An Approximate Divisia Index of Total Factor Productivity

Spencer Star; Robert E. Hall

Econometrica, Vol. 44, No. 2. (Mar., 1976), pp. 257-263.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197603%2944%3A2%3C257%3AAADIOT%3E2.0.CO%3B2-B

Econometrica is currently published by The Econometric Society.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/econosoc.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Mon Jul 31 22:54:49 2006

Econometrica, Vol. 44, No. 2 (March, 1976)

AN APPROXIMATE DIVISIA INDEX OF TOTAL FACTOR PRODUCTIVITY

BY SPENCER STAR AND ROBERT E. HALL

1. INTRODUCTION

THE DIVISIA INDEX of factor inputs is known to be, under certain conditions, the appropriate index of inputs for the measurement of total factor productivity. A firm theoretical foundation for its use has been provided by Solow [**9**], Diewert [**2**], Domar [**3**], Hulten [**5**], Jorgensen and Griliches [**6**], and Richter [**8**]. The index derived and discussed by these authors, with the exception of Diewert, treats time as continuous; the index is an integral over time. Empirical work has proceeded by calculating discrete annual changes, which are then used to obtain an approximation to the continuous index. Even when the object of the calculation is the estimation of total factor productivity over a very long time, annual data on output, inputs, and factor shares are used. However, detailed data are frequently not available annually. For example, much of the detailed data on labor inputs are available only from the *Census of Population*, which is taken every 10 years.

The question studied in this paper is whether or not a reasonable approximation to the continuous Divisia index can be calculated using data from only the beginning and end of a long period of time. Our answer is favorable. We derive a suitable approximation, calculate bounds on its errors, and suggest that in the usual cases these errors are likely to be small. We then calculate the growth in total factor productivity in the American economy from 1909 to 1958 using the conventional method based on annual changes and compare it to our approximation based on data for 1909 and 1958 only. The two calculations give almost exactly the same result. We conclude that in most cases the data for intervening years are superfluous. This conclusion makes it possible to make detailed and accurate calculations of the growth in total factor productivity from one decennial census to the next as has been recently done by Star [**10**].

2. DERIVATION OF THE DIVISIA INDEX AND THE APPROXIMATION

The usual derivation of the Divisia index runs as follows: Assume (i) that a continuous, twice differentiable production function with one output and n inputs exists, (ii) that it is homogeneous of degree one, and (iii) that all inputs are paid their marginal revenue products. The production function is written as:

$$Y(t) = F(x_1(t), x_2(t), \dots, x_n(t); t) \tag{2.1}$$

where Y is output, x_i is the ith input, and t is time. By taking the logarithm of the production function and differentiating with respect to time we get:

$$\dot{Y}(t)/Y(t) = \left[\sum_1^n F_i(t)\dot{x}_i(t)/F(t)\right] + F_t(t)/F(t) \tag{2.2}$$

where a dot over a symbol means the derivative with respect to time and $F_i = \partial F/\partial X_i$, $F_t = \partial F/\partial t$. We rearrange to get:

$$(2.3)\qquad F_t(t)/F(t) = \dot{Y}(t)/Y(t) - \sum_1^n [F_i(t)x_i(t)/F(t)][\dot{x}_i(t)/x_i(t)].$$

Assumptions (ii) and (iii) allow us to set $F_i x_i/F = \beta_i$, the share of receipts going to the ith input. We have the following partial differential equation defining the growth rate of the index:

$$(2.4)\qquad F_t(t)/F(t) = \dot{Y}(t)/Y(t) - \sum_1^n [\beta_i(t)\dot{x}_i(t)/x_i(t)].$$

This equation is set equal to $\dot{A}/A$, where $A(t)$ is the Divisia index of the residual. It is rewritten as:

$$(2.5)\qquad \dot{A}(t)/A(t) = \dot{Y}(t)/Y(t) - \sum_1^n [\beta_i(t)\dot{x}_i(t)/x_i(t)].$$

The apparent shift of the function over time, $\dot{A}/A$, is the growth of the Divisia index of total factor productivity.

We can integrate equation (2.5) to get the index of total factor productivity.

$$(2.6)\qquad A(T)/A(0) = [Y(T)/Y(0)] \exp\left(-\sum_1^n \int_0^T \beta_i(t)\dot{x}_i(t)/x_i(t)\, dt\right).$$

We normalize by setting $A(0) = 1$ and can thus speak of $A(T)$ as the index of productivity at time T. If $A(2) = 1.06$, productivity has grown six per cent over two periods or at a rate (roughly) of three per cent per period.

