

$$\sum_{m=1}^{M} \varepsilon_{m,q} = 0 \text{ for } q = 1, \dots, Q. \tag{23}$$

Restrictions (20)–(23) guarantee linear homogeneity in $\boldsymbol{y}$. The translog functional form characterised in (16)–(23) is a flexible functional form, enabling it to 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 various types of technical progress are allowed.

Under the assumptions of the short-run profit-maximising behaviour and the translog functional form, a theoretical measure $SHIFT$ coincides with a formula of price and quantity observations, as is shown in the following proposition. The proof that CCD showed the equivalence between the Malmquist and the Törnqvist productivity indices using equations (14) and (15) and the capital input counterpart still goes through for $SHIFT$. Thus, we can consider the following proposition as a corollary of CCD.[21]

Proposition 1 (Christensen, Caves and Diewert, 1982)
Assume the following: output distance functions D_O^0 and D_O^1 have the translog functional form with time-invariant second-order coefficients defined by equations (16)–(23); a firm follows short-run profit-maximising behaviour in periods $t = 0$ and 1, as in equation (11). Then, the joint effect of technical

[21] Balk (1998) derived the same result as CCD under a more general condition, allowing technical inefficiency. Balk's approach would allow us to generalise the result of the present paper so that it still holds under the existence of a technical inefficiency type.

progress and capital input growth, $SHIFT$, can be computed from observed prices and quantities as follows:

$$SHIFT = \frac{\prod_{m=1}^{M}\left(y_m^1/y_m^0\right)^{s_m}}{\prod_{q=1}^{Q}\left(x_{L,q}^1/x_{L,q}^0\right)^{\bar{s}_{L,q}}}, \qquad (24)$$

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

$s_m = \frac{1}{2}\left(\frac{p_m^0 y_m^0}{\boldsymbol{p}^0 \cdot \boldsymbol{y}^0} + \frac{p_m^1 y_m^1}{\boldsymbol{p}^1 \cdot \boldsymbol{y}^1}\right)$ 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 equation (24) 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 explicitly resorting to capital input data.

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 short-run 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 ratio of the growth rate of output to 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 endpoints 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 $f^t(x_K^t, x_L^1)/f^t(x_K^t, x_L^0)$ and the growth rate of labour input is x_L^1/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 in order that $(f^t(x_K^t, x_L^1)/f^t(x_K^t, x_L^0))/(x_L^1/x_L^0) = (f^t(x_K^t, x_L^1)/x_L^1)/(f^t(x_K^t, x_L^0)/x_L^0)$.

[Place Figure 2 appropriately here]

We generalise the growth rates of labour input and output between two points on the period t short-run production frontier 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 defined the input quantity index, which is the counterpart of the growth rate of total inputs $(\boldsymbol{x}_K, \boldsymbol{x}_L)$, by comparing the radial distances from the two input vectors to the period t production frontier. The input distance function is used for the radial scaling of total inputs $(\boldsymbol{x}_K, \boldsymbol{x}_L)$. Adapting the input distance function used by CCD, we introduce the labour input distance function that measures the radial distance from 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:

$$D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) \equiv \max_\delta \left\{\delta : F^t\left(\boldsymbol{y}, \boldsymbol{x}_K, \frac{x_{L,-1}}{\delta}\right) \leq \frac{x_{L,1}}{\delta}\right\}. \quad (25)$$

Given outputs $\boldsymbol{y}$, $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ is the maximum contraction of labour inputs $\boldsymbol{x}_L$ enabling 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 $\boldsymbol{y}$ to be 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 $\boldsymbol{x}_L$.

Furthermore, we can relate the *period t labour input distance function* to the *period t* 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 enabling the contracted labour inputs $\boldsymbol{x}_L / D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L)$ and outputs $\boldsymbol{y}$ to be 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 the radial distances from 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:

$$LABOUR(t, \boldsymbol{y}) \equiv D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1) / D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0). \quad (26)$$

If labour inputs increase between two periods $t = 0$ and 1, $\boldsymbol{x}_L^1$ moves further away from the origin than $\boldsymbol{x}_L^0$, indicating that the labour input vector $\boldsymbol{x}_L^1$ is larger than the labour input vector $\boldsymbol{x}_L^0$. The maximum contraction of labour inputs $\boldsymbol{x}_L$ for producing outputs $\boldsymbol{y}$ with the period t capital inputs $\boldsymbol{x}_K^t$ and the period t technology increases such that $D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0) \leq D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1)$. It leads 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 generalise 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 labour inputs $\boldsymbol{x}_L$ are not uniquely determined by the short-run production frontier. Let $P^t(\boldsymbol{x}_L)$ be the portion of the period t short-run production frontier conditional on labour inputs $\boldsymbol{x}_L$, consisting of the set of maximum outputs $\boldsymbol{y}$ attainable from $\boldsymbol{x}_L$ using capital inputs and technology available at period t. It is defined as follows:

$$P^t(\boldsymbol{x}_L) \equiv \{y : F^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_{L-1}) = x_{L,1}\}. \quad (27)$$

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 $P^t(\boldsymbol{x}_L)$. 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 $P^t(\boldsymbol{x}_L^0)$ and $P^t(\boldsymbol{x}_L^1)$. We begin with the reference outputs vector $\boldsymbol{y}$. We measure the distance between $P^t(\boldsymbol{x}_L^0)$ and $P^t(\boldsymbol{x}_L^1)$, comparing the radial distances from $\boldsymbol{y}$ to $P^t(\boldsymbol{x}_L^0)$ and $P^t(\boldsymbol{x}_L^1)$. It is defined as follows:

$$OUTPUT(t, \boldsymbol{y}) \equiv D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0) / D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1). \quad (28)$$

If 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 $\boldsymbol{x}_L^1$, $P^t(\boldsymbol{x}_L^1)$ shifts outward to that of outputs attainable from $\boldsymbol{x}_L^0$, $P^t(\boldsymbol{x}_L^0)$. 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 allows a firm to produce less outputs while holding capital input fixed and using the same technology, $P^t(\boldsymbol{x}_L^1)$, shifts inward to $P^t(\boldsymbol{x}_L^0)$, leading 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:[22]

$$SCALE(t, \boldsymbol{y}) \equiv \frac{OUTPUT(t,\boldsymbol{y})}{LABOUR(t,\boldsymbol{y})} = \left(\frac{D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0)}{D_O^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1)}\right) / \left(\frac{D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^1)}{D_L^t(\boldsymbol{y}, \boldsymbol{x}_K^t, \boldsymbol{x}_L^0)}\right). \quad (29)$$

Each choice of reference short-run production frontier and reference output vector $\boldsymbol{y}$ may generate a different measure of the returns to scale effect between two periods 0 and 1. We calculate two measures by using short-run production frontiers and output vectors 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:

$$SCALE \equiv \sqrt{SCALE(0, \boldsymbol{y}^0) \cdot SCALE(1, \boldsymbol{y}^1)}. \quad (30)$$

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

$$SCALE = \sqrt{\left(\frac{f^0(x_K^0, x_L^1)}{y^0} / \frac{x_L^1}{x_L^0}\right)\left(\frac{y^1}{f^1(x_K^1, x_L^0)} / \frac{x_L^1}{x_L^0}\right)}. \quad (31)$$

Given the period t short-run production frontier, the ratio of the LP associated with $\boldsymbol{x}_L^1$ to the LP associated with $\boldsymbol{x}_L^0$ represents the LP growth induced by the movement along the period t short-run production frontier. $SCALE$ is the geometric mean of those ratios conditional on the period 0 and 1 short-run production frontiers.

