The Japanese Economic Review
Vol. 49, No. 3, September 1998

AN INDEX NUMBER METHOD FOR ESTIMATING SCALE ECONOMIES AND TECHNICAL PROGRESS USING TIME-SERIES OF CROSS-SECTION DATA: SOURCES OF TOTAL FACTOR PRODUCTIVITY GROWTH FOR JAPANESE MANUFACTURING, 1964–1988*

By TAKANOBU NAKAJIMA,† MASAO NAKAMURA‡ and KANJI YOSHIOKA†

†Keio University; ‡University of British Columbia

Sample multicollinearity often makes it difficult to estimate returns to scale. We present an index number method to overcome potential multicollinearity problems when the production function is homogeneous of degree *k*. We apply our method to estimate empirically the effects of returns to scale and technical progress on growth in total factor productivity (TFP) using establishment data for Japanese manufacturing industries. We find that, while significant scale economies exist in many manufacturing industries, the TFP growth in the last twenty-five years is attributable primarily to technical progress. This finding also validates the current practice of assuming constant-returns-to-scale production functions in macroeconometric modelling.
JEL Classification Numbers: C43, D24, 030.

1. Introduction

In some policy studies, economies of scale play an essential role. For example, there is some evidence that the horizontal trade in manufactured goods among developed countries can be explained by scale economies, among other factors (Helpman and Krugman, 1985). If a nation's markets for strategically important goods were served by foreign oligolopolists enjoying scale economies, and hence were not contestable from the point of view of domestic producers (Baumol, Panzar and Willig, 1982), then policies to protect domestic producers might be justified on infant-industry grounds (Tinbergen, 1945; Kemp, 1969). In the policy debtate on the Canada–US Free Trade Agreement, Canadian proponents of the treaty based their arguments primarily on the benefits of scale economies that Canadian manufacturers would enjoy if they had access to an enlarged North American market (Harris and Cox, 1984). Estimates for scale economy parameters are important inputs to policy-oriented applied general equilibrium models such as the one used by Harris and Cox.

In principle, economies of scale could be estimated using establishment data and a production function of some flexible form. Sample multicollinearity, which normally exists among certain inputs, their prices and output in cross-section data for

* An earlier version of this paper was presented at the Econometric Society European Meeting in Uppsala, Sweden, 22–26 August 1993. We wish to thank two anonymous referees for their helpful comments on earlier versions of the paper.

Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK.

establishments, often prevents us from identifying with sufficient statistical precision scale economies and other unknown parameters such as price and substitution parameters.[1)] For the types of policy study involving scale economies which we have in mind, however, it is not essential to obtain estimates, for example, for price and substitution parameters.[2,3)]

The following simple example illustrates the relationship between multicollinearity among production inputs and estimating the returns to scale. Returns to scale for the Cobb–Douglas production function, $y = ax_1^{\alpha 1}x_2^{\alpha 2}$ or $Y = A + \alpha_1 X_1 + \alpha_2 X_2$, where $Y = \ln y$, $A = \ln a$ and $X_i = \ln x_i$ $(i = 1, 2)$, are given by $\alpha_1 + \alpha_2$. When inputs x_1 and x_2 are highly correlated, the individual coefficients α_1 and α_2 are difficult to estimate. Let $x_1 = x_2$ (i.e. $X_1 = X_2$) for simplicity. Then it is still possible to estimate returns to scale $\alpha_1 + \alpha_2$ by regressing Y on X_1 (or X_2) alone (perhaps very precisely), even though it is impossible to estimate α_1 or α_2 separately. Our estimation method could be viewed as an extension of this simple observation to the case where X_1 and X_2 are not perfectly correlated.

In the first part of this paper (Section 2) we present an econometric procedure to estimate returns to scale in production using establishment data. This procedure uses index number theory to aggregate inputs at establishments of different sizes and can be interpreted as a generalization of Frisch's (1965) approximation formula for his passus coefficient (scale elasticity).

Another problem related to scale economies in empirical productivity analysis is to distinguish between the contributions of economies of scale and technical progress in the growth of total factor productivity (TFP). Such a distinction, however, is not possible when a production function with constant returns to scale is assumed (Solow, 1957; Jorgenson and Griliches, 1967; Jorgenson, Gollop and Fraumeni, 1987; and Kuroda, Yoshioka and Jorgenson, 1984). This is because, under the assumptions of constant returns to scale and perfect competition, TFP growth coincides with technical change. The constant-returns-to-scale assumption has been imposed on the specification of production functions in many empirical studies that utilize aggregate time-series data.

Some empirical results suggest, however, that it is economies of scale, rather than

---

1) A standard method of estimating these unknown parameters is to estimate a flexible cost function using cost share equations. However, estimating scale economies using a translog cost function, for example, requires the estimation of the cost function itself as well as the share equation system (Berndt, 1991, p. 476). Since output, its squares and its cross-products with input prices are all in the cost function, multicollinearity can cause serious estimation problems. Banker *et al.* (1988, p. 40) also note that their procedure is likely to provide unreliable estimates for returns to scale if there is a collinearity problem in estimating flexible-form production functions.

2) Policy-oriented empirical studies emphasizing estimating scale effects with little attention paid to price effects include Komiya (1962), Ozaki (1969, 1976) and Giles and Wyatt (1992). Scale economy parameter estimates were stand-alone inputs to the Canada–US Free Trade model by Harris and Cox (1984).

3) Also, virtually all Japanese industrial policies in the past dealt with scale economy effects. In promoting specific manufacturing industries such as the steel, automobile and chemical industries, the Ministry of International Trade and Industry (MITI) always specified the level of production scale required for an internationally competitive production facility. Exactly the same sort of thinking underlies the policy put forward by the Ministry of Finance (MOF) for the Japanese banking industry as liberalization (deregulation) measures in finance industries are currently being implemented. Economies of scale considerations led the MOF to encourage mergers and acquisitions among Japanese banks.

technical progress, that explains TFP growth. Denison (1974), for example, concluded that economies of scale explain the TFP growth for the United States.[4] Using aggregate time-series for the USA, Berndt and Khaled (1979) also found the presence of economies of scale and relatively little technological progress. The difficulty in distinguishing between the effects of scale economies and technical progress using time-series data has long been recognized (e.g. Solow, 1959). There are two main reasons for this difficulty. First is the multicollinearity, for example among certain production inputs, their prices, output and the time parameter (a proxy for technical progress), which enter the production (or cost) function as arguments. This problem of multicollinearity is often confounded by the limited variation and ranges for values observed for aggregate time-series. Second is the question of whether an increase in aggregate output actually represents an increase in production scale at the level of individual establishments rather than an increase in the number of establishments. Regarding their findings, Berndt and Khaled (1979, p. 1221) acknowledge the inherent difficulty in separating the effects of scale economies from the effects of technical progress when using aggregate time-series data, and caution the reader by saying that "[t]hese results should be interpreted cautiously. In our judgement, more precise estimates of return to scale and rates of technical progress may require use of pooled cross-section and time-series data."

In this paper we present empirical estimates for the contributions to the Japanese TFP growth of returns to scale and technical progress for the Japanese manufacturing sector. We use time-series of cross-section data on establishments. Our estimation method, discussed in the following two sections, utilizes index number theory and is not likely to suffer from multicollinearity problems. We find in Sections 4 and 5 that, while significant scale economy effects are found for Japanese manufacturing establishments in the cross-sectional sense, they do not explain TFP gains over time. The gains in TFP over 1964–88 are due mostly to technical progress. Section 6 contains concluding remarks.

2. Elasticity of scale and its lower and upper bounds

Suppose the scalar output, x, of an establishment in period (year) t is characterized by

$$x_t = f(v_t), \qquad t = 1, 2, \ldots, T,$$

where v_t is the n-dimensional production input vector, $v = (v^1, v^2, \ldots, v^n)$, and $f(v)$ is the production function. (Unless otherwise stated, all vectors in this paper are column vectors.) Time subscript t will be omitted except when our discussion requires the explicit treatment of time. For a given fixed input vector v_0 and a positive scalar μ, the elasticity of scale, k, is defined by

$$k = (\mathrm{d}x/x)/(\mathrm{d}\mu/\mu) = \mathrm{d}\ln x/\mathrm{d}\ln \mu,$$

where $x = f(\mu v_0) = f(v)$ and k depends on v_0 in general. In many empirical applications involving cross-sectional data, we often observe high correlations among some of the inputs $(v^1, v^2, \ldots, v^n)$, their prices and output. For example, in the

4) Komiya (1962) also found that economies of scale are the primary reason for the TFP growth in the US steam power production.

application to be discussed below, we will use Japanese establishment data on three production inputs: employment (the number of workers), capital stock, and raw material ($n = 3$). Since large establishments employ more workers, own more capital stock, utilize more raw material and produce more than small establishments, we expect these inputs and output to be highly correlated (see Figure 1). A similar multicollinearity problem is found between output and workers' wages, for example, which rise as establishment size increases (Oi, 1983). When production inputs (input prices) and output are highly correlated, it is difficult to identify the returns-to-scale parameter in the production (cost) function.

2.1 Derivation of lower and upper bounds for scale elasticity

Suppose the objective of a particular empirical study is to estimate the elasticity of scale. Suppose also that, because of multicollinearity, the elasticity of scale cannot be estimated with precision using simultaneous estimation of a flexible form involving all unknown parameters including scale elasticity.[5] Under this scenario, in this section we will derive a number of empirically useful expressions for scale elasticity.

