THE AGGREGATE RATE OF RETURN TO TECHNICAL INNOVATION

By

Earl A. Thompson

Discussion Paper Number 63
July 1975

THE AGGREGATE RATE OF RETURN TO TECHNICAL INNOVATION

By

Earl A. Thompson*/

University of California at Los Angeles

The dominant economics literature presumes that more technological innovation is a good thing, despite the fact that basic economic principles tell us that it is a good thing only if it justifies the cost. This presumption has been exceptionally strong in recent years as estimates of technological progress by the "residual method" (e.g., Abramovitz and Solow) have impressed economists with the large measured returns to innovation. This study provides an estimate of the total social cost as well as the return to innovation and thus a more objective criterion for judging the net social value of investment in technological improvements.

Our estimation technique is built upon a basic defect in the usual residual approach to the estimation of the rate of technical change. The defect is based on the obvious empirical fact that some innovations are a product of labor and capital rather than manna from heaven. Given this fact, the aggregate production function admits increasing return to scale (see Thompson (1968) or Starrett), which is ruled out in the conventional residual approach to the estimation of the rate of technological improvement (see, for example, Solow).

The first problem for this study, addressed in Section I, is to derive a technique for estimating the degree of returns to scale and the annual rate of technical progress of the aggregate production function without imposing any

*/ Research support from The National Science Foundation is gratefully acknowledged. Very helpful comments were provided by Professors Kenneth Arrow and John R. Meyer.

artificial restriction on the form of the function. The technique which we develop requires only one assumption in addition to Solow's, viz., the assumption of a zero correlation between the current rate of technical change and the current changes in the quantities of the aggregate factor inputs. The second problem, addressed in Section II, is to derive the magnitude of aggregate expenditures on innovation from the degree of returns to scale. The resulting estimate of innovation costs are then combined with the estimate of returns obtained from the same regression equation to yield an average rate of return to technological innovation.

I. FIRST APPROXIMATION

A. Theory

The following three basic economic relations will be employed,

(1) $Q = F(K,L,t)$

(2) $\beta_K R \equiv F_K$, $\beta_L W \equiv F_L$

(3) $w_K + w_L \equiv 1$,

where Q = output (the numeraire), K = capital, L = labor, t = time, W = wage rate, R = return to capital, $F_L = \frac{\partial Q}{\partial L}$, $w_L = \frac{WL}{Q}$, and $w_K = \frac{RK}{Q}$. Equations (1) and (3) are those used in the classic study of Professor Solow, and (2) is a generalization of his implicit assumption that $\beta_L = \beta_K = 1$. $\frac{1}{\beta_L}$ and $\frac{1}{\beta_K}$ are the portions that the factors receive of their aggregate marginal products.[1/]

A measure of returns to scale may be derived from the first two relations provided that the production function, in a neighborhood of K, L, is homogeneous with respect to the physical inputs. The production function may then be written as

(4) $\beta(t)Q = F_K K + F_L L$,

where $\beta(t)$ is the degree of homogeneity and a measure of the degree of returns to scale at K, L. If $\beta(t) = 1$ for all (K, L), there are constant returns to scale.

1/ The existence of (1) as an aggregate of individual production function satisfying (2) for the case in which $\beta_L = \beta_K = 1$ is assured if and only if the firms possess identical production functions (Fisher, Thompson (1970)). When firms possess different functions, one must generally include each firm's specific factors as separate inputs in order to generate a linear homogeneous aggregate production function. However, since these specific factors are evaluated as part of a firm's capital stock when market value weights are used in computing the aggregate capital index, the only unavoidable bias introduced by having firms with different production functions is an ordinary index number bias. We shall assume throughout that such biases are empirically insignificant.