$SCALE$ is a theoretical measure defined by the unknown short-run distance functions, and there are several methods of implementing it. We adopt the same approach as we do for $SHIFT$. In addition to a firm's short-run profit-maximising behaviour and a translog functional form for the short-run output distance function, we also assume a translog functional form for the labour input distance function. Similar to the case of $SHIFT$, we begin by showing that the first-order derivatives of the distance functions with respect to labour input and output 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.[23] We apply the implicit function theorem to the input requirement function $F^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_{L,-1}/\delta) = x_{L,1}/\delta$ to solve for

---

[22] This formulation is a counterpart of the return to scale effect on TFP growth proposed by Lovell (2003; 450). Lovell's definition is based on the input distance function instead of the labour input distance function. We return to this point in the last section.

[23] $SCALE$ is defined using the output and labour input distance functions. Since we have already shown how to compute the first derivatives of the output distance functions in equations (14) and (15), we now focus on the short-run labour input distance functions.

$\delta = D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ around $(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t)$.[24] Its derivatives are represented by the derivatives of $F^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_{L,-1})$. We have the following equations for $t = 0$ and 1:[25]

$$\nabla_y D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{-1}{x_{L,1}^t - x_{L,-1}^t \cdot \nabla_{x_{L,-1}} F^t(y^t, x_K^t, x_{L,-1}^t)} \nabla_{\boldsymbol{y}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t), \quad (32)$$

$$\nabla_{x_L} D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = \frac{1}{x_{L,1}^t - x_{L,-1}^t \cdot \nabla_{x_{L,-1}} F^t(y^t, x_K^t, x_{L,-1}^t)} \begin{bmatrix} 1 \\ -\nabla_{x_{L,-1}} F^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_{L,-1}^t) \end{bmatrix}. \quad (33)$$

We assume that $(\boldsymbol{y}^t, \boldsymbol{x}_L^t)$ is a solution to the period t short-run profit maximisation problem (11) for $t = 0$ and 1. By substituting equations (12) and (13) obtained from the profit maximisation into equations (32) and (33), we obtain the following equations (34) and (35) for $t = 0$ and 1:

$$\nabla_{\boldsymbol{y}} D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = -\boldsymbol{p}^t / \boldsymbol{w}^t \cdot \boldsymbol{x}_L^t, \quad (34)$$

$$\nabla_{x_L} D_L^t(\boldsymbol{y}^t, \boldsymbol{x}_K^t, \boldsymbol{x}_L^t) = [1/\boldsymbol{w}^t \cdot \boldsymbol{x}_L^t] \begin{bmatrix} w_1^t \\ \boldsymbol{w}_{-1}^t \end{bmatrix} = \boldsymbol{w}^t / (\boldsymbol{w}^t \cdot \boldsymbol{x}_L^t). \quad (35)$$

Equations (34) and (35) 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 short-run 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 (34) and (35) for $t = 0$ and 1. The distance functions evaluated at the production plan in different periods such as $D_L^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $D_L^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ also constitute $SCALE$. Hence, the above equations are insufficient for obtaining $SCALE$. In addition to a firm's short-run profit maximisation, we further assume a translog functional form with time-invariant second-order coefficients for the period t labour input distance function for $t = 0$ and 1, which is defined as following:

$$\begin{aligned} &\ln D_L^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) \equiv \\ &\tau_0^t + \sum_{m=1}^M \tau_m^t \ln y_m + \left(\tfrac{1}{2}\right) \sum_{i=1}^M \sum_{j=1}^M \tau_{i,j} \ln y_i \ln y_j + \\ &\sum_{p=1}^P \upsilon_p^t \ln x_{K,p} + \left(\tfrac{1}{2}\right) \sum_{i=1}^P \sum_{j=1}^P \upsilon_{i,j} \ln x_{K,i} \ln x_{K,j} + \\ &\sum_{q=1}^Q \omega_q^t \ln x_{L,q} + \left(\tfrac{1}{2}\right) \sum_{i=1}^Q \sum_{j=1}^Q \omega_{i,j} \ln x_{L,i} \ln x_{L,j} + \\ &\sum_{m=1}^M \sum_{p=1}^P \xi_{m,p} \ln y_m \ln x_{K,p} + \sum_{m=1}^M \sum_{q=1}^Q \psi_{m,q} \ln y_m \ln x_{L,q} + \\ &\sum_{p=1}^P \sum_{q=1}^Q \zeta_{p,q} \ln x_{K,p} \ln x_{L,q} \end{aligned} \quad (36)$$

where the parameters satisfy the following restrictions:

$$\tau_{i,j} = \tau_{j,i} \text{ for all } i \text{ and } j \text{ such as } 1 \le i < j \le M; \quad (37)$$

$$\upsilon_{i,j} = \upsilon_{j,i} \text{ for all } i \text{ and } j \text{ such as } 1 \le i < j \le P; \quad (38)$$

$$\omega_{i,j} = \omega_{j,i} \text{ for all } i \text{ and } j \text{ such as } 1 \le i < j \le Q; \quad (39)$$

---

[24] It corresponds to CCD applying the implicit function theorem to the input requirement function to solve the input distance function.

[25] Equation (4) implies that equations (34) and (35) are well defined.

$$\sum_{q=1}^{Q} \omega_q^t = 1; \tag{40}$$

$$\sum_{i=1}^{Q} \omega_{i,q} = 0 \text{ for } q = 1, \ldots, Q; \tag{41}$$

$$\sum_{q=1}^{Q} \psi_{m,q} = 0 \text{ for } m = 1, \ldots, M; \text{ and} \tag{42}$$

$$\sum_{q=1}^{Q} \zeta_{p,q} = 0 \text{ for } p = 1, \ldots, P. \tag{43}$$

Equation (36) is the same functional form defined by equation (16) that we assumed for the output distance function in the discussion of $SCALE$. However, parameters in both functional forms are independent and allowed to be varied.[26] Moreover, the restrictions on parameters on the labour input distance function differ from those on the output distance function. We replace restrictions (20)–(23) with that of (40)–(43). While restrictions (20)–(23) guarantee the linear homogeneity in outputs $\boldsymbol{y}$ for the output distance function, restrictions (40)–(43) guarantee the linear homogeneity in labour inputs $\boldsymbol{x}_L$ for the labour input distance function.

The translog functional form characterised by equations (36)–(43) is a flexible functional form and 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 various types of technical progress are allowed. Under the assumptions of short-run profit-maximising behaviour and the translog functional form, a theoretical index of the returns to scale, $SCALE$, coincides with a formula of price and quantity observations as is shown in the following proposition.