Lower and upper bounds for scale elasticity are first obtained, under the assumptions of homotheticity and cost minimization. Empirical specifications for

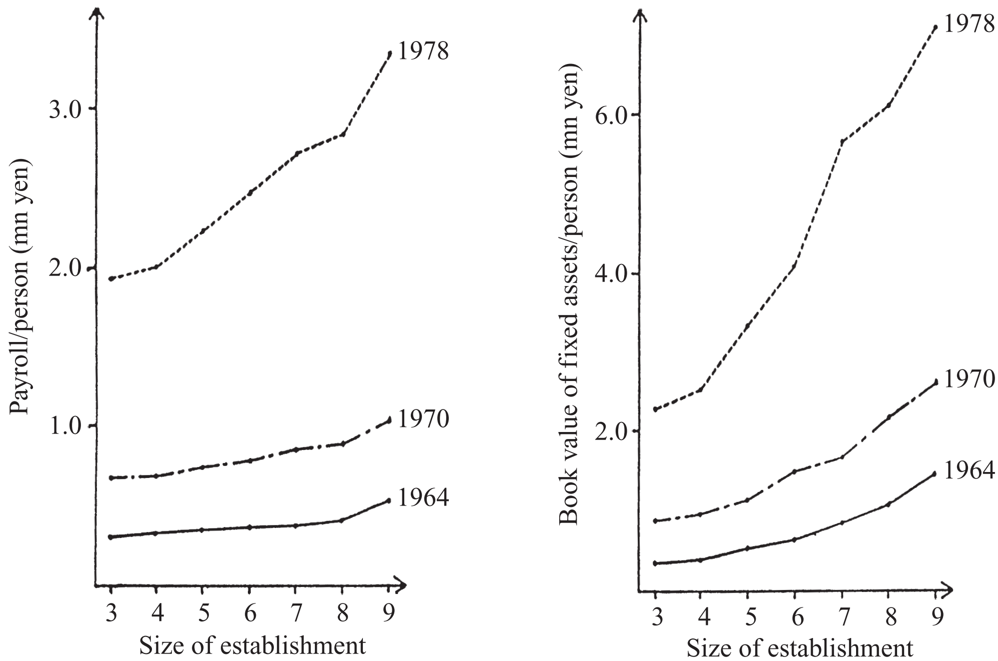

FIGURE 1. Wages and capital stock by establishment size, 1964–78

5) This situation is not uncommon in econometric practice. For example, in studying the efficiency of US manufacturing industries, Caves and Barton (1990, p. 34) note: "The idea of an intensive examination of scale economies was dropped after the results for the twelve-industry panel were analyzed. The behavior of the estimated coefficients, especially in the translog functions, did not inspire confidence in our ability to determine the minimum efficiency scales ..."

estimating scale elasticity and its bounds are then derived in Section 2.2 under the assumption that the production function is homogeneous of degree k. These empirical specifications are further extended to include technical progress in Section 3.

Suppose we observe (x_i, v_i, p_i) for small $(i = 1)$ and large $(i = 2)$ establishments, where x_i is output such that $x_2 > x_1$ and $x_i = f(v_i)$ is the production function, and where v_i denotes an n-dimensional input vector at establishment i with price vector p_i $(i = 1, 2)$.

We derive lower and upper bounds for the elasticity of scale k based on the two assumptions: (1) that the production function $x = f(v)$ is homothetic and hence the elasticity of scale depends on the size of output (x) only, and (2) that each establishment chooses its production input vector v so as to minimize the total cost of production.[6)]

Consider two rays in the input (v) space R_i $(i = 1, 2)$: R_i passes through the origin and v_i. Rays R_1 (for the small establishment, $n = 1$) and R_2 (for the large establishment, $n = 2$) are depicted in Figure 2 together with the supporting cost hyperplanes (lines) C_i at v_i and isoquants Q_i $(i = 1, 2)$. (Note that no convexity property is assumed for the production function at this point.)

We denote the intersections of isoquants Q_i with ray R_j $(i \neq j)$ by v_{*i} $(i, j = 1, 2)$.

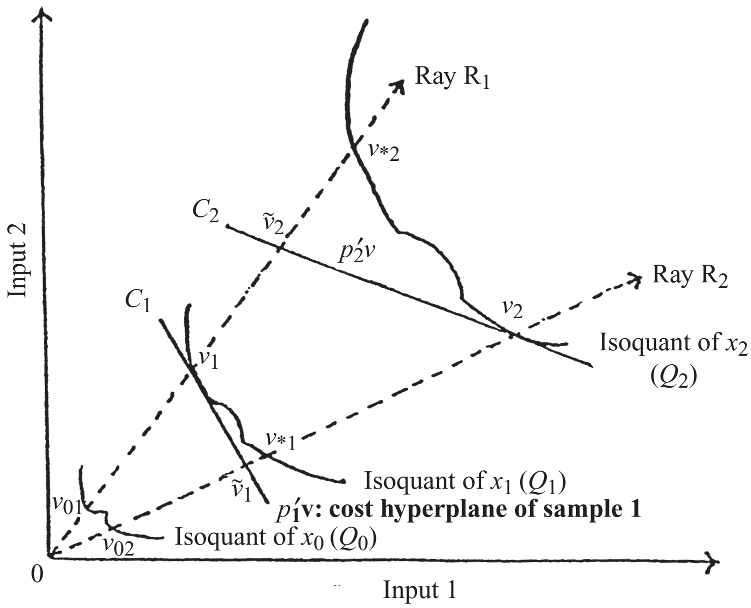

FIGURE 2. Production input space $(n = 2)$

6) Homotheticity is not empirically testable in general since testing it requires the assumption of a specific functional form for a production function. Our empirical work to follow assumes a homogeneous production function of degree k, and hence the homotheticity assumption will be satisfied. Cobb–Douglas and CES functions often used in productivity analysis are homothetic. Using translog specifications for analysing the productivity of the US meat products industry, Chambers (1988, p. 199) also tested the null hypothesis of homogeneity against the alternative hypothesis of homotheticity and could not reject homogeneity. He concluded that to assume homogeneity is not much more restrictive than homotheticity.

The intersections of cost lines C_i with ray R_j $(i \neq j)$ are denoted by $\tilde{v}_i$ $(i, j = 1, 2)$. Points v_{*i} and $\tilde{v}_i$ $(i = 1, 2)$ are shown in Figure 2. We also define another isoquant Q_0 corresponding to some fixed output x_0, and denote by v_{01} and v_{02} the intersections of Q_0 with rays R_1 and R_2, respectively (see Figure 2). The homotheticity assumption implies that isoquants Q_0, Q_1 and Q_2 are isomorphic with respect to the origin.

Since $f(v)$ is homothetic, we have, for any positive scalar μ,

$$x = f(\mu v_{01}) = f(\mu v_{02}).$$

Let μ_1 and μ_2 be defined by $v_2 = \mu_2 v_{02}$ and $v_1 = \mu_1 v_{01}$. Then we have

$$x_i = f(\mu_i v_{01}) = f(\mu_i v_{02}), \qquad i = 1, 2, \tag{1}$$

and

$$v_{*1} = \mu_1 v_{02} \qquad \text{and} \qquad v_{*2} = \mu_2 v_{01}. \tag{2}$$

Hence we have

$$v_{*2} = (\mu_2/\mu_1)v_1 \qquad \text{and} \qquad v_2 = (\mu_2/\mu_1)v_{*1}. \tag{3}$$

If v_{*1} and v_{*2} were observable, then the arithmetic mean elasticity[7] of scale, $E^*(k)$, measured on rays 1 and 2, would be given, respectively, by

$$E^*(k)^1 = (\ln x_2 - \ln x_1)/(\ln \mu_2 - \ln \mu_1) = (\ln x_2 - \ln x_1)/(\ln v^j_{*2} - \ln v^j_1)$$

and

$$E^*(k)^2 = (\ln x_2 - \ln x_1)/(\ln \mu_2 - \ln \mu_1) = (\ln x_2 - \ln x_1)/(\ln v^j_2 - \ln v^j_{*1}),$$
$$\text{any } j = 1, 2, \ldots, n,$$

where v^j is the jth element of input vector v $(j = 1, 2, \ldots, n)$. Since the denominator of $E^*(k)^i$ is $\ln(\mu_2/\mu_1)$ for $i = 1, 2$, we have $E^*(k)^1 = E^*(k)^2$. That is, the mean elasticity of scale does not depend on the ray along which it is measured.

In practice, we do not observe v_{*1} and v_{*2}. We do observe $\tilde{v}_1$ and $\tilde{v}_2$, however, since we have

$$\tilde{v}_1 = (p'_1 v_1/p'_1 v_2)v_2 \qquad \text{and} \qquad \tilde{v}_2 = (p'_2 v_2/p'_2 v_1)v_1. \tag{4}$$

Cost minimization implies[8]

$$\tilde{v}_i \leqslant v_{*i}, \qquad i = 1, 2. \tag{5}$$

Define k_l and k_u by

$$k_l = (\ln x_2 - \ln x_1)/(\ln v^j_2 - \ln \tilde{v}^j_1), \tag{6}$$

and

$$k_u = (\ln x_2 - \ln x_1)/(\ln \tilde{v}^j_2 - \ln v^j_1), \qquad \text{for any } j = 1, 2, \ldots, n. \tag{7}$$

7) $E^*(k)^i$ is the arithmetic mean elasticity of scale measured along ray i (R_i) between isoquants Q_1 and Q_2, and serves as a discrete approximation to the true elasticity of scale. This notation was originally used by Frisch (1965).