In empirical work (for example, Solow [**11**], Massell [7], and Jorgenson and Griliches [**6**]) equation (2.6) is approximated by increasing the value of the index by its annual growth, one year at a time, until one has the value $A(T)$. The results are often stated in terms of $[A(T)/A(0)]^{1/T}$, the average annual growth rate of total factor productivity. A great deal of effort often goes into the preparation of annual data, especially when a detailed disaggregation of factors is used. It appears that most of this work can be avoided by using the approximation that we will now derive.

We begin by asking what constant values for the shares (the β_i's) give the same index as the true fluctuating shares. Consider the time interval $[0, T]$. The constant weight $\tilde{\beta}_i$ must satisfy the following relation:

$$(2.7)\qquad \int_0^T \tilde{\beta}_i \frac{\dot{x}_i(t)}{x_i(t)}\, dt = \int_0^T \beta_i(t)\frac{\dot{x}_i(t)}{x_i(t)}\, dt \qquad (i = 1, 2, \ldots n).$$

The use of a constant share enables us to integrate the left-hand side of (2.7) to get:

$$(2.8)\qquad \tilde{\beta}_i \log [x_i(T)/x_i(0)] = \int_0^T [\beta_i(t)\dot{x}(t)/x_i(t)]\, dt \qquad (i = 1, \ldots n).$$

If we substitute (2.8) into (2.6), we have:

$$A(T) = [Y(T)/Y(0)] \Big/ \prod_{n}^{1} [x_i(T)/x_i(0)]\tilde{\beta}_i. \tag{2.9}$$

Thus, by using certain constant shares, we can calculate the value of the Divisia index at time T using data only from the periods 0 and T. The next step is to show how the constant shares are related to the growth rates of inputs and the varying shares.

Let $G_i(t)$ be the rate of growth of the ith input at time t:

$$G_i(t) = \dot{x}_i(t)/x_i(t). \tag{2.10}$$

Let $\bar{G}_i$ be the average growth rate over the period $[0, T]$:

$$\bar{G}_i = \frac{1}{T}\int_0^T \frac{\dot{x}_i(t)}{x_i(t)}\,dt = \frac{1}{T}\log[x_i(T)/x_i(0)]. \tag{2.11}$$

Substitute from (2.11) into (2.8) to get:

$$T\tilde{\beta}_i\bar{G}_i = \int_0^T \beta_i(t)G_i(t)\,dt. \tag{2.12}$$

Consequently:

$$\tilde{\beta}_i = \frac{1}{T}\int_0^T \beta_i(t)\frac{G_i(t)}{\bar{G}_i}\,dt. \tag{2.13}$$

Thus, the appropriate constant share $\tilde{\beta}_i$ is the weighted average over time of the fluctuating share, where the weight is the ratio of the rate of growth of the input to its average rate of growth over the period. In the special case of a factor growing at a constant rate, $G_i(t)/\bar{G}_i$ is always one, and $\tilde{\beta}_i$ is just the mean over time of $\beta_i(t)$.

Our proposal is to approximate the true $\tilde{\beta}_i$ by the simple average of $\beta_i(t)$ at the beginning and end of the period:

$$\hat{\beta}_i = \tfrac{1}{2}[\beta_i(0) + \beta_i(T)]. \tag{2.14}$$

If we put this into equation (2.9), we have our approximation to the Divisia index:

$$\hat{A}(T) = \frac{Y(T)/Y(0)}{\prod_{i=1}^{n}\left[\dfrac{x_i(T)}{x_i(0)}\right]^{\hat{\beta}_i}} \tag{2.15}$$

The only source of error is the use of the approximate shares, $\hat{\beta}_i$, in place of the weighted averages, $\tilde{\beta}_i$. The success of this approximation derives from the stability of factor shares over time, which makes $\hat{\beta}_i$ a very good estimate of $\tilde{\beta}_i$.

The index of inputs in formula (2.15) is well known in the literature on price and quantity indexes. Diewert [**2**, p. 17] has noted it corresponds to Irving Fisher's price index number 124 [**4**], and it has been advocated as a quantity index by

Törnqvist [**12**], Theil [**11**], and Hulten [**5**]. It has been referred to variously as the Divisia index, the discrete Divisia index, and the Fisher-Törnqvist-Theil index. In our framework the value of the index at time T is exactly the same as the (normalized) value of output at time T if and only if the production function is a homogeneous translog function.[1] Therefore, we propose the index be called the homogeneous translog index.