Proposition 2

Assume the following: output distance functions D_O^0 and D_O^1 have the translog functional form with time-invariant second-order coefficients defined by equations (16)–(23); labour input distance functions D_L^0 and D_L^1 have the translog functional form with time-invariant second-order coefficients defined by equations (36)–(43); a firm follows short-run profit-maximising behaviour in

[26] Beginning from an output distance function that has the translog functional form defined by equations (16)–(23), one can derive the labour input distance function that corresponds to the output distance function (the corresponding labour input distance function). This labour input distance function will not have the translog functional form. However, we assume an independent translog functional form defined by equations (36)–(43) instead of the corresponding labour input distance function. It is because we will encounter approximation errors with respect to the true labour input distance functions in either case. The translog functional form is considered the second-order approximation to the true function. Thus, the approximation errors attributed to the third and higher-order derivatives exist in output and labour input distance functions that has the translog functional form. Therefore, the corresponding labour input distance function does not disentangle itself from the influence of the approximation error of the output distance function. Both errors are very small and, to our knowledge, it is difficult to judge which is more serious. Thus, there is no reason to adopt the corresponding labour input distance function rather than a labour input distance function that has the translog functional form. See Appendix B.

periods $t = 0$ and 1, as in equation (11). Then, the returns to scale effect, $SCALE$, can be computed from observed prices and quantities as follows:

$$SCALE = \frac{\prod_{q=1}^{Q}\left(x_{L,q}^{1}/x_{L,q}^{0}\right)^{s_{L,q}}}{\prod_{q=1}^{Q}\left(x_{L,q}^{1}/x_{L,q}^{0}\right)^{\bar{s}_{L,q}}}. \tag{44}$$

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

$s_{L,q} = \frac{1}{2}\left(\frac{w_q^0 x_{L,q}^0}{\boldsymbol{p}^0 \cdot \boldsymbol{y}^0} + \frac{w_q^1 x_{L,q}^1}{\boldsymbol{p}^1 \cdot \boldsymbol{y}^1}\right)$ 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 equation (44) can be interpreted as the ratio of the quantity indexes of labour inputs. Both terms are the weighted geometric average of the growth rates for labour inputs. The numerator uses the ratio of labour compensation for a particular type of labour input to the total value added as weight, and the denominator 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 becomes small: hence, making the magnitude of $SCALE$ smaller. Conversely, if labour income share is small, the magnitude of $SCALE$ becomes larger.

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

Corollary 1

Assume the following: output distance functions D_O^0 and D_O^1 have the translog functional form with time-invariant second-order coefficients defined by equations (16)–(23); labour input distance functions, D_L^0 and $D_{L,}^1$ have the translog functional form with time-invariant second-order coefficients defined by equations (36)–(43); a firm follows short-run profit-maximising behaviour in periods $t = 0$ and 1, as in equation (11). Then, the product of $SHIFT$ and $SCALE$ can be computed from observed prices and quantities as follows:

$$SHIFT \times SCALE = \frac{\prod_{m=1}^{M}\left(y_m^1/y_m^0\right)^{s_m}}{\prod_{q=1}^{Q}\left(x_{L,q}^{1}/x_{L,q}^{0}\right)^{\bar{s}_{L,q}}}, \tag{45}$$

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-hand side of equation (45) represents growth in LP, its numerator coincides with the Törnqvist output quantity index, and the denominator is the Törnqvist labour input quantity index. Thus, we simply call the right-hand side of equation (45) the Törnqvist LP growth index. Equation (45) allows us to completely decompose LP growth into two components, $SHIFT$ and $SCALE$, when multiple inputs and outputs are employed. This decomposition is justifiable as a generalisation 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 in Figure 1.

Balk (2005) provided a general framework for decomposing productivity indexes. Balk argued that, for meaningful decomposition, each factor in decomposition should be independent of other factors.[27] Several decomposition results dealing with the Malmquist TFP index are criticised from this point of view. The difficulty in these decompositions of the Malmquist TFP index is attributed to the fact that the Malmquist TFP index itself is not transitive in input and output quantities.

On the other hand, our theoretical measures of $SHIFT$ and $SCALE$ are defined, independent of each other. Thus, seeing at a glance whether a mere multiplication of two indexes coincides with LP growth is difficult. They coincide only when the underlying distance functions have translog functional forms. Therefore, our decomposition result is immune from Balk's criticism. Moreover, we emphasise that the Törnqvist LP index, the logarithm of which appears on the right-hand side of equation (45), satisfies transitivity in labour input and output quantities for fixed shares of value added and labour compensation.

5. An Application to US Industry Data

Having discussed the theory of the decomposition, we now explore its empirical significance with industry data. The industry data covering the period 1987–2009 is taken from the Bureau of Labour Statistics (BLS) multifactor productivity data. We use a gross output, three intermediate inputs (energy, materials and purchased services) and a labour input at current and constant prices by 59 industries, which constitute the non-farm private business sector. Labour input at constant prices measures the number of hours worked.[28] These industries are categorised either as goods-producing industries (goods sector) or services-providing industries (services sector).

[Place Table 1 appropriately here]

Table 1 compares LP growth and its components across the non-farm private business, the goods and the services sectors. For the entire sample period 1987–2009, the returns to scale effect had a negative impact on LP growth of 2.19 per cent per year in the non-farm private business sector. Whereas the joint effect of technical progress and capital input growth was an annual average of 2.38 per cent, it was largely offset by the returns to scale effect of –0.19 per cent on average per year. During the same period, the returns to scale effect appeared differently in two sectors. Whereas the positive returns to scale effect raised the services sector LP by 0.36 per cent per year on average, the negative returns to scale effect lowered the goods sector LP by −0.39 per cent per year on average. During the period 1987–2009, the average growth rate of the goods sector LP was 2.43 per cent, about 0.3 per cent higher than that of the services sector LP. However, once

---

[27] 'Now, of course, every mathematical expression a can, given any other expression b, be decomposed as $a = (a/b)a$. However, not all such decompositions are meaningful'. (Balk 2005).

[28] 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 LP growth induced by changes in the characteristics of labour input.

the returns to scale effect is controlled, the order is reversed, resulting in an average growth rate of the goods sector LP of 2.07 per cent, which is about 0.4 per cent lower than that of the services sector LP.

[Place Table 2 appropriately here]

Table 2 summarises the growth in labour input for the non-farm private business, the goods and the services sectors. According to both the weighted and the unweighted average of the detailed industries, labour inputs in the goods sector decreased on average, whereas labour inputs in the services sector increased on average. The different role played by the returns to scale effect in both sectors is attributed to the difference in the growth of labour input between two sectors.

With reference to Table 1, dividing the entire sample period 1987–2009 into three periods is useful: the 'productivity slowdown' period 1987–1995; the 'productivity resurgence' period 1995–2007 and the 'great recession' period 2007–2009. A productivity slowdown in the US economy began in the early 1970s with an average annual growth rate of 1.42 per cent for the non-farm private business sector during the period 1987–1995. Productivity growth surged after 1995 with an average annual growth rate of 2.75 per cent during the period 1995–2007. During the global financial crisis, labour input used for production sharply declined at a much faster pace than real value added, leading LP growth with an average annual growth rate of 1.84 per cent in 2007–2009.

Triplett and Bosworth (2004, 2006) and Bosworth and Triplett (2007) found, in US industry data, that LP growth in the services sector was stagnant and lower than LP growth in the goods sector.[29] Our dataset also documented the difference in LP growth between the goods and the services sectors. The services sector LP grew at an average growth rate of 1.19 per cent during the period 1985–1995, much lower than an average annual rate of 1.97 per cent for the goods sector. 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 sector with an average annual rate of 1.82 per cent comes close to the goods sector with an average annual rate of 2.04 per cent. Thus, although the services sector LP grew much slower during the period 1987–1995 than the goods sector LP, the productive capacity of labour in the services sector, which is the output attainable from given labour inputs, grew at a comparable pace to the goods sector. The fact that the service sector LP grew slower than the goods sector LP reflects that the greater increase in labour input in the services sector restrained LP from increasing significantly.