8) The proof of equation (5) is as follows. Define a positive scale λ_i such that $\tilde{v}_i = \lambda_i v_{*i} (i = 1, 2)$. By definition, we have (Figure 2) $p_i v_i = p_i \tilde{v}_i$ and $x_i = f(v_{*i})$. Cost minimization implies $p_i v_i = \min_v [p_i v;\, x_i \leqslant f(v)]$, from which we have $\lambda_i p_i v_{*i} = p_i v_i = p_i \tilde{v}_i \leqslant p_i v_{*i}$. Thus we have $\lambda_i \leqslant 1$ and hence $v_i \leqslant v_{*i}$. ■

Then we have $k_1 \leqslant E^*(k)^2 = E^*(k)^1 \leqslant k_u$, or

$$k_1 \leqslant k \leqslant k_u, \tag{8}$$

where the arithmetic mean operator for k will be dropped for notational convenience in the following.

If we view establishments 1 and 2 as a base point and a compared point, respectively, then we can define the Laspeyres and Paasche input indexes, Q_L and Q_P, as follows:

$$Q_L = p_1' v_2 / p_1' v_1 \qquad \text{and} \qquad Q_P = p_2' v_2 / p_2' v_1. \tag{9}$$

Substituting (4) into the denominators of (6) and using (9), we get

$$k_l = (\ln x_2 - \ln x_1)/\ln Q_L \tag{10}$$

and

$$k_u = (\ln x_2 - \ln x_1)/\ln Q_P. \tag{11}$$

Frisch (1965, p. 68) proposed the following approximation formula for the elasticity of scale (which he called the "passus coefficient"):

$$k = (\ln x_2 - \ln x_1)/(\ln v_2^j - \ln v_1^j), \qquad \text{any } j,$$

which would be an exact formula for k under the perfect multicollinearity among production inputs ($j = 1, 2, \ldots, n$). Our lower and upper-bound formulas (10) and (11), which are exact under our assumptions, may be viewed as generalizations of Frisch's approximation formula.[9)]

2.2 Estimation of k_1, k_u and k based on flexible functional forms

For establishments 1 and 2 and the production function $x = f(v)$, (10) and (11) imply

$$k_l = [\ln f(v_2) - \ln f(v_1)]/\ln Q_L^{1,2} \tag{12a}$$

and

$$k_u = [\ln f(v_2) - \ln f(v_1)]/\ln Q_P^{1,2}, \tag{12b}$$

where $Q_L^{1,2}$ and $Q_P^{1,2}$ are Laspeyres and Paasche indexes, respectively, for establishments 1 and 2.

In our econometric specification, we assume that the error term u enters the production function in multiplicative form; i.e.,

$$x_i = f(v_i)e^{ui}, \qquad i = 1, 2, \ldots, I, \tag{13}$$

9) The inequality $k_1 \leqslant k_u$ (where k_1 and k_u are given by (10) and (11)) implies $Q_L \geqslant Q_p$; that is, the Laspeyres input quantities index is greater than or equal to the corresponding Paasche index. This follows from Bortkiewicz's index number theorem, which holds if price changes and input quantity changes are negatively correlated (Bortkiewicz, 1922; Allen, 1975). This assumption seems satisfied in our case. For example, as establishment size increases, the wage rate (price of labour) increases (Oi, 1983). A price (wage) increase is accompanied by utilizing less labour and more capital; that is, the capital–labour ratio increases as establishment size increase. This is illustrated in Figure 1 for some sample years.

where $E(u') = E(u_1, u_2, \ldots, u_I) = 0$, $E(uu') = (\sigma^2/2)F$ and F is the I-by-I unit matrix.

Using (13), we rewrite (12a–b) as follows:

$$\ln x_2 - \ln x_1 = k_l \ln Q_L^{1,2} + (u_2 - u_1). \tag{14a}$$

$$\ln x_2 - \ln x_1 = k_u \ln Q_P^{1,2} + (u_2 - u_1). \tag{14b}$$

We can generalize (14) to $I-1$ pairs of consecutive establishments ordered by size as follows:

$$\ln x_{i+1} - \ln x_i = k_l \ln Q_L^{i,i+1} + (u_{i+1} - u_i) \tag{15a}$$

$$\ln x_{i+1} - \ln x_i = k_u \ln Q_P^{i,i+1} + (u_{i+1} - u_i), \qquad i = 1, 2, \ldots, I-1. \tag{15b}$$

The unknown parameters k_l and k_u can be estimated using (15a) and (15b), respectively, by generalized least squares (GLS) with the variance–covariance matrix for the error term $(u_{i+1} - u_i)$ as follows:

$$\Omega = \sigma^2 \begin{bmatrix} 1 & -\frac{1}{2} & .. & 0 \\ -\frac{1}{2} & & & -\frac{1}{2} \\ 0 & .. & -\frac{1}{2} & 1 \end{bmatrix}.$$

Since the same error term enters both (15a) and (15b), we expect estimated values for σ^2 from both equations to be quite close. Estimates for σ^2 that are far apart would imply potential specification problems for (15a) and (15b).

Our estimation results for (15a) and (15b) (not reported here) show that k_l and k_u are very close to each other for our data sample,[10)] suggesting that it would be useful if we could directly estimate the value of k by regression by assuming a certain flexible functional form for $f(v)$. This is done in the following.[11)]

Suppose that $x = f(v)$ is translog and homogeneous of degree k:

$$k^{-1} \ln f(v) = b_0 + b_1' \ln v + \tfrac{1}{2} \ln v' R \ln v, \tag{16}$$

where the unknown parameters are scalar b_0; vector b_1, with its column sum equal to one; and non-positive definite matrix R, with all column sums equal to zero, and where the dimensions of b_1 and R conform to that of v (see Christensen, Jorgenson and Lau, 1973). If we apply the Quadratic Approximation lemma (Diewert, 1976) to (16) and evaluate it at the first and second smallest establishments, we get

$$\begin{aligned} k^{-1}(\ln x_2 - \ln x_1) &= \tfrac{1}{2}(k^{-1}\nabla \ln x_2 + k^{-1}\nabla \ln x_1)'(\ln v_2 - \ln v_1) \\ &= \tfrac{1}{2}((kx_2)^{-1} V_2 \nabla x_2 + (kx_1)^{-1} V_1 \nabla x_1)'(\ln v_2 - \ln v_1), \end{aligned} \tag{17}$$

10) We find this to be the case for our Japanese manufacturing data (see Nakajima, Nakamura and Yoshioka, 1993, for details).

11) In the following we will use the fact that scale elasticity for a homogeneous production function of degree k is given by k. This is shown as follows. Given a production function $X = f(v)$, the elasticity of scale is given by $k = \sum_{j=1}^{n}(\partial \ln X/\partial \ln v_j) = \sum_{j=1}^{n}(1/X)(\partial X/\partial v_j)v_j = (1/X)\Delta f(v)'v$, where $\Delta f(v)' = (\partial f/\partial v_1, \partial f/\partial v_2, \ldots, \partial f/\partial v_n)$. This implies $kX = \Delta f(v)'v$. On the other hand, Euler's theorem for homogeneous functions of mth order implies $mX = \Delta f(v)'v$. Thus, $k \equiv m$.

where $\nabla \ln x_1$ ($\nabla \ln x_2$) and ∇x_1 (∇x_2) are the gradients of $\ln x_1$ ($\ln x_2$) with respect to $\ln v_1$ ($\ln v_2$) and $v_1(v_2)$, respectively, and where V_1 (V_2) denotes the diagonal matrix with its (j, j)th element equal to the jth element of v_1 (v_2).

By Euler's theorem, we have $kx = \nabla f(v)'v$. Cost minimization implies that the input price vector p is proportional to $\nabla f(v)$, i.e. that $p \propto \nabla f(v)$. Thus, we have

$$\nabla f(v)'/kx = \nabla f(v)'/(\nabla f(v)'v) = p'/p'v. \tag{18}$$

By applying (18) to (17), we get

$$k^{-1}(\ln x_2 - \ln x_1) = \tfrac{1}{2}\left(\frac{p_2' V_2}{p_2' v_2} + \frac{p_1' V_1}{p_1' v_1}\right)(\ln v_2 - \ln v_1) \tag{19}$$

or

$$k = (\ln Q_T^{1,2})^{-1}(\ln x_2 - \ln x_1)$$

or

$$k = (\ln Q_T^{1,2})^{-1}(\ln f(v_2) - \ln f(v_1)).$$

In general, we have

$$\ln x_{i+1} - \ln x_i = k \ln Q_T^{i,i+1}, \tag{20}$$

where

$$\ln Q_T^{i,i+1} = \tfrac{1}{2}(w_i + w_{i+1})'(\ln v_{i+1} - \ln v_i), \qquad i = 1, 2, \ldots, I - 1 \tag{21}$$

is the log of the translog (Theil–Törnqvist) input quantity index (Theil, 1965; Törnqvist, 1936; Fisher, 1922), and w_i and w_{i+1} are the cost share vectors for the ith and $(i + 1)$st smallest establishments, given by

$$w_i = \frac{V_i p_i}{p_i' v_i}$$

and

$$w_{i+1} = \frac{V_{i+1} p_{i+1}}{p_{i+1}' v_{i+1}}. \tag{22}$$

Thus, scale elasticity can be estimated using the translog input index when $x = f(v)$ is of translog form.[12] Equation (20) can be rewritten for successive establishments i and $i + 1$ as regression equations for estimating k using the multiplicative error term specification given in (13) (see also (15a) and (15b)):

$$\ln x_{i+1} - \ln x_i = k \ln Q_T^{i,i+1} + (u_{i+1} - u_i), \qquad i = 1, 2, \ldots, I - 1, \tag{23}$$

where $f(v)$ is a translog production function.