Now write the identity,

$$F_K K + F_L L \equiv F_K K + F_L L,$$

substitute (2) into the left side, and multiply by $\frac{Q}{Q}$ to obtain

(5) $(\beta_K W_K + \beta_L W_L)Q \equiv F_K K + F_L L.$

Combining (4) and (5), we have, as a measure of returns to scale at K, L,

(6) $\beta(t) = \beta_K W_K + \beta_L W_L.$

Differentiating (1) with respect to time, dividing the resulting equation by Q, and making the appropriate substitutions from (3), we obtain

(7) $\frac{\dot{Q}}{Q} = \beta_K(w_K)\frac{\dot{K}}{K} + \beta_L(w_L)\frac{\dot{L}}{L} + \alpha_t$,

where $\alpha_t = \frac{\frac{\partial Q}{\partial t}}{Q}$ and $\dot{Q} = \frac{dQ}{dt}$, etc. If $\beta_L = \beta_K = \beta$,

(7a) $\frac{\dot{Q}}{Q} = \beta\left[(w_K)\frac{\dot{K}}{K} + (w_L)\frac{\dot{L}}{L}\right] + \alpha_t,$

where β is the degree of homogeneity or returns to scale at K, L, and α_t is the percentage shift in the aggregate production function.2/

B. Estimation

The natural thing to do at this point is to estimate β_L, β_K, and α from a group of observations on $\frac{\dot{Q}}{Q}$, $(w_K)\frac{\dot{K}}{K}$, and $(w_L)\frac{\dot{L}}{L}$. Required for the use of maximum likelihood methods are the specific observations and specifications on the

---

2/ Note that if β_L and β_K are different and constant, a production function which is also homogeneous of order β for all (K,L) must be a Cobb-Douglas function. To see this, observe from equation (6) that when β_L and β_K and β are all constant, then (recalling equation (3)) w_L and w_K are constant. Next, integrate (7) to obtain

$$Q = A(t)L^{(\beta_L w_L)}K^{(\beta_K w_K)}.$$

nature of β_L, β_K, α, and error term. In order to indicate the relative generality of the techniques used here, we now confine ourselves, with one exception, to assumptions which are no more restrictive than those employed by Professor Solow.

We shall use the same variables and observations that he uses. They are time series of w_K, K, L, and Q for the private, non-farm economy of the U.S., 1909-1949.[3/] Introducing Solow's assumption of a Hicks neutral production function and an additional assumption that there is no correlation between changes in factor inputs and the rate of technical progress, $(w_K)\frac{\dot{K}}{K}$ and $(w_L)\frac{\dot{L}}{L}$ are not dependent upon α_t.[4/] We shall first assume that $\beta_L = \beta_K = \beta$, where β is a constant. We are now prepared to fit (7a) with the time series relation,

$$(8) \quad \frac{\dot{Q}}{Q} = \beta(w_K \cdot \frac{\dot{K}}{K} + w_L \cdot \frac{\dot{L}}{L}) + \alpha + u_t$$

where u_t has a zero mean and is uncorrelated with the independent variable.

3/ 1919 and 1920 posed a problem in that the technical change index under constant returns was an extreme negative number in 1919 and an extreme positive number in 1920. We attribute this to a gestation lag, a lag in converting wartime to peacetime capital and correspondingly use average 1919 and 1920 input and output expansions for each of these years in order to remove this problem in the data. Our use of multi-year time intervals below will help reduce this bias for other years.

1945 had a similar problem with negative technical change occurring within the year, but 1946 did not have a significantly positive residual to offset it. We did our study with and without 1945 and got economically equivalent results either way. Our results reported below will all include 1945.

4/ To see this, differentiate (1), divide the equation by F, integrate the partial differential equation assuming that K and L are independent of α, and obtain $Q = A(t)f(K,L)$. Next, partially differentiate this with respect to K or L, multiply by $\frac{dK}{dt}$ or $\frac{dL}{dt}$, and observe that $\frac{F_K \frac{dK}{dt}}{F}, \frac{F_L \frac{dL}{dt}}{F} = \beta_K w_K(\frac{\dot{K}}{K}), \beta_L w_L(\frac{\dot{L}}{L})$ are not dependent upon $\frac{\dot{A}(t)}{A(t)}$ as long as $\frac{dK}{dt}$ and $\frac{dL}{dt}$ are not dependent upon $\frac{\dot{A}(t)}{A(t)}$.

Least square estimates of β and α from (8) are $\hat{\beta} = 1.13$ ($\hat{\sigma}_{\hat{\beta}} = 0.07$) and $\hat{\alpha} = 1.42$. The Durbin-Watson statistic is 1.97, indicating no significant time trend or autocorrelation in the residuals.