In reference to Table 2, labour input in the goods sector only slightly increased, leading to a modest returns to scale effect during the period 1987–1995, and even decreased, leading a positive returns to scale effect during the period 1995–2007. In contrast, labour input in the services sector steadily increased until 2007, leading to the negative returns to scale effect in the periods 1987–1995 and 1995–2007. LP growth in the goods sector is still larger than that in the services sector LP during the period 1995–2007. The gap in LP growth between two sectors is much smaller than during the period 1987–1995. However, as shown in Table 1, once we control the returns to scale effect, the order is reversed, resulting in LP

[29] Triplett and Bosworth (2006) used the term *Baumol's disease* to identify with the situation in which LP growth in the services sector is likely to stagnate. They argued that this disease was cured in the mid-1990s.

growth explained by technical progress and capital input growth at an average annual rate of 3.18 per cent for the service sector, higher than the 2.81 per cent for the goods sector. Thus, although LP growth during the period 1995–2007 was lower in the services sector than the goods sector, the productive capacity of labour increased more in the services sector than the goods sector.

The pattern that Triplett and Bosworth (2004, 2006) and Bosworth and Triplett (2007) pointed out dissolved after 2008. The services sector LP grew at an average annual growth rate of 2.09 per cent, even higher than the goods sector at an average annual rate of 1.42 per cent during 2007–2009. During this period, the declining labour inputs lead to positive returns to scale effects in both sectors. The goods sector shows a particularly large returns to scale effect with an average annual rate of 3.69 per cent, which is more than a three-fold average annual rate of 1.02 per cent for the services sector. However, even this significantly large returns to scale effect in the goods sector cannot compensates a negative effect of technical progress and capital input growth with an annual growth rate of –2.27 per cent, leading LP growth lower than the goods sector.[30]

[Place Tables 3, 4 and 5 appropriately here]

Tables 3, 4 and 5 show LP growth and its components and growth in labour input and labour income share by industry during the periods 1987–1995, 1995–2007 and 2007–2009. The pattern found in the aggregate study based on the sector data in Table 1 is also documented in the detailed industries. During the period 2007–2009, most industries in both sectors showed significantly positive returns to scale effects, reflecting sharp declines in labour inputs. Most industries in the services sector show negative returns to scale effects until 2007. They are particularly significant during the period 1987–1995. However, the returns to scale effects for most industries in the goods sector are very modest before 2007. They are negative during the period 1987–1995 and positive during the period 1995–2007.

There are exceptional industries in both the goods and the services sectors. Two industries show significant and positive returns to scale effects with an average annual rate of more than 1 per cent during the period 1987–1995: 1.99 per cent for *oil and gas extraction* industry; 1.16 per cent for *petroleum and coal products* industry. *Support activities for mining* industry shows a significant and negative returns to scale effect with an average annual rate of more than −2.17 per cent during the period 1995–2007. Conversely, three industries in the services sector show positive returns to scale effects during both 1987–1995 and 1995–2007: *utilities* industry, *rail transportation* industry and *pipeline transportation* industry. In these industries, there is a trend of decrease in labour input throughout the entire sample period, unlike other industries in the services sector. During the period 2007–2009, when decreasing labour inputs lead to positive returns to scale effects in most industries of the goods and services sector, three industries show significant and negative returns to scale effects, reflecting greatly increasing labour inputs: −3.55 per cent per year for *oil and gas extraction* industry; −5.7 per cent per year for *water transportation* industry and −2.07 per cent per year for *education services* industry.

---

[30] This negative growth rate is mainly accounted for by the decline of TFP based on our own calculation.

Equation (44) tells us that the returns to scale effect depends on labour income share as well as growth in labour input. The returns to scale effect will apparently get smaller under large labour income share. The detailed industry study reveals cases when the returns to scale effect induced by labour input growth can be greatly mitigated by the large labour income share. The *wood product* industry shows an extremely large decrease in labour input with an average annual rate of −19.15 during the period 2007–2009. However, its returns to scale effect is an annual average of 1.73 per cent, relatively small in magnitude during this period. Its large labour income share of 91.85 per cent offset the impact of a large decline of labour input for this industry. The *rental and leasing services and lessors of intangible assets* industry shows one of the largest returns to scale effect with an average annual rate of 5.45 per cent. Growth in its labour input is an average annual rate of –6.57 per cent, which is a comparable scale. The returns to scale effect of this industry was amplified by its smallest labour income share of 16.91 per cent among all industries in both sectors.

6. Conclusion

This paper 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 growth in LP induced by the shift in the short-run production frontier. The returns to scale effect appears as the LP growth induced by 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 US industry data for the period 1987–2009. A large part of the difference in LP growth between the goods sector and the services sector can be attributed to the difference in the returns to scale effect.

It is possible to give our decomposition result a different interpretation. Once we make labour inputs include all the capital inputs, $SHIFT$ becomes merely the technical progress effect, measuring the shift in the underlying production frontier. Similarly, under the same setting, the returns to scale effect $SCALE$ represents the TFP induced by the movement along the underlying production frontier. Note that $SCALE$ coincides with the geometric mean of Lovell's (2003) scale effect.[31] Thus, Corollary 1 means that the TFP growth calculated by Törnqvist quantity indexes is fully decomposed into the product of the two effects. Our result differs from the previous studies in that we give an exact interpretation of the index number formula for the returns to scale effect in Proposition 2. Although CCD derived the same index number formula that is a function of the degree of returns to scale, they did not explain what this formula itself measures. On the other hand, while Lovell (2003) proposed a measure of the returns to scale effect representing the TFP growth induced by movement along the underlying production frontier, Lovell did not offer any index number formula that equals or approximates it. The contribution of the present paper to the literature of TFP growth decomposition is

[31] In a strict sense, $SCALE(t, \boldsymbol{y}^t)$ coincides with the period t *activity effect* in Lovell (2003), which is the returns to scale effect evaluated on the period t production frontier. Thus, when we consider the returns to scale effect on the TFP growth between two periods, the geometric mean of the period 0 and 1 activity effect $SCALE = \sqrt{SCLAE(0, \boldsymbol{y}^0) \cdot SCALE(1, \boldsymbol{y}^1)}$ is an appropriate measure.

in showing the equivalence between CCD's index number formula and Lovell's theoretical measure.

This paper assumes the firm's profit-maximising behaviour and ruled out inefficient production processes. If we relax the firm's profit-maximising 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 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 implementing the decomposition of LP growth without assuming a firm's profit-maximising behaviour, we can estimate the distance function using econometric and linear programming techniques. However, we leave this exercise for future research.

Appendix A

Proof of Proposition 2

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

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

Since the firm's short-run profit maximisation 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 equations (16) and (36)

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

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

from equations (14), (15), (34) and (35).

Appendix B

In this paper, we assume the translog functional form for the output and labour input distance functions (translog output distance function and translog labour input distance function), allowing parameters to be independent. However, since distance functions are defined with reference to the same production frontier, there should be some mathematical relationship in the functional form and parameters between two distance functions. Beginning with the period t translog output distance function D_O, we can derive the corresponding period t labour input distance function D_L as follows:

$$D_L(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) = \max_{\delta} \left\{ \delta : D_O\left(\boldsymbol{y}, \boldsymbol{x}_K, \frac{\boldsymbol{x}_L}{\delta}\right) \leq 1 \right\}. \tag{B.1}$$

Clearly, this labour input distance function corresponding to the translog output distance function (the corresponding labour input distance function) does not have the translog functional form.

However, we should not necessarily assume the corresponding labour input distance function in addition to the translog output distance function. In this appendix, we consider the problem of choosing the translog labour input distance function or the corresponding labour distance function when we assume the translog output distance function.