12) Scale elasticity can also be estimated using Fisher's ideal index when $f(v)$ is of Diewert's quadratic form (see Nakajima, Nakamura and Yoshioka, 1993).

3. Estimation of returns to scale and technical progress effects using pooled cross-section and time-series data

In this section we present a method to estimate both scale elasticity and technical progress using time-series of cross-section data. Our aim is to obtain statistically significant and consistent estimates for both scale economies and technical progress. We will use cross-sectional information to estimate scale elasticity as before, while time-series information is used to estimate technical progress. The parsimonious nature of our econometric model has proved helpful for estimating the two parameters of our policy interest, given the serious multicollinearity problems we encountered in estimating from our data flexible-form production functions with many unknown parameters. (Such multicollinearity problems were also reported by other researchers who used either cross-sectional or time-series data.[13])

In order to estimate the effects of scale economies and technical progress using time-series of cross-section data, let us suppose that the production function for the ith establishment in period t ($i = 1, 2, \ldots, I,\ t = 1, 2, \ldots, T$) is given by

$$x_{it} = f(v_{it}, t), \tag{24}$$

where x_{it} and v_{it} are, respectively, a scalar output and the production input vector for the ith establishment in period t. It will be assumed that the $f(v_{it}, t)$ is homogeneous of degree k in v_{it}.

We define production input index numbers over two successive time periods, $t = S$, $S+1$ ($1 \leqslant S \leqslant T-1$), and over establishments of different sizes, $i = 1, 2, \ldots, I$, as follows. Using the smallest establishment ($i = 1$) as a base, we define chain index numbers for any time period t, q_{it}, by

$$\begin{aligned} q_{1t} &= 1 \\ q_{2t} &= q_{1t}Q_t^{1,2} = Q_t^{1,2} \\ q_{3t} &= q_{2t}Q_t^{2,3} = Q_t^{1,2}Q_t^{2,3} \\ &\cdots \\ q_{It} &= q_{I-1,t}Q_t^{I-1,I} = \prod_{i=1}^{I-1} Q_t^{i,i+1} \end{aligned} \tag{25}$$

Also, using the translog input quantity index, $Q^i_{S,S+1}$, defined for two consecutive time periods S and $S+1$ ($1 \leqslant S \leqslant T-1$) for the ith establishment by

$$\ln Q^i_{S,S+1} = \tfrac{1}{2}(w_{iS} + w_{i,S+1})'(\ln v_{i,S+1} - \ln v_{i,S}), \tag{26}$$

we can define chain index numbers, $q_{i,S+1}$, for time period $S+1$ as follows:

---

13) Efficient estimation based on fully simultaneous estimation of all unknown parameters is desirable, but our own computational experiences as well as others' suggest it is not always possible to implement it where serious multicollinearity problems exist. For example, in a study to estimate scale economies and technical progress (approximated by time) using time-series data and a translog production function, Chan and Mountain (1983, p. 665) state that "All these problems point towards the difficulty of distinguishing between scale economies and time at such an aggregate level."

$$q_{1,S+1} = q_{1S} Q^1_{S,S+1} = Q^1_{S,S+1}$$

$$q_{2,S+1} = q_{1,S+1} Q^{1,2}_{S+1} = Q^1_{S,S+1} Q^{1,2}_{S+1}$$

$$\ldots$$

$$q_{I,S+1} = q_{I-1,S+1} Q^{I-1,I}_{S+1} = Q^1_{S,S+1} \prod_{i=1}^{I-1} Q^{i,i+1}_{S+1}. \qquad (27)$$

Note that, because our regressions will involve two consecutive years S and $S+1$, we can set $q_{1S} = 1$ above.

We denote by $a(S)$ and $a(S+1)$ the amounts of theoretical output in years S and $S+1$ $(1 \leqslant S \leqslant T-1)$ for the smallest establishment $(i = 1)$ corresponding to input v_{1S} as follows:

$$a(S) = f(v_{1S}, S) \qquad (28a)$$

$$a(S+1) = f(v_{1S}, S)e^{r(S)} = x_{1S}e^{r(s)}, \qquad (28b)$$

where $x_{1S} = f(v_{1S}, S)$ by (24) and $r(S)$, time-varying rate of technical progress in time period S, is defined in Appendix 1.[14)]

Then it is shown (see Appendix 1) that the log of output for the ith establishment for two consecutive time periods $t = S,\ S+1$ satisfies the following equations:

$$\ln x_{iS} = \ln a(S) + k(S) \ln q_{iS} \qquad (29a)$$

$$\ln x_{i,S+1} = \ln a(S+1) + k(S) \ln q_{i,S+1} \qquad i = 1, 2, \ldots, I,\ S = 1, 2, \ldots, T-1. \qquad (29b)$$

Since we have by (28b) $\ln a(S+1) = \ln a(S) + r(S)$, we can write (29a) and (29b) in combined regression form:

$$\ln x_{it} = b_0 + b_1 D_{it} + b_2 \ln q_{it} + \varepsilon_{it}, \qquad i = 1, 2, \ldots, I,\ t = S, S+1, (1 \leqslant S \leqslant T-1), \qquad (30)$$

where

$$D_{it} = 1 \text{ if } t = S+1$$

$$= 0 \text{ if } t = S$$

and $b_0 \equiv \ln a(S)$, $b_1 \equiv r(S)$, $b_2 \equiv k(S)$. The ε_{it} in (30) is an added regression error term which satisfies the following:

14) Hicks neutrality as a first approximation is assumed here because, first, severe multicollinearity among input factors and time makes it difficult to estimate technical change bias and, second, the conceptual relationship of the technical change bias for input factors to the technical change bias for our aggregate input index is unclear. The change over time of technical change (i.e. $\partial/\partial t(\partial \ln X/\partial t)$) will, however, be identified in our procedure by estimating the period-to-period technical change. (In fact, our procedure allows the parameters of the production function to vary over time.) It should also be noted that testing for technical change bias generally relies on specific (production, cost or profit) functional forms and empirical results do not carry over from one functional form to another. Even non-parametric methods do not generally solve this problem. (For example, Chavas and Cox (1990) non-parametrically test Hicks neutrality assuming constant returns to scale. In our framework constant returns to scale cannot be assumed.)

$$E(\varepsilon_{it}) = 0 \tag{31a}$$

$$E(\varepsilon_{it}\varepsilon_{jt}) = \begin{cases} \sigma^2 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \tag{31b}$$

$$E(\varepsilon_{it}\varepsilon_{j,t+k}) = \begin{cases} \rho^2 \text{ if } i = j,\ k = 1, & i, j = 1, 2, \ldots, I, \\ 0 \quad \text{otherwise} & t = 1, 2, \ldots, T,\ k = 1, 2, \ldots, T - t. \end{cases} \tag{31c}$$

Thus, the ε_{it} have mean zero, have common variance σ^2, are uncorrelatedly distributed over establishments, and have autocovariance ρ^2 between two successive periods.

Using data on I establishments pooled over two consecutive years, we can estimate equation (30). The constant term b_0 gives an estimate for $\ln a(S)$ while b_1 and b_2 provide estimates, respectively, for time-varying technical change $r(S)$ and elasticity of scale $k(S)$. By repeating this estimation process for $S = 1, 2, \ldots, T - 1$, we will obtain estimates for $\ln a(S)$, $r(S)$ and $k(S)$ for each of the consecutive years: years 1 and 2, years 2 and 3, …, years $T - 1$ and T.

There are only three unknown parameters to estimate in our econometric specification (30).[15)]

Furthermore, the year dummy D_{it} and the translog input quantity chain index number q_{it} in (30) are not expected to be highly correlated.[16)] This will allow us to identify empirically both $r(S)$ and $k(S)$ without the sample problem of multi-collinearity. Since we allow the error term ε_{it} in (30) to obey a first-order autoregressive process, we estimate b_0, b_1 and b_2 using generalized least squares (GLS).

4. Empirical estimates for technical change and economies of scale: Japanese manufacturing industries, 1964–88

The Japanese Ministry of International Trade and Industry (MITI) annually conducts the Census of Manufacturing by Industry. (Twenty manufacturing industries are involved.) This Census consists of a cross-section of establishments chosen on the basis of numbers of employees. Typical size groups (numbers of employees) used are: (1) 30–49, (2) 50–99, (3) 100–199, (4) 200–299, (5) 300–499, (6) 500–999 and (7) 1000 and more. (The number of these groups, and hence the specific size range in each group, varies somewhat over time, however.) Henceforth the “size” refers to the size of establishment measured in terms of the number of employees. (See Appendix 4 for details on the sources of data used.) MITI publishes only average figures for each of the size groups by industry.