The primary concern of those considering the homogeneous translog index has been to show it is a good approximation to some aggregate of the inputs.[2] In contrast to this concern with the approximation properties of the index for small changes, we are interested in the large changes that take place for periods of 10 years or longer. For this purpose a rather different analysis of the approximation error is required and is the subject of the next section.

3. THE APPROXIMATION ERROR

We define the relative error per time period, ε, as:

$$\varepsilon = \frac{1}{T}(\log \hat{A}(T) - \log A(T)). \tag{3.1}$$

If ε is 0.01, for example, the approximate index is growing one per cent per year faster than the true Divisia index. Equation (3.1) can be written (see equation (2.6)) as:

$$\varepsilon = \frac{1}{T}\int_0^T \sum_1^n [\beta_i(t) - \hat{\beta}_i]G_i(t)\,dt. \tag{3.2}$$

Let $\bar{\beta}_i$ be the unweighted average of $\beta_i(t)$:

$$\bar{\beta}_i = \frac{1}{T}\int_0^T \beta_i(t)\,dt. \tag{3.3}$$

From the definition of $\bar{G}_i$ we have:

$$0 = \frac{1}{T}\int_0^T \sum_i^n \bar{\beta}_i G_i(t)\,dt + \sum_1^n \bar{\beta}_i \bar{G}_i, \tag{3.4}$$

and

$$0 = \frac{1}{T}\int_0^T \sum_1^n \hat{\beta}_i G_i(t)\,dt - \sum_1^n \hat{\beta}_i \bar{G}_i. \tag{3.5}$$

Further, from the definition of $\bar{\beta}_i$,

$$0 = -\frac{1}{T}\int_0^T \sum_1^n (\beta_i(t) - \bar{\beta}_i)\bar{G}_i\,dt. \tag{3.6}$$

[1] See Diewert [**2**] for the proof of the general case and the conditions under which it holds.

[2] Theil [**11**, Ch. 6] shows its error involves the third and higher powers of the log-changes of the inputs. Diewert [**2**] shows that under certain conditions it is exact for a homogeneous translog production function; therefore, it is also a second order approximation to an arbitrary linear homogeneous production function, since that is one of the properties of the homogeneous translog function.

By adding together equations (3.2), (3.4), (3.5), and (3.6), we have:

$$(3.7) \qquad \varepsilon = \frac{1}{T}\int_0^T \sum_1^n [\beta_i(t) - \bar{\beta}_i][G_i(t) - \bar{G}_i]\,dt - \sum_1^n (\hat{\beta}_i - \bar{\beta}_i)\bar{G}_i.$$

We define the covariance across inputs as

$$(3.8) \qquad \operatorname{cov}(v, w) = \frac{1}{n}\sum_1^n (v_i - \mu)w_i$$

where μ is the average across all inputs:

$$(3.9) \qquad \mu = \frac{1}{n}\sum_1^n v_i.$$

Then,

$$(3.10) \qquad \varepsilon = \frac{n}{T}\int_0^T \operatorname{cov}(\beta(t) - \bar{\beta}, G(t) - \bar{G})\,dt - n\operatorname{cov}(\hat{\beta} - \bar{\beta}, \bar{G});$$

note that the means of $\beta(t) - \bar{\beta}$ and of $\hat{\beta} - \bar{\beta}$ are zero.

This formula partitions the error into two components. The first is positive if deviations from the average shares are positively correlated with deviations from the average factor growth rates, while the second is positive if over-estimates of the average share are negatively correlated with the average factor growth rates. The homogeneous translog (HT) index is exact when the two terms in (3.10) are offsetting or are zero. The Cobb-Douglas production function is a specific type of HT function; if the underlying function is Cobb-Douglas, the shares are constant and the error is zero—the HT index is exact.

We can get a rough idea of the possible magnitude of ε by using the Cauchy-Schwartz inequality:

$$(3.11) \qquad |\varepsilon| \leqslant \frac{n}{T}\int_0^T S(\beta(t) - \bar{\beta})S(G(t) - \bar{G})\,dt + nS(\hat{\beta} - \bar{\beta})S(\bar{G})$$

where $S(v)$ is the standard deviation of v:

$$(3.12) \qquad S(v) = \frac{1}{n}\sqrt{\sum (v_i - \mu)^2}.$$

Equality occurs only when the variables are perfectly correlated. Now if there are four factors ($n = 4$), the standard deviation of factor shares from their historical averages might be one percentage point ($S[\beta(t) - \bar{\beta}] = 0.01$), the standard deviation of growth rates from their averages might be two percentage points ($S[G(t) - \bar{G}] = 0.02$), the standard deviation of the error in estimating $\bar{\beta}$ might be 0.5 percentage points ($S[\hat{\beta} - \bar{\beta}] = 0.005$), and the standard deviation of average growth rates might be one percentage point ($S(\bar{G}) = 0.01$). Then:

$$(3.13) \qquad |\varepsilon| \leqslant 0.0012.$$

The largest possible error in the growth rate of productivity is 0.12 per cent of the actual growth per year. In most cases, the correlation between factor shares and growth rates will be weak, and the actual error will be much smaller than this bound.