The translog labour input distance function approximates an arbitrary labour input distance function to the second order at an arbitrary point of approximation $(\boldsymbol{y}^*, \boldsymbol{x}_K^*, \boldsymbol{x}_L^*)$.[32] Thus, irrelevant of the type of functional form the true distance function has, the translog functional form is a good local approximation to it. Needless to say, since the translog distance function does not consider the further approximation based on the third- and higher-order derivatives, the values of the translog and the true labour input distance functions diverge as the point at which the functions are evaluated moves from the approximation point $(\boldsymbol{y}^*, \boldsymbol{x}_K^*, \boldsymbol{x}_L^*)$ (**Type 1 error**). From the same reasoning, the translog output distance function suffers from the approximation error, due to neglecting third- and higher-order derivatives. Thus, if we derive the labour input distance function from the translog output distance function as in (B.1), the corresponding labour input distance function reflects such an approximation error of the translog output distance function. Thus, the values of the corresponding labour input distance functions also differ from that of the true labour input distance function as the evaluation point moves from the approximation point (**Type 2 error**).

Since we cannot analytically compare the magnitude of these two types of approximation errors, we use a numerical example to discuss how to choose the labour input distance function. To implement $SCALE$, it is necessary to evaluate the period 0 output and labour input distance functions at $(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0)$ and $(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$, and the period 1 output and labour input distance functions at $(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ and $(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1)$. Suppose that the period 0 and 1 translog output and labour input distance functions are the local approximations at $(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0)$ and $(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1)$.[33] Then, $D_O^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$, $D_O^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$, $D_L^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and

[32] Strictly speaking, it provides the second-order log approximation.

[33] The following argument can be applied to the case for which the period 0 and 1 translog short-run labour input distance functions are the local approximations at $(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$.

$D_L^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ calculated from the translog as well as the corresponding distance functions differ from those calculated from the true distance functions.

We consider the simple case consisting of two outputs, one capital input and two labour inputs $(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) = (y_1, y_2, x_{K,1}, x_{L,1}, x_{L,2})$ so that production took place at $t = 0$ and 1 as follows:

$$(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0) = (11,2,1,2,3)$$

$$(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1) = (22,4,2,4,6).$$

The period t input requirement function $F^t(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L)$ for $t = 0$ and 1 is defined as follows:

$$F^1(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) = \frac{y_1 + 2y_2^2 - 6x_{K,1}}{x_{L,2}} - 3 \tag{B.2}$$

$$F^2(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) = \frac{y_1 + 4y_2^2 - 5x_{K,1}}{2x_{L,2}} - 2. \tag{B.3}$$

The period t output and labour input distance functions are constructed with reference to the period t short-run production frontier described by the above equations as follows:

$$D_O^0(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) = 0.5\big(6x_{K,1} + (x_{L,1} + 3)x_{L,2}\big)^{-1}\Big(y_1 + \Big(y_1^2 + 8y_2^2\big(6x_{K,1} + (x_{L,1} + 3)x_{L,2}\big)\Big)^{0.5}\Big) \tag{B.4}$$

$$D_O^1(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) = 0.5\big(5x_{K,1} + 2(x_{L,1} + 2)x_{L,2}\big)^{-1}\Big(y_1 + \Big(y_1^2 + 16y_2^2\big(5x_{K,1} + 2(x_{L,1} + 2)x_{L,2}\big)\Big)^{0.5}\Big) \tag{B.5}$$

$$D_L^0(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) = 0.5\big(y_1 + 2y_2^2 - 6x_{K,1}\big)^{-1}\Big(3x_{L,2} + \Big(9x_{L,2}^2 + 4x_{L,1}x_{L,2}\big(y_1 + 2y_2^2 - 6x_{K,1}\big)\Big)^{0.5}\Big) \tag{B.6}$$

$$D_L^1(\boldsymbol{y}, \boldsymbol{x}_K, \boldsymbol{x}_L) = 0.5\big(y_1 + 4y_2^2 - 5x_{K,1}\big)^{-1}\Big(4x_{L,2} + \Big(16x_{L,2}^2 + 8x_{L,1}x_{L,2}\big(y_1 + 4y_2^2 - 5x_{K,1}\big)\Big)^{0.5}\Big). \tag{B.7}$$

Since the period t production takes place at the period production frontier, $D_O^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0) = D_O^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1) = 1$ and $D_L^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0) = D_L^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1) = 1$.

The period 0 translog output and labour input distance functions are derived by applying the second-order Taylor expansion to the period 0 output and labour input distance functions in (B.4) and (B.6) at $(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0)$. Similarly, the period 1 translog output and labour input distance functions are derived by applying the second-order Taylor expansion to the period 1 output and labour input distance functions in (B.5) and (B.7) at $(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1)$. Once we derive the translog output distance functions, we can also derive the corresponding distance function following equation (B.1). Thus, we compute $SHIFT$ under three difference approaches in this example, as shown in Table B.1.

[Place Table B.1 appropriately here]

Approach 1 assumes the independent translog functional form for output and labour input distance functions. This paper takes this approach. Approach 2 uses the corresponding labour input distance function instead of the translog one. Approach 3 uses the true distance functions of (B.3)–(B.6), and is a reference for comparison of Approaches 1 and 2. The approach that gives us the value of $SHIFT$ closer to that of Approach 3 is the better one. Because $(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0)$ and $(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1)$ are approximation points, $D_L^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0)$, $D_L^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1)$, $D_L^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^0)$ and $D_L^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^1)$ equal 1 across Approaches 1, 2 and 3.

[Place Table B.2 appropriately here]

Table B.2 compares the values of $SHIFT$, $D_L^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $D_L^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ under three different approaches.[34] First, the differences between the true values based on Approach 3 and the estimates based on Approaches 1 and 2 are very small, at approximately a 0.2 per cent difference. This means that Approaches 1 and 2 are capable of approximating the true value with great accuracy. Second, if we compare Approaches 1 and 2 for estimating $D_L^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $D_L^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ by constraint, Approach 2 is better. Although the difference between them is extremely small, at less than 0.002 per cent, the corresponding labour input distance function can approximate the true labour input distance function above $(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$ more accurately than the translog labour input distance function in this example. However, for estimating $SHIFT$, Approach 1 is better than Approach 2. The estimate based on Approach 1 is closer to the true value based on Approach 3 than that on Approach 2, by about 0.003 per cent. Thus, the translog labour input distance function can approximate the true value of $SCALE$ more accurately than the corresponding labour input distance function in this example.

Thus, we conclude the following from our numerical example. First, both Type 1 and Type 2 errors are rather small. This means that assuming the translog labour input distance function is as good as assuming the corresponding labour input function. Second, even though Type 1 error (which comes from Approach 1) might be more serious than Type 2 error (which comes from Approach 2) in some cases, this does not necessarily mean that the estimate of $SHIFT$ based on Approach 2 is more accurate than that on Approach 1. Therefore, it is hardly possible to judge Approaches 1 and 2 from our numerical example. In our present knowledge, there is no reason to adopt the corresponding labour input distance function instead of the translog labour input distance function.

[34] We also need to calculate $D_O^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$ and $D_O^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$. However, since the same translog output distance function is used in Approaches 1 and 2, we omit their values from Table B.2.

References

[1] Balk, B.M. (1998), *Industrial Price, Quantity, and Productivity Indices*, New Boston: Kluwer Academic Publishers.

[2] Balk, B.M. (2001), “Scale Efficiency and Productivity Change”, *Journal of Productivity Analysis* 15, 159–183.

[3] Balk, B.M. (2005), “The Many Decompositions of Productivity Change”, Presented at the North American Productivity Workshop, Toronto, 2004 and at the Asia-Pacific Productivity Conference, Brisbane, 2004 (available at www.rsm.nl/bbalk).

[4] 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.

[5] 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.