In the following, these grouped data on establishments will be viewed as ordered cross-sectional observations ($i = 1, 2, \ldots, I$); that is, establishments are ordered in the

---

15) In estimating scale economies and technical change using aggregate time series, Berndt and Khaled (1979) and Chan and Mountain (1983), for example, both had to estimate 22 unknown parameters using 25 annual observations.

16) For the particular data set that we used, the correlation coefficients calculated for the 18 manufacturing industries are quite small and range between 0.009 and 0.025. We should also point out that, because of the way our input index q was created, it behaves like an instrument itself and hence is not likely to be affected by the errors-in-variables problem in estimating (30). Our preliminary tests of endogeneity suggest little correlation between the regressors and the error term in (30).

ascending order of size: $i = 1$ and $i = I$ correspond to the smallest and largest size groups, respectively. The production inputs included are: the number of workers (v^1) as labour, the fixed assets at the beginning of each year (v^2) as capital, and the intermediate goods (v^3) as raw materials, all measured per establishment.[17,18] The corresponding input prices used are: the average annual cash earnings per worker (p^1) and v^1, and the depreciation rate for fixed assets plus the average interest rate for one-year term deposits (p^2) for v^2. Intermediate goods price p_3 is assumed to be one, since it is common to all observations for each industry and for each year. Output (x) is measured as net sales plus net increases in the inventories of final products.

In order to estimate equation (30) using our data pooled over time periods and establishments, it is necessary to deflate some of the quantities defined above (1964 = 100). The Bank of Japan output price index by industry is used to deflate our output variable x (sales). In computing the capital stock v^2, new investment in fixed assets is deflated using the investment goods deflator by industry published by the Economic Planning Agency. The input price of capital (p^2) is also adjusted by the investment goods deflator. The input of intermediate goods (v^3) is now deflated by the Bank of Japan input price deflator, which is also used as the price of intermediate goods (p^3). Because of the lack of correct industry-specific deflators, two manufacturing industries—Printing and Other—will be excluded from our empirical analysis. Thus, the following empirical analysis will be carried out using data for eighteen manufacturing industries for the period 1968–88.[19]

Table 1 and Figure 3 present GLS estimation results for r, k and $\ln a(S)$ in equation (30) averaged over the period 1964–88 for the eighteen Japanese manufacturing industries included in this study. The largest estimated values for the elasticity of scale and the rate of technical progress are found for the Food and Precision industries, respectively. Year-to-year GLS estimates for these two industries are also given in Table A1 in Appendix 3. (The complete version of this table containing estimation results for all eighteen industries is available on request from the authors.)

Estimates for the elasticity of scale, k, are statistically highly significant and stable over time. The null hypothesis H_0: $k = 1$ is rejected decisively in favour of the alternative H_1: $k > 1$ for many industries for many time periods, showing the presence of increasing returns to scale. Constant returns to scale (i.e. $k = 1$) cannot be rejected, however, for the textile industry. In addition, the Pulp, Nonferrous metals, Petroleum, and Iron and Steel industries, exhibit only modest economies of scale. However, this

17) Establishment data of this sort exist, for example, for Japan and Norway. In the following the input price vector is denoted by $p = (p^1, p^2, \ldots, p^n)$ where p^j is the price for the jth input v^j, $j = 1, 2, \ldots, n$.

18) It is possible that the costs of capital, for example, facing establishments systematically differ depending on firm size. Since our database does not allow identification of the sizes of firms that own establishments in our sample, we did not attempt to use size-based costs of capital in this paper.

19) In this paper we follow the standard practice in the productivity analysis literature that productivity is measured for the levels of all inputs that are currently utilized by establishments. This practice allows us to combine all inputs into an aggregate input index. It is also consistent with firms' optimization when all input levels adjust instantaneously to and hence reflect their desired (optimal) levels. To the extent that our interest is in measuring productivity associated with utilized levels of inputs, this practice, at least in our estimation, would not be seriously affected by the presence of lags in implementing optimal levels of certain fixed or quasi-fixed inputs such as capital. This issue, however, requires further investigation. We are indebted to a referee for bringing this point to our attention.

TABLE 1

EMPIRICAL ESTIMATES FOR THE RATE OF TECHNICAL PROGRESS AND THE ELASTICITY OF SCALE: MANUFACTURING INDUSTRIES, 1964–88[a]

Industry	Technical change r	Elasticity of scale k	Constant $\ln a(S)$
Food/kindred products	−0.0001 (0, 3)[b]	1.080 (23, 1)[c]	5.61[d]
Textiles	0.0164 (19, 2)	1.004 (16, 4)	5.22
Apparels	0.0040 (11, 1)	1.019 (16, 0)	4.70
Lumber/wood products	0.0056 (2, 1)	1.018 (16, 1)	5.42
Furniture/fixture	0.0090 (9, 2)	1.047 (18, 3)	5.17
Pulp/paper products	0.0118 (7, 2)	1.008 (13, 0)	5.55
Chemicals	0.0206 (8, 3)	1.046 (17, 2)	6.20
Petroleum/coal products	0.0088 (5, 0)	1.012 (6, 2)	6.59
Rubber/plastic products	0.0124 (7, 4)	1.047 (24, 0)	5.23
Leather/leather products	0.0065 (4, 1)	1.016 (8, 2)	5.47
Pottery/glass products	0.0135 (9, 3)	1.073 (23, 0)	5.37
Iron/steel	0.0036 (14, 2)	1.012 (12, 2)	6.01
Non-ferrous metals	−0.0014 (9, 2)	1.008 (9, 0)	6.18
Metal products	0.0147 (8, 1)	1.030 (24, 0)	5.38
General machinery	0.0187 (12, 1)	1.019 (24, 0)	5.29
Electrical machinery	0.0260 (17, 2)	1.044 (24, 0)	4.96
Transportation machinery	0.0245 (16, 2)	1.016 (16, 3)	5.29
Precision	0.0316 (13, 2)	1.021 (14, 4)	5.04

[a]This table reports the mean for each of the estimated regression coefficients calculated from 24 year-to-year regressions for the period 1964–88 using equation (30). Examples of year-to-year regression results (for Food and Precision industries) are found in Table A1 in the Appendix. (Values of R^2 for all of those regressions exceed 0.99.) Complete year-to-year regression results for each of the 18 manufacturing industries are available on request from the authors.
[b]Numbers in parentheses represent the numbers of estimated values of r that are significant at 1% and 5% levels, respectively. For Chemicals, for example, the mean for 24 estimated values of r is 0.0206. Of these 24 estimates, 8 and 3 estimates are significant at 1% and 5% levels, respectively.
[c]Numbers in parentheses represent the numbers of estimated values of k that are significantly different from one ($k = 1$) at 1% and 5% levels, respectively. For Chemicals, for example, the mean for 24 estimated values of k is 1.046. Of these 24 estimates, 17 and 2 estimates are significantly different from one at 1% and 5% levels, respectively.
[d]All estimated coefficients are highly significant, with t-ratios of at least 40.

does not necessarily imply that there are small increasing returns to scale in these industries. All of these industries contain sub-industries which were classified at some time during the sample period as depressed industries by MITI and which were subject to the Law of Extraordinary Measures for the Stabilization of Specific Depressed Industries. It is likely that the excess capacity of large establishments in these depressed industries has tended to lower our estimates for k.[20] It is not possible to control for capacity utilization in depressed sub-industries, however, since no published data are available for capacity utilization.

20) During business downturns in Japan, it is typical that small establishments suffer from excess capacity much more than large establishments, resulting in an overestimation of scale elasticity. During the 1970s and the 1980s, however, when depressed industries were restructured, many of the small establishments in these depressed industries dropped out of our data sample. When a data sample has a relatively large number of large establishments with idle capacity, scale elasticity is usually underestimated.

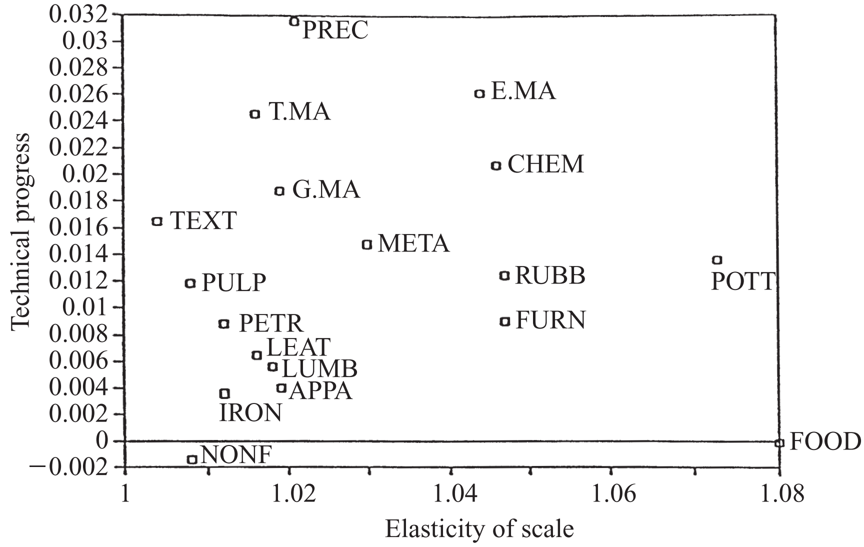

FIGURE 3. Technical progress *v*. Scale elasticity: Japanese manufacturing, 1964–88

Compared with our estimates for economies of scale, our estimates for r, the rate of technical progress, are smaller in magnitude, fluctuate more over time and often are not statistically significant for many industries. This is not, however, the case for Iron/Steel, Metal products, Machineries and Precision; some of these industries, particularly Steel, Machineries and Precision, became highly competitive in the global market during the sample period.