Dropping the $\beta_L = \beta_K$ assumption and estimating equation (7), $\hat{\beta}_L$ and $\hat{\beta}_K$ were significantly different than one another. We also ran the regressions on (7a) and (7) using longer time intervals to reduce the impact of measurement errors, using two, four, and five year intervals.

The average $(\hat{\beta}_L, \hat{\beta}_K)$ over these four sets of estimates was (1.43, 0.84). While there was a great deal of variation in the four different estimates of both β_L and β_K, there was relatively little variation in the four different estimates of the <u>sum</u>, $\beta_K\bar{w}_K + \beta_L\bar{w}_L$. In all four cases, this estimate differed by less than 0.02 from the corresponding estimate of returns to scale generated by the assumption that $\beta_L = \beta_K$.

This pattern indicates that there is little stability in the relationship between marginal products and factor prices for particular factors relative to the relationship between the marginal product of the overall input index and the overall factor cost (i.e., the degree of aggregate returns to scale). This peculiarity is explained in Thompson (1970), where it is shown that the presence of non-pecuniary rewards prevents us from identifying the marginal product of separate inputs with their respective factor prices but still allows us to identify the marginal product of the overall input index with its factor cost under neoclassically competitive conditions. The large variations in the (β_L, β_K) estimate may therefore be due to variations in the extent to which additional inputs provide non-pecuniary rewards to owners of other inputs.

In any case, the returns to scale estimates increased gradually as the time intervals expanded until $\hat{\beta} = 1.22$ ($\hat{\alpha} = 1.27$) for five year intervals. This pattern indicates the familiar, downward least-squares bias due to measurement errors present in the annual observations on the independent variable.

There being little a priori information on the constancy of the degree of returns to scale, we computed returns to scale estimates using (7a) within each of the successive five-year intervals and found significantly different estimates of the degrees of returns to scale in the various five-year intervals, with a definite trend toward lower returns to scale in the later years. A revised estimation technique, based on an economic rationalization of the preceeding pattern of estimates, will be developed in the following section.

II. ECONOMIC RATIONALIZATION AND SECOND APPROXIMATION

A. Theory

We assumed above that $\beta(t)$ was possibly different than unity without providing an economic justification. The standard competitive model with technical change which is manna from heaven will justify $\beta = 1$, but we have not specified what model justifies $\beta > 1$.

In the justification which follows, which implements the general outline of Thompson (1965) or Starrett (1975), we allow for the fact that some resources are devoted to the production of innovations rather than the production of current output, Q. We let I_Q and I_t represent the respective amounts of aggregate input devoted to the production of current output and innovations. We thus write a subaggregate production function:

$$(9) \quad Q = f(I_Q, T_t),$$

where T_t represents the technology at time t. We cannot directly observe I_Q or I_t. We can only observe their sum,

$$(10) \quad I = I_Q + I_t.$$

The subaggregate production function can be assumed to be linearly homogeneous in I_Q by the usual duplication argument which is the basis of the standard technological argument for constant returns to scale. Thus, by Euler's Theorem,

$$(11) \quad Q = \frac{\partial f}{\partial I_Q} \cdot I_Q.$$

Differentiating (9), and then dividing by Q,

$$(12) \quad \dot{Q} = \frac{\partial f}{\partial I_Q} \cdot \dot{I}_Q + \frac{\partial f}{\partial T_t} \cdot \dot{T}_t \text{ and } \frac{\dot{Q}}{Q} = \frac{\partial f}{\partial I_Q} \cdot \frac{I_Q}{Q} \cdot \frac{\dot{I}_Q}{I_Q} + \frac{\partial f}{\partial T_t} \quad \frac{\dot{T}_t}{Q}.$$

Using (11),

$$(13) \quad \frac{\dot{Q}}{Q} = \frac{\dot{I}_Q}{I_Q} + \frac{\partial f}{\partial T_t} \frac{\dot{T}_t}{Q}.$$

To simplify the second term in (13), we now make a strong assumption, which we will empirically verify later in this section, regarding both the subaggregate production function and the production function for current innovations, or

(14) $\dot{T}_t = g(T_t, I_t, I_{t-1}, I_{t-2}, \ldots)$.