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

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

[8] Diewert, W.E. and H. Mizobuchi (2009), “Exact and Superlative Price and Quantity Indicators”, *Macroeconomic Dynamics* 13(S2), 335–380.

[9] Diewert, W.E. and K.J. Fox (2010), “Malmquist and Törnqvist Productivity Indexes: Returns to Scale and Technical Progress with Imperfect Competition”, *Journal of Economics* 101, 73–95.

[10] 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.

[11] 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.

[12] Fare, R. and D. Primont (1995), *Multiple-Output Production and Duality: Theory and Applications*, Boston: Academic Publishers.

[13] 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.

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

[15] 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.

[16] 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.

[17] Jones, C.I. (2002), “Sources of U.S. Economic Growth in a World of Idea”, *American Economic Review* 92, 220–239.

[18] 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.

[19] Lovell, C.A.K. (2003), “The Decomposition of Malmquist Productivity Indexes”, *Journal of Productivity Analysis* 20, 437–458.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

Table 1: Sources of Sectoral Labour Productivity Growth

	1987-2009	*1987-1995*	*1995-2007*	*2007-2009*
Non-farm private business sector				
Labour productivity growth	2.19	1.42	2.75	1.84
Technical progress and capital input growth	2.38	1.87	3.08	0.25
Returns to scale effect	-0.19	-0.45	-0.32	1.60
Goods sector				
Labour productivity growth	2.43	1.97	2.90	1.42
Technical progress and capital input growth	2.07	2.04	2.81	-2.27
Returns to scale effect	0.36	-0.07	0.09	3.69
Services sector				
Labour productivity growth	2.10	1.19	2.71	2.09
Technical progress and capital input growth	2.49	1.82	3.18	1.08
Returns to scale effect	-0.39	-0.63	-0.47	1.02

Note: All figures are average annual percentages.

Table 2: Labour Input Growth and Labour Income Share

	1987-2009	*1987-1995*	*1995-2007*	*2007-2009*
Labour input growth (un-weighted average)				
Private non-farm business	0.62	1.41	0.96	-4.63
Goods-producing industries	-1.01	0.21	-0.39	-9.65
Services-producing industries	1.24	1.97	1.47	-3.06
Labour input growth (weighted average)				
Private non-farm business	0.15	1.22	0.24	-4.69
Goods-producing industries	-1.65	-0.14	-1.51	-8.59
Services-producing industries	1.30	2.09	1.36	-2.20
Labour income share				
Private non-farm business	67.86	68.27	68.00	65.38
Goods-producing industries	66.87	68.79	66.44	61.73
Services-producing industries	68.20	68.02	68.59	66.59

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 labour.

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

Industry		*Labour productivity growth*	*Technical progress and capital input*	*Returns to scale effect*	*Labour input growth*	*Labour income ratio*
Goods sector						
Forestry, fishing, and related activities	113-115	-4.45	-3.36	-1.09	2.28	53.22
Oil and gas extraction	211	3.11	1.12	1.99	-2.93	30.33
Mining, except oil and gas	212	7.35	6.62	0.73	-1.81	58.58
Support activities for mining	213	1.60	1.04	0.56	-1.96	69.54
Construction	23	0.50	0.64	-0.14	0.85	85.42
Food and beverage and tobacco products	311,312	1.61	1.93	-0.33	0.72	52.66
Textile mills and textile product mills	313,314	2.81	2.61	0.20	-0.94	75.97
Apparel and leather and applied products	315,316	5.77	5.14	0.63	-2.67	76.43
Wood product	321	-1.57	-1.47	-0.10	0.34	72.25
Paper products	322	0.49	0.59	-0.10	0.26	58.33
Printing and related support activities	323	0.76	0.95	-0.20	1.41	86.15
Petroleum and coal products	324	5.45	4.29	1.16	-1.64	28.92
Chemical products	325	-0.40	-0.15	-0.25	0.53	48.66
Plastics and rubber products	326	2.01	2.70	-0.69	2.02	66.22
Nonmetallic mineral products	327	1.07	1.17	-0.10	0.12	69.21
Primary metals	331	1.07	0.95	0.12	-0.35	73.11
Fabricated metal products	332	1.40	1.64	-0.24	0.79	72.60
Machinery	333	-1.17	-0.85	-0.32	1.33	73.21
Computer and electronic products	334	19.98	19.68	0.30	-1.40	76.69
Electrical equipment, appliances, and components	335	-3.57	-4.01	0.45	-1.09	59.65
Transportation equipment	336	-1.85	-1.92	0.07	-0.49	78.37
Furniture and related products	337	0.61	0.62	-0.01	-0.08	82.06
Miscellaneous manufacturing	339	2.41	2.84	-0.43	1.38	67.67
Services sector						
Utilities	22	3.47	3.18	0.29	-0.39	25.90
Wholesale trade	42	2.50	2.83	-0.33	1.15	72.75
Retail trade	44,45	3.38	3.71	-0.32	1.41	77.65
Air transportation	481	2.90	4.29	-1.39	5.11	75.51
Rail transportation	482	5.14	4.51	0.63	-2.50	73.19
Water transportation	483	8.53	8.15	0.39	-0.80	53.05
Truck transportation	484	3.84	4.10	-0.27	1.17	77.49
Transit and ground passenger transportation	485	-0.87	-0.20	-0.67	3.22	79.29
Pipeline transportation	486	2.86	2.22	0.63	-1.27	50.13
Other transportation and support activities	487,488,492	-3.38	-2.17	-1.21	6.34	81.10
Warehousing and storage	493	4.68	4.86	-0.18	0.88	81.20
Publishing industries	511,516	2.42	2.91	-0.49	1.71	71.38
Motion picture and sounds recording industries	512	-3.80	-2.36	-1.43	5.19	72.44
Broadcasting and telecommunications	515,517	4.63	4.80	-0.16	0.24	38.12
Information and data processing services	518,519	-2.60	-1.66	-0.95	4.37	78.22
Federal reserve banks, credit intermediation, and related activities	521,522	0.18	-0.12	0.30	-0.69	57.02
Securities, commodity contracts, and investments	523	5.90	6.19	-0.29	1.69	78.21
Insurance carriers and related activities	524	2.29	2.50	-0.21	1.21	82.09
Funds, trusts, and other financial vehicles	525	0.12	1.55	-1.42	1.87	22.12
Real estate	531	0.29	1.10	-0.81	0.96	15.58
Rental and leasing services and lessors of intangible assets	532,533	1.61	4.02	-2.41	3.19	24.35
Legal services	5411	-0.47	-0.32	-0.15	1.46	89.86
Computer systems design and related services	5415	2.04	2.60	-0.56	5.43	89.29
Miscellaneous professional, scientific, and technical services	5412-5414,5416-5419	1.11	1.39	-0.28	2.50	88.73
Management of companies and enterprises	55	-0.77	-0.80	0.03	0.08	94.27
Administrative and support services	561	0.32	0.98	-0.67	4.74	85.81
Waste management and remediation services	562	1.14	1.90	-0.76	2.01	62.58
Educational services	61	0.76	1.52	-0.76	2.55	65.30
Ambulatory health care services	621	-1.74	-1.17	-0.57	3.69	84.48
Hospitals and nursing and residential care facilities	622,623	-2.75	-2.08	-0.66	3.33	79.69
Social assistance	624	2.27	2.42	-0.15	2.12	92.95
Performing arts, spectator sports, museums, and related activities	711,712	2.49	2.72	-0.23	1.57	85.77
Amusements, gambling, and recreation industries	713	-0.48	1.08	-1.57	5.47	70.81
Accommodation	721	1.73	2.30	-0.56	1.87	71.12
Food services and drinking places	722	-0.64	-0.28	-0.36	1.94	81.68
Other services, except government	81	-0.18	0.10	-0.28	2.25	87.48

Note: All figures are average annual percentages.