4. APPLICATION TO PRODUCTIVITY MEASUREMENT

We now measure total factor productivity in the United States from 1909 to 1958. Table I compares the (almost) continuous Divisia index to two discrete indexes: the HT index and a naïve Laspeyres index. Since data are available only

TABLE I
COMPARISON OF DIFFERENT MEASURES OF THE GROWTH OF TOTAL FACTOR PRODUCTIVITY IN THE UNITED STATES 1910 TO 1958[a]

	Almost Continuous Divisia Index	Homogeneous Translog Index	Naïve Laspeyres Index
A (1958)	1.34615	1.33585	1.90986
Average annual growth of $A(t)$	0.0061	0.0059	0.0133
ε	0.0	-0.0002	$+0.0071$

[a] Output from [1, p. 146, column 2], inputs from [1, p. 141, columns 3 and 5], shares from [13, series 50–60, p. 141].

for discrete periods, a truly continuous Divisia index cannot be estimated. For the purpose of comparison, annual data from 1909 to 1958 have been used to approximate the continuous index, while data only from the endpoints 1909 and 1958 have been used for the discrete indexes. The naïve Laspeyres index is as follows:

$$\hat{A}(T) = 1 + \frac{Y(T) - Y(0)}{Y(0)} - \sum_{i=1}^{n} \beta_i(0)\frac{X_i(T) - X_i(0)}{X_i(0)}. \tag{4.1}$$

The Laspeyres approximation is often used with annual data, the base period being changed every 5 or 10 years. We have exaggerated the effect of using this index (and the HT index) by considering endpoints 50 years apart. The HT index understates the growth of productivity by 0.02 per cent per year, probably well within the margin of error of the data. The naïve Laspeyres overstates the annual productivity growth by more than 100 per cent.

We conclude that quite accurate estimates of the growth in total factor productivity can be calculated from data at the beginning and end of a period, even when the period is extremely long. The investigator must be careful, however, to choose a well-behaved approximation, especially over periods of more than a few years. For this purpose our approximation seems well suited.

University of British Columbia
and
Massachusetts Institute of Technology

Manuscript received October, 1973; revision received November, 1974.

REFERENCES

[1] DENISON, E. F.: *The Sources of Economic Growth in the United States*. Supplementary Paper No. 13. New York: Committee for Economic Development, 1962.

[2] DIEWERT, W. E.: "Homogeneous Weak Separability and Exact Index Numbers," Technical Report No. 122, Institute for Mathematical Studies in the Social Sciences, Stanford, California, January, 1974.

[3] DOMAR, E.: "On the Measurement of Technological Change," *Economic Journal*, 71 (1961), 709–729.

[4] FISHER, F.: *The Making of Index Numbers*. Boston: Houghton Mifflin, 1922.

[5] HULTEN, C. R.: "Divisia Index Numbers," *Econometrica*, 41 (1973), 1017–1025.

[6] JORGENSON, D., AND Z. GRILICHES: "The Explanation of Productivity Change," *Review of Economic Studies*, 34 (1967), 249–283.

[7] MASSELL, B. F.: "Capital Formation and Technical Change in U.S. Manufacturing," *Review of Economics and Statistics*, 42 (1960), 182–188.

[8] RICHTER, M. K.: "Invariance Axioms and Economic Indexes," *Econometrica*, 34 (1966), 739–755.

[9] SOLOW, R. M.: "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 39 (1957), 312–320.

[10] STAR, S.: "Accounting for the Growth of Output," *American Economic Review*, 64 (1974), 123–134.

[11] THEIL, H.: *Economic and Information Theory*. Chicago: Rand McNally, 1967.

[12] TÖRNQVIST, L.: "The Bank of Finland's Consumption Price Index," *Bank of Finland Monthly Bulletin*, 10 (1936), 1–8.

[13] UNITED STATES DEPARTMENT OF COMMERCE: *Historical Statistics of the United States*. Washington, D.C.: Government Printing Office, 1960.