In particular, we assume that these functions are super-Hicks-neutral in the following sense:

(15) $Q = f(I_Q, T_t) = T_t f^*(I_Q)$, and

(16) $\dot{T}_t = g(T_t, I_{T_t}, I_{T_{t-1}}, I_{T_{t-2}}, \ldots) = T_t g^*(I_{T_t}, I_{T_{t-1}}, \ldots)$.

Using equations (15) and (16), equation (13) can be written as

(17) $\frac{\dot{Q}}{Q} = \frac{\dot{I}_Q}{I_Q} + g^*(I_{T_t}, I_{T_{t-1}}, \ldots)$

With only simple Hicks neutrality, g^* would have a coefficient which is a function of T_t. In general, that is without Hicks neutrality, g^* would have a coefficient which depends on both T_t and I_t.

Our first approximation indicated that the rate of annual technical change is roughly constant over time. Temporarily adopting this assumption without empirical justification, (13) and (17) indicate that the amount of inputs devoted to the production of innovation, I_{T_t}, can be assumed to be constant over time. Then, using (10), $\dot{I}_Q = \dot{I}$ so that (17) can be rewritten as

(18) $\frac{\dot{Q}}{Q} = \beta_t \frac{\dot{I}}{I} + \alpha_t$, where

(19) $\beta_t = \frac{I}{I_Q}$.

Equations (18) and (19) give us an economic interpretation of the $\beta(t)$ appearing in Section I. For the estimates using annual data, where $\hat{\beta} = 1.13$, the implied average proportion of inputs going to technical innovation, $\frac{I_T}{I} = \frac{I - I_Q}{I} = 1 - \frac{I_Q}{I} = 1 - \frac{1}{\beta_t}$, is about 11 1/2 percent. However, since β_t is not constant over time while the β of Section I was assumed constant, this interpretation of the $\hat{\beta}$ of Section I is of a highly approximate nature. Since I generally grows over time at a rate which is slower than I_Q if I_T remains constant, equations (18) and (19) explain why the short-period degree of returns to scale estimates in the first approximation decreased over time.

Moreover, equations (18) and (19) yield a guide as to what form of linear regression equation to estimate using the entire time series. In particular, since

$$(20) \quad \beta_t = \frac{I}{I - I_T} = \frac{1}{1 - \frac{I_T}{I}} = 1 + \frac{I_T}{I} + \left(\frac{I_T}{I}\right)^2 + \dots,$$

and $\left(\frac{I_T}{I}\right)^r$ for $r \geq 2$ is economically insignificant for all reasonable values of $\frac{I_T}{I}$,

$$(21) \quad \beta_t \approx 1 + \frac{I_T}{I}.$$

Using (18), (19), and (21),

$$(22) \quad \frac{\dot{Q}}{Q} - \frac{\dot{I}}{I} \approx I_T\left(\frac{\dot{I}}{I^2}\right) + \alpha_t$$

B. <u>Estimation</u>

Since I_T is constant, a linear regression of $\frac{\dot{Q}}{Q} - \frac{\dot{I}}{I}$ on $\frac{\dot{I}}{I^2}$ is justified by the above theory. We ran a least-squares regression on the time series used in Section I. We again found a coefficient, 0.16, almost two standard deviations above unity and an insignificant Durbin Watson statistic (D.W. = 1.96). The estimated rate of average annual technical change was 1.43% per year (compared to the 1.42% for the regression of y on x). The implied estimate of the average degree of returns to scale, $\frac{\bar{I}}{\bar{I} - \hat{I}_T}$, is 1.12 (compared to 1.13 for the regression of y on x.) I is indexed to equal unity in 1909 and $\bar{I} = 1.29$. Thus the implied estimate of the average proportion of all inputs devoted to technical change, $\frac{\hat{I}_T}{\bar{I}}$, is 12.4% per year. That is, an annual investment averaging 12.4% of foregone output yielded an annual increment of 1.43% to our gross national product.