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

Industry		*Labour productivity growth*	*Technical progress and capital input*	*Returns to scale effect*	*Labour input growth*	*Labour income ratio*
Goods sector						
Forestry, fishing, and related activities	113-115	2.35	2.52	-0.18	0.49	63.48
Oil and gas extraction	211	-3.42	-3.63	0.22	-0.62	21.10
Mining, except oil and gas	212	-0.23	-0.39	0.16	-0.55	50.28
Support activities for mining	213	-1.55	0.62	-2.17	4.83	62.80
Construction	23	-2.10	-1.72	-0.37	2.46	85.45
Food and beverage and tobacco products	311,312	-0.52	-0.57	0.05	-0.11	50.98
Textile mills and textile product mills	313,314	5.83	4.26	1.56	-5.95	72.98
Apparel and leather and applied products	315,316	8.58	6.31	2.28	-9.55	76.02
Wood product	321	1.98	1.77	0.21	-1.43	78.57
Paper products	322	2.74	1.52	1.21	-2.97	58.08
Printing and related support activities	323	2.08	1.76	0.33	-2.53	86.73
Petroleum and coal products	324	4.80	3.79	1.01	-1.37	19.48
Chemical products	325	4.37	3.55	0.82	-1.52	44.88
Plastics and rubber products	326	3.51	3.02	0.48	-1.34	62.47
Nonmetallic mineral products	327	-0.10	-0.13	0.03	-0.18	63.74
Primary metals	331	0.14	-0.63	0.77	-2.88	67.29
Fabricated metal products	332	1.18	1.14	0.04	-0.28	69.43
Machinery	333	2.52	2.06	0.46	-2.00	72.82
Computer and electronic products	334	27.30	27.15	0.14	-2.79	79.05
Electrical equipment, appliances, and components	335	0.88	-0.10	0.98	-2.90	65.63
Transportation equipment	336	4.67	4.20	0.47	-1.56	72.19
Furniture and related products	337	3.13	2.85	0.28	-1.17	76.31
Miscellaneous manufacturing	339	5.48	5.25	0.23	-0.69	66.56
Services sector						
Utilities	22	2.53	1.28	1.25	-1.76	29.69
Wholesale trade	42	5.74	5.93	-0.19	0.62	71.36
Retail trade	44,45	3.37	3.47	-0.10	0.42	76.64
Air transportation	481	10.35	9.99	0.36	-2.35	75.87
Rail transportation	482	2.17	1.74	0.43	-1.56	69.04
Water transportation	483	3.58	5.12	-1.54	2.93	45.89
Truck transportation	484	2.37	2.61	-0.24	1.07	79.17
Transit and ground passenger transportation	485	2.08	2.31	-0.23	1.08	77.97
Pipeline transportation	486	5.60	4.71	0.89	-1.70	41.71
Other transportation and support activities	487,488,492	2.07	2.21	-0.14	0.64	77.10
Warehousing and storage	493	1.38	2.04	-0.66	3.12	78.95
Publishing industries	511,516	6.07	5.98	0.09	-0.17	66.94
Motion picture and sounds recording industries	512	1.32	1.70	-0.38	1.30	66.79
Broadcasting and telecommunications	515,517	6.89	7.10	-0.20	0.26	43.40
Information and data processing services	518,519	4.75	4.81	-0.06	2.21	65.16
Federal reserve banks, credit intermediation, and related activities	521,522	1.09	1.82	-0.72	1.60	54.21
Securities, commodity contracts, and investments	523	9.26	9.63	-0.37	2.84	86.52
Insurance carriers and related activities	524	1.40	1.55	-0.15	0.64	77.70
Funds, trusts, and other financial vehicles	525	5.05	7.23	-2.18	2.62	15.78
Real estate	531	0.93	1.98	-1.05	1.28	18.20
Rental and leasing services and lessors of intangible assets	532,533	3.27	4.11	-0.84	1.10	19.90
Legal services	5411	1.05	1.11	-0.06	0.78	92.09
Computer systems design and related services	5415	3.53	4.09	-0.56	5.87	88.85
Miscellaneous professional, scientific, and technical services	5412-5414,5416-5419	1.57	1.95	-0.37	2.36	83.48
Management of companies and enterprises	55	1.13	1.26	-0.13	0.96	86.61
Administrative and support services	561	2.05	2.36	-0.31	2.73	89.76
Waste management and remediation services	562	0.07	0.99	-0.92	2.23	57.97
Educational services	61	-1.00	0.25	-1.25	3.62	61.23
Ambulatory health care services	621	0.23	0.72	-0.49	3.21	85.00
Hospitals and nursing and residential care facilities	622,623	-2.62	-2.11	-0.52	2.54	78.66
Social assistance	624	2.88	3.09	-0.21	2.82	92.71
Performing arts, spectator sports, museums, and related activities	711,712	2.03	2.19	-0.16	0.94	85.10
Amusements, gambling, and recreation industries	713	0.59	0.92	-0.32	1.06	71.74
Accommodation	721	1.06	1.38	-0.32	0.88	64.62
Food services and drinking places	722	2.03	2.40	-0.37	1.73	78.57
Other services, except government	81	-0.29	-0.18	-0.11	0.85	88.23

Note: All figures are average annual percentages.