We see in Figure 3 that most of the eighteen manufacturing industries exhibit economies of scale. The Pottery and Food/Kindred industries have particularly high estimates (1.07 or higher) of elasticity of scale. The Precision and Transportation machinery industries, on the other hand, have modest scale elasticities (about 1.02) but very high rates of technical progress (2.5%–3% per year). The Electric machinery industry enjoys both a high rate of technical progress and a high elasticity of scale.

5. Contributions of scale economies and technical change to aggregate total factor productivity

We found in the previous section that economies of scale, more than technical progress, characterize production activities in many of the Japanese manufacturing industries at the establishment level. This does not imply, however, that gains in TFP at the aggregate industry level are due primarily to economies of scale rather than technical change.

In order to decompose TFP growth at the aggregate industry level into scale economy and technical change components, we first aggregate input indexes and predicted outputs over all establishments in each industry to derive aggregate input and output indexes at the industry level ($D(t)$ in equation (A15) and $X(t)$ in (A13a) in Appendix 2). The difference between the log of changes in the aggregate output index and in the input index (A14) is the industry TFP growth, which is then decomposed into the effects of scale economy and technical change (A18).[21)]

Using (A18) and our estimates for the elasticity of scale and the rate of technical change, we can decompose TFP gains at the industry level into scale economy and technical change components. (Year-by-year decomposition results for each of the eighteen manufacturing industries are available on request from the authors.) Gains in TFP and the decomposition of these gains into scale economy and technical change components averaged over 1964–88 are presented in Table 2. The estimated effects of scale economies on TFP gains are generally very small,[22)] compared with the large contribution of technical change. More than 90% of the gains in TFP during this

TABLE 2

DECOMPOSITION OF AVERAGE ANNUAL GAINS IN TFP 1964–88

Industry	TFP gains[a]	Technical change (E_2)[b]	Scale economies (E_1)[b]
Food/kindred products	0.02167	0.02072 (96%)	0.00095 (4)
Textiles	0.02281	0.02279 (100)	0.00002 (0)
Apparels	0.05027	0.05058 (101)	−0.00031 (−1)
Lumber/wood products	0.02313	0.02213 (96)	0.00100 (4)
Furniture/fixture	0.04135	0.04016 (97)	0.00119 (3)
Pulp/paper products	0.02961	0.02925 (99)	0.00036 (1)
Chemicals	0.02956	0.02874 (97)	0.00082 (3)
Petroleum/coal products	0.05489	0.05478 (100)	0.00011 (0)
Rubber/plastic products	0.04627	0.04207 (91)	0.00420 (9)
Leather/leather products	0.03996	0.03836 (96)	0.00160 (4)
Pottery/glass products	0.03192	0.03292 (103)	−0.00100 (−3)
Iron/steel	0.04480	0.03434 (77)	0.01046 (23)
Non-ferrous metals	0.03759	0.03691 (98)	0.00068 (2)
Metal products	0.03218	0.03224 (100)	−0.00006 (0)
General machinery	0.02707	0.02479 (92)	0.00228 (8)
Electrical machinery	0.05186	0.04667 (90)	0.00519 (10)
Transportation machinery	0.04332	0.03983 (92)	0.00349 (8)
Precision	0.04324	0.04163 (96)	0.01161 (4)

[a]These are based on (A14) in the text.
[b]These are based on (A18) in the text. Numbers in parentheses are percentage contributions.

21) A standard way to decompose TFP growth at an aggregate level into scale economy and technical change effects is to use:

$$\underbrace{\mathrm{d}\ln TFP/\mathrm{d}t}_{\text{TFP growth}} = \underbrace{\{1-(1/k)\}(\mathrm{d}\ln X/\mathrm{d}t)}_{\text{scale economy}} + \underbrace{\{-(\partial \ln C/\partial T)(\partial \ln T/\partial T\}}_{\text{technical change}},$$

where X is output, $C = C(p_1(t), \ldots, p_n(t), X(t), T(t))$ is a cost function and T is technology (often approximated by time t). The first term on the right-hand side vanishes under constant returns to scale ($k = 1$). In addition to the potential problem of non-existence of an aggregate production function when constant returns to scale cannot be assumed, we may have a serious multicollinearity problem, as we have argued in this paper, in estimating both scale economy and technical change parameters, k and $(\partial \ln C/\partial T)$, using aggregate data.

22) Our year-to-year decomposition results (not reported here) show that for Chemicals (1972–3), Rubber/plastics (1987–8), Pottery/glass products (1987–8) Iron/steel (1987–8), Electric machinery (1983–4) and Transportation machinery (1987–8), TFP gains exceeded 0.01 and the contributions of scale economies to TFP gains exceeded those of technical change. It is of interest to note that these TFP gains arising from scale economies can be traced back to the (identifiable) addition of new production capacity which became available in these industries in these specific calendar years.

period is due to technical change. These results also support the standard practice in macroeconometric modelling (e.g. Solow, 1957, and Jorgenson, Gollop and Fraumeni, 1987) that attributes gains in TFP at the aggregate level to technical progress by specifying an aggregate production function to be homogeneous of degree one.

6. Concluding remarks

We have presented an index number method for estimating the elasticity of scale and the rate of technical change using establishment data grouped by size and pooled over time. Our estimation method may be viewed as an extension of the approximation method proposed by Frisch (1965) for estimating scale elasticities when cross-sectional data exhibit serious multicollinearity problems. Our Japanese manufacturing data suffer from such multicollinearity problems.

Because of the small number of parameters to be estimated, and because of the explanatory variables included in our model, which are generally not highly correlated, our estimation results for Japanese manufacturing industries are quite stable and satisfactory.

We have found empirical evidence for the presence of substantial scale economies and modest technical progress for the period 1964–88 at the establishment level. Estimates for the sources of growth in aggregate (macro) TFP were calculated by aggregating estimation results derived at the establishment (micro) level. The change over time in aggregate TFP is explained primarily by technical progress.

These findings provide a justification for the standard practice of using a homogeneous production function of degree one (which attributes gains in TFP to technical change) in macroeconometric modelling.

Our findings that the effects of scale economies exist at the establishment level but disappear at the aggregate level imply, among other things, that establishment size does not adjust rapidly within the time period we consider. That is, large establishments do not grow at the expense of small establishments. It is the slowly changing technical level that explains most of the gains in aggregate TFP in the Japanese manufacturing sector. Using aggregate time series data for the period 1961–80, Tsurumi, Wago and Ilmakunnas (1986) also find that Japanese manufacturers spend relatively long periods of time (up to ten years) adjusting their production methods to incorporate new technological requirements. Their findings are consistent with our empirical results, suggesting the presence of slow but steady technical progress for the Japanese manufacturing sector.

Final version accepted 7 May 1996.

Appendix 1: Derivation of equations (29a–b)

A1. Derivation of equation (29a)

Equation (24) for $i = 1$ and $t = s$ with (28a) and (25) gives (29a) for $i = 1$:

$$\ln x_{1S} = \ln a(S) + k \ln q_{1S}, \text{ where } \ln q_{1S} \equiv 0. \tag{A1}$$

From (20) we also obtain

$$\ln x_{2S} - \ln x_{1S} = k(\ln Q_S^{1,2}). \tag{A2a}$$

$$\ln x_{3S} - \ln x_{2S} = k(\ln Q_S^{2,3}). \tag{A2b}$$

$$\ldots$$

$$\ln x_{iS} - \ln x_{i-1,S} = k(\ln Q_S^{i-1,i}). \tag{A2c}$$

$Q_S^{i-1,i}$ denotes the translog input quantity index for time period S.

Adding (A2a) to (A1) and using (25), we get

$$\ln x_{2S} = \ln a(S) + k \ln q_{2S}. \tag{A3}$$

Adding (A2b) to (A3) and using (25), we get

$$\ln x_{3S} = \ln a(S) + k \ln q_{3S}. \tag{A4}$$

Continuing in this manner, we obtain (29a) in general.

A2. Derivation of equation (29b)

From (24), we have

$$\ln x_{it} = \ln f(v_{it},\ t).$$

It follows that

$$\begin{aligned} \mathrm{d} \ln x_{it}/\mathrm{d}t &= \Sigma_j(\partial \ln f/\partial \ln v_{itj})(\partial \ln v_{itj}/\partial t) + (\partial \ln t/\partial t) \\ &= k\Sigma_j w_{ij}(\partial \ln v_{itj}/\partial t) + (\partial \ln f/\partial t) \\ &= k(\partial \ln q_{it}/\partial t) + (\partial \ln f/\partial t), \end{aligned} \tag{A5}$$

where $v_{it,j}$ denotes the jth component of vector v_{it}, w_i is the cost share vector for the ith establishment and $w_{i,j}$ is the jth component of w_i given by (22). Cost minimization and the assumption that $f(v_{it},\ t)$ is homogeneous of degree k in v_{it} gives the second equality. The definition of a Divisia quantity index provides the last equality.