To test the assumption of super-Hicks-Neutral technical change, we estimated equation (22) with a quadratic time trend and again with a quadratic function of I as extra variables. All the coefficients on these variables were far from statistically significant and had no effect whatsoever on our estimates of returns to scale or $\hat{I}_T$.

We again adjusted our annual data for observation error and gestation lags by using longer time intervals, averaging over consecutive sets of years rather than accepting annual data. Again the values of the regression coefficient, as well as the values of R^2, increased with the length of the intervals. Our longest interval was again five years. (We felt that using longer intervals would be throwing away too much data for a small additional reduction in observation error.) Our five year interval estimates of I_T and α were 0.24 and .127 respectively, with the standard errors .14 and .005 respectively and D.W. = 1.77. These imply an estimate of the average fraction of inputs devoted to technical innovation, $\frac{\hat{I}}{\bar{I}}$, of 18.5%. The corresponding estimate using the first approximation and five-year intervals, $1 - \frac{1}{\hat{\beta}}$, is 18%. And the average rate of technical change was also 1.27%.

C. Rate of Return

The rate of return to technical innovation is computed in the following fashion: The cost of 1909's expenditure and innovation in the above model was .24, or 24% of 1909's inputs. Assuming only a one-year lag in the implementation of new technologies, the return in 1910 to the investment in 1909 was .0127 Q_{1909} = .0127. The returns in 1911 was .0127 Q_{1910}, etc.

The return thus increases over time when Q_t grows over time so that corresponding social discount rate applied to the original return of .0127 must be reduced by the growth rate of output to arrive at a correct calculation of the social value of the investment. Alternatively, we may merely add the growth rate of output to the initial rate of return to investment to produce a rate of return which should be compared to the social discount rate. Adopting the latter approach, our estimate of the rate of return to innovation in 1909 is $\frac{.0127}{.24} + \frac{\dot{Q}}{Q} = \frac{.0127}{.24} + .032 \approx .085$. From the above discussion, the return to a given investment in innovation generally increases with the date of the investment because later innovation efforts generally occur at greater levels of outputs. On the other hand, the alternative value of the inputs devoted to the investment rises with the rate of technical progress. Thus, given that the average output for the entire period is 1.64 times the 1909 output, and the average technology level is 1.25 of the 1909 level, the average rate of return to investment in innovation at the average levels of output and technology for the period is $\left(\frac{.0127}{.24}\right)\left(\frac{1.64}{1.25}\right) + .032 = 10.1\%$.

III. CONCLUSION

Our final estimate of the average percentage of inputs devoted to innovation during the 1909-1949 period is 18.5%, and the corresponding annual return to these investments is 1.27% of current output.

Our final estimate of the average of the average rates of return to the investments in innovation over the period is 10.1%. This estimate, however, is overly generous for at least two reasons. First, it assumes an unrealistically short lag between investment in an innovation and its return. Second, it is obtained by only partially adjusting for errors of measurement.

Our interpretation of the results is that while it is likely that there has been a significant average rate of return to investments in innovation, extravagant claims of a great net social value to innovation are difficult to justify, at least for the U.S., 1909-1949.

References

1. Abramovitz, M., "Resource and Output Trends in the U.S. since 1870," American Economic Review, Papers and Proceedings, May 1956, 5-23.
2. Fisher, F., "The Existence of Aggregate Production Functions," Econometrica, October 1969, 553-57.
3. Solow, R. M., "Technical Progress and the Aggregate Production Function," Review of Economics and Statistics, August 1957, 312-320.
4. Starrett, D., "Measuring Returns to Scale in the Aggregate, and the Scale Effect of Public Goods," Technical Report #163, March 1975, Institute for Mathematical Studies in the Social Sciences, Stanford University, p. 27.
5. Thompson, E. A., "The Perfectly Competitive Allocation of Public Goods," MR 49, September 1965, Institute of Government and Public Affairs, UCLA, p. 60.
6. Thompson, E. A., "The Perfectly Competitive Production of Collective Goods," Review of Economics and Statistics, 1968, 1-12.
7. Thompson, E. A., "Non Pecuniary Rewards and the Aggregate Production Function," Review of Economics and Statistics, November 1970, 395-404.