Table 5: Sources of Industry Labour Productivity Growth, 2007–2009

Industry		*Labour productivity growth*	*Technical progress and capital input*	*Returns to scale effect*	*Labour input growth*	*Labour income ratio*
Goods sector						
Forestry, fishing, and related activities	113-115	3.35	1.83	1.52	-5.48	72.37
Oil and gas extraction	211	9.91	13.46	-3.55	4.19	16.66
Mining, except oil and gas	212	11.85	8.58	3.27	-5.36	38.54
Support activities for mining	213	17.57	14.06	3.51	-6.64	47.97
Construction	23	1.47	-0.06	1.53	-12.90	88.62
Food and beverage and tobacco products	311,312	-7.67	-8.58	0.91	-1.91	52.15
Textile mills and textile product mills	313,314	-1.38	-6.34	4.95	-14.46	65.47
Apparel and leather and applied products	315,316	0.99	-1.48	2.47	-12.96	82.85
Wood product	321	-20.18	-21.91	1.73	-19.15	91.85
Paper products	322	-4.67	-7.50	2.83	-6.56	56.86
Printing and related support activities	323	-2.42	-3.71	1.29	-9.52	86.36
Petroleum and coal products	324	15.88	14.56	1.31	-1.55	13.36
Chemical products	325	-15.62	-17.79	2.17	-3.79	42.96
Plastics and rubber products	326	-9.36	-13.21	3.84	-11.78	68.14
Nonmetallic mineral products	327	-2.14	-5.74	3.60	-12.49	70.79
Primary metals	331	1.39	-4.08	5.47	-13.07	58.85
Fabricated metal products	332	-17.35	-20.78	3.42	-10.72	67.91
Machinery	333	0.53	-2.53	3.06	-9.55	67.97
Computer and electronic products	334	11.94	10.00	1.94	-5.07	61.60
Electrical equipment, appliances, and components	335	6.26	2.77	3.48	-8.43	58.31
Transportation equipment	336	-11.77	-14.95	3.18	-12.09	73.67
Furniture and related products	337	-24.10	-25.57	1.47	-9.17	85.18
Miscellaneous manufacturing	339	4.17	0.74	3.43	-9.01	61.92
Services sector						
Utilities	22	-0.71	-0.16	-0.55	0.76	28.11
Wholesale trade	42	7.50	6.30	1.20	-4.00	70.02
Retail trade	44,45	-0.48	-1.36	0.89	-3.98	77.69
Air transportation	481	11.60	9.70	1.89	-6.95	71.34
Rail transportation	482	-4.01	-5.98	1.97	-4.62	57.30
Water transportation	483	1.03	6.73	-5.70	9.57	40.57
Truck transportation	484	-2.48	-4.06	1.58	-7.35	78.81
Transit and ground passenger transportation	485	-4.59	-4.75	0.16	-1.16	85.07
Pipeline transportation	486	-6.92	-5.78	-1.13	1.72	35.04
Other transportation and support activities	487,488,492	-0.98	-2.18	1.20	-4.82	75.06
Warehousing and storage	493	-3.03	-3.11	0.08	-0.42	78.76
Publishing industries	511,516	3.43	1.09	2.35	-5.99	60.73
Motion picture and sounds recording industries	512	-0.72	-1.50	0.78	-2.42	68.42
Broadcasting and telecommunications	515,517	4.84	3.28	1.56	-2.50	37.34
Information and data processing services	518,519	6.81	4.86	1.95	-3.63	46.10
Federal reserve banks, credit intermediation, and related activities	521,522	6.76	4.34	2.42	-5.10	52.11
Securities, commodity contracts, and investments	523	-3.49	-4.15	0.66	-4.41	85.13
Insurance carriers and related activities	524	8.54	8.23	0.30	-1.33	77.02
Funds, trusts, and other financial vehicles	525	7.68	5.68	2.00	-2.24	10.76
Real estate	531	1.07	-2.14	3.21	-3.79	14.64
Rental and leasing services and lessors of intangible assets	532,533	2.93	-2.51	5.45	-6.57	16.91
Legal services	5411	-1.81	-1.97	0.16	-1.99	91.70
Computer systems design and related services	5415	2.09	2.34	-0.25	1.84	86.33
Miscellaneous professional, scientific, and technical services	5412-5414,5416-5419	8.30	6.74	1.55	-7.46	79.09
Management of companies and enterprises	55	-0.35	-0.23	-0.12	0.83	87.49
Administrative and support services	561	-1.15	-1.68	0.52	-4.63	88.60
Waste management and remediation services	562	-1.80	-2.55	0.75	-1.80	58.10
Educational services	61	-1.62	0.45	-2.07	4.83	57.31
Ambulatory health care services	621	1.31	1.64	-0.33	2.06	83.68
Hospitals and nursing and residential care facilities	622,623	2.95	3.22	-0.27	1.09	72.48
Social assistance	624	1.88	1.85	0.03	-0.47	92.90
Performing arts, spectator sports, museums, and related activities	711,712	-1.45	-1.89	0.43	-2.68	82.98
Amusements, gambling, and recreation industries	713	-3.29	-4.12	0.84	-3.56	77.04
Accommodation	721	-1.31	-2.80	1.49	-4.14	63.92
Food services and drinking places	722	-6.41	-6.76	0.35	-1.77	80.16
Other services, except government	81	-4.59	-4.82	0.23	-2.24	89.86

Note: All figures are average annual percentages.

Table B1: Different Approaches to Calculating *SHIFT*

	Output distance function	*Labour input distance function*
Approach 1	Translog	Translog
Approach 2	Translog	Corresponding
Approach 3	True function	True function

Table B2: Comparison of Accuracy of Different Approaches

	$SCALE$	$D_L^0(\boldsymbol{y}^0, \boldsymbol{x}_K^0, \boldsymbol{x}_L^1)$	$D_L^1(\boldsymbol{y}^1, \boldsymbol{x}_K^1, \boldsymbol{x}_L^0)$
Approach 1	0.816458	1.758342	0.558346
Approach 2	0.816440	1.758287	0.558252
Approach 3	0.814507	1.758306	0.558258

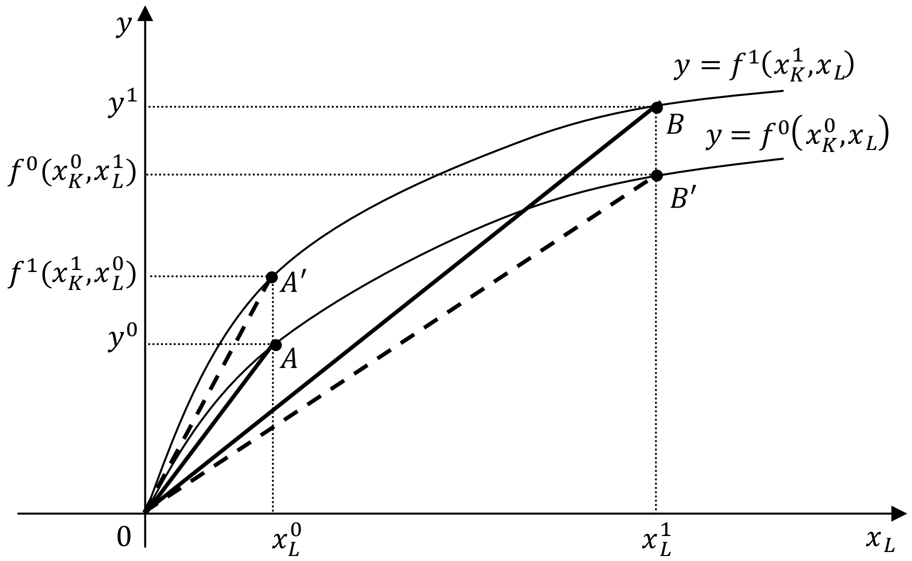

Figure 1: Labour Productivity Growth and Shift in the Short-run Production Frontier

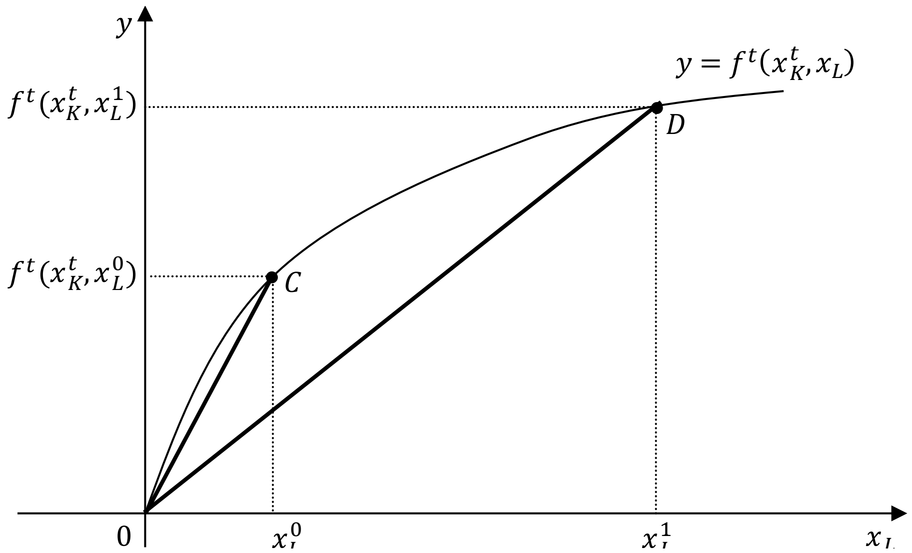

Figure 2: Returns to Scale Effect and Movement along the Short-run Production Frontier