Assuming for simplicity that time t is continuous, we integrate (A5) for $i = 1$ from $t = S$ to $t = S + 1$ to get

$$\ln x_{1,S+1} - \ln x_{1,S} = k(\ln q_{1,S+1} - \ln q_{1,S}) + r(S),$$

or

$$\ln x_{1,S+1} = \ln x_{1,S} + k \ln q_{1,S+1} + r(S), \tag{A6}$$

where $\ln q_{1,S} = 0$ and $r(S)$ is the definite integral of $\partial \ln f/\partial t$ from S to $S + 1$; $r(S)$ is assumed to be a function of time only, representing the rate of technical progress between time periods S and $S + 1$.

From (28b), we also have

$$\ln a(S + 1) = \ln x_{1S} + r(S). \tag{A7}$$

From (A6) and (A7), we have

$$\ln x_{1,S+1} = \ln a(S + 1) + k \ln q_{1,S+1}, \tag{A8}$$

which is (29b) for $i = 1$.

From (20), for $i = 1, 2, 3, \ldots$ and for time period $S + 1$, we get

$$\ln x_{2,S+1} - \ln x_{1,S+1} = k(\ln Q^{1,2}_{S+1}), \tag{A9a}$$

$$\ln x_{3,S+1} - \ln x_{2,S+1} = k(\ln Q^{2,3}_{S+1}), \tag{A9b}$$

$$\ldots$$

$$\ln x_{i,S+1} - \ln x_{i-1,S+1} = k(\ln Q^{i-1,i}_{S+1}). \tag{A9c}$$

Adding (A9a) to (A8) and using (27), we get

$$\ln x_{2,S+1} = \ln a(S+1) + k \ln q_{1,S+1} + k \ln Q^{1,2}_{S+1}$$

or

$$\ln x_{2,S+1} = \ln a(S+1) + k \ln q_{2,S+1}, \tag{A10}$$

which is (29b) for $i = 2$. Adding (A9b) to (A10) and using (27), we get

$$\ln x_{3,S+1} = \ln a(S+1) + k \ln q_{3,S+1}, \tag{A11}$$

which is (29b) for $i = 3$. Continuing in a similar manner, (29b) is shown to hold for any $i (i = 1, 2, \ldots, I)$ in general.

Appendix 2: Decomposition of total factor productivity growth into the effects arising from scale economies and technical change

A standard way to measure the change in TFP is

$$\ln \frac{TFP(t+1)}{TFP(t)} = \ln \frac{X(t+1)}{X(t)} - \sum_{j=1}^{n} \frac{1}{2}[w_j(t) + w_j(t+1)] \ln \frac{v_j(t+1)}{v_j(t)}, \tag{A12}$$

where aggregate output X and aggregate production input indices v_j are defined by

$$X(t) = N(t) \int_0^\infty x \, \mathrm{d}F_t(x), \tag{A13a}$$

and

$$v_j(t) = N(t) \int_0^\infty v_j d\phi_t(v_j), \tag{A13b}$$

and where $w_j(t)$ denotes the scalar cost share for the jth aggregate production input. (Recall that in (26) w_i denotes the cost share vector for the ith establishment.) In (A13a, b) $N(t)$ is the total number of establishments in period t; x and v_j are random variables representing, respectively, the output and the jth production input for a representative establishment; and $F_t(x)$ and $\phi_t(v_j)$ denote the distribution functions for random variables x and v_j in period t. (Strictly speaking, (A12) is a valid measure of TFP growth under the assumption of constant returns to scale. This assumption seems satisfied at the industry level in this study. See Chan and Mountain (1983) for a modification of (A12) when constant returns to scale cannot be assumed.)

Another way to compute the change in TFP, which is more consistent with our model for individual establishments, is the following:

$$\ln \frac{\Phi(t+1)}{\Phi(t)} = \ln \frac{X(t+1)}{X(t)} - \ln \frac{D(t+1)}{D(t)}, \tag{A14}$$

where $\Phi(t+1)/\Phi(t)$ denotes a new measure for the change in TFP defined by (A14) and $D(t)$ represents an aggregate production input index defined by

$$D(t) = N(t)\int_0^\infty q \, \mathrm{d}G_t(q). \tag{A15}$$

In (A15) q is a random variable representing a production input index for an establishment and $G_t(q)$ is the distribution function for q in period t.

From (29a), we have

$$x = a(t)q^k. \tag{A16}$$

Substituting (A16) and (A15) into (A14), we obtain

$$\ln \frac{\Phi(t+1)}{\Phi(t)} = \ln \frac{N(t+1)\int_0^\infty a(t+1)q^k \, \mathrm{d}G_{t+1}(q)}{N(t)\int_0^\infty a(t)q^k \, \mathrm{d}G_t(q)} - \ln \frac{N(t+1)\int_0^\infty q \, \mathrm{d}G_{t+1}(q)}{N(t)\int_0^\infty q \, \mathrm{d}G_t(q)}. \tag{A17}$$

Rewriting (A17), we obtain

$$\ln \frac{\Phi(t+1)}{\Phi(t)} = E_1 + E_2, \tag{A18}$$

where

$$E_1 = \ln \frac{N(t+1)\int_0^\infty a(t+1)q^k \, \mathrm{d}G_{t+1}(q)}{N(t)\int_0^\infty a(t+1)q^k \, \mathrm{d}G_t(q)} - \ln \frac{N(t+1)\int_0^\infty q \, \mathrm{d}G_{t+1}(q)}{N(t)\int_0^\infty q \, \mathrm{d}G_t(q)}, \tag{A19}$$

and

$$E_2 = \ln \frac{N(t)\int_0^\infty a(t+1)q^k \, \mathrm{d}G_t(q)}{N(t)\int_0^\infty a(t)q^k \, \mathrm{d}G_t(q)} - \ln \frac{N(t)\int_0^\infty q \, \mathrm{d}G_t(q)}{N(t)\int_0^\infty q \, \mathrm{d}G_t(q)}. \tag{A20}$$

E_1 measures the gains in TFP arising from the economies of scale, while E_2 measures the gains in TFP arising from technical progress. Both E_1 and E_2 can be evaluated using establishment data and estimated parameters $a(t)$ and k, both of which are statistically highly significant.

Finally, it is of interest to compare TFP growth calculated by (A12) and (A14). The estimates we get from (A12) and (A14) for TFP growth are quite close. Absolute deviations between these two types of estimate range from 0.06% to 2.5% (Nakajima *et al.*, 1993). In our present application, the calculation based on (A14) leads to a natural decomposition of TFP growth given by (A18). Yet, because our estimates based on (A14) are very close to the estimates based on the standard formula (A12), we conclude that our decomposition results reported in Table 2 also hold for the TFP growth measured by (A12).

Appendix 3: Some GLS estimates

Table A1 presents year-to-year GLS estimates for the Food and Precision industries.

Appendix 4: Data

The Census of Manufacturing by Industry published annually by the Ministry of International Trade and Industry gives, for each of the seven size groups distinguished by the number of employees, the average figures for: the number of employees, labour compensation, the cost of intermediate input, the value of output, investment expenditure, depreciation, and the book value of capital stock. Because of the new additions of establishments as well as the closures and mergers of existing establishments, the numbers of establishments included in the seven establishment groups do change over time. The manufacturing industries covered by the Census are: Food/kindred products, Textiles, Apparels, Lumber/wood products, Furniture/fixture, Pulp/paper, Printing, Chemicals, Petroleum/coal products, Rubber/plastic products, Leather/leather products, Pottery/glass products, Iron/steel, Non-ferrous metals, Metal products, General machinery, Electrical machinery, Transportation machinery, Precision and Other.

Table A1

Year-to-year Estimates for the Rate of Technical Progress and the Elasticity of Scale: Food Products and Precision, 1964–88[a]

	Food/kindred products						Precision					
	ln $a(s)$	r	k	F^{c}	R^2	Sample size	ln $a(s)$	r	k	P^{c}	R^2	Sample size
1964–65	3.33	0.0074	1.047	5.0	0.9966	18	2.95	0.0143	1.008	0.9	0.9993	18
	40.3[b]	0.60	44.2				95.6	1.05	101.2			
1965–66	3.34	0.0253	1.056	6.5	0.9967	18	2.98	0.0586	1.015	3.0	0.9993	18
	40.7	1.74	43.3				94.4	4.33	103.1			
1966–67	3.51	−0.0045	1.042	15.7	0.9977	18	3.11	0.0494	1.018	5.3	0.9995	18
	59.8	−0.20	87.2				112.4	4.80	116.3			
1967–68	4.33	−0.0195	1.064	13.0	0.9980	16	3.88	0.0248	1.020	44.0	0.9997	16
	73.6	−1.48	53.7				220.6	1.75	292.8			
1968–69	4.51	0.0214	1.060	9.4	0.9977	16	4.11	0.0524	1.018	40.3	0.9999	16
	71.6	1.94	47.8				495.1	6.21	320.4			
1969–70	4.58	0.0401	1.092	35.2	0.9981	16	4.25	−0.0167	1.018	13.6	0.0009	16
	82.9	2.45	61.6				288.1	−2.63	186.4			
1970–71	4.71	−0.0064	1.082	25.9	0.9985	16	4.37	0.0249	1.015	8.8	0.9997	16
	91.1	−0.52	58.8				298.4	2.06	179.7			
1971–72	4.87	0.0164	1.056	14.1	0.9979	16	4.94	0.0157	1.029	21.4	0.9994	14
	93.0	0.86	62.3				180.9	0.69	139.1			
1972–73	4.94	0.0442	1.094	31.6	0.9980	16	5.03	0.0996	1.024	53.2	0.9996	14
	87.9	2.06	57.4				399.0	4.02	265.2			
1973–74	5.23	−0.0256	1.070	51.0	0.9989	16	4.66	0.0556	1.011	2.1	0.9994	16
	129.4	−1.47	95.9				188.7	3.10	110.8			
1974–75	5.41	0.0121	1.060	40.2	0.9993	16	4.90	0.0003	1.035	22.2	0.9995	16
	166.4	0.75	97.6				188.2	0.02	121.6			
1975–76	6.22	0.0220	1.063	43.3	0.9987	14	5.37	0.0554	1.031	36.7	0.9995	14
	230.8	0.95	94.5				256.0	1.91	173.7			
1976–77	6.37	−0.0097	1.082	42.5	0.9987	14	5.46	0.0647	1.021	7.3	0.9996	14
	164.5	−0.95	73.1				321.2	4.34	113.1			

continued overleaf

Table A1
(*continued*)

	Food/kindred products						Precision					
	ln $a(s)$	r	k	F^c	R^2	Sample size	ln $a(s)$	r	k	P^c	R^2	Sample size
1977–78	6.49	0.0114	1.074	30.8	0.9988	14	5.62	0.0259	1.016	9.8	0.9997	14
	176.9	0.79	69.7				261.1	2.82	168.3			
1978–79	6.49	0.0108	1.114	43.0	0.9984	14	5.65	0.0511	1.019	9.8	0.9997	14
	145.8	0.90	55.0				304.6	4.86	143.7			
1979–80	6.49	−0.0044	1.119	43.7	0.9984	14	5.73	0.0996	1.018	5.3	0.9996	14
	132.5	−0.33	53.1				235.0	14.72	112.3			
1980–81	6.67	−0.0315	1.118	45.4	0.9976	14	5.89	0.0406	1.015	2.2	0.9994	14
	150.3	−1.28	45.3				192.0	3.13	86.9			
1981–82	6.74	0.0223	1.097	19.3	0.9953	14	5.90	0.0374	1.021	5.3	0.9994	14
	114.0	1.29	42.2				193.9	1.97	96.3			
1982–83	6.82	−0.0108	1.084	13.8	0.9946	14	5.89	0.0170	1.033	15.2	0.9995	14
	100.7	−1.06	40.6				214.5	1.08	105.5			
1983–84	6.81	−0.0094	1.108	14.6	0.9959	14	5.91	0.0175	1.034	18.9	0.9996	14
	96.5	−0.54	33.3				234.8	1.30	114.0			
1984–85	6.89	−0.1115	1.087	38.4	0.9972	14	6.00	−0.0836	1.030	12.7	0.9824	14
	146.1	−2.27	65.8				228.0	−0.58	104.6			
1985–86	6.64	−0.0090	1.089	29.9	0.9985	14	6.09	0.0127	1.019	1.0	0.9983	12
	156.8	−1.47	56.8				117.0	0.85	43.3			
1986–87	6.65	−0.0076	1.076	36.4	0.9989	14	6.09	−0.0163	1.021	1.7	0.9985	12
	196.8	−0.39	73.1				132.0	−2.86	51.6			
1987–88	6.64	0.0153	1.075	33.9	0.9990	14	6.11	0.0571	1.023	4.6	0.9989	12
	212.8	1.06	70.9				175.1	1.38	77.3			

[a]These estimates are based on equation (30). Complete estimation results for 18 manufacturing industries are available from the authors on request.
[b]Numbers under regression coefficients estimates denote t-statistics for these estimates.
[c]F denotes the F-statistic for testing the null hypothesis $H_0:\ k = 1$.

REFERENCES

Allen, R. G. D. (1975) *Index Numbers in Theory and Applications*, London: Macmillan.

Banker, R. D., A. Charnes, W. W. Cooper and A. Maindiratta (1988) "A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations form a Known Technology", in A. Dogramaci and R. Färe (eds.), *Applications of Modern Production Theory: Efficiency and Productivity*, New York: Kluwer, pp. 33–55.

Baumol, W. J., J. C. Panzar and R. D. Willig (1982) *Contestable Markets and the Theory of Industry Structure*, New York: Harcourt Brace Jovanovich.

Berndt, E. R. (1991) *The Practice of Econometrics: Classic and Contemporary*, Reading, Mass.: Addison Wesley.

—— and M. S. Khaled (1979) "Parametric Productivity Measurement and Choice among Flexible Functional Forms", *Journal of Political Economy*, vol. 87, pp. 1220–45.

Bortkiewicz, L. von (1922) "Zweck und Struktur einer Preisindexah", *Nordisk Statistiks Tidskrift*, vol. 5, pp. 1–29.

Caves, R. and D. Barton (1990) *Efficiency in the US Manufacturing Industries*, Cambridge, Mass.: MIT Press.

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

Chan, N. W. L. and D. C. Mountain (1983) "Economies of Scale and the Törnqvist Discrete Measure of Productivity Growth", *Review of Economics and Statistics*, vol. 70, pp. 663–7.

Chavas, J.-P. and T. Cox (1990) "A Non-Parametric Analysis of Productivity: the Case of US and Japanese Manufacturing", *American Economic Review*, vol. 80, pp. 450–64.

Christensen, L. R., D. W. Jorgenson and L. J. Lau (1973) "Transcendental Logarithmic Production Frontiers", *Review of Economics and Statistics*, vol. 60, pp, 28–45.

Denison, E. F. (1974) *Accounting for United States Economic Growth 1929–1969*, Washington, DC: Brookings Institution.

Diewert, W. E. (1976) "Exact and Superlative Index Numbers", *Journal of Econometrics*, vol. 4, pp. 221–56.

Fisher, I. (1922) *The Making of Index Numbers*, New York: Macmillan.

Frisch, R. (1965) *Theory of Production*, Dordrecht: D. Reidel.

Giles, D. E. A. and N. S. Wyatt (1992) "Economies of Scale in the New Zealand Electricity Distribution", in P. C. B. Phillips (ed.), *Models, Methods and Applications of Econometrics*, Oxford: Blackwell.

Harris, R. G. and D. Cox (1984) *Trade, Industrial Policy, and Canadian Manufacturing*, Ontario: Ontario Economic Council.

Helpman, E. and P. R. Krugman (1985) *Increasing Returns, Imperfect Competition, and International Trade*, Cambridge, Mass.: MIT Press.

Jorgenson, D. W., F. M. Gollop and B. M. Fraumeni (1987) *Productivity and US Economic Growth*, Cambridge, Mass.: Harvard University Press.

—— and Z. Griliches (1967) "The Explanation of Productivity Change", *Review of Economic Studies*, vol. 34, pp. 349–83.

Kemp, M. C. (1969) Chapter 8 in *The Pure Theory of International Trade and Investment*, Englewood Cliffs, NJ: Prentice-Hall, pp. 154–80.

Komiya, R. (1962) "Technological Progress and Production Function in the United States Steam Power Industry", *Review of Economics and Statistics*, vol. 44, pp. 156–66.

Kuroda, M., K. Yoshioka and D. W. Jorgenson (1984) "Relative Price Changes and Biases of Technical Change in Japan", *Economic Studies Quarterly*, vol. 35, pp. 116–38.

Nakajima, T., M. Nakamura and K. Yoshioka (1993) "An Index Number Method for Estimating Scale Economies and Technical Progress Using Time Series of Cross Section Data: Sources of Total Factor Productivity for Japanese Manufacturing 1964–1988", paper presented at the Econometric Society European Meeting, Uppsala, Sweden.

Oi, W. Y. (1983) "The Fixed Employment Costs of Specialized Labor", in J. E. Triplett (ed.), *The Management of Labor Cost*, Chicago: University of Chicago Press, pp. 63–122.

Ozaki, I. (1969) "Economies of Scale and Input Output Coefficients", in A. Carter and A. Brody (eds.), *Applications of Input–Output Analysis*, Amsterdam: North-Holland, pp. 280–302.

—— (1976) "The Effects of Technological Change on the Economic Growth of Japan", in K. Polenske and J. Skolka (eds.), *Advances in Input–Ouput Analysis*, Cambridge, Mass.: Ballinger, pp. 93–111.

Solow, R. M. (1957) "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics*, vol. 39, pp. 312–20.

—— (1959) "Investment and Economic Growth", *Productivity Measurement Review*, vol. 21, pp. 28–53.

Theil, H. (1965) "The Information Approach to Demand Analysis", *Econometrica*, vol. 33, pp. 288–93.

Tinbergen, J. (1945) *International Economic Cooperation*, Amsterdam: Elsevier.

Törnqvist, L. (1936) "The Bank of Finland's Comsumption Price Index", *Bank of Finland Monthly Bulletin*, vol. 10, pp. 23–55.

Tsurumi, H., H. Wago and P. Ilmakunnas (1986) "Gradual Switching Multivariate Regression Models with Stochastic Cross-Equational Constraints and an Application to the KLEM Translog Production Model", *Journal of Econometrics*, vol. 31, pp. 235–53.